Conformal Geometry and Dynamics of Quadratic Polynomials, vol I-II

Mikhail Lyubich

Contents

Preface

In the last quarter of XXth century the *complex and real quadratic family*

$$f_c : z \mapsto z^2 + c$$

was recognized as a very rich and representative model of chaotic dynamics. In the complex plane it exhibits fractal sets of amazing beauty. On the real line, it contains regular and stochastic maps intertwined in an intricate fashion. It also has remarkable universality properties: you can see small pieces that look exactly the same as the whole family. This *Universality* is related to a profound *Renormalization* idea originated in the particle and statistical physics. Interplay between *real* and *complex* worlds provides us with deep insights into both. These ideas eventually led to a complete picture of dynamics in the real quadratic family and a nearly complete picture in the complex family.

In this series of books we attempt to present this picture beginning from scratch and supplying all needed background (beyond the basic graduate education). We hope to fit it into four volumes dedicated to the following themes:

I: Background in Conformal and Quasiconformal Geometry;

II: Basic Holomorphic Dynamics;

III: Complete Picture in the Real Quadratic Family;

IV: Advances in the MLC;

Several prominent ideas that will be highlighted throughout the book are rigidity, puzzle, combinatorial models, topological & measure-theoretic attractors, geometric bounds, and renormalization. Let us overview them in more detail.

Conformal and quasiconformal geometry. Volume I of the book contains a necessary background in conformal and quasiconformal geometry. Main analytical and topological tools of Holomorphic Dynamics are collected here in the form suitable for dynamical applications. Classical themes include principles of hyperbolic metric and extremal length, the classical Uniformization Theorem, Measurable Riemann Mapping Theorem (including holomorphic dependence on parameters), and the Carathéodory boundary theory. More contemporary themes include a general (non-dynamical) introduction to the theory of geodesic laminations, puzzles, thin-thick decomposition for bordered Riemann surfaces, holomorphic motions (which probably provides the biggest feedback from the contemporary Holomorphic Dynamics to Analysis), and elements of Teichmüller theory.

Dynamics. Volume II is dedicated to the Basic Holomorphic Dynamics developed from the mid-XIXth century through the early 1990s (adapted for the quadratic family).

Dynamical plane I: basic objects. Chapter 3 covers most of the classical local theory and Fatou-Julia global theory: basic properties of the Fatou and Julia sets, classification of periodic motions and the associated remarkable functional equations. (These equations were one of the original motivations for the classical theory). It is completed with a more contemporary material on the landing properties of rational external rays (preparing a foundation for the combinatorial theory of Julia sets) and the Yoccoz Inequality for the multipliers of periodic points.

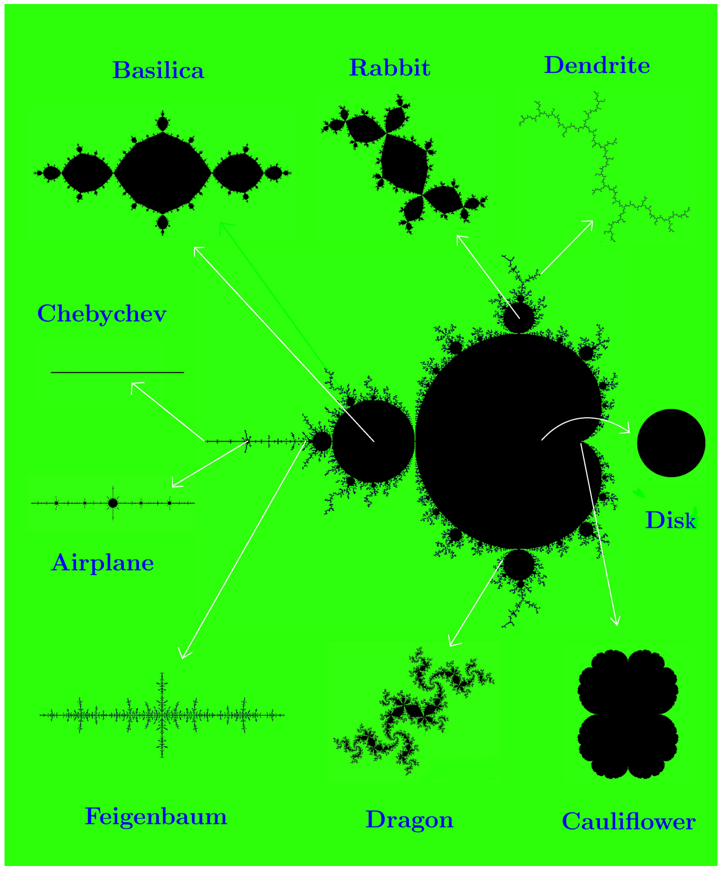

FIGURE 0.1. Mandelbrot set $\mathcal{M}$. It encodes in one picture all beauty and subtlety of the complex quadratic family. Every point $c \in \mathcal{M}$ represents some Julia set $\mathcal{J}(f_c)$. A handful of popular ones is depicted.

Dynamical plane II: fine structures and models. Chapter 4 covers good part of the dynamical theory developed in the 1980s. Central themes here are:

• Global dynamics of important special classes of maps: hyperbolic, parabolic, and postcritically non-recurrent.

• Problem of local connectivity of Julia sets ("JLC Problem") and building of their precise topological models (Douady & Hubbard and Thurston).

• Idea of quadratic-like maps and their renormalizations. This theory was designed by Douady & Hubbard as a tool of explaining presence of baby Mandelbrot copies inside the Mandelbrot set. It became foundational for the Complex Renormalization Theory which will be a central theme in the upcoming volumes.

• Sullivan's No Wandering Domains Theorem, which completed description of the dynamics on the Fatou set. The proof is based on the method of quasiconformal deformations adapted by Sullivan from the Ahlfors-Bers Deformation Theory for Kleinian groups. It supplied the first line in Sullivan's Dictionary between the dynamics of rational maps and Kleinian groups, which was largely responsible (along

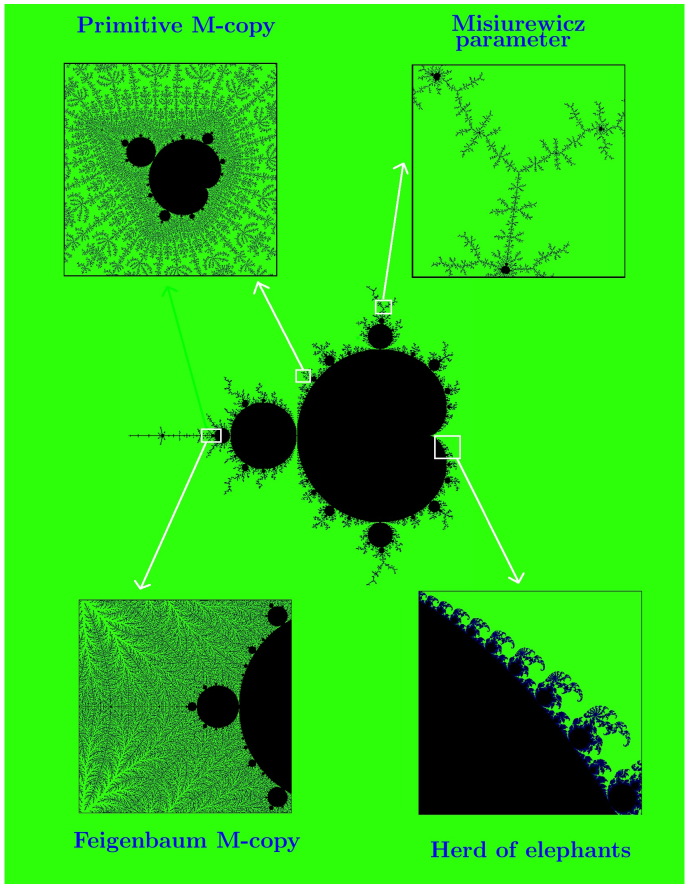

FIGURE 0.2. Blow-ups of the Mandelbrot set.

with appearance of beautiful computer images of Julia sets and the Mandelbrot set) for the spectacular revival of Holomorphic Dynamics after 60 years of stagnation.

• Topological structure of real quadratic maps. The topological exploration of interval maps originated in the work of Sharkovskii in the 1960s. In the 1970s a great interest to this area was sparked by the Milnor-Thurston Kneading Theory. The topological/combinatorial theory was completed in the 1980s, due to the effort of many researchers. Guchenheimer's No Wandering Intervals Theorem and real *a priori* bounds for solenoidal maps (due to Guckenheimer, Blokh-Lyubich and Sullivan) are key geometric ingredients needed for this picture.

• Combinatorial theory of Yoccoz Puzzle, a powerful tool of contemporary Holomorphic Dynamics, followed by a discussion of various combinatorial models for Julia sets.

Parameter plane. In Chapter 5 we pass to the parameter plane, introducing the Mandelbrot set $\mathcal{M}$ which encodes in one picture the whole richness of the quadratic family. After analyzing elementary properties of $\mathcal{M}$, we prove first two breakthrough results about it from the early 1980s: the Connectivity and the Multiplier Theorems (due to Douady and Hubbard). A new remarkable tool, Quasiconformal

Surgery, was introduced to the field along the lines. Then we proceed with the following themes:

• Structural Stability Theory (by Mañé-Sad-Sullivan and the author) and a quasiconformal classification of quadratic polynomials.

• Limb Decomposition of the Mandelbrot set. (Limbs are the pieces of $\mathcal{M}$ attached to the main cardioid (clearly visible on the pictures) and to other hyperbolic components of $\mathcal{M}$.) It implies, in particular, that any abstract superattracting Hubbard tree (which encodes certain combinatorial data) is realizable by some superattracting parameter.

• Proof of the Milnor-Thurston Entropy Monotonicity Conjecture that gives the first illustration of the power of complex methods in real dynamics.

• Discussion of central conjectures in this area and the interplay between them. Most famous conjecture is known as MLC (local connectivity of the Mandelbrot set). It is equivalent to the Combinatorial Rigidity Conjecture, which is very similar in spirit to the Mostow-Thurston Rigidity Phenomenon in 3D hyperbolic geometry. In turn, these conjectures imply the Fatou Conjecture asserting that the set of hyperbolic maps is dense in the quadratic family (which sounds particularly prominent for the dynamics community). The real counterparts of these conjectures were established in the 1990s: they are formulated in this section, but the proofs are postponed until Vol III (except that we give a proof of Rigidity for real Feigenbaum maps).

• A fundamental Thurston's Realization Theorem (in the context of superattracting quadratic polynomials). It allows one to realize (in an appropriate sense) any topological self-map of S^2 that "looks like a superattracting quadratic polynomial" as an actual quadratic polynomial. (An equivalent version of this result, in terms of Hubbard trees, was mentioned above.)

Straightening, puzzle geometry, and attractors. Let us pass to Chapter 6, the final chapter of the second volume.

One of the most fascinating features of the Mandelbrot set, clearly observed on computer pictures, is the presence of the little copies of itself ("baby M-sets"), which look almost identical to the original set (except for possible absence of the main cusp). The complex renormalization theory is designed to explain this phenomenon. We present the Douady-Hubbard theory of quadratic-like maps and complex renormalization that justifies presence of the baby M-sets, and classify them. (The geometric theory that explains why these babies have a universal shape will be developed in the forthcoming volumes.) Note that though this theory is widely known and used, it has never appeared in a complete form (to the best of our knowledge).

Other themes covered in this chapter are:

• A proof of Yoccoz's Theorem on local connectivity of the Julia sets for at most finitely renormalizable maps with all periodic points repelling. A proof that these Julia sets have zero area (due to Shishikura and the author).

• Measurable Dynamics of real quadratic polynomials developed by Blokh and the author in the mid-1980s. The main outcome is that such a map has a unique measure-theoretic attractor that attracts almost all orbits, and this attractor can be of four possible types: an attracting or parabolic cycle, a cycle of intervals, a

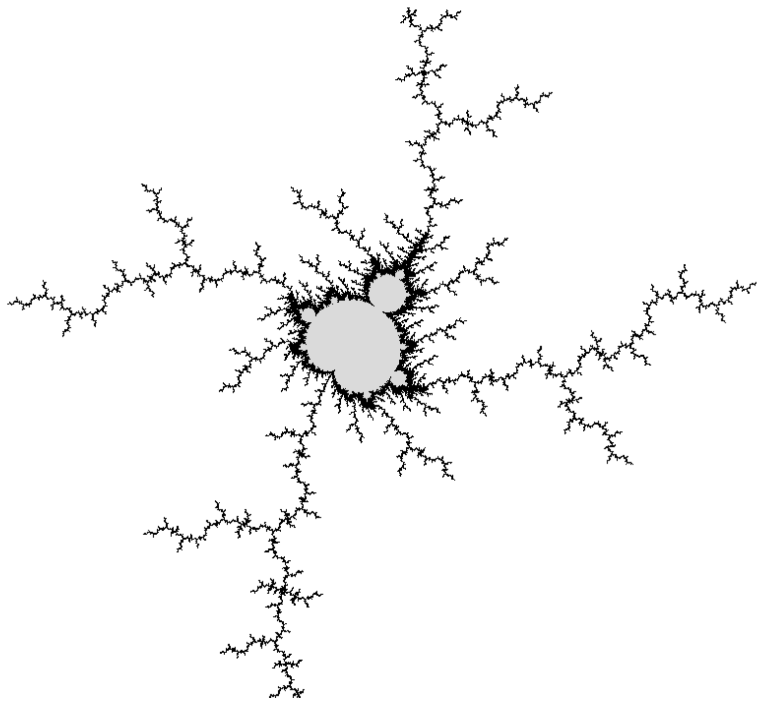

FIGURE 0.3. Baby M-set.

solenoid, or a wild Cantor attractor. (The problem of existence of the latter will be discussed in the third volume).

In conclusion, we discuss properties of stochastic maps (i.e., maps that have an absolutely continuous invariant measure (*acim*)), and give Misiurewicz's condition for stochasticity.

• Combinatorial Parapuzzle Theory. It provides us with a hierarchical structure of the Mandelbrot set, leading to the partition of it into hyperbolic components (accompanied with their boundaries), Yoccoz parameters, and infinitely renormalizable parameters. This sets the stage for further advances in the MLC Conjecture (in the forthcoming volumes).

• We conclude this chapter by completing a proof of local connectivity of $\mathcal{M}$ at the boundaries of hyperbolic components and describing a topological model for $\mathcal{M}$ (which is homeomorphic to $\mathcal{M}$ as long as the latter is locally connected – this was the original motivation for formulating the MLC Conjecture).

This roughly constitutes the first two volumes of the series.

Projected volumes. In the third volume we plan to prove the Feigenbaum-Coullet-Tresser Renormalization Conjecture (by Sullivan [**S3**], McMullen [**McM2**], and the author [**L12**]), Density of Hyperbolic Maps in the real quadratic family [**L10, GS**]), and the Regular and Stochastic Theorem [**L11, L13**] asserting that almost any real quadratic map is either regular (i.e., has an attracting cycle that attracts almost all orbits) or stochastic (i.e., it has an acim that governs behavior of almost all orbits). These results were obtained in 1990's, but more recent insights,

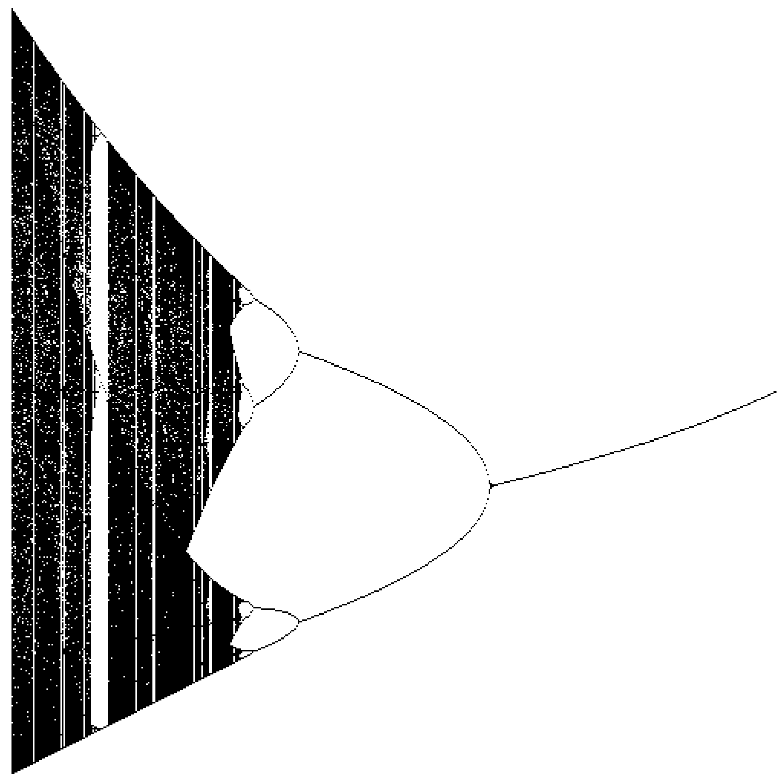

FIGURE 0.4. Real quadratic family $f_c : x \mapsto x^2 + c$ as a model of chaos. This picture presents how the limit set of the orbit $\{f_c^n(0)\}_{n=0}^{\infty}$ bifurcates as the parameter c changes from $1/4$ on the right to -2 on the left. Two types of regimes are intertwined in an intricate way. The gaps correspond to the regular regimes. The black regions correspond to the stochastic regimes (though of course there are many narrow invisible gaps therein). In the beginning (on the right) you can see the cascade of doubling bifurcations. This picture became symbolic for one-dimensional dynamics.

particularly by Avila, Kahn, Moreira, and Shen deepened and further advanced our understanding of the phenomena (see [**AKLS, AL1, ALS, AM1, AM2**]).

We plan to dedicate the fourth volume to recent advances in the MLC Conjecture, mostly based on the work of Kahn and the author [**K, KL1**]–[**KL3**].

Interplay between Complex and Real worlds. Throughout this book, Real Dynamics is largely treated as a special $\mathbb{R}$-symmetric case of Complex Dynamics. While this gives an elegant view for the initial real theory (collected in vol II), most of it can still be developed by purely real methods. This will not be the case anymore for more advanced real theory that will be developed in vol III: most of it will rely on complex methods in a crucial way.

Missing themes. Let us mention several important themes of Quadratic Dynamics that are not covered by this project.

• Advanced topics of Neutral Dynamics: structure of Siegel and Cremer maps, Parabolic and Siegel Renormalization Theories, see section "Neutral Phenomenon" in the bibliography.

• Holomorphic Ergodic Theory: see selected references in the corresponding section of bibliography, in particular, the book [**PrU**].

• Problem of Hausdorff dimension and Lebesgue measure of Julia sets and the Mandelbrot set, see [**Sh2, BC, AL3**].

However, our project may eventually expand in one of these directions...

Remarks on more general theories. Dynamics of quadratic polynomials is quintessential for the one-dimensional dynamics, both complex and real. With luxury of being globally holomorphic, it raises in the simplest (albeit, already highly nontrivial) combinatorial setting, some of the deepest geometric problems. Much of the further theory is modeled on this setting, though new very interesting phenomena eventually emerge. Let us mention some of these developments (see Notes in the main bulk of the book for further leads):

Much of the dynamical theory developed in Chapters 3–4 can be generalized to polynomials of higher degree. In fact, some of it is needed even in the quadratic case, as the iterates of f are higher degree polynomials. We formulate relevant pieces of the theory as exercises.

Note, however, that the basic parameter theory (Chapter 5) is less amenable to generalizations, since the parameter spaces of higher degree polynomials are higher dimensional.

Even more generally, good part of the basic theory can be generalized to rational maps, though the combinatorial theory is much less developed in this generality. We touch on this theme only briefly (for instance, we need Blaschke products). There is a plenty of contemporary introductory sources to this area: [**Be1, Bl, CG, EL1, L1, M2, St**] (with Milnor's book [**M2**] being particularly popular).

Going further, one can study the dynamics of transcendental maps, but it becomes technically very difficult quite fast. Nevertheless, it is a flourishing research area, with an important feedback to the polynomial dynamics (see [**EL1, Ber**] for introductory surveys and section "Transcendental Dynamics" in the bibliography for further references).

In a different direction, one can also generalize much of the real theory to smooth maps. In fact, we develop the theory for a class of real analytic maps which is needed for the quadratic dynamics, but otherwise we do not pursue this direction in depth. An interested reader can consult text books by Collet & Eckmann [**CE**] and de Melo & van Strien [**MvS**].

Let us finally mention that various deep geometric issues (including local connectivity and area problems, existence of wild attractors, density of hyperbolicity, and regular or stochastic dichotomy) are degree-sensitive even in the unicritical case $z \mapsto z^d + c$. However, our exposition in vol III is planned so that it can be easily adapted for the general unicritical setting.

A major development in the general multimodal case was a proof in the 2000s of Density of Hyperbolicity by Kozlovski, Shen and van Strien [**KSS**]. However, the multimodal generalization of Regular or Stochastic Dichotomy still remains unsettled.

How the book can be used. This book is being designed as an "educational monograph" in contemporary analytic one-dimensional dynamics, so it can be used in many ways:

• As the first introduction to the analytic one-dimensional dynamics, complex and real. Then the reader should begin with Chapter 3 consulting the background material from vol I as needed.

• As an introduction to advanced themes of contemporary one-dimensional dynamics for the reader who knows basics and intends to do research in this field. Such a reader can go through selected pieces of Chapter 3 proceeding fairly fast to more advanced topics.

• For a graduate class in topics of conformal and quasiconformal geometry illustrated with dynamical examples. This would cover vol I with selected pieces from vol II.

• Of course, the book can also be used for reference.

As we have already mentioned, we have made an effort to collect all necessary background in the user-friendly form. What we assume from the reader is just the basic knowledge of real & complex analysis, and topology & geometry, roughly corresponding to the core graduate curriculum at a US University. For instance, the following collection of text books covers most of the needed background: Munkres [**Mu1**] (Topology), Shabat [**Shab**] (Complex Analysis), Do Carmo [**DoC**] (Differential Geometry), Halmos [**Ha1**] (Measure Theory), Kolmogorov and Fomin [**KolF**] (Real Functional Analysis), and Spivak [**Spiv**, vol. 1] (Global Analysis).

Various remarks. a) The text is supplied with many "Exercises" and "Projects". Mostly, they constitute an intrinsic part of the discussion, an invitation to the reader to think through some technical details or to develop a piece of the theory him/herself (in the spirit of the book by I.M. Glazman and Yu.I. Lyubich [**GL**] that develops an advanced theory as a carefully organized series of not so difficult problems). More challenging exercises are called "Problems". Open problems and conjectures are marked as such.

e) There are very few references in the main bulk of the text. However, each chapter (and some sections) are concluded with *Notes* supplying historical background, references, generalizations, and leads for further reading.

f) For the reader's convenience, a list of notations (as well as some basic definitions) is provided at the end. Most of these notations and definitions are also introduced in the main bulk of the book, but there are some exceptions.

g) The bibliography is roughly classified according to the topics: "Real One-Dimensional Dynamics", "General Holomorphic Dynamics", etc., so it may take a few extra seconds to scroll through several sections in order to find a desired reference. We hope this inconvenience would be compensated by an extra orientation that such classification provides.

h) The numeration of the chapters and sections is uniform throughout the series of volumes.

Acknowledgement. This project has been gradually developed since the European Lecture Series I gave in Copenhagen–St Petersburg–Barcelona in 1999. Besides these lectures, it is based upon my graduate classes in Stony Brook and Toronto since the early 1990s, as well as mini-courses in Trieste, Kyoto, and Cullera in the 2000s. Over all these years, I have received invaluable feedback from my colleagues, postdocs and students, and more recently, from anonymous referees. Some specific acknowledgments are spread over the *Notes* throughout the book, but of course, they are far from being complete. Special thanks a due to Sergei Gelfand for his patient but persistent encouragement over the years.

I am grateful to Joshua Bowman and especially to Sabyasachi Mukherjee for their help with computer pictures. I also thank the NSF and NSERC, as well as the Sloan and Guggenheim Fellowships, for their support over various periods during this time.

An abbreviated version of the book, available as "Six Lectures on dynamics of quadratic polynomials " [**L3**], reflects the status of the area through the 1990s. It is summarized in the survey "The quadratic family as a qualitatively solvable model of chaos" [**L2**]. Recent surveys "Forty years of unimodal dynamics: on the occasion of Artur Avila winning the Brin prize " [**L4**] and "Analytic low dimensional dynamics: from dimension one to two" [**L5**] give an updated overview of the area.

1. Preliminaries: Topological background

In this section we collect some preliminary material, mostly topological. It can be reviewed briefly and then consulted as the corresponding objects and results appear in the text.

In what follows, all topological spaces (except Banach spaces $L^\infty(X)$) are assumed to satisfy the *Second Countability Axiom*, i.e., they have a countable basis of open sets. We also assume that all topological spaces in question are *metrizable*, unless otherwise is explicitly said. Recall that a compact space is metrizable iff it satisfies the Second Countability Axiom (and iff it is separable), so in the compact case our two conventions exactly match. We will also follow a convention that *compactness* includes being Hausdorff.

1.1. First encounter with "wild" creatures.

1.1.1. *Cantor set.* The $(1/3)$-*Cantor set* was perhaps the first example of a "wild fractal" object. This famous construction goes as follows. By removing from the unit interval $\mathbb{I} \equiv I^0$ the middle open 1/3-subinterval, we are left with the union of two closed intervals, $I^1_0 = [0, 1/3]$ and $I^1_1 = [1/3, 1]$. By removing from each of them the middle $(1/3)$-interval, we are left with the union of four closed intervals $I^2_{00} = [0, 1/9]$, $I^1_{01} = [2/9, 1/3]$, $I^2_{10} = [2/3, 5/9]$, and $I^2_{11} = [8/9, 1]$. Repeating this procedure over and over again, we obtain a hierarchy of 2^n intervals $I^n_{\varepsilon_1 \dots \varepsilon_n}$, $\varepsilon_n \in \{0, 1\}$, of level n, naturally labeled by dyadic sequences of length n. Let

$$\mathbb{I}^n := \bigcup_{\varepsilon_1 \dots \varepsilon_n} I^n_{\varepsilon_1 \dots \varepsilon_n} \quad \text{and} \quad K_{1/3} := \bigcap_{n=1}^{\infty} \mathbb{I}^n.$$

EXERCISE 1.1. *The* $(1/3)$-*Cantor set is a perfect totally disconnected set of zero length.*

This prompts an intrinsic definition of a *Cantor set* as a totally disconnected perfect set.

EXERCISE 1.2. *All Cantor sets are homeomorphic.*

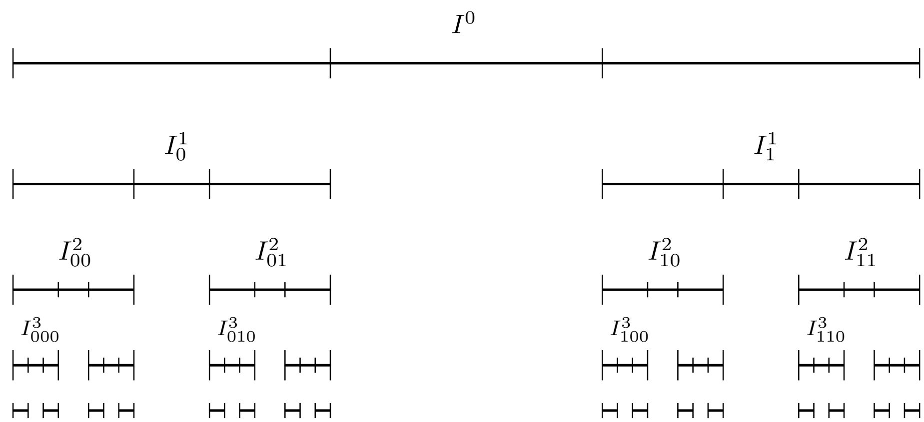

FIGURE 1.1. Hierarchy of intervals generating the (1/3)-Cantor set.

One can generalize the classical Cantor construction in the following obvious way. Consider a rooted tree $\mathcal{T}$ with the root v^0 and vertices of level $n \geq 1$ labeled as $v^n_{\varepsilon_1\ldots\varepsilon_n}$, where each index ε_n runs through a finite set, and any two vertices $v^n_{\varepsilon_1\ldots\varepsilon_n}$ & $v^{n+1}_{\varepsilon_1\ldots\varepsilon_n\varepsilon_{n+1}}$ are connected by and edge.

REMARK 1.3. For $n = 0$, this statement should be read as "v^0 is connected to $v^1_{\varepsilon_1}$". Such a convention will be assumed without mentioning under similar circumstances throughout the book.

Let us associate to $\mathcal{T}$ a hierarchical family of closed intervals $I^n_{\varepsilon_1\ldots\varepsilon_n} \subset I^0$ such that the intervals of the same level are disjoint, while $I^n_{\varepsilon_1,\ldots,\varepsilon_n} \supset I^n_{\varepsilon_1,\ldots,\varepsilon_n,\varepsilon_{n+1}}$. It is also convenient to put an order on the range of each index ε_n, and to assume that for given $(\varepsilon_1 \ldots \varepsilon_{n-1})$, the correspondence $\varepsilon_n \mapsto I^n_{\varepsilon_1\ldots\varepsilon_n}$ is monotonic.

Assume that

$$\operatorname{diam} I^n_{\varepsilon_1,\ldots,\varepsilon_n} \to 0 \text{ as } n \to \infty$$

along any infinite branch $(\varepsilon_1\varepsilon_2\varepsilon_3\ldots)$ of the tree. Then

$$K := \bigcap_{n=0}^{\infty} \bigcup_{\varepsilon_1\ldots\varepsilon_n} I^n_{\varepsilon_1\ldots\varepsilon_n}$$

EXERCISE 1.4. *Show that under the above circumstances, K is a Cantor set.*

The tree $\mathcal{T}$ encodes the *combinatorics* of this Cantors set K.[1] We say that K has an $N-$*bounded combinatorics* if the number of branches emanating from any vertex of $\mathcal{T}$ to the next level is bounded by N.

EXERCISE 1.5. *Any two Cantor sets $K, \tilde{K} \subset \mathbb{R}$ with the same combinatorics $\mathcal{T}$ are topologically equivalent by a homeomorphism $h : (\mathbb{R}, K) \to (\mathbb{R}, \tilde{K})$ respecting the combinatorics, i.e., inducing* id *on the tree $\mathcal{T}$.*

[1]Of course, a given Cantor set K can be encoded by various trees. but in practice K appears together with the coding tree.

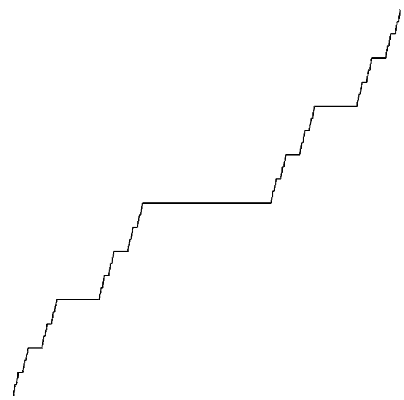

FIGURE 1.2. Devil staircase.

We say that K has a *C-bounded geometry* if any interval $I^n_{\varepsilon_1\dots\varepsilon_n}$ is $C-$comparable with all the intervals $I^n_{\varepsilon_1\dots\varepsilon_n\varepsilon_{n+1}}$ of the next leves and with all the gaps in between.[2] For instance, the $1/3-$Cantor set has a $2-$bounded combinatorics and $3-$bounded geometry.

Clearly, a Cantor set with a bounded geomerty has a bounded combinatorics, but not necessarily the other way around. Interplay between combinatorics and geometry for various fractal sets will be one of the main themes of this book.

1.1.2. *Devil's Staircase.* Let $K \subset J$ be a compact subset of a closed topological interval J. Connected components L_j of $J \smallsetminus K$ are called *gaps* in K, or *complementary intervals* of K. If any two gaps have disjoint closures (i.e. K does not have isolated points) then we can consider the equivalence relation $\sim$ on J by declaring the closures $\bar{L}_j$ to be equivalence classes, while other classes to be singletons. Then the quotient $J/\sim$ is a topological interval as well, as the *Devil 's Staircase* construction shows.[3] This is a continuous monotone function $h : J \to I$ onto another interval I whose fibers are the above equivalence classes (so, h is constant on the gaps L_j, while $h(x) \neq h(y)$ if $x \not\sim y$).

The inverse construction is called *blow-up of points*. Given any countable set $\{x_j\} \subset \operatorname{int} I$ and a summable series $\bar{\varepsilon} := \sum \varepsilon_j < \infty$ with $\varepsilon_j > 0$, one can "blow-up" points x_j to closed intervals $\bar{L}_j$ of length ε_j to obtain a new interval J of length $1 + \bar{\varepsilon}$. The natural projection $\pi : J \to I$ is a Devil Staircase. Let $\hat{I} := J \smallsetminus \bigcup L_j$. Then the projection $\pi : \hat{I} \to I$ is one-to-one over all points except the x_j, where it is two-to-one.

EXERCISE 1.6. *Work out details of the Devil's Staircase and blow-up constructions.*

EXERCISE 1.7. *For the* $(1/3)$-*Cantor set* $K \equiv K_{1/3}$,

[2]To avoid dependence on a particular choice of the hierarchy of intervals, we should allow *adjustments* of the intervals $I^n_{\varepsilon_1\dots\varepsilon_n}$ without changing their slices by K.

[3]This term usually applies to the case when K is nowhere dense, i.e., it is a Cantor set.

a) The devil's staircase can be constructed by means of the map

$$K \to \mathbb{I}, \quad \sum \frac{\varepsilon_n}{3^n} \mapsto \sum \frac{\varepsilon_n}{2^n}, \quad \text{where } \varepsilon_n \in \{0,1\}.$$

b) Its graph is affinely self-similar, namely, it is invariant under the transformation

$$\mathbb{I}^2 \to \mathbb{I}^2, quad (x,y) \mapsto (3x, 2x) \mod 1.$$

This devil's staircase provides the standard example of a monotone function whose derivative vanishes almost everywhere (so the Newton-Leibniz Formula fails — no absolute continuity).

The above discussion applies as well to a topological circle in place of an interval. Moreover, it can be also extended to a disk:

Exercise 1.8. *Let $\bar{L}_j$ be a family of disjoint closed arcs on the unit circle $\mathbb{T}$. Consider an equivalence relation $\sim$ on the unit disk $\bar{\mathbb{D}}$ whose classes are $\bar{L}_j$ and singletons. Then the quotient $\bar{\mathbb{D}}/\sim$ is a topological disk.*

1.2. Local connectivity. Notion of local connectivity is crucial for Holomorphic Dynamics: it makes even fractal objects fairly "tame".

1.2.1. *Paths and arcs.* A *path* and a *curve* in a topological space X mean the same: a continuous map γ of an interval (of any type) or a circle to X. In the latter case we also refer to it as a *closed* curve or a *loop*. Abusing terminology, we often refer to the image of γ as a path/curve as well.

An *arc* is an embedding of an interval into X. A *simple closed curve* is the embedding of the circle into X.

Exercise 1.9. *Any path parametrized by a closed interval contains an arc with the same endpoints.*

Thus, *path connectivity* of a space X is equivalent to its *arcwise connectivity*.

1.2.2. *Definition and basic properties.* A topological space X is called *locally connected* ("*lc*") *at a point* $x \in X$ if x has a local basis of connected neighborhoods.[4] A space X is called locally connected if it is locally connected at every point.

Exercise 1.10. *A space X is locally connected iff connected components of any open subset $U \subset X$ are open.*

There is a convenient weaker notion: A space X is called *weakly locally connected at a point* $x \in X$ if any neighborhood $U \ni x$ contains a connected set P such that $x \in \operatorname{int} P$. (Such spaces are also called *connected im kleinen.*) There is a subtle difference between local connectivity and weak local connectivity at an individual point, but fortunately it disappears globally:

Exercise 1.11. *If a space is weakly locally connected* (*at every point*) *then it is locally connected. However, a space can be weakly locally connected at some point x without being locally connected at this point* (see Figure 1.4).

For a metric space X, a *lc modulus* at $x \in X$ is a function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$, $\omega(\varepsilon) \to 0$ as $\varepsilon \to 0$, such that if $d(x,y) < \varepsilon$ then there exists a connected set Y containing both x and y such that $\operatorname{diam} Y < \omega(\varepsilon)$. If ω works for all points $x \in X$ then it is called *lc modulus for X*.

[4]According to our convention, these neighborhoods are open. Sometimes "local connectivity" is defined in terms of closed neighborhoods, which corresponds to the notion of weak lc below.

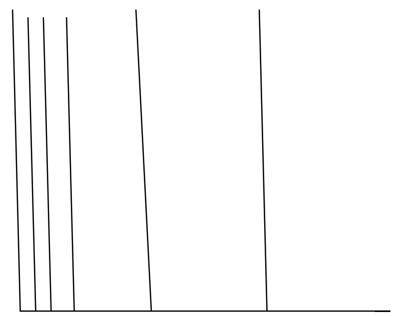

FIGURE 1.3. A topological *comb* is a typical cause for non-local-connectivity. To establish non-local-connectivity of a subset set in $\mathbb{R}^2$, *chercher le peigne*. Notice that this comb is path connected and is locally connected at the corner point.

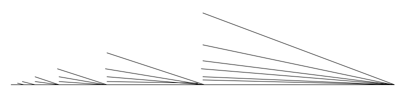

FIGURE 1.4. A *witch's broom* is weakly locally connected, but not locally connected, at the tip.

EXERCISE 1.12. *Show that a metric space X is weakly lc at some point x iff it has an lc modulus at this point. Conclude that a compact metric space X is locally connected iff it has an lc modulus.*

EXERCISE 1.13. *a) Show that curves are locally connected.*
b) More generally, the image of an lc continuum is an lc continuum.

PROBLEM 1.14. *An lc continuum $K \subset \mathbb{R}^n$ is arcwise connected.*

A space X is called a *path/arcwise locally connected at a point $x \in X$* if there exists a *path/arcwise lc modulus* $\omega(\varepsilon)$ such that any point $y \in X$ which is ε-close to x can be connected to x with a path/arc of diameter less than $\omega(\varepsilon)$. As usual, *path/arcwise lc of the whole space* means path/arcwise lc at every point.

EXERCISE 1.15. *For the whole space, properties of being arcwise lc, path lc, and locally connected are all equivalent.*

EXERCISE 1.16. *Let K be a compact subset of $\mathbb{R}^n$, and let $J = \partial K$. If J is locally connected then so is K.*

Quite remarkably, local connectivity gives a characterization of curves:

THEOREM 1.17 (Hahn-Mazurkevich). *Let X be a compact space. Then X is a lc continuum iff there is a space-filling curve $\gamma : [0,1] \to X$ ("Peano curve").*

PROBLEM 1.18. *Prove this theorem.*

1.3. Plane topology. A *plane domain* is a domain in $\bar{\mathbb{C}}$.

1.3.1. *Hulls and their cellular approximations.* A *Jordan curve* γ is a simple closed curve in the 2-sphere S^2. It is called *polygonal* if it is contained in $\mathbb{R}^2$ and composed of finitely many straight intervals.

Lemma 1.19 (Polygonal Jordan Theorem). *A polygonal Jordan curve bounds a domain P whose closure (the "polygon" $\bar{P}$) is homeomorphic to $\bar{\mathbb{D}}$.*

Proof. Take a line L that does not pass through any vertex of γ and is not parallel to any edge of γ. Components of $\gamma \smallsetminus L$ admit a checkerboard coloring, depending on which side of L they lie. Since $\gamma \approx \mathbb{T}$, there is an even number of intersection points between γ and L. Hence $L \smallsetminus \gamma$ admits a checkerboard coloring such that both unbounded components are colored white. (This can be also done, with an appropriate adjustment of the notion of checkerboard coloring, in the case when L is allowed to pass through a vertex.)

Take now the foliation of all lines parallel to L and declare P to be the set of black points of this foliation.

To see that $\bar{P}$ is a topological disk, triangulate it and carry induction in the number of triangles. $\square$

A compact subset K in $\mathbb{R}^2$ is called *full* or *non-separating* if $\mathbb{R}^2 \smallsetminus K$ is connected. (Intuitively, K "does not have holes"). A full non-trivial continuum is called a *hull*.

A point $z \in \partial K$ is called *peripheral* if it belongs to the boundary of some component of $\operatorname{int} K$.

A subset K' of a hull K is called a *subhull* if it is a hull such that $\operatorname{int} K'$ is the union of some components of $\operatorname{int} K$.

A compact set $K \subset \mathbb{R}^2$ is called *cellular* if there exists a nest of closed topological disks $\bar{D}_i$ (where $D_i = \operatorname{int} \bar{D}_i$) shrinking to K:

$$D_1 \Supset D_2 \Supset \cdots \Supset K; \quad \bigcap D_i = K. \tag{1.1}$$

Proposition 1.20. *Any hull $K \subset \mathbb{R}^2$ is cellular.*

Proof. For any ε-neighborhood U of K, we can construct a polygonal Jordan curve $\gamma \subset U \smallsetminus K$ that bounds a polygon P containing K (for instance, by covering K with a union of small grid boxes). We can then organize the corresponding polygons into a nest $P_1 \Supset P_2 \Supset \dots$, and take their intersection P_∞. It is a compact set containing K.

Let us show that $P_\infty \subset K$. Take some $a \in \mathbb{R}^2 \smallsetminus K$. Since K is a hull, there is a path δ in $\mathbb{R}^2 \smallsetminus K$ connecting a to ∞. It stays some positive distance away from K and hence does not intersect the curves ∂P_n for n sufficiently big. It follows that δ does not intersect the polygons $\bar{P}_n$ either, and in particular, $a \notin P_\infty$. $\square$

Corollary 1.21. *If $K \subset \mathbb{R}^2$ is a hull then there exists a continuous map $h : (\mathbb{R}^2, K) \to (\mathbb{R}^2, 0)$ whose restriction to $\mathbb{R}^2 \smallsetminus K$ is a homeomorphism onto $\mathbb{R}^2 \smallsetminus \{0\}$.*

Proof. Consider a cellular approximation (1.1) of K. Let $A_i := \bar{D}_i \smallsetminus D_{i+1}$ and $\tilde{A}_i := \mathbb{A}[2^{-(i+1)}, 2^{-i}]$. Construct consecutively homeomorphisms $A_i \to \tilde{A}_i$ matching on the common boundaries. We obtain a homeomorphism $\bar{D}_0 \smallsetminus K \to \bar{\mathbb{D}} \smallsetminus \{0\}$ that extends continuously to a desired map. $\square$

Thus, the space obtained by *collapsing a hull $K \subset \mathbb{R}^2$ to a single point*[5] is a topological plane $\mathbb{R}^2$. Of course, we can also interpret this result on the sphere S^2, where a "hull" $K \subset S^2$ should be understood as a continuum with connected (and

[5] i.e., taking the quotient by the equivalence relation that identifies all points of K to a single point.

non-empty) complement $S^2 \setminus K$. We see that the space obtained by collapsing such a K to a point is a topological sphere S^2.

Doing it inductively, we conclude:

PROPOSITION 1.22. *Let $K_i \subset S^2$ be a finite family of disjoint hulls. The space obtained by collapsing each K_i to a single point is a topological sphere S^2.*

1.3.2. *Jordan Theorem.* The following result gives an intrinsic characterization of hulls in $\mathbb{R}^2$ in terms of vanishing Alexander cohomology:

THEOREM 1.23. *A continuum $K \subset \mathbb{R}^2$ is a hull iff $H^1_{\mathrm{A}}(K) = 0$.*

PROOF. It is enough to show that for r sufficiently small, any two r-chains C and C' in K are discretely homotopic in K rel the endpoints. By Proposition 1.20, any neighborhood of K contains a Jordan disk $D \supset K$. By Example 1.124, C and C' are discretely homotopic in D. The chains that appear in the course of this homotopy can be projected to K with a small error, providing us with the desired homotopy in K.

Vice versa, assume $\mathbb{R}^2 \setminus K$ contains a bounded component D. Then for a point $a \in D$, the interaction

$$\phi_a(z, \zeta) := \arg(z/\zeta), \quad z, \zeta \in K,\ |z - \zeta| < \mathrm{dist}(a, K),$$

defines a non-trivial 1-cocycle on K (compare Example 1.125). □

Thus, being a hull is an intrinsic property of K, independent of the embedding of K into the plane.

COROLLARY 1.24 (Jordan Theorem: weaker version). *Any Jordan curve $\gamma \subset \mathbb{R}^2$ separates the plane.*

PROOF. Otherwise γ would be a hull. Then $H^1_{\mathrm{A}}(\gamma) = 0$ contradicting Example 1.125. □

PROBLEM 1.25. *For a continuum $K \subset \mathbb{R}^2$, the rank of $H^1_{\mathrm{A}}(K)$ is equal to the number of bounded components of $\mathbb{R}^2 \setminus K$.*

Putting this together with Example 1.125, we obtain:

JORDAN THEOREM. *The complement of a Jordan curve γ consists of two components D_1 and D_2 with the common boundary γ.*

These components are called *(open) Jordan disks*. Their closures $\bar{D}_i = D_i \cup \gamma$ are called *closed Jordan disks.*

A closed curve γ on a manifold S is called *trivial* if it is nul-homotopic, i.e., there is a continuous family of closed curves γ_t, $0 \le t \le 1$, such that $\gamma_0 = \gamma$ while γ_1 is a single point x. A connected manifold S is called *simply connected* if any closed curve on S is nul-homotopic. In other words, its fundamental group $\pi_1(S, x)$ is trivial (which is independent of the choice of the base point x).

EXERCISE 1.26. *Show that any open Jordan disk is simply connected.*

When S^2 is realized as one-point compactification of $\mathbb{R}^2$, $S^2 = \mathbb{R}^2 \cup \{\infty\}$, and a Jordan curve γ lies in $\mathbb{R}^2$, then one of the corresponding Jordan disks is bounded in $\mathbb{R}^2$, while another contains ∞. They are called the *inner* and *outer* Jordan disks respectively. If a point z belongs to the inner Jordan disk, we say that "γ goes around z" or "γ surrounds z".

1.3.3. *Embedded trees.* Let us orient the plane $\mathbb{R}^2$, and let X be a *planar space* (i.e., X can be embedded into $\mathbb{R}^2$). Two embeddings, $i_j : X \to \mathbb{R}^2$, $j = 1, 2$, are called *ambient equivalent* if there is an orientation preserving homeomorphism $h : \mathbb{R}^2 \to \mathbb{R}^2$ such that $h \circ i_1 = i_2$.

An *(abstract) closed star* $\mathcal{S}$ *of valence* $\mathfrak{q}$ is $\mathfrak{q}$ copies of the interval $[0, 1]$ glued at 0. (In other words, it is a tree with $\mathfrak{q}$ edges that have a common *root* of valence $\mathfrak{q}$.) Similarly, we can consider an (*abstract*) *open star* S° by taking $\mathfrak{q}$ rays $[0, +\infty)$ and gluing them at 0.

Recall from §50.2 that $e(\theta) \equiv e^{2\pi i\theta}$.

Exercise 1.27. (i) *Any properly embedded open star of valence* $\mathfrak{q}$ *is ambient equivalent to the standard open star*

$$\mathcal{S}^\circ_{\mathrm{st}} := \bigcup_{k=0}^{\mathfrak{q}-1} e(k/\mathfrak{q}) \cdot [0, +\infty).$$

(ii) *Any ambient self-homeomorphism* $h : (\mathbb{R}^2, S^\circ_{\mathrm{st}}) \to (\mathbb{R}^2, S^\circ_{\mathrm{st}})$ *of the standard open star is the composition of a rotation by* $\mathfrak{p}/\mathfrak{q}$, $\mathfrak{p} \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$, *and a homeomorphism* $h : (\mathbb{R}^2, S_{\mathrm{st}}) \to (\mathbb{R}^2, S_{\mathrm{st}})$ *preserving all the rays.*

In particular, *the rays of a properly embedded star are cyclicly ordered and any ambient orientation preserving self-homeomorphism induces a rotation on this ordered set.* (See Appendix 1.11 on the notion of cyclic order.) Moreover, a proper embedding of an open star is determined (up to ambient equivalence) by a cyclic order on its rays. Note also that if the rays have distinct asymptotic slopes $\theta_i \in \mathbb{R}/\mathbb{Z}$ at the root of $\mathcal{S}$ or at ∞, the cyclic order of the rays coincides with the cyclic order of the angles θ_i (induced from $\mathbb{R}/\mathbb{Z}$).

Similarly,

$$S_{\mathrm{st}} := \bigcup_{k=0}^{\mathfrak{q}-1} e(k/\mathfrak{q}) \cdot [0, 1].$$

provides a standard embedded model for closed stars. It provides us with a local version of the above conclusion: the rays of an embedded star are cyclically ordered, and any ambient orientation preserving local self-homeomorphism (near the start vertex) induces a rotation of this ordered set. An embedded star is determined (up to local ambient equivalence) by a cyclic order on the edges.

More generally, we can consider an abstract tree $\mathcal{T}^{\mathrm{abs}}$, and its embeddings into the oriented plane $\mathbb{R}^2$ up to the ambient equivalence. Any vertex v of $\mathcal{T}^{\mathrm{abs}}$ is the root of the attached *local star* S^{abs}_v (comprising the edges e such that $v \in \partial e$).

Exercise 1.28. *An embedding* $i : \mathcal{T}^{\mathrm{abs}} \to \mathbb{R}^2$ *is determined, up to ambient equivalence, by a cyclic order on all the local stars* S^{abs}_v.

1.3.4. *Finer structure of hulls.*

Exercise 1.29. (i) *If* K *is a hull, then any component of* $\operatorname{int} K$ *is simply connected.*
(ii) *Let* J *be a compact subset of* $\mathbb{R}^2$, *and let* U_i *be the bounded components of* $\mathbb{R}^2 \smallsetminus J$. *Then* $K := J \cup \bigcup U_i$ *is a hull.*

This procedure is called *filling-in the holes of* J.

LEMMA 1.30. *Let $K \subset \mathbb{R}^2$ be a lc hull, and let U be a component of* $\operatorname{int} K$. *Take a $z \in K \smallsetminus \bar{U}$ and connect it with an arc $\alpha \subset K$ to some point in $\bar{U}$. Let $\pi_\alpha(z) \equiv \pi_{U,\alpha}(z)$ be the the first point of intersection of α with $\bar{U}$. Then $\pi_\alpha(z)$ is independent of α.*

PROOF. Assume we have two arcs α_1 and α_2 in K connecting z to $\bar{U}$ such that $\zeta_1 := \pi_{\alpha_1}(z) \neq \pi_{\alpha_2}(z) =: \zeta_2$. Without loss of generality we can assume that the α_i end at ζ_i. Let $(u, \zeta_1]$ be the maximal subarc of α_1 that contains ζ_1 and does not cross α_2, and let $\alpha_i' = [u, \zeta_i]$ be the closed subarcs of the α_i bounded by u and ζ_i $(i = 1, 2)$. Then u is the only common point of the latter arcs. Moreover, $u \notin \bar{U}$.

Let us take some points $w_1, w_2 \in U$ that are ε-close to ζ_1, ζ_2 respectively. By Lemma 1.9 and Exercise 1.12, w_i can be connected to the respective point ζ_i by an arc $\gamma_i \subset K$ with $\operatorname{diam} \gamma_i < \omega(\varepsilon)$. So, for ε small enough, γ_1 is disjoint from $\delta_2' := \alpha_2' \cup \gamma_2$ and γ_2 is disjoint from $\delta_1' := \alpha_1' \cup \gamma_1$.

Applying Lemma 1.9 again, we can straighten the curves δ_i' to arcs $\delta_i \subset \delta_i'$ connecting u to w_i. Then u is the only one common point of these arcs as well.

Let us now connect w_1 to w_2 with an arc $\sigma \subset U$ disjoint from $\delta_1 \cup \delta_2$ (except for the endpoints).

The union of three arcs, δ_1, δ_2 and σ, form a Jordan curve in K. Let D be the open Jordan disk bounded by this curve. Since K is full, $D \subset K$. Moreover, D intersects U, and hence $U \cup D$ is contained in a component of $\operatorname{int} K$, so $D \subset U$. Hence $u \in \bar{U}$ – contradiction. ☐

So, under the above circumstances we have a well defined projection:

$$\pi_U : K \to \bar{U}. \tag{1.2}$$

EXERCISE 1.31. *The projection π_U is continuous and locally constant on $K \smallsetminus \bar{U}$.*

Together with Exercise 1.13 b), this implies:

COROLLARY 1.32. *If $K \subset \mathbb{R}^2$ is a lc hull and U is a component of* $\operatorname{int} K$ *then $\bar{U}$ is a hull as well.*

EXERCISE 1.33. *Let K be a lc hull in $\mathbb{R}^2$ whose complement has infinitely many components D_i. Then* $\operatorname{diam} D_i \to 0$.

An *external neighborhood* of a hull $K \subset \mathbb{C}$ is a set $U \smallsetminus K$ where U is a neighborhood of K.

Further study of plane hulls will require analytic methods (see §9.2).

Let us conclude with a useful criterion that ensures that a piecewise homeomorphism is actually a homeomorphism.

EXERCISE 1.34. *Let K and $\tilde{K}$ be two compact subsets of $\mathbb{C}$, and let D_i, $\tilde{D}_i$ be the components of their interior. Assume* $\operatorname{diam} D_i \to 0$ *and* $\operatorname{diam} \tilde{D}_i \to 0$ *(which holds automatically when K and $\tilde{K}$ are lc hulls). Let $h : K \to \tilde{K}$ be a bijection that restricts to homeomorphisms $\partial K \to \partial \tilde{K}$ and $\operatorname{cl} D_i \to \operatorname{cl} \tilde{D}_i$. Then h is a homeomorphism. Show by example that the shrinking condition for D_i and $\tilde{D}_i$ cannot be dropped.*

A nowhere dense lc hull is called a *dendrite*.

1.3.5. *Accessibility.* Let $D \subset \hat{\mathbb{C}}$ be a simply connected domain whose boundary contains more than one point. Let $\gamma : [0,1) \to D$ be a curve. We say that γ *lands* at some point $\zeta \in \partial D$ if $\gamma(t) \to \zeta$ as $t \to 1$. A boundary point $\zeta \in \partial D$ is called *accessible* from D if there is a curve γ landing at ζ.

EXERCISE 1.35. *The set of accessible points is dense in ∂D.*

Let γ_0 and γ_1 be two curves in D landing at ζ. We say that γ_0 and γ_1 *represent the same access* to ζ if they are homotopic in D rel ζ (i.e., there is a family of curves $\gamma_\tau : [0,1) \to D$, $\tau \in [0,1]$, deforming γ_0 to γ_1, all landing at ζ). Thus, an *access* at ζ is a class of homotopic curves in D landing at ζ.

Note that there exist $\mathfrak{q}$ accesses to a vertex of valence $\mathfrak{q}$ of an embedded graph.

1.3.6. *Disk isotopies rel boundary.* The following basic topological fact shows that in the disk isotopy classes rel boundary are determined by the boundary values.

ALEXANDER TRICK. *Any two homeomorphisms ϕ and ψ of the closed unit disk $\mathbb{D}$ that coincide on $\mathbb{T} = \partial\mathbb{D}$ are isotopic rel $\mathbb{T}$. The same is true for the punctured disk $\mathbb{D}^*$.*

PROOF. Let us first consider the case $\psi = \mathrm{id}$, and hence $\phi | \mathbb{T} = \mathrm{id}$.

EXERCISE 1.36. *Show that ϕ is isotopic rel $\mathbb{T}$ to a map fixing 0.*

In what follows we assume that $\phi(0) = 0$. Let $\phi_0 = \mathrm{id}$, while for $t \in (0,1]$, let us define a homeomorphism $\phi^t : \bar{\mathbb{D}} \to \bar{\mathbb{D}}$ as follows

$$\phi^t(z) = t\,\phi(z/t) \text{ for } z \in \bar{\mathbb{D}}_t = \{z : \ |z| \leq t\}, \quad \phi^t(z) = z \text{ otherwise.}$$

This is the desired isotopy.

The general case is reduced to $\psi = \mathrm{id}$ by replacing ϕ with $\psi^{-1} \circ \phi$. □

EXERCISE 1.37. *Let $K \subset \mathbb{T}$ be a closed subset of the circle. Then two homeomorphisms ϕ and ψ of the closed unit disk $\mathbb{D}$ that coincide on K are isotopic rel K.*

1.3.7. *Annulus (cylinder) twists.* Next, we will classify homeomorphisms of the cylinder $\mathrm{Cyl} = (\mathbb{R}/\mathbb{Z}) \times \mathbb{I}$ up to homotopy rel the boundary. Let us consider the standard *twist*

$$\tau : (\theta, x) \mapsto (\theta + x, x).$$

Notices that it pointwise fixes the boundary and it maps the vetrical interval $\{0\} \times \mathbb{I}$ to the spiral $\{x = \theta\}$.

EXERCISE 1.38. (i) *The twists τ^n, $n \in \mathbb{Z}$, represent distinct homotopy classes of cylinder homeomorphisms rel $\partial\,\mathrm{Cyl}$ fixing the boundary pointwise.*

(ii) *Any cylinder homeomorphism which fixes the boundary pointwise is homotopic rel $\partial\,\mathrm{Cyl}$ to some twist τ^n, $n \in \mathbb{Z}$.*

Of course, this discussion can be immediately transported to any (compact) *topological annulus.* We will use the same name and notation for the corresponding topological twists.

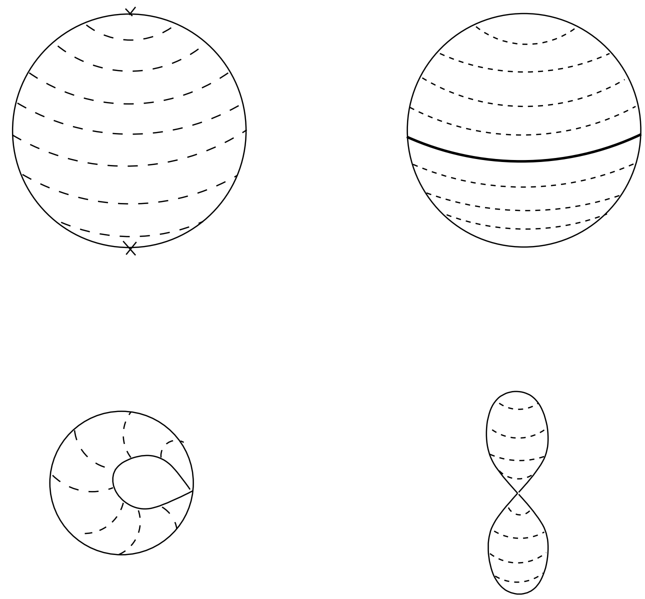

FIGURE 1.5. On the left-hand side picture, one equivalence class is disconnected (consists of two points), on the right-hand side one, it is not full (a circle).

1.3.8. *Quotients of the sphere.*

MOORE'S THEOREM. Let $\mathcal{P}$ be a partition of the sphere S^2, and let X be the corresponding quotient space. Then X is a topological sphere if and only if $\mathcal{P}$ is closed and every non-singleton class of $\mathcal{P}$ is a hull.

See [**Tim**] for a proof.

This is a deep topological result. However, its conclusion can be verified directly in all occasions that we will encounter in this book. Figure 1.5 illustrates what can go wrong if the conditions of the theorem are violated.

1.4. Zoo of wild creatures.

1.4.1. *Koch snowflake.* Let us consider an equilateral triangle Δ (viewed as a polygonal curve). Divide each of its sides L_j into three sub-interval and attach an equilateral triangle to each of the three middle intervals I_j; then erase the I_j. We obtain an 12-gon (the first level approximation to the snowflake). Again, divide each of its sides into three sub-interval and attach an equilateral triangles to each of the twelve middle intervals, erasing those intervals afterwards. We obtain a 48-gon (the second level approximation to the snowflake). Proceed inductively and pass to a limit (in the Hausdorff topology on the space of sets). The limiting set is called the *Koch snowflake* S.

EXERCISE 1.39. (i) *Justify existence of the limit. Parameterizing the approximating polygons in a piecewise linear way, show that the convergence holds in the uniform topology on the space of curves.*

(ii) *Show that S is self-similar: Each piece S_j attached to the intervals I_j can be affinely amplified to the piece of S attached $\Delta \setminus L_j$.*

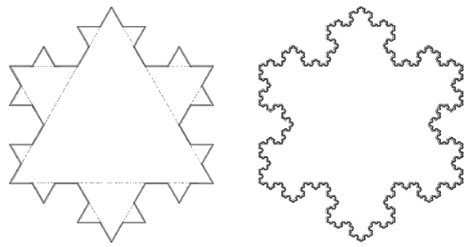

FIGURE 1.6. The Koch snowflake (on the right) and the 2nd order approximation to it (on the left).

(iii) *Show that the Koch snowflake is a nowhere differentiable Jordan curve (it does not admit tangent lines anywhere).*

1.4.2. *Jordan curve of positive area.* Take a sequence of numbers $\varepsilon_n > 0$, $n = 1, 2, \dots,$ with $\sum \varepsilon_n < \infty$. Start with a standard annulus A^0. Inscribe into it a narrow oscillating annulus A^1 such that area $A^1 \geq (1-\varepsilon_1)$ area A^0. Then inscribe into A^1 a much more narrow annulus A^2 with much higher oscillating frequency rel A^1 such that area $A^2 \geq (1-\varepsilon_2)$ area A^1; etc.

EXERCISE 1.40. *Show that this construction can be arranged so that in the end we obtain a Jordan curve $\Gamma := \bigcap A^n$ of positive area.*

1.4.3. *Knaster continuum.* Let us consider the standard (1/3)-Cantor set $K \subset [0,1]$. It is the union of two "Cantor intervals" $J := K \cap [0,1/3]$ and $I := K \cap [2/3,1]$, permuted by the reflection $\sigma : x \mapsto 1-x$ with respect to 1/2. Let us connect each pair of symmetric points x and $\sigma(x)$ ($x \in J$) with the *upper* half-circle arc (see Figure 1.8) .

Let us now decompose the Cantor interval I into the union of two Cantor intervals of the next level, $I = L^1 \cup R^1$, where $L^1 := K \cap [2/3, 7/9]$, $R^1 := K \cap [8/9, 1]$. These intervals are permuted by the reflection $\sigma_0 : I^0 \to I^0$ with respect to the middle of $I^0 \equiv I$. Connect each pair of symmetric points x and $\sigma_0(x)$ ($x \in L^1$) with the *lower* half-circle arc.

Let us now consider the rescaled Cantor intervals

$$I^n := 3^{-n} \cdot I^0, \ L^{n+1} := 3^{-n} \cdot L^1, \ R^{n+1} := 3^{-n} \cdot R^1, \quad n = 0, 1, 2, \dots.$$

Then

$$K = \{0\} \cup \bigcup_{n=0}^{\infty} I^n, \quad I^n = L^{n+1} \cup R^{n+1}.$$

The intervals L^{n+1} and R^{n+1} are permuted by the reflection $\sigma_n : I^n \to I^n$ with respect to the middle-point of I^n. Connect each pair of symmetric points x and $\sigma_n(x)$, $x \in L^{n+1}$, with the lower half-circle arc. The union of all these arcs constitute the *Knaster continuum*.

A continuum is called *indecomposable* if it cannot be represented as the union of two proper subcontinua.

EXERCISE 1.41. (i) *The Knaster continuum K is connected but not path connected. All path connected components of K, except one, are densely immersed real lines* (*the exceptional component, through* 0, *is a densely immersed ray* $[0, +\infty)$).

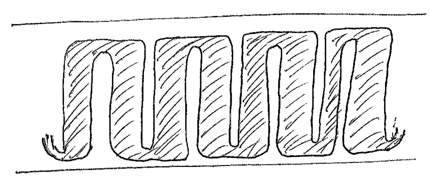

FIGURE 1.7. A thin oscillating annulus occupying substantial area (a local picture).

(ii) *The Knaster continuum is indecomposable.*

1.4.4. *Pseudo-arcs.* A continuum X is called *hereditary indecomposable* if any subcontinuum of X is indecomposable. So, *it does not contain any arcs.*

A *chain* in X is a covering of X by a sequence of open sets, $\mathcal{C} = (C_0, C_1, \dots, C_n)$, in such a way that $C_i \cap C_j \neq \emptyset$ iff $|i - j| \leq 1$. If moreover, $\operatorname{diam} C_i < \varepsilon$ for all $i = 0, \dots n$, then $\mathcal{C}$ is called an ε-chain. A continuum X is called *chainable* if it admits an ε-chain for any $\varepsilon > 0$.

A *pseudo-arc* is a chainable hereditary indecomposable continuum.

Let us briefly outline a construction of a pseudo-arc. A chain $\mathcal{D}$ *refines* $\mathcal{C}$ if the closure $\bar{D}_i$ of any element of $\mathcal{D}$ is contained in some element C_m of $\mathcal{C}$. Under these circumstances, a chain $\mathcal{D}$ is *crooked* in $\mathcal{C}$ if for any elements C_m, C_n of $\mathcal{C}$ with $m < n - 2$, and any elements D_i, D_j of $\mathcal{D}$ such that $D_i \cap C_m \neq \emptyset$ and

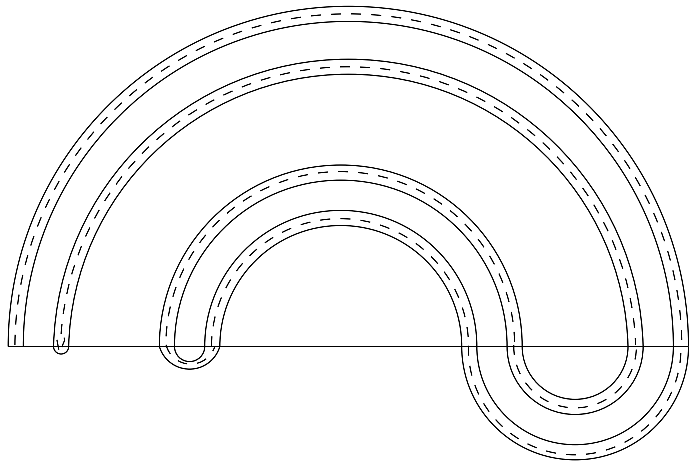

FIGURE 1.8. Knaster continuum.

$D_j \cap C_n \neq \emptyset$, there exist elements $D_k \subset C_{n-1}$ and $D_l \subset C_{m+1}$, where $i < k < l < j$ or $i > k > l > j$. So, $\mathcal{D}$ *oscillates* in all scales associated with $\mathcal{C}$.

We say that a plane chain $\mathbb{R}^2$ "begins" at $a \in \mathbb{R}^2$ and "ends"at $b \in \mathbb{R}^2$ if the first element of $\mathcal{C}$ contains a while the last element contains b.

EXERCISE 1.42. (i) *Let $\varepsilon_k \to 0$ and let $\mathcal{C}^k$ be a sequence of plane ε_k-chains beginning at* 0 *and ending at* 1 *such that C^{k+1} is crooked inside C^k. Then $\bigcap \mathcal{C}^k$ is a pseudo-arc.*

(ii) *Construct a pseudo-arc.*

(iii) *Show that a pseudo-arc is a nowhere dense hull.*

Similarly, one can define and construct a *pseudo-circle* by using *cyclic chains* $\mathcal{C} = (C_i)_{i \in \mathbb{Z}/n\mathbb{Z}}$ in place of chains.

EXERCISE 1.43. *Prove that a pseudo-circle divides the plane into two components.*

1.4.5. *Lakes of Wada. Lakes of Wada* is the union of three (or more) disjoint topological disks in $\mathbb{R}^2$ that share a common boundary. To construct such domains, select a sequence $\varepsilon_n \to 0$ and start with three open topological disks D_i^0, $i = 1, 2, 3$, in $\mathbb{D}$, with disjoint closures. Then select an ε_1-net X_1 in $\mathbb{D} \smallsetminus \bigcup \bar{D}_i^0$, and dig out fjords from the D_i^0 that pass ε_1-close to each point of X_1. Call these new disks $D_i^1 \supset D_i^0$. Take now an ε_2-net X_2 in $\mathbb{D} \smallsetminus \bigcup \bar{D}_i^1$ and and dig out fjords out of the D_i^1 that pass ε_2-close to each point of X_2. Proceed inductively.

EXERCISE 1.44. *Go through details of the above construction. Show that similarly one can construct arbitrary many (including countably many) lakes of Wada.*

1.4.6. *Cantor bouquet.* Let us consider a compact subset $K \subset \mathbb{R}^2$ which is the union of straight intervals $T_\theta = [0, t_\theta]\, e(\theta)$, $\theta \in \mathbb{R}/\mathbb{Z}$ (some of which may be degenerate), with the following properties:

- the set $\Theta := \{\theta : t_\theta > 0\}$ is dense in $\mathbb{R}/\mathbb{Z}$ (the corresponding points $t_\theta\, e(\theta)$ are called *tips* of K);

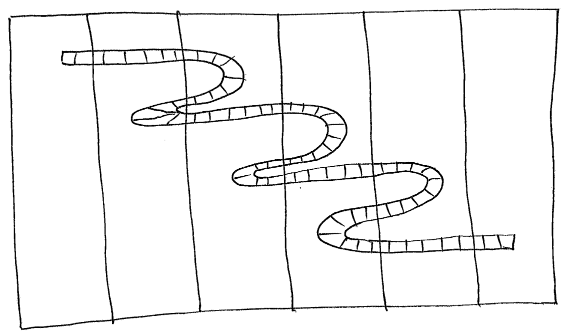

FIGURE 1.9. A partly crooked chain. It is crooked in the union of any four consecutive rectangles. (The rectangles should be slightly enlarged to form a chain.)

• for any $\theta \in \Theta$ there exist sequences $\alpha_n \nearrow \theta$ and $\beta_n \searrow \theta$ such that $t_{\alpha_n} e(\alpha_n) \to t_\theta e(\theta)$ and $t_{\beta_n} e(\beta_n) \to t_\theta e(\theta)$.

Let $T := \{t_\theta e(\theta)\}_{\theta\in\Theta}$ be the set of tips.

EXERCISE 1.45. (i) *Construct a Cantor bouquet K;*

(ii) *Show that K is path connected but not locally connected;*

(iii) *Show that T is connected while $T^* := T \smallsetminus \{0\}$ is totally disconnected.*

The last property makes 0 an *explosion point* for T.

1.5. Group actions and foliations.

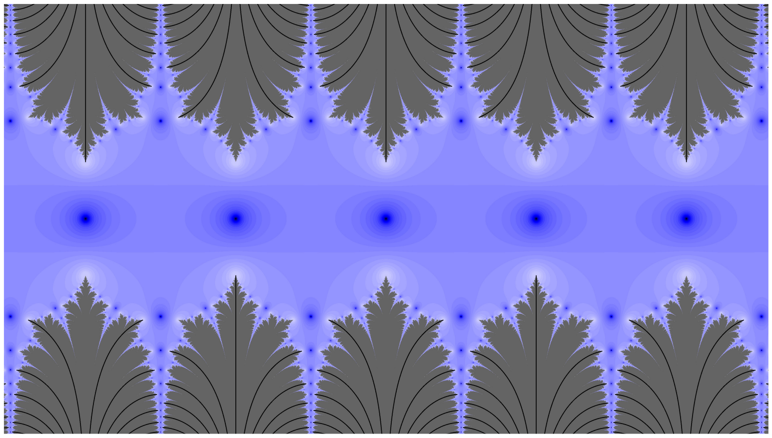

FIGURE 1.10. A Cantor bouquet generated by dynamics of $z \mapsto \lambda \sin z$ (courtesy of Lasse Rempe-Gillen).

1.5.1. *Groups actions.* Let us consider a general action of a group G on a space X (denoted $G \curvearrowright X$). A point $\alpha \in X$ is called *fixed* for the action if it is fixed by *some non-trivial* element $\gamma \in G$, i.e., there exists $\gamma \neq \mathrm{id}$ such that $\gamma(\alpha) = \alpha$. An action is called *free* if it does not have fixed points, i.e., all points are moved under non-trivial elements of G, The space X is called *homogeneous* for G if G acts transitively on X.

An action of a discrete group Γ on a locally compact space X is said to be *properly discontinuous* if any two points $x, y \in X$ have neighborhoods $U \ni x$, $V \ni y$ such that $\gamma(U) \cap V = \emptyset$ for all but finitely many $\gamma \in \Gamma$.

EXERCISE 1.46. (i) *The orbits of a properly discontinuous action are discrete.*
(ii) *The quotient of X by a properly discontinuous group action is a Hausdorff locally compact space.*
(iii) *Vice versa, if the quotient X/G is Hausdorff then the action is properly discontinuous.*
(iv) *Consider the $\mathbb{Z}$-action on the punctured plans $(\mathbb{R}^2)^*$ generated by the linear hyperbolic transformation $f : (x, y) \mapsto (2x, y/2)$. Show that it is not properly discontinuous, albeit all its orbits are discrete in $(\mathbb{R}^2)^*$.*

The *stabilizer* (or, the *isotropy group*) $\mathrm{Stab}(Y)$ of a subset $Y \subset X$ is the subgroup $\{\gamma \in \Gamma : \gamma(Y) = Y\}$. A set Y called *completely invariant* under some subgroup $G \subset \Gamma$ if $G = \mathrm{Stab}(Y)$ and $\gamma(Y) \cap Y = \emptyset$ for any $\gamma \in \Gamma \smallsetminus G$.

EXERCISE 1.47. *For a properly discontinuous action, the stabilizer of any point (or, more generally: of any compact subset) is finite.*

A group element γ is called *primitive* if it generates a maximal cyclic group.

Isometries of a metric space are also called *motions* (e.g., *Euclidean motions, hyperbolic motions*, etc.).

Let us consider two group actions, $G \curvearrowright X$ and $\tilde{G} \curvearrowright \tilde{X}$. A map $h : X \to Y$ is called *equivariant* (with respect to these actions) if there is a homomorphism $A : G \to \tilde{G}$ (*induced by* h) such that

$$h \circ \gamma = A(\gamma) \circ h \quad \forall \gamma \in G.$$

An equivariant homeomorphism $h : X \to \tilde{X}$ inducing an isomorphism $A : G \to \tilde{G}$ is also called a *conjugacy* between the actions.

A connected set $\Delta \subset X$ is called a *fundamental domain* for a group action $G \curvearrowright X$ if

- Δ is closed[6] and $\operatorname{int}\Delta$ is dense in Δ;
- $\gamma(\operatorname{int}\Delta) \cap \Delta = \emptyset$ for any $\gamma \in G$;
- $\bigcup_{\gamma\in G} \gamma(\Delta) = X$.

PROPOSITION 1.48. *For any properly discontinuous group action, there is a fundamental domain.*

The idea is to take a base boint $x_\circ \in X$, to consider its orbit $O := \operatorname{orb} x_\circ$ under G, and to let

$$\Delta := \{y \in X : \ \operatorname{dist}(y, x_\circ) = \operatorname{dist}(y, O) \equiv \min_{x\in O} \operatorname{dist}(y, x).\} \tag{1.3}$$

Norice that since the orbits of a properly discontinuous action are discrete, $\operatorname{dist}(y, O)$ is attained at finitely many points of O.

EXERCISE 1.49. *Check that the set Δ (1.3) is indeed a fundamental domain.*

A fundamental domain constructed in this way is called *Dirichlet*.

A fundamental domain gives a concrete idea of the quotient space X/G. Indeed, X/G is obtained from Δ by identifying boundary points $x, x' \in \partial\Delta$ related by the group action: $x' = \gamma x$ for some $\gamma \in \Gamma$.

1.5.2. *Foliations.* A (topological) k-dimensional *foliation* $\mathcal{F}$ of an n-dimensional manifold M is a partition of M into immersed k-dimenional subminafolds $\mathcal{L}_\alpha$ (called the *(global) leaves* of $\mathcal{F}$) such any point $x \in M$ has a neighborghood U with the following property:

There is a homeomorphism ϕ from U onto $\mathbb{D}^k \times \mathbb{D}^{n-k}$ such that for any $t \in \mathbb{D}^{n-k}$ the pullback $\phi^{-1}(\mathbb{D}^k \times \{t\})$ is a component of $\mathcal{L}_\alpha \cap U$ for some leaf $\mathcal{L}_\alpha$.

Under these circumstances, U is called a *flow box*, ϕ is the corresponding *foliated local chart*, $\mathcal{L}_\alpha \cap U$ is a *local leaf* or a *plague*.

A graph of a continuous function $\phi : \mathbb{D}^{n-k} \to \mathbb{D}^k$ is called a *transversal* in our flow box. More genrally, a *transversal* to a foliation is a curve which is locally (near any point $x \in \mathcal{T}$) is a transversal in some flow box.

If all the manifolds and the local charts involved are *smooth/holomorphic*, etc., then the foliation inherits the corresponding name. In the smooth setting, transversals are assumed (with out saying) to be smooth and also transverse to the leaves (in the usual smooth sense).

For instance, the fibers of a smooth map $\phi : M^n \to N^{n-k}$ between two manifolds form a smooth foliation k-dimensional away from the critical points of ϕ (by the IFT). Trajectories of a differential equation $x' = v(x)$ on a manifold M form a smooth 1D foliation away from zertos (also called "singular points") of the vector field v (by the Straightening Theorem for ODE). On a 2D manifold M, we can also consider a smooth 1D foliation whose leaves are tangent to a line field given by an

[6]So, Δ is not a "domain" in the usual sense, but this traditional abuse of terminology is commonly accepted.

equation $\omega(x) = 0$, where $\omega = u\,du + v\,dv$ is a 1-form (the line field and the foliation live away from zeros of ω).

In the above examples, if we keep critical/singular points or zeros of the corresponding objects (functions/vector fields or differential forms) then we can still talk about corresponding *singular foliations* (e.g., a foliation on "the whole M" given by a smooth function $\phi : M \to N$ has singularities at the critical points of ϕ).

All of the above notes remain valid in the holomorphic category as well.

EXERCISE 1.50. *Decsribe two singular foliations* $\operatorname{Re} z^m = 0$ *and* $\operatorname{Im} z^m = 0$ *on* $\mathbb{C}$*,* $m \geq 1$*. The same question for foliations* $\operatorname{Re} dz^l = 0$ *and* $\operatorname{Im} dz^l = 0$*,* $l \geq 2$*.*

Given two transversals, $\mathcal{T}_1$ and $\mathcal{T}_2$, in some flow box, we can consider a homeomorphism $h : \mathcal{T}_1 \to \mathcal{T}_2$ by sliding along the local leaves of the foliation. Such a homeomorphism is called a *holonomy.* Given two transversals, $\mathcal{T}_1$ and $\mathcal{T}_2$, and a path γ lying in some leaf $\mathcal{L}$ and connecting a point $x_1 \in \mathcal{T}_1$ to $x_2 \in \mathcal{T}_2$, we can *develop* a *local holonomy* $(\mathcal{T}_1, x_1) \dashrightarrow (\mathcal{T}_2.x_2)$ *along* γ*.*

A transversal $\mathcal{T}$ is called *global* if it intersects each leaf of the foliation. Sometimes the self-holonomy to $\mathcal{T}$ can be represented by a nice transformation (called the *monodromy map*):

EXERCISE 1.51. *For* $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$*, sonsider the foliation of the torus* $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ *be the lines* $y = \alpha x + c$ *with slope* α*. Show thar all the leaves of this foliation are dense in the torus.*

1.6. Coverings.

1.6.1. *Definition and first observations.* In this section we summarize for reader's convenience necessary background in the theory of covering spaces.

Let E and B be topological manifolds (possibly with boundary), where B is connected. A continuous map $p : E \to B$ is called a *covering of degree* $d \in \mathbb{Z}_+ \cup \{\infty\}$ (with *base* B and covering space E) if any point $b \in B$ has a neighborhood V such that

$$p^{-1}(V) = \bigsqcup_{i=1}^{d} U_i,$$

where each U_i is mapped homeomorphically onto V. The preimages $p^{-1}(b)$ are called *fibers* of the covering. The inverse maps $p_i^{-1} : V \to U_i$ are called the *local branches* of p^{-1}. Let us make a couple of simple observations:

- A covering of degree one is a homeomorphism;
- Restriction of a covering $p : E \to B$ to any connected component of E is also a covering.
- If V is a domain in B, $U = p^{-1}(V)$ then the restriction $p : U \to V$ is also a covering.

Coverings $p : E \to B$ and $p' : E' \to B'$ are called *equivalent* if there exist homeomorphisms $H : E \to E'$ and $h : B \to B'$ such that $h \circ p = p' \circ H$, i.e., the following diagram

$$\begin{array}{ccc} E & \xrightarrow{H} & E' \\ p\downarrow & & \downarrow p' \\ B & \xrightarrow[h]{} & B' \end{array} \tag{1.4}$$

is commutative.

1.6.2. *Lifting.* Given two coverings, $p : E \to B$ and $p' : E' \to B'$, and a continuous map $h : B \to B'$, a continuous map $H : E \to E'$ is called a *lift* of h (via p and p') if H makes diagram (3.2) commutative, i.e., $h \circ p = p' \circ H$. Under these circumstances, h is called *liftable via* p *and* p', or just $(p, p')-$*liftable.* (In case when $E = B$ and $p = id$, the lift of $h : B \to B'$ to E' means a map $H : B \to E'$ such that $p' \circ H = h$.) Similarly, one defines a *lift of a homotopy.*

Theory of covering spaces is based on the following fundamental property:

Path Lifting Property. *Let γ be a path in B that begins at $b \in B$, and let $e \in p^{-1}(b)$. Then there is a unique lift $\tilde{\gamma}$ of γ (i.e., $p \circ \tilde{\gamma} = \gamma$) that begins at e. If γ is homotopic to γ' (rel the endpoints) then the corresponding lifts $\tilde{\gamma}$ and $\tilde{\gamma}'$ are also homotopic rel the endpoints.*

It implies, in particular, that the induced homomorphism

$$p_* : \pi_1(E, e) \to \pi_1(B, b)$$

is injective; let $G \equiv G_p \subset \pi_1(B, b)$ be its image. If E is connected then replacing e with another point in the fiber $p^{-1}(b)$ replaces G with a conjugate subgroup. In this way, to any covering $p : E \to B$ (with connected E) and a base point $b \in B$, we associate a subgroup of the fundamental group $\pi_1(B, b)$, up to conjugacy.

The Path Lifting Property implies a general

Lifting Criterion. *A continuous map $h : (B, b) \to (B', b')$ admits a lift $H : (E, e) \to (E', e')$, where $e \in p^{-1}(b)$) and $e' \in p^{-1}(b')$, if and only if*

$$h_* \circ p_*(\pi_1(E, e)) \subset p'_*(\pi_1(E', e')).$$

In particular, if E is simply connected, then all maps $h : B \to B'$ are liftable.

Exercise 1.52. *Prove the Lifting Criterion.*

In what follows we *assume that E is connected.*

1.6.3. *Universal covering and monodromy.* A covering is called *Galois* or *regular* if there is a group G acting freely and properly discontinuously on E whose orbits are fibers of the covering. In this case $B \approx E/G$. The group G is called the *group of deck transformations*, or the *covering group* for p.

Vice versa, if a group G acts freely and properly discontinuous on a manifold E then the quotient $B := E/G$ is a manifold, and the natural projection $p : E \to B$ is a covering.

Exercise 1.53. *Let $\pi : X \to X'$ be a regular covering with a group Γ of deck transformations. Let $Y' \subset X'$ be connected, and let $Y \subset X$ be a connected component of $\pi^{-1}(Y')$. Then Y is completely invariant under the action of Γ.*

A covering $u : \mathcal{U} \to B$ is called *Universal* if the space $\mathcal{U}$ is connected and simply connected. This covering is Galois, with the fundamental group $\pi_1(T)$ acting by deck transformations. *Any manifold has a unique Universal Covering up to covering equivalence.*

Remark 1.54. We suppress the base point in the notation for the fundamental group, unless it can lead to confusion. (On most occasions, our statements are invariant under conjugacies (inner automorphisms) in the fundamental group, and hence are base point independent.)

The Universal Covering allows us to recover any covering p from the a subgroup $\Gamma \subset \pi_1(B)$ as $p : \mathcal{U}/\Gamma \to B$. Moreover, the universal covering factors through p since $p \circ q = u$, where $q : \mathcal{U} \to \mathcal{U}/\Gamma$. This provides us with a natural one-to-one correspondence between classes of conjugate subgroups of $\pi_1(B)$ and classes of equivalent coverings $p : E \to B$. Moreover, the covering p is Galois if and only if the corresponding subgroup Γ is normal. In this case, the group of deck transformations of p is $\pi_1(B)/\Gamma$.

In particular, a *simply connected manifold B does not admit any non-trivial coverings*: any covering $p : E \to B$ with connected E is a homeomorphism. Putting this together with the above observations, we obtain the following important statement:

Monodromy Theorem. *If $p : E \to B$ is a covering and $V \subset B$ is a simply connected domain, then $p^{-1}(V)$ is a disjoint union of domains U_i, $i = 1, \dots, d$, such that each restriction $p : U_i \to V$ is a homeomorphism. Thus, on any simply connected domain there exist d well defined inverse branches $p_i^{-1} : V \to U_i$.*

Given a base point $b \in B$, there exists a natural *monodromy action* of the fundamental group $\Gamma := \pi_1(B, b)$ on the fiber $F := \pi^{-1}(b)$. Namely, let an element $A \in \Gamma$ is represented by a loop α in B based at b. Lift α to a path $\tilde{\alpha}$ in E based at some $e \in F$. Then $A(e)$ is defined as the endpoint of $\tilde{\alpha}$. The stabilizer of this action is the subgroup $G_e \subset \Gamma$ corresponding to p (well defined up to conjugacy), which gives yet another viewpoint on the relation between coverings over B and subgroups of $\pi_1(B)$.

Remark 1.55. We see that the coverings over B are classified purely algebraicly, so the list depends only on the homotopy type of B.

Exercise 1.56. (i) *For any $d \in \mathbb{Z}_+ \cup \{\infty\}$ there is a unique covering $p_d : E_d \to S^1$ of degree d over the topological circle S^1. The p_d constitute the full list of coverings (up to covering equivalence) over S^1. Moreover, $E_d \approx S^1$ for any finite d, while $E_\infty \approx \mathbb{R}$ (which is the universal covering). All these coverings are Galois.*
(ii) *Any homeomorphism $h : S^1 \to S^1$ admits d lifts to a homeomorphism $H : E_d \to E_d$. The lift is determined by the value $H(e)$ (selected arbitrary in the appropriate fiber) at any point $e \in E_d$.*

Exercise 1.57. *Let $f : S^1 \to S^1$ be a continuous map of degree $d \in \mathbb{Z}$ (i.e., f induces the π_1-endomorphism $n \mapsto d \cdot n$). Let $e : \mathbb{R} \to S^1$ be the universal covering for which the lattice $\mathbb{Z} \subset \mathbb{R}$ serves as the group of deck transformations. By the Lifting Criterion, f admits a lift $F : \mathbb{R} \to \mathbb{R}$. Then the action of F is $\mathbb{Z}$-equivariant:*

$$F(x + 1) = F(x) + d, \quad x \in \mathbb{R}.$$

Exercise 1.58. *Let $p : E \to B$ be a covering of degree d. Then there exists a Galois covering $q : L \to B$ of degree at most $d!$ that factors through p, i.e., $q = p \circ r$ (where $r : L \to E$ is also (automatically) a Galois covering).*

1.6.4. *Essential submanifolds.* One says that a connected submanifold $V \subset B$ (possibly with boundary) is *essential* in B if the induced (by the embedding) homomorphism $\pi_1(V) \to \pi_1(B)$ is injective. In other words, any non-trivial loop in V remains non-trivial in B.

PROPOSITION 1.59. *Let $u : \mathcal{U} \to B$ be the universal covering. If $V \subset B$ is an essential submanifold then any component $\hat{V}$ of $u^{-1}(V)$ is simply connected. Moreover, the stabilizer Γ of $\hat{V}$ in the group G of deck transformations is the covering group for the restriction $p : \hat{V} \to V$ (and thus, Γ is isomorphic to $\pi_1(V)$).*

PROOF. By the above observations, restriction $u : \hat{V} \to V$ is a covering. If $\hat{V}$ was not simply connected, then there would be a non-trivial loop $\tilde{\alpha}$ in $\hat{V}$. Then the loop $\alpha = p_*(\tilde{\alpha})$ would be non-trivial in V (since p_* is injective) but trivial in B (since $\tilde{\alpha}$ is trivial in $\mathcal{U}$).

Since $p^{-1}(V)$ is invariant under G, each deck transformation $\gamma : \mathcal{U} \to \mathcal{U}$ permutes the components of $p^{-1}(V)$. Hence for any $\gamma \in G$, $\hat{V}$ is either invariant under γ or else $\gamma(\hat{V}) \cap \hat{V} = \emptyset$. It follows that the stabilizer Γ of $\hat{V}$ acts transitively on the fibers of $p|\,\hat{V}$, and the conclusion follows. □

COROLLARY 1.60. *Let γ be an essential simple closed curve in B. Then each lift $\tilde{\gamma}$ to the universal covering $\mathcal{U}$ is a topological line whose stabilizer is an infinite cyclic group. Different lifts have conjugate stabilizers.*

Thus, to each (oriented) simple closed curve in B we can associate a conjugacy class in the fundamental group $\pi_1(B)$ (the generators of the above stabilizers).

EXERCISE 1.61. *There is a natural one-to-one correspondence between classes of freely homotopic* (*oriented*) *closed curves* (*not necessarily simple*) *and conjugacy classes in $G = \pi_1(B)$.*

LEMMA 1.62. *Let V be an essential submanifold in B. Then there is a covering $q : E \to B$ with $\pi_1(E) = \pi_1(V)$ and such that one of the components U of $q^{-1}(V)$ projects homeomorphically onto V.*

PROOF. In the notation of Lemma 1.59, let $E = \mathcal{U}/\Gamma$, $U = \hat{V}/\Gamma$. □

Informally speaking, we unwind all the loops in B except those that are essentially confined to V.

COROLLARY 1.63. *Let $\gamma \subset B$ be a non-trivial simple closed curve. Then there is a covering space E with $\pi_1(E) \approx \mathbb{Z}$ containing a simple closed curve $\hat{\gamma}$ that projects homeomorphically onto γ.*

1.7. Topological surfaces.

1.7.1. *Definitions and examples.*

DEFINITION 1.64. A (*topological) surface* S (without boundary) is a two-dimensional topological manifold with countable base. It means that S is a topological space with a countable base such that any $z \in S$ has a neighborhood $U \ni z$ homeomorphic to an open subset V of $\mathbb{R}^2$. The corresponding homeomorphism $\phi : U \to V$ is called a (*topological*) *local chart* on S. Such a local chart assigns to any point $z \in U$ its *local coordinates* $(x, y) = \phi(z) \in \mathbb{R}^2$.

A family of local charts whose domains cover S is called a *topological atlas* on S. Given two local charts $\phi_1 : U_1 \to V_1$ and $\phi_2 : U_2 \to V_2$, the composition

$$\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)$$

is called the *transition map* from one chart to the other.

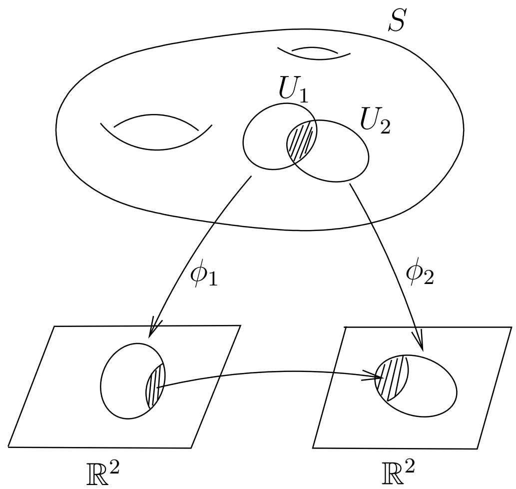

FIGURE 1.11. A transition map between two local charts.

A surface is called *orientable* if it admits an atlas with orientation preserving transition maps. Such a surface can be oriented in exactly two ways. *In what follows we will only deal with orientable* (*and naturally oriented*) *surfaces.*

Unless otherwise is explicitly said, we will assume that the surfaces under consideration are *connected*. The simplest (and most important for us) surfaces are:

• The whole plane $\mathbb{R}^2$; it is homeomorphic to the open unit disk $\mathbb{D} \subset \mathbb{R}^2$.

• The unit *sphere* S^2 in $\mathbb{R}^3$ (homeomorphic via the stereographic projection to the one-point compactification of the plane); it is also called a "closed surface of genus 0" (in this context "closed" means "compact without boundary").

• A *cylinder* or *topological annulus* $\mathrm{Cyl} = S^1 \times (0,1)$. It can also be represented as the quotient of the strip $\mathbb{S} = \mathbb{R} \times (0,1)$ modulo the cyclic group of translations $z \mapsto z+n$, $n \in \mathbb{Z}$. It is homeomorphic to any annulus $\mathbb{A}(r,R)$, as well as to the punctured disk $\mathbb{D}^*$ and to the punctured plane $\mathbb{C}^*$.

• The torus $\mathbb{T}^2 = \mathbb{T} \times \mathbb{T}$, also called a "closed surface of genus 1". It can also be represented as the quotient of $\mathbb{R}^2$ modulo the action of a rank 2 abelian group $z \mapsto z + \alpha m + \beta n$, $(m,n) \in \mathbb{Z}^2$, where $\{\alpha, \beta\}$ is an arbitrary basis in $\mathbb{R}^2$.

If we have a certain standard surface S (say, the unit disk or the unit sphere), a *topological* S (say, a "topological disk" or a "topological sphere") refers to a surface homeomorphic to the standard one.

One can also consider *bordered surfaces*, or *surfaces with boundary.* The local model for such a surface near a boundary point is given by a relative neighborhood of a point $(x,0)$ in the closed upper half-plane $\bar{\mathbb{H}}$. The orientation of a surface naturally induces an orientation of its boundary (locally corresponding to the positively oriented real line).

For instance, we can consider cylinders with boundary: $S^1 \times [0,1]$ or $S^1 \times [0,1)$. They will be still called "cylinders" or "topological annuli". Cylinders without boundary will be also called "open", while cylinders of other type will be called "closed" and "semi-closed" respectively. We will use the same notation, Cyl, for a topological cylinder of any type.

More generally, non-compact surfaces without boundary are called *open*. Compact surfaces without boundary are called *closed*.

REMARK 1.65. We are bound to live with inconsistency in using the notion of "open" and "closed", reflecting different traditions of point-set and algebraic topology. Hopefully, it will not lead to confusion.

1.7.2. *New surfaces from old ones.* There are two basic ways of building new surfaces out of old ones: making holes and gluing their boundaries. Of course, any open subset of a surface is also a surface. In particular, one can make a (closed) *hole* in a surface, that is, remove a closed Jordan disk. A topologically equivalent operation is to make a *puncture* in a surface. By removing an open Jordan disk (an open hole) we obtain a surface with boundary.

If we have two open holes (on a single surface or on two different surfaces S_i) bounded by Jordan curves γ_i, we can *glue* (or: *paste*) these boundaries together by means of an orientation reversing homeomorphism $h : \gamma_1 \to \gamma_2$ (where orientation of the γ_i is induced from S_i). We denote this operation by $S_1 \sqcup_h S_2$. For instance, by gluing together two closed disks we obtain a topological sphere: $\mathbb{D} \sqcup_h \mathbb{D} \approx S^2$.

EXERCISE 1.66. *Justify the last assertion.*

Combining the above operations, we obtain operations of taking *connected sums* and attaching a *handle*. To take a connected sum of two surfaces S_1 and S_2, make an open hole in each of them and glue together the boundaries of these holes. To attach a handle to a surface S, make two open holes in it and glue together their boundaries.

If we attach a handle to a sphere, we obtain a topological torus. If we attach g handles to a sphere, we obtain a *closed surface of genus g*.

FUNDAMENTAL THEOREM OF 2D TOPOLOGY. *Any orientable closed surface S is homeomorphic to a surface of genus g. Moreover, the genus is a complete topological invariant of a such a surface.*

See e.g., [**Mu1**] for a proof. Note that there are several ways to see that that g is a topological invariant of a closed surface:

• Identify it with $b_1/2$, where $b_1 = \operatorname{rk} H_1(S)$ is the first Betti number of S.

• Characterize it as the maximal cardinality of a family of disjoint simple closed curves γ_i on S with the property that by cutting S along these curves we obtain a topological sphere with several holes. (Such a family can be obtained by taking simple closed curves "going across (or along) the handles".)

Topological surfaces are very flexible: one can move points around at will:

EXERCISE 1.67. (i) *Let S be a (connected) topological surface with two configurations of N points marked in $\operatorname{int} S$: x_i and y_i, $i = 1, \dots, N$. Then there exists a homeomorphism $h : S \to S$ that moves x_i to y_i.*

(ii) *Make a similar assertion for boundary components of a bordered surface.*

(iii) *Let S be bordered, and let h be a homeomorphism of some boundary circles to themselves. Then h extends to a homeomorphism of S.*

We are now ready to classify all compact surfaces (closed or bordered):

Corollary 1.68. *Any orientable compact surface S with boundary is homeomorphic to a closed surface of genus g with N open holes removed. Moreover, the pair (g, N) is a complete topological invariant of a such a surface.*

Proof. Let S has N boundary components γ_i. Each of them is a topological circle. By attaching disks D_i to the γ_i ("caps"), we obtain a closed orientable surface $\mathbf{S}$.

If two compact surfaces with boundary, S and S', are homeomorphic, then they have the same number N of boundary components, and the corresponding closed surfaces $\mathbf{S}$ and $\mathbf{S}'$ are also homeomorphic. Hence the pair (g, N) is a topological invariant of compact surfaces. Exercise 1.67(ii) implies that this invariant is complete. □

Note that the first Betti number of a surface of genus g with $N > 0$ holes is equal to $2g + N - 1$.

Corollary 1.69. *The closed topological disk is the only simply connected surface with boundary. The closed topological cylinder is the only orientable compact surface with $\pi_1(S) \approx \mathbb{Z}$.*

One says that a surface S (with or without boundary) has a *finite topological type* if its fundamental group $\pi_1(S)$ is finitely generated (e.g., any compact surface is of finite type). We will see later (Theorem 1.87) that it is equivalent to saying that S is *tame*, i.e., is homeomorphic to a closed surface with finitely many open or closed holes. Clearly such a surface admits a decomposition

$$S = \mathcal{C} \cup \bigsqcup \mathrm{Cyl}_i ,$$

where $\mathcal{C}$ is a compact surface and the Cyl_i are cylinders attached to some boundary components of $\mathcal{C}$. The set $\mathcal{C} \equiv \mathcal{C}_S$ is called the *compact core* of S. Note that it is obviously a deformation retract for S. The cylinders Cyl_i represent "tame ends" of S. If a cylinder Cyl_i is not closed, we can add the outer boundary to it. Doing this to all such cylinders, we obtain the *ideal circle compactification* $\hat{S}$ of S. (Compare with §1.7.8 below.)

1.7.3. *Triangulations.* Recall that *triangulation* of a surface S is a tiling of S by topological triangles (i.e., by closed topological disks with three marked points on the boundary) such that any two triangles are either disjoint or share exactly one vertex, or share exactly one side. In case of a bordered surface, a similar requirement is imposed on any triangle Δ and any boundary circle γ: if the intersection $\Delta \cap \gamma$ is non-empty, then it is equal to an edge or to a vertex of Δ.

A triangulation is finite if and only if S is compact.

Theorem 1.70. *Any surface* (*with countable base*) *can be triangulated.*

See [**Mu2**] for a proof. In fact, all surfaces that appear in Conformal Dynamics (at least, in this book) are plane domains or their (branched) coverings, or a sphere or a torus, when the Triangulation Theorem is a simple exercise. Note also that the Triangulation Theorem is easy for smooth surfaces, in particular, for Riemann surfaces (see Proposition 2.2 below).

Given a triangulation on S, a closed subset $K \subset S$ is called *simplicial* if it is composed of triangles.

1.7.4. *Euler characteristic.* Let S be a compact surface (with or without boundary). Its *Euler characteristic* is defined as

$$\chi(S) = f - e + v,$$

where f, e and v are respectively the numbers of faces, edges and vertices in any triangulation of S. It is a topological invariant equal to the alternating sum of the Betti numbers, so for a closed surface of genus g, we have: $\chi(S) = 2 - 2g$.

The Euler characteristic is obviously additive for connected sums:

$$\chi(S_1 \sqcup_h S_2) = \chi(S_1) + \chi(S_2).$$

Since the cylinder $\mathbb{T} \times [0,1]$ has zero Euler characteristic, $\chi(\hat{S}) = \chi(\mathcal{C}_S)$ for a tame surface. (Recall from §1.7.2 that $\hat{S}$ is the ideal circle compactification of S and $\mathcal{C}_S$ is the compact core of S). We can use this as a definition of $\chi(S)$ in this case.

Making a hole in a surface drops its Euler characteristic by one; attaching a handle does not change it. Hence $\chi(S) = 2 - 2g - N$ for a surface of genus g with N holes (which also follows from the basic algebraic topology).

Note that the above list of simple surfaces is the full list of orientable surfaces of finite type without boundary with *non-negative* Euler characteristic:

$$\chi(\mathbb{R}^2) = 1, \quad \chi(S^2) = 2, \quad \chi(\mathrm{Cyl}) = \chi(\mathbb{T}^2) = 0.$$

1.7.5. *Topological ends and compactification.* A non-compact domain $F \subset S$ bounded by a simple closed curve γ in S is called a *fjord*. A *nest of fjords*, $F_0 \supset F_1 \supset \dots$, is *subordinated* to a fjord F if $F_n \subset F$ for some n. A nest of fjords F_n is *subordinated* to another nest of fjords F'_m if (F_n) is subordinated to any F'_m. Two nests are *equivalent* if each of them is subordinated to the other.

A nest (F_n) is called *escaping* if

$$\bigcap F_n = \emptyset \tag{1.5}$$

A (*topological*) *end* E of S is a class of equivalent escaping nests of fjords.

LEMMA 1.71. *Any non-compact Riemann surface S has at least one end.*

PROOF. Let us triangulate S and exhaust it by an increasing nest of compact simplicial subsets $K_0 \subset K_1 \subset \dots$, $\bigcup K_n = S$. One of the components of $S \smallsetminus K_0$ is unbounded; call it U_0. The boundary of U_0 is a polygonal curve contained in ∂K_0. Approximate it with a simple closed curve σ_0 contained in U_0, and let $F_0 \equiv F_{n_0}$ be the unbounded component of $S \smallsetminus \sigma_0$ contained in U_0.

Since the sets K_n exhaust S, there is a set K_n containing σ_0. Then one of its complementary components, U_n, is contained in F_0. Approximating ∂U_n with a simple closed curve $\sigma_n \subset U_n$, we construct a new fjord $F_n \equiv F_{n_1} \subset F_0$.

Proceeding this way, we construct an escaping nest of fjords

$$F_{n_0} \supset F_{n_1} \supset \dots$$

representing an end of S. □

COROLLARY 1.72. *Any non-compact domain $U \subset S$ with compact boundary ∂U contains an escaping nest of fjords, $F_0 \supset F_1 \supset \dots$, representing an end of S.*

We let $\partial^T S$ be the set of ends of S and $\mathrm{cl}^T S = S \sqcup \partial^T S$. We call them the *topological end boundary* and the *topological end compactification* respectively. Endow $\mathrm{cl}^T S$ with the following topology. For a fjord F, let $\mathcal{U}(F)$ be the union of

F and the set of ends E represented by nests of fjords (F_n) that are subordinated to F. The family of sets $\mathcal{U}(F)$, together with the family of open sets of S, form a base of neighborhoods of $\mathrm{cl}^T(S)$.

EXERCISE 1.73. *Show that:*

(i) *This is indeed a base of topology.*

(ii) $\mathrm{cl}^T S$ *is compact and metrizable.*

(iii) *The boundary* $\partial^T S$ *is totally disconnected.*

(iv) *The end compactification is topologically natural: any homeomorphism* $h : S \to S'$ *extends to a homeomorphism* $h_T : \mathrm{cl}^T S \to \mathrm{cl}^T S'$.

We say that a map $f : S \to S'$ *properly maps* an end E of S to an end E' of S' if $f(z) \to E'$ as $z \to E$. In other words, f extends continuously to E by mapping it to E'.

EXERCISE 1.74. *Any proper map* $f : S \to S'$ *extends continuously to a map* $f_T : \mathrm{cl}^E S \to \mathrm{cl}^E S'$.

1.7.6. *Remark on a general framework for the notion of end.* The above notion of end can be embedded in a quite general framework as follows. Let S be a locally compact topological space, and let $\mathcal{F}$ be a family of open non-precompact subsets called *fjords* satisfying the following properties:

• The intersection of any two fjords is either precompact or contains a fjord;

• For any compact subset $Q \subset S$, there is a finite family of fjords F_j such that $S \smallsetminus \bigcup F_j$ is a compact set containing Q.

For *nests of fjords*, define *escaping*, *subordination*, and *equivalence* as in §1.7.5.

Then we define an *end* E of $\mathcal{F}$ as a class of equivalent escaping nests of fjords.

We let $\partial^{\mathcal{F}} S$ be the set of ends of S and $\mathrm{cl}^{\mathcal{F}} S = S \sqcup \partial^{\mathcal{F}} S$. It is endowed with a natural topology as above, which provides us with the $\mathcal{F}$*-end compactification* of S (associated with the family $\mathcal{F}$).

EXERCISE 1.75. *Check that* $\mathrm{cl}^{\mathcal{F}} S$ *is a compact topological space.*

If S is a domain in some manifold $\hat{S}$, then there can be a good relation between $\mathrm{cl}^{\mathcal{F}} S$ and the closure $\bar{S}$ of S in $\hat{S}$. A relevant notion is the *impression* $I(E)$ of an end E,

$$I(E) = \bigcap \bar{F}_n,$$

where (F_n) is any nest of fjords representing E (the outcome is obviously independent of the choice of the nest),

EXERCISE 1.76. (i) *If all the end impressions* $I(e)$ *are pairwise disjoint, then there is a natural continuous projection* $\pi : \bar{S} \to \mathrm{cl}^{\mathcal{F}} S$ *extending the identical map* $S \to S$. *Fibers of* π *over* $y \in \partial^{\mathcal{F}} S$ *are end impressions.*

(ii) *Under the above circumstances, if all the end impressions* $I(e)$ *are singletons then* π *is a homeomorphism.*

The simplest application of this scheme is the construction of *one-point compactification* for a locally compact space S, where fjords are defined as arbitrary non-precompact open subsets. Note also that the topological end compactification described in §1.7.5 can be generalized to arbitrary manifolds S, using as fjords non-precompact domains whose topological boundaries in S are compact manifolds. In this book, we will encounter two more important occasions: *prime end* (*Carath'eodory*) and *puzzle end* compactifications.

1.7.7. *Simply connected surfaces.*

LEMMA 1.77. *Let S be an open simply connected surface. Then:*
(i) *S has one end;*
(ii) *Any simple closed curve σ on S is dividing;*
(iii) *Any proper arc γ on S is dividing.*

PROOF. (i) Being open, S must have at least one end (by Lemma 1.71). Assume it has two different ends, E_0 and E_1. Let F_0 and F_1 be fjords with disjoint closures representing the corresponding ends, and let $\sigma_i := \partial F_i$. Construct a proper arc $\gamma : (0, 1) \to S$ landing at E_0 as $t \to 0$ and landing at E_1 as $t \to 1$, and crossing each curve σ_i at a single point. Since the intersection number between closed curves and proper arcs is invariant under proper homotopy, the curves σ_i cannot be trivial – contradiction.

(ii) The proof is similar: If σ were non-dividing then there would be a proper arc γ on S crossing σ at a single point (with both ends of γ landing at the end of S). This would imply that σ were non-trivial.

(iii) Similarly: if a proper arc γ was non-dividing then there would exist a simple closed curve σ crossing γ at a single point. □

We can define a *hull* in S as in the case of $\mathbb{R}^2$, as a continuum with connected complement.

LEMMA 1.78. *Any simplicial continuum $J \subset S$ can be filled-in to a simplicial hull K, which is the smallest hull containing J.*

PROOF. Since S has only one end, all but one components of $S \smallsetminus J$ are bounded (by Corollary 1.72) . Call the unbounded component U. Its boundary $\sigma := \partial U$ is a polygonal curve that can be approximated by a simple closed curve $\sigma \subset U$. For the same reason, there is only one unbounded complementary component of σ; call it F. For the same reason, the complement of $\bar{F}$ is bounded, so it comprises only finitely many triangles. Adding all these triangles to J, we obtain the desired hull K. □

LEMMA 1.79. *For a given triangulation of S, any simplicial hull $K \subset S$ is simply connected.*

PROOF. Any simplicial hull is composed of finitely many triangles, so let us do induction in the number of triangles. The statement is obvious for one triangle. To pass from $\leq n$ to $n + 1$, let us consider two cases:

Case 1. Assume there is a cut-point in K, i.e., a vertex $a \in K$ whose removal disconnects K. Then there are at least two accesses from $S \smallsetminus K$ to a. Since $S \smallsetminus K$ is connected, there are two proper topological rays $\mathcal{R}_\pm : \mathbb{R}_\pm \to S \smallsetminus K$ landing at a as $t \to 0$ and landing at ∞ as $t \to \pm\infty$ (respectively). The union $\gamma := \mathcal{R}_- \cup \mathcal{R}_+ \cup \{a\}$ is a proper arc. By Lemma 1.77 (iii), $S \smallsetminus \gamma$ consists of two components, U_1 and U_2.

Let $\mathcal{L}_i^* := K \cap U_i$ and $\mathcal{L}_i := \bar{\mathcal{L}}_i$ $(i = 1, 2)$. Since γ cuts through K at a, the sets $\mathcal{L}_i^*$ are non-empty. Hence each $\mathcal{L}_i$ is a simplicial hull composed of $\leq n$ triangles.

By induction assumption, each of them is simply connected. Then so is the whole hull K.

Case 2. Assume K does not have cut-points. Take any boundary triangle $\Delta \subset K$, and let $K' := \mathrm{cl}(K \smallsetminus \Delta)$. The latter is a simplicial hull composed of n triangles. Moreover, by the *no-cut-points assumption*, $I := \Delta \cap K'$ is the union of one or two edges of Δ. In either case, I is a deformation retract for Δ, and hence K' is a deformation retract for K. The conclusion follows. $\square$

LEMMA 1.80. *An open simply connected surface S can be exhausted by topological disks, i.e., there exist an increasing nest of closed topological disks $\bar{D}_n \subset S_n$ such that $D_n \Subset D_{n+1}$ and $\bigcup D_n = S$.*

PROOF. Let us triangulate S and select a base triangle Δ_0. Let us define Δ_{n+1} inductively as the union of Δ_n with all triangles attached to it. In this way, we exhaust S with an increasing nest of simplicial continua.

By Lemma 1.78, Δ_n can be filled-in to a simplicial hull $K_n := \hat{\Delta}_n$. The K_n provide us with an increasing nest of hulls exhausting S. The unbounded components U_n of $S \smallsetminus K_n$ form an escaping nest of domains in S.

Let $\sigma_n \subset U_n$ be a simple closed curve approximating ∂U_n, and let $F_n \subset U_n$ be the unbounded component of $S \smallsetminus \sigma_n$. They form an escaping family of fjords. Selecting a converging subsequence if needed, we turn the F_n into a decreasing nest.

The complement $\bar{D}_n := S \smallsetminus F_n$ is a compact surface with boundary σ_n (where $D_n = \mathrm{int}\, \bar{D}_n$) . Moreover, $\bar{D}_n$ can be retracted onto K_n, so it is simply connected by Lemma 1.79. By the Fundamental Theorem of 2D Topology (Corollary 1.69), $\bar{D}_n$ is a closed topological disk. $\square$

COROLLARY 1.81. *An open simply connected surface is a topological disk.*

Applying the Fundamental Theorem of 2D Topology once again, we conclude:

THEOREM 1.82. *There are only two simply connected topological surfaces without boundary: a topological disk and a topological sphere.*

1.7.8. *Tame ends.* An end E is called *tame* if eventually all the fjords F_n (in some and hence in any nest (F_n) representing E) are cylinders. Any of these cylinders uniquely determines the corresponding end.

A tame end compactification results in completion each of the cylinders F_n to a topological disk $D_n = F_n \cup \{E\}$. Thus, $\mathrm{cl}^T S$ is a topological surface near E. Under these circumstances, E is referred to as a *puncture at infinity* for S.

A tame end can also be compactified in a different way by attaching a topological circle at infinity called the *ideal circle* (at infinity) (compare §1.7.2). However, this compactification is not topologically or smoothly natural:

EXERCISE 1.83. *Construct a diffeomorphism $f : \mathbb{D} \to \mathbb{D}$ that does not extend continuously to any point of $\mathbb{T} = \partial\mathbb{D}$.*

REMARK 1.84. In the conformal category, compactifying an end with an ideal circle is a natural operation, and in fact, a conformal end knows exactly what should be attached to it: see Theorem 2.59 and §§2.6.1, 5.6. below.

1.7.9. *Open surfaces of finite type.*

LEMMA 1.85. *For a surface S of finite type, the boundary $\partial^T S$ is finite.*

PROOF. Assume $\partial^T S$ is infinite. Since it is totally disconnected and compact, for an arbitrary big $N \in \mathbb{Z}_+$ there is a family of N disjoint fjords F_n whose closures in $\mathrm{cl}^T S$ cover $\partial^T S$. Then the complement $S \smallsetminus \bigcup F_n$ is a compact surface with boundary $\bigsqcup \partial F_n$ comprising N simple closed curves. The first Betti number of such a surface is at least $N-1$. On the other hand, it is bounded by $b_1(S)$, which is finite by assumption – contradiction. □

LEMMA 1.86. *For an open surface S of finite type, any end is tame.*

PROOF. Let $F_0 \Supset F_1 \Supset \dots$ be a nest of fjords representing an end $E \in \partial^T S$. Since E is isolated in $\partial^T S$, eventually the $\bar{F}_n$ do not have any other ends, so we can assume this holds already for F_0. Then the bordered surfaces $A_n := \bar{F}_n \smallsetminus F_{n+1}$ are compact. If infinitely many of them had positive genus, the genus of compact bordered surfaces $\bar{F}_n \smallsetminus F_0$ would grow to ∞, contradicting to boundedness of their first Betti number. Hence eventually the A_n have zero genus. As they have two ends, these A_n are topological cylinders (Corollary 1.69 from the Fundamental Theorem of 2D Topology). Hence the fjords $\bar{F}_n = \bigcup_{m \geq n} A_m$ are topological cylinders as well. □

Now the Fundamental Theorem of 2D Topology and its Corollary 1.68 for compact surfaces can be refined for general surfaces of finite type:

THEOREM 1.87. *Any surface S of finite type (with or without boundary) is homeomorphic to a compact surface with finitely many punctures.*

PROOF. Indeed, the compactification $\mathrm{cl}^T S$ amounts to adding to S finitely many tame ends (punctures at infinity), resulting in a compact surface. □

Let us conclude with characterization of open topological annuli (compare with Corollary 1.69):

COROLLARY 1.88. *An orientable open surface S is a topological annulus iff*

$$\pi_1(S) \approx \mathbb{Z}.$$

1.7.10. *Multicurves and pair of pants decompositions.* Note that a simple closed curve γ on a surface S[7] is trivial iff one of the components of $X \smallsetminus \gamma$ is a topological disk.

A simple closed curve on $\operatorname{int} S$ is called *peripheral* if it is either trivial or dividing in such a way that one of the components (call it F) of $S \smallsetminus \gamma$ is a topological cylinder. This cylinder can be open or semi-closed. In the former case, F is a fjord of a tame end. In the latter case, γ is homotopic to a boundary component of S (which is the boundary component of F).

For instance, if $S = S^2 \smallsetminus \{x_i\}$ is a sphere with finitely many punctures then γ is non-peripheral iff each component of $S^2 \smallsetminus \gamma$ contains at least two punctures.

[7]For purposes of this section, any simple closed curve $\gamma : \mathbb{T} \to S$ can be assumed extendable to an embedding $\mathbb{A}(r^{-1}, r) \to S$, $r > 1$. It ensures, without using the Schönflies Theorem, that γ can be attached as a boundary circle to any complementary component.

Exercise 1.89. *If S contains a non-peripheral curve γ, then $\chi(S) < 0$. Moreover, each component of $S \smallsetminus \gamma$ has negative Euler characteristic as well.*

A *pair of pants* is a surface S homeomorphic to the disk with two holes. We can consider open, closed, and semi-closed pairs of pants. For instance, an open one is modeled on $S \approx \mathbb{D} \smallsetminus (\bar{\mathbb{D}}(a,r) \cup \bar{\mathbb{D}}(-a,r))$, where e.g., $a = 1/2$, $r = 1/4$, while a closed pair of pants is obtained by adding the boundary (three circles) to it. Notice that the Euler characteristic of a pair of paints is equal to -1.

Lemma 1.90. *Let S be a surface of finite type with negative Euler characteristic. If S is not a topological pair of pants, then it contains a non-peripheral curve γ.*

Proof. Completing the ends of S with circles, we obtain a compact Riemann surface. Then

$$S \approx \mathbf{S} \smallsetminus \bigsqcup_{i=1}^{N} D_i,$$

where $\mathbf{S}$ is a closed surface of genus g and D_i are open topological disks with disjoint closures. Let $\sigma_i := \partial D_i$.

If $\mathbf{S}$ has a positive genus, then we can let γ be a non-dividing curve on $\mathbf{S}$ going along a handle avoiding the $\bar{D}_i$. So, assume $\mathbf{S}$ has zero genus, i.e., $\mathbf{S} \approx S^2$. Then $N \geq 4$ since $\chi(S) < 0$. Consider two of the holes, say D_1 and D_2, and connect them with an arc $\alpha \subset S$. We obtain a graph $\Gamma \subset S$ comprising α and two circles, σ_1 and σ_2. Thicken this graph slightly to obtain a Jordan disk Δ. It is bounded by a desired non-peripheral curve γ. $\square$

A *multicurve* $\mathcal{G} = \bigsqcup \gamma_i$ is a collection of pairwise disjoint non-peripheral simple closed curves γ_i on S.

Theorem 1.91. *Let us S be a surface of finite type (g, N) with negative Euler characteristic: $\chi < 0$. Then any multicurve on S has at most*

$$l := (3|\chi| - N)/2 = 3g - 3 + N$$

components. Moreover, any multicurve is contained in a maximal one comprising exactly l components. They divide S into $|\chi|$ pairs of pants.

Proof. By Lemma 1.90, if S is not a pair of pants, then it contains a non-peripheral curve γ. Let S' be a component of $S \smallsetminus \gamma$. If γ is non-dividing then

$$g(S') = g(S) - 1, \quad \chi(S') = \chi(S).$$

Otherwise

$$g(S') \leq g(S), \quad |\chi(S')| < |\chi(S)|,$$

where the second inequality follows from the additivity of the Euler characteristic and Exercise 1.89. In either case,

$$g(S') + |\chi(S')| < g(S) + |\chi(S)|. \tag{1.6}$$

If S' is not a pair of pants, then we can apply to it the same cutting procedure, and so on. Proceeding this way, we construct an increasing sequence of multicurves $\mathcal{G}_n$ that cut S into pieces S_i^n such that $g(S_i^n) + |\chi(S_i^n)|$ decays with n. Hence this process must stop in finite time, producing a maximal multicurve $\mathcal{G} \equiv \mathcal{G}_m$. At this final moment, all the components of $S \smallsetminus \mathcal{G}$ must be pairs of pants.

Of course, we can turn on this procedure starting with any given multicurve $\mathcal{G}_0$. It produces a maximal multicurve $\mathcal{G}$ containing $\mathcal{G}_0$.

By additivity of the Euler characteristic, the maximal multicurve $\mathcal{G}$ divides S in exactly $|\chi|$ pairs of pants. Their disjoint union has $3|\chi|$ boundary components. Out of them, $3|\chi|-N$ are glued in pairs, producing $(3|\chi|-N)/2$ curves comprising $\mathcal{G}$. $\square$

Corollary 1.92. *Let $S \approx S^2 \smallsetminus \mathcal{Z}$ where $\mathcal{Z}$ is a set of $N \geq 3$ punctures. Then any multicurve on S has at most $N-3$ components. Moreover, any multicurve is contained in a maximal one comprising exactly $N-3$ components.*

We can reverse the above cutting procedure. Take a finite family of pairs of pants, and paste together some pairs of their boundary components. We obtain a surface of finite type with negative Euler characteristic. Theorem 1.91 shows that in this way we can obtain an arbitrary such a surface.

In fact, it can be generalized to surfaces of infinite type:

Theorem 1.93. *Any surface S with negative Euler characteristic can be obtained by pasting together some family (finite or countable) of pairs of pants.*

This result applies to a surface of any type: closed, open, or bordered, as long as we make use of pairs of pants of various types.

Proof. Since $\partial^T S$ is compact and totally disconnected, there is a nest of neighborhoods $U_0 \supset U_1 \supset \cdots \supset \partial^T S$ such that $\bigcap U_n = \partial^T S$ and each $U_n \smallsetminus \partial^T S$ is a finite union of fjords. Then compact bordered surfaces $S_n := S \smallsetminus U_n$ form an increasing nest exhausting S. Decomposing S_0 and each $S_{n+1} \smallsetminus S_n$ into pairs of pants (by Theorem 1.91), we obtain a desired decomposition for S. $\square$

Note in conclusion that we will also need more general *pants* defined as topological disks with finitely many (more than one) holes.

1.7.11. *Arc diagrams.* A similar theory can be developed for proper arcs instead of simple closed curves. Let us first assume that S is an open surface of finite type. According to general definitions, a *proper arc* on S is an embedding $\alpha : (-\infty, \infty) \to S$ such that $\alpha(t) \to \partial^T S$ as $|t| \to \infty$. More generally, we can talk about *proper curves* on S. Both ends of a proper curve land at some points of the ideal boundary $\partial^T S$, so a proper curve "connects" points at infinity.

Two proper arcs/curves on S are called *properly homotopic* if they are homotopic through a family of proper curves. A proper arc/curve is called *trivial* if it is homotopic to an arc contained in an arbitrary small neighborhood of some point at infinitely.

A *multi-arc* on a surface S is a family of disjoint non-trivial proper arcs representing different proper homotopy classes. The corresponding family of proper homotopic classes in called an *arc diagram* on S.

In case of a bordered surface S of finite type, we will slightly modify this notion. A *proper arc/curve* $\alpha : I \to S$ will mean not only that $\alpha : I \to S$ is a proper map, but also that $\alpha(\operatorname{int} I) \subset \operatorname{int} S$ and $\alpha | \operatorname{int} I$ is a proper arc/curve in $\operatorname{int} S$ (in the above sense); in particular, $f(\partial I) \subset \partial S$.[8] A *proper homotopy* is then understood as a homotopy through proper curves, (so the endpoint of the curve can slide along the boundary of S).

Similarly to Theorem 1.91, we have:

[8]Here $\operatorname{int} I$ and ∂I are understood in the intrinsic way, e.g., $\partial(0, 1] = \{1\}$.

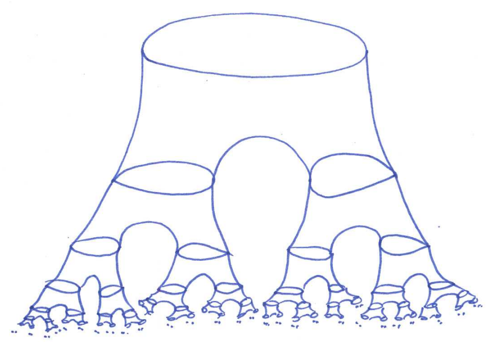

FIGURE 1.12. Pair of pants decomposition for a disk with a Cantor set removed.

PROPOSITION 1.94. *Let S be a bordered surface of finite type with negative Euler characteristic: $\chi < 0$. Then any multi-arc on S contains at most $3|\chi|$ arcs. In particular, for a disk with $N \geq 2$ holes we obtain at most $3(N-1)$ arcs.*

PROOF. Topologically, $\operatorname{int} S$ is a closed surface $\mathbf{S}$ with $k > 0$ punctures. The canonical arc diagram $\mathcal{A}$ can be realized as a net of disjoint edges on $\mathbf{S}$ connecting the punctures. This net can be completed to a triangulation of $\mathbf{S}$ whose only vertices are punctures. Let t and $e \geq |\mathcal{A}|$ be respectively the number of triangles and edges of this triangulation. Then $3t = 2e$, and by the Euler Formula,

$$-\frac{1}{3}e = t - e = \chi(\mathbf{S}) - k = \chi(S).$$

$\square$

1.7.12. *Plane domains.* A *plane domain* U is a domain in the sphere S^2 which is different from the whole sphere. Realizing S^2 as the one-point compactification of $\mathbb{R}^2$, we can place U inside $\mathbb{R}^2$.

LEMMA 1.95. *For a plane domain U, there is a natural one-to-one correspondence between the ends of U and connected components of the complement $K := S^2 \smallsetminus U$. Moreover, if Q is the component corresponding to an end E then ∂Q is equal to the impression $I((E)$.*

PROOF. Let (F_n) be a nest of fjords of U representing some end $E \in \partial^T U$, and let $\sigma_n := \partial F_n$. Each σ_n is a Jordan curve, so it bounds a Jordan disk D_n containing F_n. Moreover, these disks are strictly nested: $D_0 \Supset D_1 \Supset \dots$, so $Q := \bigcap D_n = \bigcap \bar{D}_n$ is a hull or a singleton. We have

$$Q \cap U = \bigcap (D_n \cap U) = \bigcap F_n = \emptyset,$$

so Q is contained in K. Since any point of $K \smallsetminus Q$ is separated from Q by some curve $\sigma_n \subset U$, Q is a connected component of K.

Vice versa, any connected component Q of K is a hull or a singleton, so it is cellular by Proposition 1.20. Hence there is a sequence of Jordan curves $\sigma_n \subset U$ that bound a nest of Jordan disks shrinking to K. Fjords $F_n := D_n \cap U$ represent the end E of U corresponding to Q.

We let the reader to verify the last assertion. $\square$

In particular, a bounded topological annulus $A \subset \mathbb{R}^2$ has two complementary components, the unbounded component K^o called *outer*, and the bounded component K^i called *inner*. Respectively, it has two boundary components, the *outer boundary* $\partial^0 A \subset K^o$, and the *inner boundary* $\partial^i A \subset K^i$.

The following statement shows that the end compactification of a plane domain is a sphere obtained by collapsing all the complementary components to singletons (compare Exercise 1.76):

PROPOSITION 1.96. *For any plane domain $U \subset S^2$, there is a continuous surjection $h : S^2 \to S^2$ that restricts to a homeomorphism $h : U \to h(U)$ and maps the complement $K := S^2 \smallsetminus U$ onto a totally disconnected set $h(K)$ so that the fibers of $h : K \to h(K)$ are connected components of K.*

PROOF. Consider a pair of pants decomposition (P_i^n) for U. Construct a pair of pants family (Δ_i^n) with the same combinatorics, where the Δ_i^n are bounded by circles or points (with the points corresponding to tame ends of the P_i^n). Moreover, the construction can be arranged so that $\operatorname{diam} \Delta_i^n \to 0$ as $n \to \infty$. Then the pairs of pants Δ_i^n tile a plane domain V whose complement is totally disconnected.

Construct now a homeomorphism $h : U \to V$ that maps P_i^n to the corresponding Δ_i^n, tile by tile. It induces a homeomorphism $\hat{h}$ between the end compactifications of U and V. By Lemma 1.95, each end e of U corresponds to a connected component Q_e of $S^2 \smallsetminus U$. The corresponding end $\hat{h}(e)$ corresponds to a connected component of $S^2 \smallsetminus V$, which is a singleton z_e. Extends h to the whole sphere by collapsing each Q_e to the corresponding point z_e. This provides us with a desired map. $\square$

COROLLARY 1.97. *For a plane domain U, the end compactification $\operatorname{cl}^T U$ is a topological sphere S^2.*

Corollary 1.98. *Let U and $\tilde{U}$ be two plane domains with totally disconnected complements. Then any homeomorphism $h : U \to \tilde{U}$ extends to a homeomorphism $\hat{h} : S^2 \to S^2$.*

Exercise 1.99. *Let $U = \mathbb{D} \setminus K$ where $K \subset \mathbb{D}$ is a Cantor set. Show that there is a homeomorphism $h : \mathbb{D} \to \mathbb{D}$ that maps K to the standard triadic Cantor set.*

1.7.13. *Surface coverings.* Specializing Lemma 1.62 and its Corollary to the surface case, we obtain:

• To any essential domain $V \subset S$ corresponds a surface covering $q : \hat{S} \to S$ such that $\pi_1(\hat{S}) = \pi_1(V)$ and one of the components $\hat{V}$ of $q^{-1}(V)$ projects homeomorphically onto V.

• If S is orientable, then to any non-trivial simple closed curve $\gamma \subset S$ corresponds an annulus covering $q : \mathbb{A}_\gamma \to S$ containing a non-trivial simple closed curve $\hat{\gamma} \subset \mathbb{A}_\gamma$ (an "equator") that projects homeomorphically onto γ. (We make use of Corollary 1.88.)

Exercise 1.100. (i) *Let $S_d \to \mathrm{Cyl}$ be a covering of finite degree $d \in \mathbb{Z}_+$ over a topological cylinder* Cyl. *Then S_d is a topological cylinder as well. For each degree $d \in \mathbb{Z}_+$, there is only one such a covering (up to covering equivalence). Write down a model for each of these coverings.*

(ii) *There is only one infinite degree covering $S_\infty \to \mathrm{Cyl}$ over* Cyl, *which is the Universal covering. Here $S_\infty \approx \mathbb{S} \equiv \{0 < \operatorname{Im} z < 1\}$ is a topological strip. Write down a model for this covering.*

(iii) *The above coverings constitute the full list of coverings over* Cyl *(up to covering equivalence). All of them are Galois.*

(iv) *For a finite degree $d \in \mathbb{Z}_+$, any homeomorphism $h : \mathrm{Cyl} \to \mathrm{Cyl}$ admits d lifts to a homeomorphism $H : S_d \to S_d$. The lift is determined by the value $H(e)$ (selected arbitrary in the appropriate fiber) at any point $e \in S_d$. Similarly, for the Universal covering there are infinitely many lifts.*

1.7.14. *Surface branched coverings.* Topological proper maps are defined in §50.3.2.

Exercise 1.101. *Assume that S and T are precompact domains in some ambient surfaces and that $f : S \to T$ admits a continuous extension to the closure $\bar{S}$. Then f is proper if and only if $f(\partial S) \subset \partial T$.*

Exercise 1.102. *Let $V \subset T$ be a domain and $U \subset S$ be a component of $f^{-1}V$. If $f : S \to T$ is proper, then the restriction $f : U \to V$ is proper as well.*

Let now S and T be topological surfaces, and f be a *topologically holomorphic map.* The latter means that for any point $a \in S$, there exist local charts $\phi : (U, a) \to (\mathbb{C}, 0)$ and $\psi : (V, f(a)) \to (\mathbb{C}, 0)$ such that $\psi \circ f \circ \phi^{-1}(z) = z^d$, where $d \in \mathbb{N}$. The number $d \equiv \deg_a f$ is called the (*local*) *degree* of f at a. If $\deg_a f > 1$, then a is called a *branched* or *critical point* of f, and $f(a)$ is called a branched or critical *value* of f. We also say that d is the *multiplicity* of a as a preimage of $b = f(a)$.

Exercise 1.103. *Show that a continuous map $f : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ that restricts to a covering $\mathbb{D}^* \to \mathbb{D}^*$ is topologically holomorphic.*

A basic property of topologically holomorphic proper maps is that they have a *global degree*:

Proposition 1.104. *Let $f : S \to T$ be a topologically holomorphic proper map between two surfaces. Assume that T is connected. Then all points $b \in T$ have the same (finite) number of preimages counted with multiplicities. This number is called the degree of f,* $\deg f$.

Proof. Since the fibers of a topologically holomorphic map are discrete, they are finite. Take some point $b \in T$, and consider the fiber over it, $f^{-1}b = \{a_i\}_{i=1}^{l}$. Let $d_i = \deg_{a_i} f$. Then there exists a neighborhood V of b and neighborhood U_i of a_i such that any point $z \in V$, $z \neq b$, has exactly d_i preimages in U_i, and all of them are unbranched.

Let us show that if V is sufficiently small then all preimages of $z \in V$ belong to $\bigcup U_i$. Otherwise there would exist sequences $z_n \to b$ and $\zeta_n \in S \smallsetminus \bigcup U_i$ such that $f(\zeta_n) = z_n$. Since f is proper, the sequence $\{\zeta_n\}$ would have a limit point $\zeta \in S \smallsetminus \bigcup U_i$. Then $f(\zeta) = b$ while ζ would be different from the a_i – contradiction.

Thus, all points close to b have the same number of preimages (counted with multiplicities) as b, so that this number is locally constant. Since T is connected, this number is globally constant. □

Corollary 1.105. *Topologically holomorphic proper maps are surjective.*

The above picture for proper maps suggests the following generalization. A topologically holomorphic map $f : S \to T$ between two surfaces is called a *branched covering* of degree $d \in \mathbb{N} \cup \{\infty\}$ if any point $b \in T$ has a neighborhood V with the following property. Let U_i be the components of $f^{-1}V$. Then each U_i contains a single preimage a_i of b and and there exist maps $\phi_i : (U_i, a_i) \to (\mathbb{C}, 0)$ and $\psi : (V, b) \to (\mathbb{C}, 0)$ such that $\psi \circ f \circ \phi_i^{-1}(z) = z_i^d$. Moreover, $\sum d_i = d$. (A branched covering of degree 2 will be also called a *double branched covering*.)

We see that *a topologically holomorphic map is proper if and only if it is a branched covering of finite degree.* All terminology developed above for proper maps immediately extends to arbitrary branched coverings.

Note that if $V \subset T$ is a domain which does not contain any critical values, then the *map f is unbranched over* V, i.e., its restriction $f^{-1}V \to V$ is a covering map. In particular, if V is simply connected, then $f^{-1}V$ is the union of d disjoint domains U_i each of which homeomorphically projects onto V. In this case we have d well-defined *branches* $f_i^{-1} : V \to U_i$ of the inverse map. (When it does not lead to confusion, we will often use notation f^{-1} for the inverse branches.)

Exercise 1.106. *Let $f : (U, a) \to (V, b)$ be a branched covering of local degree d at a. Assume U is connected and V is simply connected. If* $\deg f > d$ *then f has a critical value $v \neq b$.*

Remark 1.107. We will follow the following terminological convention. Given a map $f : S \to T$ and two open subsets $U \subset S$, $V \subset T$, we say that $f(U)$ *covers V with degree d* if the map

$$f : U \cap f^{-1}(V) \to V \tag{1.7}$$

is a degree d branched covering. In particular, if $d = 2$ the "$f(U)$ double (or, two-to-one) covers V"; if $d = 1$ then "$f(U)$ univalently covers V". If $U \cap f^{-1}(V)$ does not have critical points, then we say that " $f(U)$ covers V without branching". If

$\overline{U}$ and $\overline{V}$ are compact (e.g., closed Jordan disks), we will also say that $f(\overline{U})$ *covers* $\overline{V}$ *with degree* d.

Similarly to the unbranched situation (see §1.6.1), a branched covering $f : S \to T$ is called *Galois* or *regular* if there is a discrete group Γ of *deck transformations* faithfully acting on S such that the fibers of f coincide with the orbits of Γ. A point $c \in S$ is branched for f iff it is fixed for some non-trivial deck transformation $\gamma \in \Gamma$. Moreover, the local degree at c is equal to the order of $\mathrm{Stab}(c)$.

EXERCISE 1.108. *Show that* Γ *acts on* S *properly discontinuous and that* $\mathrm{Stab}(x)$ *is cyclic for any* $x \in S$. *Vice versa, if* Γ *is a discrete group acting properly discontinuously on* S *then the natural projection* $S \to S/\Gamma$ *is a Galois branched covering.*

As in the unbranched case, Γ is called the *covering group.*

1.7.15. *Marking.* A surface S can be *marked* with an extra topological data. It can be either several marked points $a_i \in S$, or several closed curves $\gamma_i \subset S$ up to homotopy (usually but not always they form a basis of $\pi_1(S)$), or a parametrization of several boundary components $\gamma_i \subset \partial S$, $\phi_i : \mathbb{T} \to \gamma_i$.

The marked objects may or may not be distinguished (for instance, two marked points or the generators of π_1 may be differently "colored"). Accordingly, the marking is called *colored* or *uncolored.*

A homeomorphism $h : S \to \tilde{S}$ between marked surfaces should respect the marked data: marked points should go to the corresponding points ($h(a_i) = \tilde{a}_i$), marked curves γ_i should go to the corresponding curves $\tilde{\gamma}_i$ up to homotopy ($h(\gamma_i) \simeq \tilde{\gamma}_i$), and the boundary parametrizations should be naturally related ($h \circ \phi_i = \tilde{\phi}_i$).

1.8. Orbifolds.

1.8.1. *General notion.* This notion accounts to varieties with simple singularities represented as local quotients of manifolds by finite group actions. More precisely, an *n-dimensional orbifold* $\mathcal{O}$ is a topological space M covered with a base $\mathcal{U}$ of neighborhoods U_i such that:

(i) The atlas $\mathcal{U}$ is closed under finite intersections;

(ii) For each U_i, there exists a homeomorphism $\phi_i : U_i \to \hat{U}_i/G_i$, where $\hat{U}_i$ is a neighborhood in $\mathbb{R}^n$ and G_i is a finite group of homeomorphisms acting on $\hat{U}_i$;

(iii) If $U_i \subset U_j$ then the group G_i embeds into G_j so that there is an equivariant embedding $\phi_{ij} : \hat{U}_i \to \hat{U}_j$, called a *transit map*, that induces the natural embedding $U_i \subset U_j$.

The space M is called the *underlying space*, the groups G_i are called the *(local) orbifold groups*, while the natural projections $\pi_i : \hat{U}_i \to U_i$ are called the *(inverse) local charts* of $\mathcal{O}$. Sometimes, we will informally refer to neighborhoods $\hat{U}_i$ themselves (endowed with the orbifold group actions) as "local charts". The singular set $S \subset M$ is the union of $\pi_i(\hat{S}_i)$, where $\hat{S}_i$ is the set of fixed points of the G_i-action.

A basic example is a global quotient $\hat{M}/G$ of a manifold $\hat{M}$ by a *global* properly discontinuous group action, but not all orbifolds are obtained this way.

A *morphism* $\mathcal{O} \to \mathcal{O}'$ between orbifolds is a continuous map $f : M \to M'$ between the corresponding underlying spaces that locally lifts to an equivariant continuous map $\hat{U}_i \to \hat{U}_j$ between local charts. The notion of an *orbifold homeomorphism* naturally follows. Note that an orbifold homeomorphism induces a homeomorphism between the singular sets that locally lifts to a conjugacy between the orbifold group actions. In particular, the local degrees of the singular points are preserved under orbifold homeomorphisms.

Note also that for any neighborhood $U \subset M$, there is a natural orbifold restriction $\mathcal{O}|\,U$.

An orbifold morphism $\mathcal{O} \to \mathcal{O}'$ is called a *covering of degree* $d \in \mathbb{Z}_+ \cup \{\infty\}$ if any point $y \in M'$ has a neighborhood (V, y) whose full preimage is a disjoint union of neighborhoods (U_i, x_i) such that each restriction $f : (U_i, x_i) \to (V, y)$ induces an orbifold homeomorphism $\mathcal{O}|\,U_i \to \mathcal{O}'|\,V$ such that

$$\sum [G_i : \Gamma] = d,$$

where G_i are the orbifold groups for $\hat{U}_i$ and Γ is the orbifold group for $\hat{V}$. In this case, the $\hat{U}_i$ and $\hat{V}$ can be identified (by some homeomorphisms) and the groups G_i can be embedded into Γ so that the map $\hat{U}_i/G_i \to \hat{V}/\Gamma$ induced by f becomes the natural projection $\hat{U}_i/G_i \to \hat{U}_i/\Gamma$. The index $[\Gamma : G_i]$ is called the *local degree* $\deg_{x_i} f$. For instance, if $\mathcal{O}$ is a manifold, then a covering $f : \mathcal{O} \to \mathcal{O}'$ locally looks like the quotient $\hat{U} \to \hat{U}/G$ with the local degree equal to the order of G. A *Universal orbifold covering* $p : \mathcal{U} \to \mathcal{O}$ is an orbifold covering such that any other orbifold covering $\pi : \mathcal{W} \to \mathcal{O}$ with the same base is a factor of p, i.e., there exists an orbifold covering $q : \mathcal{U} \to \mathcal{W}$ such that $p = \pi \circ q$.

THEOREM 1.109. *Any orbifold admits the canonical*[9] *Universal covering.*

Below we will supply a construction in dimension two.

If the universal covering of an orbifold $\mathcal{O}$ coincides with itsels, we will refer to it as *universal*.

EXERCISE 1.110. *The underlying space of a universal orbifold is simply connected.*

An orbifold is called *good* if its Universal covering is a manifold. It is exactly the case of a global quotient of a manifold by a group action.

An orbifold is called *oriented* if the neighborhoods $\hat{U}_i$ are oriented and the groups G_i as well as the transit maps ϕ_{ij} are orientation preserving.

1.8.2. 2D *orbifolds.* If $\mathcal{O}$ is an oriented 2D orbifold then its local charts can be selected as the disk $\mathbb{D}$ endowed with cyclic rotations groups G_i, $z \mapsto e(m/\mathfrak{q}_i)\, z$, $m \in \mathbb{Z}/\mathfrak{q}_i\mathbb{Z}$. As the quotients $\mathbb{D}/G_i$ are homeomorphic to $\mathbb{D}$, the underlying space M is an (oriented) topological surface as well. So, in this case, we can think of $\mathcal{O}$ in a simply minded way, as an oriented surface M with a set X of isolated points x_i endowed with "ramification indices" $\mathfrak{q}_i \in \mathbb{Z}_+$, $\mathfrak{q}_i \geq 2$, $i = 1, \dots, n$. The data $(M; \{\mathfrak{q}_1, \dots, \mathfrak{q}_n\})$ is called the *signature* of $\mathcal{O}$. We can also let $\mathfrak{q}(x) = 1$ for any regular point $x \in M$, so we make the ramification function $\mathfrak{q} : M \to \mathbb{Z}_+$ defined everywhere.

REMARK 1.111. If $\mathcal{O}$ is a smooth 2D Riemannian orbifold, then the above picture becomes too simplistic, as the underlying space M develops *conic singularities* with angles $2\pi/\mathfrak{q}_i$ at points x_i. See §2.1.4.

[9] Meaning "unique up to equivalence between coverings".

EXERCISE 1.112. *Let $\mathcal{O}_\mathfrak{q}$ be the orbifold with the underlying space $U = \mathbb{D}$, and the local chart $\hat{U} = \mathbb{D}$ endowed with the cyclic rotation group G,*

$$z \mapsto e(m/\mathfrak{q})z, \quad m \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$$

(*so that all other charts are obtained by restricting this one*). *Let*

$$f_n : \mathbb{D} \to \mathbb{D}, \quad z \mapsto z^n.$$

(i) *The map f_n induces a morphism $\mathcal{O}_\mathfrak{q} \to \mathcal{O}_\mathfrak{p}$ if and only if $n\mathfrak{q}$ is a multiple of $\mathfrak{p}$. If so, this morphism is an orbifold covering of degree $d = n\mathfrak{q}/\mathfrak{p}$. In particular, $f_\mathfrak{p}$ induces the universal covering $\mathbb{D} \to \mathcal{O}_\mathfrak{p}$ (where $\mathbb{D}$ is identified with $\mathcal{O}_1$).*

(ii) *If $n\mathfrak{q}$ divides $\mathfrak{p}$ then the multi-valued inverse map f_n^{-1} lifts to the orbifold universal covering $\hat{f}_n^{-1} : \hat{\mathcal{O}}_\mathfrak{p} \to \hat{\mathcal{O}}_\mathfrak{q}$. In the $\mathbb{D}$-model for the universal coverings, it becomes $z \mapsto z^k$ with $k = \mathfrak{p}/(n\mathfrak{q})$.*

EXERCISE 1.113. *A morphism $f : \mathcal{O} \to \mathcal{O}'$ between two 2D orbifolds is a covering of degree d iff it induces a Galois branched covering $f : M \to M'$ of degree d between the underlying surfaces such that $\mathfrak{q}(fx) = \deg_x f \cdot \mathfrak{q}(x)$ for any $x \in M$*

Construction of the Universal covering in 2D. Let us puncture out the singular points x_i from M, setting $M^* = M \smallsetminus X$, and let $p^* : \mathcal{U} \to M^*$ be its universal covering. Let V_i be small disk neigborhoods of the x_i, and let $V_i^* := V_i \smallsetminus \{x_i\}$. For each x_i, let us select a component $\hat{V}_i^*$ of $p^{-1}(V_i^*)$. By Proposition 1.59, each restriction $p^* : \hat{V}_i^* \to V_i^*$ is a universal covering, so it is equivalent to the exponential map $e : \mathbb{H} \to \mathbb{D}^*$.

Let Γ_i be the stabilizer of $\hat{V}_i^*$ in the group G of deck transformations, and let Γ be the subgroup of G generated by the Γ_i. Let $\mathcal{W}^* := \mathcal{U}/\Gamma$ and let $\pi : \mathcal{W}^* \to M^*$ be the corresponding covering. Then each restriction $p_i^* : \hat{V}_i^*/\Gamma_i \to V_i^*$ is equivalent to the covering $f_i : \mathbb{D}^* \to \mathbb{D}^*$, $z \mapsto z^{\mathfrak{r}_i}$, with some $\mathfrak{r}_i$ dividing $\mathfrak{q}_i$. Hence, by adding one ideal point $\hat{x}_i$ to each $\hat{V}_i^*$, we can extend p_i^* to a branched covering equivalent to $f_i : \mathbb{D} \to \mathbb{D}$. It provides us with an extension of p^* to a branched covering $p : \mathcal{W} \to M$, which can be interpreted as an orbifold covering.

EXERCISE 1.114. *Verify that this orbifold covering is universal* (*and canonical*).

EXAMPLE 1.115. *Let us consider the orbifold $\mathcal{O}$ with signature $(S^2; \{2, 3\})$. Let us realize it as the Riemann sphere $\hat{\mathbb{C}}$ with the singular set $X = \{0, \infty\}$. According to the above construction, its universal covering $\mathcal{U}$ is obtained by taking the exponential coivering $e : \mathbb{C} \to \mathbb{C}^*$ and quotening it by the translation group generated by $T_2 : z \mapsto z + 2$ and $T_3 : z \mapsto z + 3$. This group generates the whole group $\mathbb{Z}$ of deck transformations for e, so we get $\mathcal{O}$ back as its own universal covering. We see that this orbifold is bad.*

EXERCISE 1.116. *Describe the Universal covering of the orbifolds with signatures $(S^2; \mathfrak{p})$ and $(S^2; \{\mathfrak{p}, \mathfrak{q}\})$,*

The *Euler characteristic* of a 2D orbifold $\mathcal{O}$ with signature $(M; \{\mathfrak{q}_i\})$ is defined as

$$\chi(\mathcal{O}) = \chi(M) - \sum (1 - \frac{1}{\mathfrak{q}_i}). \tag{1.8}$$

This definition secures the standard behavior of the Euler characteristic under covering and connected sum operations. Indeed, since the orbifold $\mathcal{O}_\mathfrak{q}$ is covered with degree $\mathfrak{q}$ by the disk $\mathbb{D}$, we want $\chi(\mathcal{O}_\mathfrak{q}) = 1/\mathfrak{q}$. If $\mathcal{O}$ has n singular points x_i, then the underlying space of $\mathcal{O}$ is obtained by gluing n orbifolds $\mathcal{O}_\mathfrak{q}$ into $M \smallsetminus \bigsqcup U_i$, where the U_i are Jordan disk neighborhoods of the q_i (with disjoint closures). Then the desired additivity of the Euler characteristic requires

$$\chi(\mathcal{O}) = \chi(M \smallsetminus \bigsqcup U_i) + \sum \chi(\mathcal{O}_{\mathfrak{q}_i}) = \chi(M) - n + \sum \frac{1}{\mathfrak{q}_i},$$

yielding (1.8).

Exercise 1.117. *Let $\mathcal{O}$ and $\mathcal{O}'$ be two* 2D *orbifolds.*

(i) *Show that if $f : \mathcal{O} \to \mathcal{O}'$ is a degree d covering then $\chi(\mathcal{O}) = d\,\chi(\mathcal{O}')$.*

(ii) *If $O \cup_h \mathcal{O}'$ is obtained from $\mathcal{O}$ and $\mathcal{O}'$ by gluing the underlying surfaces along boundary Jordan curves, then $\chi(\mathcal{O} \sqcup_h \mathcal{O}') = \chi(\mathcal{O}) + \chi(\mathcal{O}')$.*

One says that a 2D orbifold $\mathcal{O}$ is of *finite topological type* if its underlying surface M is of finite type and it has only finitely many singular points.

Exercise 1.118. *A* 2D *orbifold is of finite topological type iff it has a finite Euler characteristic.*

It is convenient to consider a puncture in a surface as an orbifold point of infinite index (modeled on the $\mathbb{Z}$-action of $\mathbb{H}$ generated by the translation $z \mapsto z+1$). With this convention, an orbifold of finite type has a *compact* underlying surface with finitely many singular points (including punctures).

Exercise 1.119. *Check that the above discussion carries through with this more general interpretation of* 2D *orbifolds.*

It turns out that in dimension two almost all orbifolds are good: the only bad once are $(S^2; \mathfrak{p})$ and $(S^2; \{\mathfrak{p}, \mathfrak{q}\})$, with $\mathfrak{p} \neq \mathfrak{q}$. See §2.8.2.

1.9. Appendix 1: Hausdorff metric. Let Z be a compact metric space, and let $\mathfrak{S}(Z)$ be the space of its closed subsets. The *Hausdorff distance* between two subsets $X, Y \in \mathfrak{S}(Z)$ is defined as follows:

$$\operatorname{dist}_{\mathrm{H}}(X, Y) = \max\{\sup_{x \in X} \operatorname{dist}(x, Y),\ \sup_{y \in Y} \operatorname{dist}(X, y)\} \tag{1.9}$$

(where $\operatorname{dist}(x, Y)$ is defined in §50.3.4). Note that $\operatorname{dist}_{\mathrm{H}}(X, Y) < \varepsilon$ means that X is contained in the ε-neighborhood of Y and the other way around.

Exercise 1.120. *For a compact metric space, we have:*

(i) $\operatorname{dist}_{\mathrm{H}}$ *defines a metric on $\mathfrak{S}(Z)$ (called "Hausdorff").*

(ii) *$X_n \to X$ iff all limits points $\lim x_k$, $x_{n_k} \in X_{n_k}$, $n_k \to \infty$, belong to X, and for any point $x \in X$ there is a sequence $x_n \in X_n$ converging to x.*

(iii) *$\mathfrak{S}(Z)$ is compact.*

- Given a sequence of sets $X_n \subset Z$, we let:
- $\limsup_{n\to\infty} X_n$ be the set consisting of all limits $\lim x_k$, where $x_k \in X_{n_k}$ for some subsequence $n_k \to \infty$.
- $\liminf_{n\to\infty} X_n$ be the set consisting of all limits $\lim x_n$, where $x_n \in X_n$.

As Exercise 1.120 shows, $X_n \to X$ in the Hausdorff metric iff

$$\limsup X_n = \liminf X_n = X.$$

If we have a closed set $X_\lambda \subset Z$ depending on some parameter $\lambda \in \Lambda$, we say that it depends *upper semicontinuously* on λ at $\lambda_\circ$ if $\limsup X_{\lambda_n} \subset X_{\lambda_\circ}$ for any $\lambda_n \to \lambda_\circ$. In other words, $\forall\ \varepsilon > 0\ \exists\ \delta > 0$ such that if $|\lambda - \lambda_\circ| < \delta$ then X_λ is contained in the ε-neighborhood of $X_{\lambda_\circ}$ (i.e., $X_{\lambda_\circ}$ cannot "blow-up" under perturbations).

Exercise 1.121. *Let Z be compact, and let $\phi : Z \to \Lambda$ be a continuous map. Then the fibers $X_\lambda := \phi^{-1}(\lambda)$ depend upper semicontinuously on $\lambda \in \Lambda$.*

Similarly, *lower semicontinuity* means that $\liminf X_{\lambda_n} \supset X_{\lambda_\circ}$ for any $\lambda_n \to \lambda_\circ$ (so $X_{\lambda_\circ}$ cannot "collapse" under perturbations). Thus, we have the usual property: X_λ depends continuously on λ with respect to the Hausdorff distance iff it depends simultaneously upper and lower semicontinuously.

1.10. Appendix 2: Alexander cohomology. This is a very general theory that can be applied to fractal sets (unlike the homotopy theory). In particular, it allows us to characterize hulls in $\mathbb{R}^2$ (which are intuitively "simply connected") in terms of vanishing cohomology. We will not try to develop this theory in a regular way but will rather give a quick account (for the first cohomology only) suitable for our purposes. To make it more intuitive, we will use the physical language.

Let K be a compact metric space whose points are viewed as "particles". An *interaction energy* is a continuous function $\phi(x, y)$ defined for all pairs $x, y \in X$ with $\operatorname{dist}(x, y) < r$ (for some $r > 0$) such that

- $\phi(x, x) = 0$;
- $\phi(x, y) = -\phi(y, x)$;
- *Chain Rule*: $\phi(x, z) = \phi(x, y) + \phi(y, z)$ (as long as all three distances are bounded by r).

Such a function is also called 1-*cocycle*.

We can extend the interaction to the pairs of particles connected by an r-*chain*, i.e., a sequence of points $C = (x = x_0, x_1, \dots, x_n = y)$ such that $d(x_i, x_{i+1}) < r$. Namely, let

$$\phi_C(x, y) = \sum_{i=0}^{n-1} \phi(x_i, x_{i+1}).$$

Of course, the result may depend on the chain.

Exercise 1.122. *Show that $\phi_{C'}(x, y) = \phi_C(x, y)$ if C' is a small perturbation of C with the same endpoints.*

The interaction is called *potential* if there exists a continuous function $u : X \to \mathbb{R}$ such that $\phi(x, y) = u(y) - u(x)$. In this case, the cocycle ϕ is called *trivial*.

Exercise 1.123. (i) *An interaction is potential if and only if $\phi_C(x, y)$ is independent of the choice of the chain C.*

(ii) *Any interaction is locally potential.*

Let us first define vanishing cohomology:

$$H^1_{\mathrm{A}}(K) \equiv H^1_{\mathrm{A}}(K,\mathbb{R}) = 0$$

if all interactions with sufficiently small $r > 0$ are potential.

EXAMPLE 1.124. $H^1_{AK}(\bar{\mathbb{D}}) = 0$. To see it, notice that any two r-chains C and C' connecting x to y are *discretely homotopic* rel the endpoints.[10] It means that for any $\varepsilon > 0$ there is a sequence of r-chains $C = C^0, C^1, \dots, C^k = C'$ with the same endpoints such that $\operatorname{dist}(C^i, C^{i+1}) < \varepsilon$. The latter means that the chains in question can be concatenated into the same number of subchains C^i_k, C^{i+1}_k $(k = 1, \dots, m)$ with diameter $< r/2$ and $\operatorname{dist}_{\mathrm{H}}(C^i_k, C^{i+1}_k) < \varepsilon$ (where $\operatorname{dist}_{\mathrm{H}}$ stands for the Hausdorff distance, see Appendix 1 below). Since the energy of a chain is homotopy invariant, it depends only on the endpoints. Hence it is potential.

EXAMPLE 1.125. $H^1_{\mathrm{A}}(\mathbb{T}) \neq 0$. For points z and ζ which are not antipodal on $\mathbb{T}$, we have a well defined interaction energy $\phi(z,\zeta) = \arg(\zeta/z) \in (-\pi, \pi)$. It satisfies the Chain Rule for any three points that lie on the same side of some diameter. However, it is not potential since the energy of a closed chain that goes around the circle is equal to 2π.

In general, the *first Alexander cohomology group*, $H^1_{\mathrm{A}}(K) \equiv H^1_{\mathrm{A}}(K,\mathbb{R})$, is defined as the space of 1-cocycles modulo trivial ones.

EXERCISE 1.126. *Show that $H^1_{\mathrm{A}}(\mathbb{T}) \approx \mathbb{R}$.*

1.11. Appendix 3: Cyclic order. A *cyclic order* on a finite set Θ with $\mathfrak{q}$ elements can be defined in one of the following equivalent ways:

(i) A cyclic permutation $\sigma : \Theta \to \Theta$;

(ii) A bijection $o : \Theta \to \mathbb{Z}/\mathfrak{q}\mathbb{Z}$, up to translation $o(\theta) + k$, $k \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$;

(iii) An oriented graph Γ supported on Θ (as the set of vertices) which is a single cycle;

(iv) An assignment to any point $\theta \in \Theta$ the *next* one, $\sigma(\theta)$, so that the corresponding oriented graph is a single cycle.

Any subset $\Theta' \subset \Theta$ naturally inherits a cyclic order from Θ (whose cycle Γ' is obtained from Γ by concatenating arrows of Γ through vertices of $\Theta \smallsetminus \Theta'$).

Note that a 2-point set has only one cyclic order, while a 3-point set supports exactly two different cyclic orders.

EXERCISE 1.127. *Let $\Theta \approx \mathbb{Z}/\mathfrak{q}\mathbb{Z}$ be a finite cyclically ordered set, and let $g : \Theta \to \Theta$ be an order preserving permutation (i.e., if θ' is next to θ then $g(\theta')$ is next to $g(\theta)$). Then g is conjugate to a translation $n \mapsto n + \mathfrak{p}$ on $\mathbb{Z}/\mathfrak{q}\mathbb{Z}$. In particular, all points of Θ have the same period.*

Such a permutation g is called a *rotation* of Θ, with *rotation number* $\mathfrak{p}/\mathfrak{q}$. If $\mathfrak{q}$ and $\mathfrak{p}$ are mutually co-prime, then g is called a *cyclic rotation*.

Any finite subset Θ of an oriented topological circle S^1 is endowed with a natural cyclic order corresponding to the positive motion around the circle. So, a point $\theta' \in \Theta$ is *next* to $\theta \in \Theta$ if the interval $(\theta, \theta') \in S^1$ (whose orientation from θ to θ' is positive in S^1) does not contain points of Θ.

---

[10][In fact, it can even be done continuously by projecting the chains in question orthogonally to the interval connecting x and y.

More generally, an infinite set Θ is called *cyclically ordered* if any its finite subset is, and these orders are compatible: If $\Theta_0 \subset \Theta_1$ then the cyclic order of Θ_0 is induced from Θ_1.

Any subset of a cyclically ordered set Θ naturally inherits a cyclic order. We say that a triple $(\theta_1, \theta_2, \theta_3) \in \Theta^3$ is *positively ordered* if its natural cyclic order coincides with the one induced from Θ.

Also, selecting any two points $\theta_1, \theta_2 \in \Theta$, the set of points $\theta \in \Theta$ such that the triple $(\theta_1, \theta, \theta_2)$ is positively ordered forms a *linearly ordered* "interval" (θ_1, θ_2) (so for any two distinct points θ, θ' in this interval, we have $\theta < \theta'$ or the other way around). Of course, we can add endpoints to (θ_1, θ_2) to obtain a "closed interval" $[\theta_1, \theta_2]$ or "semi-closed intervals" $[\theta_1, \theta_2)$ and $(\theta_1, \theta_2]$.

Notes. The text book by Munkres [**Mu1**] serves well as a basic reference in topology. Orsay Notes by Douady and Hubbard [**DH2**] and Milnor's book [**M2**] can also serve as an efficient introduction to the topology of plane continua.

Various takes on the Jordan Theorem can be found in [**Mu1, KaC**], [**Hat**, §2.B] The most general approach goes through the *Alexander Duality* (see [**Hat**, Thm.3.44]). A general treatment of the Alexander cohomology can be found in [**Spen**]. Note that it was independently introduced by A.N. Kolmorogov [**Ko**].

Nadler's book [**Na**] goes in depth into the topological structure of continua. (See also [**Lew**] for a discussion of pseudo-arcs.) All wild creatures that one can imagine appear naturally in dynamics, albeit not necessarily for polynomials, see [**De3, He3, KY, Mayer, Re3**].

The Triangulation Theorem for 2D manifolds was proven by Rado in the 1920s [**Rado**]. Moore's Theorem appeared in 1925 [**Moo**].

The notion and a basic theory of orbifolds is due to Thurston, see [**Sc**].

Acknowledgment. The author thanks Oleg Viro for reading this chapter of the manuscript and making helpful comments, and Lasse Rempe-Gillen for useful comments on wild creatures.

Part 1

Conformal and quasiconformal geometry

CHAPTER 1

Conformal geometry

2. Riemann surfaces

2.1. Analytic and geometric structures on surfaces.

2.1.1. *Smooth surfaces.* Rough topological structure can be refined by requiring that the transition maps belong to a certain "structural pseudo-group", which often means: "have certain regularity". For example, a *smooth* structure on S is given by a family of local charts $\phi_i : U_i \to \mathbb{R}^2$ such that all the transition maps are smooth (with a prescribed order of smoothness). A surface endowed with a smooth structure is naturally called a *smooth surface*. A local chart $\phi : U \to V$ smoothly related to the charts ϕ_i (i.e., with smooth transition maps) is referred to as a "smooth local chart". A family of smooth local charts covering S is called a *smooth atlas* on S. A smooth structure comes together with affiliated notions of smooth functions, maps and diffeomorphisms.

There is a smooth version of the connected sum operation in which the boundary curves are assumed to be smooth and the boundary gluing map h is assumed to be an orientation reversing diffeomorphism. To get a feel for it, we suggest the reader to do the following exercise:

Exercise 2.1. *Consider two copies D_1 and D_2 of the closed unit disk $\overline{\mathbb{D}} \subset \mathbb{R}^2$. Glue them together by means of an orientation reversing diffeomorphism $h : \partial D_1 \to \partial D_2$ of the boundary circles. You obtain a topological sphere S^2. Show that it can be endowed with a unique smooth structure compatible with the smooth structures on D_1 and D_2 (that is, such that the tautological embeddings $D_i \to S^2$ are smooth). The boundary circles ∂D_i become smooth Jordan curves on this smooth sphere. Show that this sphere is diffeomorphic to the* standard *"round sphere" in $\mathbb{R}^3$.*

Using a partition of unity, any smooth surface can be endowed with a *Riemannian metric.* This makes the Triangulation Theorem (1.70) easy in the smooth category:

Proposition 2.2. *Any smooth surface S can be triangulated.*

Proof. Take a fine net of points on S (including ∂S) in a general position, and connect each of them with nearby points by geodesic arcs. We obtain a tessellation of S by geodesic polygons. Triangulating these polygons, we obtain the desired. □

Real analytic structures would be the next natural refinement of smooth structures.

2.1.2. *Riemann surfaces.* If $\mathbb{R}^2$ is considered as the complex plane $\mathbb{C}$ with $z = x + iy$, then we can talk about *complex analytic* $\equiv$ *holomorphic* transition maps and corresponding complex analytic structures and surfaces. Such surfaces are known under a special name of *Riemann surfaces.* A holomorphic diffeomorphism between two Riemann surfaces is often called an (*biholomorphic*) *isomorphism.* Accordingly a holomorphic diffeomorphism of a Riemann surface onto itself is called its (*biholomorphic*) *automorphism.*

For instance, the one-point compactification $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ of $\mathbb{C}$ is a topological sphere endowed with the natural complex structure, with two local charts, $\mathrm{id} : \mathbb{C} \to \mathbb{C}$ and $\phi : \hat{\mathbb{C}} \smallsetminus \{0\} \to \mathbb{C}$, $\phi(z) = 1/z$. This Riemann surface is called the *Riemann sphere.*

Connected sum operation still works in the category of Riemann surfaces. In its simplest version the boundary curves and the gluing diffeomorphism should be taken real analytic. Here is a representative statement:

Exercise 2.3. *Assume in* Exercise 2.1 *that* $\mathbb{R}^2 \equiv \mathbb{C}$ *and that the gluing diffeomorphism h is real analytic. Then S^2 can be supplied with a unique complex analytic structure compatible with the complex analytic structure on the disks $D_i \subset \mathbb{C}$. The boundary circles ∂D_i become real analytic Jordan curves on this Riemann sphere.*

More generally, we can attach handles to the sphere by means of real analytic boundary map, and obtain an example of a *Riemann surface of genus g.* It is remarkable that, in fact, it can be done with only smooth gluing maps, or even with singular maps of a certain class. This operation (with singular gluing maps) has important applications in Teichmüller theory, theory of Kleinian groups, and Dynamics.

If $\mathbb{R}^2$ is supplied with the standard Euclidean metric, then we can consider *conformal* transition maps, i.e., diffeomorphisms preserving angles between curves. The first thing students usually learn in complex analysis is that the class of orientation preserving conformal maps coincides (in dimension 2!) with the class of invertible complex analytic maps. Thus the notion of a *conformal structure* on an oriented surface is equivalent to the notion of a *complex analytic structure$\equiv$holomorphic* (though it is worthwhile to keep in mind a conceptual difference between them: one comes from geometry, the other comes from analysis).

We say that a Riemann surface S is a *conformal disk/annulus etc* if it is isomorphic to the standard disk $\mathbb{D}$, a round annulus $\mathbb{A}(r, R)$ etc. A conformal sphere naturally bears the same name as a holomorphic one: the *Riemann sphere.*

2.1.3. *Fine geometric structures and rough structures.* One can go further to *projective*, *affine*, *Euclidean/flat* or *hyperbolic* structures. We will refine this discussion momentarily.

One can also go in the opposite direction and consider *rough structures* on a topological surface whose structural pseudo-group is bigger then the pseudo-group of diffeomorphisms, e.g., *bi-Lipschitz structures*. Even rougher, *quasiconformal*, structures will play an important role in our discussion.

Let $h : S \to S'$ be a surface homeomorphism. Then any structure μ' of T can be obviously *pulled back* to a structure $h^*\mu'$ of the same kind on S, and any structure μ on S can be *pushed forward* to a structure $h_*\mu$ on S'. In fact, the *pullback* $h^*\mu$ is well defined as long as h is a covering map. Namely, for any μ'-local chart $\phi | U'$ and any neighborhood $U \subset S$ such that $h : U \to S'$ is an embedding into U', the composition $\phi \circ h| U$ is declared to be a local chart on S. (Note that the transition maps for μ' and $h^*\mu'$ are the same.)

To comfort a rigorously-minded reader, let us finish this brief excursion with a definition of a *pseudo-group* on $\mathbb{R}^2$ (in the generality adequate to the above discussion). It is a family of local homeomorphisms $f : U \to V$ between open subsets of $\mathbb{R}^2$ (where the subsets depend on f) which is closed under taking inverse maps and taking compositions (on the appropriately restricted domains). The above structures are related to the pseudo-groups of all local (orientation preserving) homeomorphisms, local diffeomorphisms, locally biholomorphic maps, local isometries (Euclidean or hyperbolic) etc.

The *uniformization* of a Riemann surface S is a holomorphic parametrization $\phi : S_\circ \to S$ of S by some model Riemann surface $S_\circ$, e.g., by the complex plane, the Riemann sphere, or the hyperbolic plane. These three models lead to three geometries that will be discussed below.

2.1.4. *Geometric structures on orbifolds.* Let us now refine the discussion from §1.8.

An orbifold $\mathcal{O}$ is called *smooth* if the local group elements, as well as the transit maps are diffeomorphisms. (Of course, the underlying space M of a smooth orbifold still has singularities.)

A *smooth orbifold morphism* $\mathcal{O} \to \mathcal{O}'$ (between smooth orbifolds) is a continuous map $f : M \to M'$ between the underlying spaces which is smooth in local charts, (i.e., it locally lifts to equivariant smooth maps $f_{ij} : \hat{U}_i \to \hat{U}_j$) between local charts). The notions of an *orbifold diffeomorphism* naturally follows.

The orbifold is called *Riemannian* if all the U_i are endowed with the Riemannian metrics ds_i^2 that turn all the group elements and the transit maps into isometries. The notion of an *orbifold isometry* naturally follows.

Similarly, we can define a general notion of *orbifold geometric structure* (conformal $\equiv$ Riemann, Euclidean, spherical, hyperbolic, etc.) and associated (*iso-*)*morphisms*.

For 2D orientable manifolds, the local orbifold groups are cyclic, $\mathbb{Z}/\mathfrak{q}\mathbb{Z}$. In the Riemannian case, the local quotients are *cones* with angle $2\pi/\mathfrak{q}$, $\mathfrak{q} \in \mathbb{Z}_+$ (with singular points corresponding to $\mathfrak{q} \geq 2$). Obviously, there is no way to turn the underlying surface near cone singularities into a smooth Riemannian one (compare with Remark 1.111.)

2.2. Flat (Euclidean) and affine geometries. Consider the complex plane $\mathbb{C}$. Holomorphic automorphisms of $\mathbb{C}$ are *complex affine maps* $A : z \mapsto az + b$, $a \in \mathbb{C}^*$, $b \in \mathbb{C}$. They form a group $\operatorname{Aff}(\mathbb{C})$ acting freely bi-transitively on the plane: any pair of points can be moved in a unique way to any other pair of points. Moreover, it acts freely transitively on the tangent bundle of $\mathbb{C}$.

Thus the complex plane $\mathbb{C}$ is endowed with the *affine structure* canonically affiliated with its complex analytic structure. Of course, the plane can be also endowed with a Euclidean metric $|z|^2$. However, this metric can be multiplied by any scalar $t > 0$, and there is no way to normalize it in terms of the complex structure only. All these Euclidean structures have the same group $\operatorname{Euc}(\mathbb{C})$ of (orientation preserving) *Euclidean motions* $A : z \mapsto az + b$ with $|a| = 1$. This group acts transitively on the plane with the group of rotations $z \mapsto e(\theta)\, z$, $\theta \in \mathbb{R}/\mathbb{Z}$, stabilizing the origin. Moreover, it acts freely transitively on the unit tangent bundle of $\mathbb{C}$ (corresponding to any Euclidean structure).

The group Aff has very few discrete subgroups acting freely on $\mathbb{C}$: rank 1 lattice $z \mapsto z + an$, $n \in \mathbb{Z}$, where $a \in \mathbb{C}^*$, and rank 2 lattice $z \mapsto an + bm$, $(m, n) \in \mathbb{Z}^2$, where (a, b) is an arbitrary basis in $\mathbb{C}$ over $\mathbb{R}$. All rank 1 lattices are conjugate by affine transformations, so that the quotients modulo their actions are all isomorphic. Taking $a = 1$ we realize these quotients as the *bi-infinite cylinder* $\mathbb{C}/\mathbb{Z}$. It is isomorphic to the punctured plane $\mathbb{C}^*$ by means of the exponential map $e : \mathbb{C}/\mathbb{Z} \to \mathbb{C}^*$. The rank 2 quotients are all topological tori. However, they generically represent different Riemann surfaces. Indeed, by a complex affine transformation, any rank 2 lattice can be brought to t he form

$$\mathbb{L} \equiv \mathbb{L}_\tau := \{m + n\tau : \ ((m, n) \in \mathbb{Z}^2, \quad \operatorname{Im} \tau > 0\},$$

so we obtain a complex one-prameter family of tori $\mathbb{T}^2 \equiv \mathbb{T}^2_\tau := \mathbb{C}/L$, $\tau \in \mathbb{H}$. However, not all of these tori are conformally distinct, and in fact, the space of various complex tori is the quotient of $\mathbb{H}$ by some discrete group (see §2.6.3 below).

Note that the above discrete groups preserve the Euclidean structures on $\mathbb{C}$. Hence these structures can be pushed down to the quotient Riemann surfaces. Moreover, now they can be canonically normalized: in the case of the cylinder we can normalize the length of the simple closed geodesic to be 1. In the case of the torus we can normalize its total area. Thus, the complex tori and the bi-infinite cylinder are endowed with the canonical Euclidean structures.

By the Geometric Uniformization Theorem (see Appendix 2), we have exhausted the list of complete Euclidean surfaces:

THEOREM 2.4. *Any complete Euclidean surface is isometric to either the Eucliuidean plane $\mathbb{R}^2$, or to the flat cylinder $\mathbb{T} \times \mathbb{R}$, or to the torus $\mathbb{T}^2$.*

EXERCISE 2.5. *Let us consider a torus $\mathbb{T}^2 \equiv \mathbb{T}^2_\tau := \mathbb{C}/L_\tau$, $\tau \in \mathbb{H}$.*

(i) *Any holomorphic endomorphism f of $\mathbb{T}^2$ is affine, i.e., it is induced by an affine map $\hat{f} : z \mapsto \rho z + b$ of $\mathbb{C}$ (such that $\rho \cdot L \subset L$). In particular, we can take*

$$\hat{f} = A_n : z \mapsto nz \quad \text{with } n \in \mathbb{Z}. \tag{2.1}$$

(ii) *$\deg f = |\rho|^2$; so f is an automorphism iff $|\rho| = 1$. In particular, any torus admits a holomorphic involution σ unduced by $A_{-1} : z \mapsto -z$. This involution has four fixed points in $\mathbb{T}^2$.*

(iii) *The involution σ is the only (up to conjugacy by translations) non-trivial automorphism of any torus, except two special ones corresponding to extra symmetric lattices, $\mathbb{L}_i$ and $\mathbb{L}_{e(1/6)}$. Describve the group of symmetries for these two.*

(iv) *If $\mathbb{T}^2_\tau$ admits an endomorphism $z \mapsto \rho z + b$ with non-integer ρ, then ρ and τ are quadratic irrationals. So, there are only countably many such endomorphisms. They are called "complex multiplications".*

By taking quotients of $\mathbb{C}$ by *non-free* actions of discrete subgroups Γ of $\operatorname{Euc}(\mathbb{C})$, we obtain 2D *Euclidean orbifolds* (also called *flat* or *parabolic*).

EXERCISE 2.6. (i) *Make a list of such Euclidean orbifolds and the corresponding branched coverings $\mathbb{C} \to M$ of their underlying spaces.*

(ii) *Note that all of them have zero Euler characteristic.*

(iii) *Pay attention to three special ones related to a checker-board tesselation $\mathcal{T}$ of $\mathbb{C}$ by white and black triangles with angles $\pi/\mathfrak{p}$, $\pi/\mathfrak{q}$, $\pi/\mathfrak{r}$, where*

$$\frac{1}{\mathfrak{p}}+\frac{1}{\mathfrak{q}}+\frac{1}{\mathfrak{r}}=1, \quad \mathfrak{p},\mathfrak{q},\mathfrak{r}\in\{2,3,\dots\}.$$

Each of these triangles is as fundamental domain for the full group of symmetries $\hat{\Gamma}\supset\Gamma$ of $\mathcal{T}$ (including reflections), which is the index two extension of Γ.

Notice that these properties are (naturally) in agreement with the Gauss-Bonnet Formula (see Appendix 2 below). For instance, orbifolds in (iii) are endowed with flat metric with three cone singularities. By (2.33) these singularities support curvatures $2\pi(-1/\mathfrak{p})$, $2\pi(1-1/\mathfrak{q})$, and $2\pi(1-1/\mathfrak{r})$, so the total curvature of this metric is equal to

$$2\pi\left(3-\left(\frac{1}{\mathfrak{p}}+\frac{1}{\mathfrak{q}}+\frac{1}{\mathfrak{r}}\right)\right)=4\pi=2\pi\cdot\chi(S^2).$$

2.3. Projective and spherical geometries.

2.3.1. *Möbius group.* Consider now the Riemann sphere $\hat{\mathbb{C}}$. Its biholomorphic automorphisms are *Möbius transformations*

$$\phi: z\mapsto\frac{az+b}{cz+d};\quad \det\begin{pmatrix} a & b\\ c & d\end{pmatrix}\neq 0.$$

We will denote this *Möbius group* by $\text{Möb}(\hat{\mathbb{C}})$. It acts freely triply transitive on the sphere: any (ordered) triple of points (a,b,c) on the sphere can be moved by a unique Möbius transformation to any other triple (a',b',c').

Note that the Riemann sphere is isomorphic to the complex projective line $\mathbb{CP}^1$. For this reason Möbius transformations are also called *projective*. Algebraicly, the Möbius group is isomorphic to the linear projective group $\mathrm{PSL}(2,\mathbb{C})=\mathrm{SL}(2,\mathbb{C})/\{\pm I\}$ of 2×2 matrices A with $\det A=1$ modulo reflection $A\mapsto -A$.

2.3.2. *Classification of Möbius transformations.* Any Möbius transformation A has a fixed point $\alpha\in\hat{\mathbb{C}}$, i.e. $A(\alpha)=\alpha$. Hence *there are no Riemann surfaces whose universal covering is* $\hat{\mathbb{C}}$ (except $\hat{\mathbb{C}}$ itself). In fact, any non-identical Möbius transformations has either one or two fixed points, and can be classified depending on their nature.

To this end, let us bring a Möbius transformation to a simplest *normal form* by means of a conjugacy $\phi^{-1}\circ f\circ\phi$ by some $\phi\in\text{Möb}(\hat{\mathbb{C}})$. Since $\text{Möb}(\hat{\mathbb{C}})$ acts double transitively, we can find some ϕ which sends one fixed point of f to ∞ and the other (if exists) to 0. This leads to the following classification:

(i) A *hyperbolic* Möbius transformation A has an *attracting* and *repelling* fixed points, α_+ and α_-, with *multipliers*[1] ρ and ρ^{-1} respectively, where $0<|\rho|<1$. Its *normal form* is a global linear contraction $A: z\to\rho z$ (with possible spiralling if ρ is unreal). These fixed points are called *hyperbolic* as well.

Hyperbolic Möbius transformations with unreal ρ are also called *loxodromic*.

(ii) An *elliptic* Möbius transformation has two fixed points $\alpha_\pm$ with multipliers ρ and ρ^{-1} where $\rho=e(\theta)$, $\theta\in\mathbb{R}/\mathbb{Z}$. Its normal form is the rotation $z\to e(\theta)z$.

[1]The multiplier of a fixed point α is the derivative $A'(\alpha)$ calculated in *any* local chart around α, compare §§19.5, 21.1.

(iii) A *parabolic* Möbius transformation has a single fixed point α with multiplier 1. Its normal form is a translation $z \mapsto z+1$.

Exercise 2.7. (i) *Verify those of the above statements that look new to you.*
(ii) *Show that dilations $z \mapsto \rho z$, $0<\rho<1$, rotations $z \mapsto e(\theta)\, z$, $\theta \in \mathbb{R}/\mathbb{Z}$, and translations $z \mapsto z+a$, $a \in [0,1)$, generate the whole group* $\mathrm{Aff}(\mathbb{C})$.
(iii) *Show that* $\mathrm{Aff}(\mathbb{C})$ *together with the involution $z \mapsto 1/z$, generate the whole group* $\mathrm{M\ddot{o}b}(\hat{\mathbb{C}})$.

Exercise 2.8. *Classify Möbius transformations in terms of the representing matrices $A \in \mathrm{SL}(2,\mathbb{R})$. Namely, A is elliptic, parabolic (including $A=\mathrm{id}$), or hyperbolic according to whether $\operatorname{tr} A \in (-2,2)$, $|\operatorname{tr} A| = 2$, or $\operatorname{tr} A \in \mathbb{C} \smallsetminus [-2,2]$, respectively. Moreover, A is loxodromic iff $\operatorname{tr} A$ is unreal.*

2.3.3. *Dynamics.* Using the above normal forms, it is easy to describe the dynamics of Möbius transformations:

(i) If A is *hyperbolic* then its forward orbits $\{A^n z\}_{n\in\mathbb{N}}$ converge to the attracting fixed point α_+, while backward orbits $\{A^{-n} z\}_{n\in\mathbb{N}}$ converge to the repelling fixed point α_-. Moreover, A preserves the foliation of circular arcs ("separatrices") passing through α_+ and α_-.

(ii) If A is *parabolic* then both forward and backward orbits converge to the α-fixed point. Moreover, A preserves a foliation of circles (" horocycles") passing through α, all tangent to one line.

(iii) If A is *elliptic*, then it acts as a rotation by θ around points $\alpha_\pm$. Moreover, it preserves a foliation of circles separating these two points.

Exercise 2.9. *Justify the above description.*

2.3.4. *Compactness and degeneration.*

Exercise 2.10. *Show that topology of* $\mathrm{PSL}(2,\mathbb{C})$ *and topology of uniform convergence on the sphere coincide. Given an $\varepsilon>0$, let us consider the set $\mathcal{K}(\varepsilon)$ of Möbius transformations ϕ such that the triple $(\phi^{-1}(0,1,\infty)$ is ε-separated in the spherical metric (i.e., the three points stay at least distance ε apart). Show that $\mathcal{K}(\varepsilon)$ is compact in* $\mathrm{M\ddot{o}b}(\hat{\mathbb{C}})$.

Let us now describe the way how Möbius maps can degenerate:

Proposition 2.11. *If a family $\mathcal{F}$ of Möbius maps is not precompact in* $\mathrm{M\ddot{o}b}(\hat{\mathbb{C}})$ *then it contains a sequence $\{A_n\}$ such that $A_n(z) \to a$ uniformly on compact subsets of $\hat{\mathbb{C}} \smallsetminus \{b\}$, while $A_n^{-1}(z) \to b$ uniformly on compact subsets of $\hat{\mathbb{C}} \smallsetminus \{a\}$ (for some points $a, b \in \hat{\mathbb{C}}$ depending on the sequence).*

Proof. Without loss of generality we can assume that $\mathcal{F}$ is a sequence escaping to infinity in $\mathrm{PSL}(2,\mathbb{C})$.

For any two points $a, b \in \hat{\mathbb{C}}$ we can select a subsequence $\{M_n\}$ from $\mathcal{F}$ such that the limits $\alpha = \lim M_n(a)$ and $\beta = \lim M_n(b)$ exist. Assume first it can be done so that $\alpha \neq \beta$. Then the family of Möbius maps

$$\phi_n(z) = \frac{z - M_n(a)}{z - M_n(b)}$$

is precompact, so it is enough to verify the statement for the family

$$f_n = \phi_n \circ M_n \circ \psi, \quad \text{where } \psi(z) = \frac{(b-a)z}{z+1} + a$$

But $f_n(0) = 0$, $f_n(\infty) = \infty$, so $f_n(z) = \rho_n z$ for some $\rho_n \in \mathbb{C}^*$.

Since this sequence escapes to infinity in $\mathrm{PSL}(2, \mathbb{C})$, it contains a subsequence with $\rho_n \to 0$ or $\rho_n \to \infty$. In either case the conclusion is obvious.

Assume now that $\alpha = \beta$ for any choice of two points a, b and any subsequence $\{M_n\}$ as above. It implies that $M_n(a) \to \alpha$ pointwise along the whole sequence $\mathcal{F}$. Without loss of generality we can assume that $\alpha = \infty$. Similarly to the above argument, we can make a change of variable $f_n = \phi_n \circ M_n$ such that $f_n(\infty) = \infty$, so $f_n(z) = \rho_n z + c_n$, and $\phi_n \to \mathrm{id}$ uniformly on $\hat{\mathbb{C}}$.

Note that $c_n = f_n(0) \to \infty$. If $\rho_n = o(c_n)$ then $f_n \to \infty$ uniformly on compact subsets of $\mathbb{C}$. Otherwise, $f_n(z) = \rho_n(z - b_n)$ with $\rho_n \to \infty$, $b_n = O(1)$ along a subsequence. These affine maps have fixed points

$$\alpha_n = \frac{\rho_n b_n}{\rho_n - 1} = O(1).$$

Hence the translations $\psi_n : z \mapsto z - \alpha_n$ form a precompact family. Moreover, they conjugate the f_n to complex rescalings $g_n : z \mapsto \rho_n z$ for which the conclusion is obvious. $\square$

2.3.5. *Uniqueness of the sphere.* By the Geometric Uniformization Theorem (see Appendix 2), the standard sphere is the only Riemann surface endowed with spherical structure:

THEOREM 2.12. *The only (up to isometry) Riemann surface emdowed with a complete spherical structure is the standard sphere $S^2 \subset \mathbb{R}^3$.*

2.3.6. *Platonic orbifolds.* By taking quotients of the unit sphere $S^2 \subset \mathbb{R}^3$ by actions of finite groups of rotation, we obtain 2D *spherical orbifolds* (also called *elliptic*). The full list is provided by the serious of orbifolds with signatures $(S^2, \mathfrak{q}, \mathfrak{q})$, $\mathfrak{q} = 2, 3, \dots$ (corresponding to the cyclic groups of rotations), and three *Platonic orbifolds* corresponding to the five Platonic bodies.[2] Projecting each of these bodies B to the sphere, we obtain a spherical polygonal tiling $\mathcal{T} \equiv \mathcal{T}_B$. Let

$$\mathrm{Sym} \equiv \mathrm{Sym}(B) \subset \mathrm{SO}(3)$$

be the group of rotations preserving $\mathcal{T}$, and let

$$\widehat{\mathrm{Sym}} \equiv \widehat{\mathrm{Sym}}(B) \subset \mathrm{O}(3)$$

be the full group of symmetries of $\mathcal{T}$ (including reflections), so Sym is the index two normal subgroup of $\widehat{\mathrm{Sym}}$. Each tiling $\mathcal{T}$ can be further refined to a a *checker-board tessellation* of S^2 by black and white triangles (see Figure)), so that the full symmetry group $\widehat{\mathrm{Sym}}$ acts freely and transitively on the family of triangles, which makes each of these triagles *fundamental* for $\widehat{\mathrm{Sym}}$. The fundamental domain for the subgroup Sym is a spherical rectangle composed of two triangles, black and white. This description makes the groups $\mathrm{Sym} \subset \widehat{\mathrm{Sym}}$ examples of (spherical) *triangle groups.*

[2]The orbifolds corresponding to the dual bodies are the same.

EXERCISE 2.13. *For each Platonic body B,*

(i) *Justify the above tessellation picture.*

(ii) *Identify the corresponding symmetry groups,* $\mathrm{Sym} \subset \widehat{\mathrm{Sym}} \subset \mathrm{O}(3)$.

(iii) *Show that the Platonic orbifold $\mathcal{O} \equiv \mathcal{O}_B$ corresponding to the group* Sym *has signature* $(S^2; \{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\})$ *with*

$$\frac{1}{\mathfrak{p}} + \frac{1}{\mathfrak{q}} + \frac{1}{\mathfrak{r}} > 1. \tag{2.2}$$

(iv) *Show that*

$$\chi(\mathcal{O}) = \frac{2}{|\mathrm{Sym}|} > 0, \quad \text{while} \quad \mathrm{area}(\mathcal{O}) = \frac{4\pi}{|\mathrm{Sym}|} = 2\pi\, \chi(\mathcal{O}). \tag{2.3}$$

(v) *Conclude that the spherical area of the fundamental triangle Δ is equal to*

$$\mathrm{area}(\Delta) = \pi \cdot \left(\frac{1}{\mathfrak{p}} + \frac{1}{\mathfrak{q}} + \frac{1}{\mathfrak{r}} - 1 \right). \tag{2.4}$$

(vi) *Show that all signatures $(S^2; \{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\})$ satisfying (2.2) are realized by Platonic orbifolds.*

Formula (2.4) is a special case of the dollowing result:

GAUSS-BONNET FORMULA (FOR SPHERICAL TRIANGLES). For a spherical triangle Δ with angles α, β, γ we have:

$$\mathrm{area}(\Delta) = \alpha + \beta + \gamma - \pi.$$

EXERCISE 2.14. *Verify this formula.*

2.4. Hyperbolic geometry.

2.4.1. *Conformal disk and its automorphisms.* Let us consider a conformal disk S. It is a Riemann surface S conformally equivalent to the unit disk $\mathbb{D}$, or equivalently, to the upper half plane $\mathbb{H}$, or equivalently, to the strip $\mathbb{S}$. Using the isomorphism $S \approx \mathbb{D}$, it can be naturally compactified by adding to it the *ideal boundary* $\partial^I S \approx \mathbb{T}$ also called the *ideal circle* or the *absolute* (compare §1.7.8).

The group $\mathrm{Aut}(S)$ of conformal automorphisms of S in the the upper half-plane model consists of Möbius transformations with real coefficients:

$$M : z \mapsto \frac{az+b}{cz+d}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}).$$

Hence $\mathrm{Aut}(S) \approx \mathrm{SL}(2,\mathbb{R})/\{\pm I\} = \mathrm{PSL}(2,\mathbb{R})$. In the unit disk model, it is realized as the group $\mathrm{PSL}^{\#}(2,\mathbb{R})$:

$$M : z \mapsto \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}} = \lambda \frac{z-a}{1-\bar{a}z}, \quad \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathrm{PSL}^{\#}(2,\mathbb{R}),$$

where $\lambda = \alpha/\bar{\alpha} \in \mathbb{T}$, $a = -\beta/\alpha \in \mathbb{D}$ (see §50.8).

The above classification of Möbius transformations (see §2.3.2) has a clear meaning in terms of their action on S:

(i) A *hyperbolic* transformation $A \in \mathrm{PSL}(2,\mathbb{R})$ has two fixed points on the absolute $\partial^I S$ (and does not have fixed points in S). Its normal form in the $\mathbb{H}$-model is a dilation $z \mapsto \lambda z$ $(0 < \lambda < 1)$, and is a translation $z \mapsto z + a$ in the $\mathbb{S}$-model, where $a = \log \lambda$.

(ii) A *parabolic* transformation has a single fixed point on $\partial^I S$ (and does not have fixed points in S). Its normal form in the $\mathbb{H}$-model is the translation $z \mapsto z+1$.

(iii) An *elliptic transformation* $A \neq \mathrm{id}$ has a single fixed point $a \in S$ (and does not have fixed points on $\partial^I S$). Its normal form in the $\mathbb{D}$-model is a rotation $z \mapsto e(\theta)z$, $\theta \in \mathbb{R}/\mathbb{Z}$.

EXERCISE 2.15. (i) *Verify those of the above statements that look new to you.*
(ii) *Show that dilations* $z \mapsto \rho z$, $0 < \rho < 1$, *translations* $z \mapsto z + a$, $a \in [0, 1]$ (*in the* $\mathbb{H}$-*model*), *and rotations* $z \mapsto e(\theta)\, z$, $\theta \in \mathbb{R}/\mathbb{Z}$ (*in the* $\mathbb{D}$-*model*) *generate the whole group* $\mathrm{Aff}(\mathbb{R})$.
(iii) *Show that* $\mathrm{Aff}(\mathbb{R})$ *and the involution* $z \mapsto -1/z$ *generate the whole group* $\mathrm{Aut}(S)$ *of* $S \approx \mathbb{H}$.

2.4.2. *Hyperbolic metric.* A remarkable discovery by Poincaré is that a conformal disk S is endowed with the intrinsic hyperbolic structure, that is, there exists a Riemannian metric ρ_S on S of constant curvature -1 invariant with respect $\mathrm{PSL}(2, \mathbb{R})$-action. In the $\mathbb{H}$-, $\mathbb{D}$- and $\mathbb{S}$-models, the length element of ρ_S is given respectively by the following expressions:

$$d\rho_{\mathbb{D}} = \frac{2|dz|}{1-|z|^2}, \quad d\rho_{\mathbb{H}} = \frac{|dz|}{y}, \quad d\rho_{\mathbb{S}} = \frac{|dz|}{\sin y}, \tag{2.5}$$

where $z = x + iy$. This metric is called *hyperbolic* .

REMARK 2.16. Yet another useful model for the hyperbolic plane, the slit plane, will appear in §2.4.5.

EXERCISE 2.17. *Verify that the above three expressions correspond to the same metric on* S*, which has curvature* -1 *and is invariant under* $\mathrm{PSL}(2, \mathbb{R})$. *Show that the group of orientation preserving hyperbolic motions of* S *is equal to* $\mathrm{Aut}(S) \approx \mathrm{PSL}(2, \mathbb{R})$.

A conformal disk S endowed with the hyperbolic metric is called the *hyperbolic plane*.

2.4.3. $\mathbb{H}$ *as a symmetric space.* In this way, $\mathrm{PSL}(2, \mathbb{R})$ assumes the meaning of the group of (orientation preserving) *hyperbolic motions* of the hyperbolic plane. It acts freely transitively on the unit tangent bundle of $\mathbb{H}$, so the latter can be identified with $\mathrm{PSL}(2, \mathbb{R})$. The *isotropy group* of $i \in \mathbb{H}$ coincides with the group $\mathrm{PSO}(2)$ of hyperbolic rotations

$$z \mapsto \frac{z\cos\theta - \sin\theta}{z\sin\theta + \cos\theta}, \quad \theta \in \mathbb{R}/\pi\mathbb{Z}.$$

Thus, the hyperbolic plane gets identified with the *symmetric space*

$$\mathrm{PSL}(2, \mathbb{R})/\mathrm{PSO}(2) \approx \mathbb{H}. \tag{2.6}$$

REMARK 2.18. The symmetricity of $\mathbb{H}$ is reflected by the properties that it is homogeneous for the group pf motions, and for any pointed geodesic (γ, z) there is an isometric involuion $M : (\mathbb{H}, \gamma, z) \to (\mathbb{H}, \gamma, z)$ flipping γ around z (e.g., $z \mapsto -1/z$ flips $i \cdot \mathbb{R}_+$ around i, or in the $\mathbb{D}$-model: rotations by π flip the geodesics around 0).

From the Lie Theory point of view, the hyperbolic metric on $\mathbb{H}$ can be interpreted as follows. Let us consider the Lie algebra $\mathrm{sl}(2,\mathbb{R})$ of trace free 2×2 real matrices. It is endowed with the inner product $< a, b >= 2 \operatorname{tr} ab$ (the *Killing form*) which is invariant under the adjoint action

$$a \mapsto gag^{-1}, \quad a \in \mathrm{sl}(2,\mathbb{R}),\ g \in \mathrm{SL}(2,\mathbb{R}),$$

of $\mathrm{SL}(2,\mathbb{R})$ on $\mathrm{sl}(2,\mathbb{R})$.

Viewed as the linear space, $\mathrm{sl}(2,\mathbb{R})$ is the tangent space to $\mathrm{SL}(2,\mathbb{R})$ at the identity. By the left action of $\mathrm{SL}(2,\mathbb{R})$ on itself, the Killing form can be promoted to a left-invariant Riemannian metric on $\mathrm{SL}(2,\mathbb{R})$. Moreover, it descends to a metric on the symmetric space $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ invariant under $\mathrm{SL}(2,\mathbb{R})$-action.

EXERCISE 2.19. *Verify that this metric coincides* (*via the identification* (2.6)) *with the hyperbolic metric on* $\mathbb{H}$ *(subject of appopriate normalizations).*

REMARK 2.20. The Lie Theory discussion can be extended further to provide a general underlying principle for the hyperbolicty of the symmetric space $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ (without *a priori* familiarity with the hyperbolic plane $\mathbb{H}$). Namely, any irreducible non-compact symmetric space H of dimenion ≥ 2 is hyperbolic in the sense that the curvature is negative at any point in all two-dimensional directions (see [**W**, Cor, 8.4.6]). Since $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ is two-dimensional, there is only one direction at any point, so the curvature is constant by homegenuity.

2.4.4. *Hyperbolic geodesics and horocycles.* Hyperbolic geodesics in the $\mathbb{D}$-model of the hyperbolic plane are arcs of Euclidean circles orthogonal to the absolute $\mathbb{T}$. If we want to emphasize that we consider the full geodesic rather than a geodesic arc, we sometimes call it *complete*. For any hyperbolic unit tangent vector $v \in \mathrm{T}\mathbb{D}$, there exists a unique oriented complete hyperbolic geodesic tangent to v. For any two points x and y on the absolute, there exists a unique complete hyperbolic geodesic $\gamma = \gamma_{xy}$ with endpoints x and y.

EXERCISE 2.21. *Let* $\mathfrak{G}^+ \equiv \mathfrak{G}^+(\mathbb{D})$ *be the space of oriented complete hyperbolic geodesics in* $\mathbb{D}$ *endowed with the Hausdorff metric* (*associated with the Euclidean metric on* $\overline{\mathbb{D}}$). *Show that:*

(i) *The space* $\mathfrak{G}^+$ *is homeomorphic to* $(\mathbb{T}\times\mathbb{T})\smallsetminus\mathrm{diag}$*, where* $\mathrm{diag} = \{(x,x) \in \mathbb{T}\times\mathbb{T}\}$.

(ii) *The space* $\mathfrak{G} \equiv \mathfrak{G}(\mathbb{D})$ *of non-oriented geodesics is the quotient of* $\mathbb{T}\times\mathbb{T}\smallsetminus\mathrm{diag}$ *modulo the involution* $(x,y) \mapsto (y,x)$.

(iii) *The space* $\mathfrak{Gp}^+$ *of pointed oriented geodesics* (γ, z)*,* $z \in \gamma$*, is homeomorphic to* $\mathfrak{G}^+ \times \mathbb{R}$.

(iv) *If* $(\gamma_n, z_n) \to (\gamma, z)$ *in* $\mathfrak{Gp}^+$ *then the naturally parametrized geodesics* γ_n C^1*-converge to the parametrized* γ*. Modify this statement appropriately for non-oriented pointed geodesics.*

The stabilizer $\mathrm{Stab}_+(\gamma)$ of this geodesic preserving its orientation is the one-parameter group of hyperbolic transformations with the endpoints x and y fixed. (By normalizing $x = 0$ and $y = \infty$ in $\mathbb{H}$, we can bring it to the normal form

$$\mathrm{Stab}_+(\gamma) = \{z \mapsto \lambda z : \ \lambda \in \mathbb{R}_+\}.$$

Moreover, γ is called the *axis* of any $A \in \mathrm{Stab}_+(\gamma)$. The group $\mathrm{PSL}(2,\mathbb{R})$ acts freely and transitively on the space of pointed oriented hyperbolic geodesics.

FIGURE 2.1. Slit plane model for the hyperbolic plane and the dipole electric field.

EXERCISE 2.22. (i) *Verify the above assertions if they are not familiar to you.*
(ii) *Show that a hyperbolic R-neighborhood of the vertical axis $i \cdot \mathbb{R}_+$ in $\mathbb{H}$ is the sector $\{|\arg z - \pi/2| < \theta\}$ with $\theta = \theta(R) \in (0, \pi/2)$.*
(iii) *Show that for $M \in \mathrm{Stab}_+(\gamma)$ and for any $\zeta \in \gamma$, we have*

$$\Delta_M := \inf_{z \in \mathbb{H}_+} \mathrm{dist}_{\mathrm{hyp}}(z, M(z)) = \mathrm{dist}_{\mathrm{hyp}}(\zeta, M(\zeta)). \tag{2.7}$$

The quantity Δ_M (2.7) is called the *displacement* of M. A sector described in (ii) is called a *Stolz sector* centered at $0 \in \mathbb{H}$. In general, a Stolz sector centered at $a \in \partial\mathbb{H}$ is an R-hyperbolic neighborhood of a hyperbolic geodesic landing at a.

A *horocycle in* $\mathbb{D}$ *centered at* $x \in \mathbb{T}$ is a Euclidean circle $\gamma \subset \mathbb{D}$ tangent to $\mathbb{T}$ at x. A *horodisk* $D \subset \mathbb{D}$ is the disk bounded by the horocycle. In purely geometric terms, horocycles centered at x form a foliations orthogonal to the foliation of geodesics landing at x. The stabilizer of any horocycle (and the corresponding horoball) is the parabolic group fixing its center.

In fact, the $\mathbb{H}$-model fits better for describing horocycles: in this model the horocycles centered at $x = \infty$ are horizontal lines $\mathbb{L}_h \equiv \mathbb{L}_h(\infty) = \{\mathrm{Im}\, z = h\}$, the corresponding horoballs are the upper half-planes $\mathbb{H}_h \equiv \mathbb{H}_h(\infty) = \{\mathrm{Im}\, z > h\}$, which are stabilized by the one-parameter group of parabolic translations $z \mapsto z+t$, $t \in \mathbb{R}$. Similarly we let

$$\mathbb{L}_r(a) := \{z : |z - (a + ir/2)| = r/2\}, \quad \mathbb{L}_r(a) := \{z : |z - (a + ir/2)| < r/2\}$$

be horocycles and horoballs centered at $a \in \mathbb{R}$.

2.4.5. *Slit plane.* There is one more model of the hyperbolic plane which is useful in the real dynamics. Namely, let us consider an open interval $L \subset \mathbb{R}$ and the corresponding *slit plane*

$$\mathbb{C}(L) := \mathbb{C} \setminus (\mathbb{R} \setminus L) \tag{2.8}$$

(slit along two real rays).

For an angle $\theta \in [0, \pi)$, we let[3] $\mathbb{D}_\theta(L)$ be the $\mathbb{R}$-symmetric domain intersecting $\mathbb{R}$ along L whose upper half $\mathbb{D}^+_\theta(L) := \mathbb{D}_\theta(L) \cap \overline{\mathbb{H}}_+$ is bounded by a circle arc that meets the real line at angle θ, together with the interval L. Note that $\mathbb{D}_0(L) \equiv \mathbb{C}(L)$, while $\mathbb{D}_{\pi/2}(L)$ is the *Euclidean disk* $\mathbb{D}(L)$ *based upon* L *as a diameter.*

EXERCISE 2.23. (i) *Write down an explicit conformal map $\phi : \mathbb{C}(L) \to \mathbb{H}$.*
(ii) *Show that the interval L is a hyperbolic geodesic in $\mathbb{C}(L)$.*
(iii) *Show that $\mathbb{D}_\theta(L)$ is a hyperbolic $R(\theta)$-neighborhood of L.*

EXERCISE 2.24. *The above circle arcs represent the flow lines of the electrostatic field generated by the dipole of two opposite charges placed at ∂L.*

[3]A slight notational ambiguity ($\mathbb{D}_r$ vs $\mathbb{D}_\theta(L)$) hopefully will not cause confusion.

2.4.6. *Hyperbolic convexity.* A subset $Q \subset \overline{\mathbb{D}}$ is called (*hyperbolically*) *convex* if for any two points $x, y \in X$, the hyperbolic geodesic arc connecting x and y is also contained in Q. The boundary of Q in $\overline{\mathbb{D}}$ consists of two disjoint parts: the *relative boundary* $\partial^{\mathrm{rel}} Q$ in $\mathbb{D}$ and the *ideal boundary* $\partial^I Q$ in $\mathbb{T}$. We say that Q has a *totally geodesic boundary* if $\partial^{\mathrm{rel}} Q$ is the union of complete geodesics (in other words, there are no corners in $\partial^{\mathrm{rel}} Q$). Such sets appear naturally as follows.

The *hyperbolic convex hull* $\hat{X}$ of a subset $X \subset \overline{\mathbb{D}}$ is the smallest convex set containing X. For instance, let X be a closed subset of $\mathbb{T}$, and let $I_j \subset \mathbb{T}$ be the complementary intervals ("gaps") of X. Let us consider open (in $\overline{\mathbb{D}}$) hyperbolic half-planes $H_j \supset I_j$ based on the I_j (they are bounded the hyperbolic geodesics Γ_j that share the endpoints with I_j). Then

$$\hat{X} = \overline{\mathbb{D}} \smallsetminus \bigcup H_j. \tag{2.9}$$

Note that $\hat{X}$ is closed in $\overline{\mathbb{D}}$ and $\hat{X} \cap \mathbb{T} = X$. In fact, we have:

LEMMA 2.25. *For a closed non-singleton $X \subset \mathbb{T}$, the convex hull $\hat{X}$ is a closed Jordan disc.*

PROOF. The boundary of $\hat{X}$ can be homeomorphically retracted onto $\mathbb{T}$ by projecting the boundary geodesics Γ_j onto the ideal intervals I_j. To be definite, one can take the orthogonal projection along geodesics orthogonal to Γ_j. (To see it explicitly, move Γ_j by a Möbius automorphism to a diameter of $\mathbb{D}$.) □

COROLLARY 2.26. *Let $x \in \hat{X}$ and $y \in \mathbb{T} \smallsetminus X$. Then any path connecting x to y crosses some boundary geodesic Γ_j.*

EXERCISE 2.27. *Let Q be a hyperbolically convex subset of $\overline{\mathbb{D}}$ which is a closure of its interior. Then Q has a totally geodesic boundary iff $Q = \hat{X}$ for some closed subset $X \subset \mathbb{T}$. Moreover, under these circumstances, $X = \partial^I Q$.*

EXERCISE 2.28. *Let X_n be a sequence of closed subsets of $\mathbb{T}$ converging in the Hausdorff metric to a set X (i.e., in the space $\mathfrak{S}(\mathbb{T})$). Then $\hat{X}_n \to \hat{X}$ in the space $\mathfrak{S}(\overline{\mathbb{D}})$.*

2.4.7. *Hyperbolic triangles.* Let us consider three points A, B, C in the hyperbolic plane $\mathbb{D}$ or on the absolute $\mathbb{T}$. Connecting them with arcs of hyperbolic geodesics, we obtain a *hyperbolic triangle* Δ with vertices A, B, C. Its boundary inherits the orientation from the complex plane, giving the cyclic order to the vertices. Relabeling the vertices if necessary, we can assume that the cyclic order (A, B, C) is positive. Let α, β, γ be the angles at the vertices A, B, C. Notice that an angle, say α, vanishes iff the corresponding vertex A is ideal: $A \in \mathbb{T}$.

In a striking contrast with Euclidean geometry, the vertices determine the triangle:

THEOREM 2.29. *For any cyclically ordered set (α, β, γ) of non-negative angles satisfying*

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} < 1,$$

there exists a unique, up to a hyperbolic motion, oriented hyperbolic triangle $\Delta \equiv \Delta(\alpha, \beta, \gamma)$ with these angles.

FIGURE 2.2. The left hand-side shows an ideal triangle with angles $(\alpha, \beta, 0)$. The right-hand one illustrates how it can be deformed to a triangle with with arbitrary admissible angles (α, β, γ).

FIGURE 2.3. Representation of an arbitrary triangle as a difference of two ideal triangles.

PROOF. A realization of the triangle $\Delta(\alpha, \beta, 0)$ with one ideal vertex is shown on the left-hand side of Figure 2.2. It can be then deformed to the triangle $\Delta(\alpha, \beta, \gamma)$ with arbitrary $\gamma \in [0, \pi - \alpha - \beta)$, see the right-hand side of that figure. One can also readily see that any triangle with given angles can be moved to a form depicted on the figure. We leave the reader to fill in details. $\square$

As the angles determine the triangle up to a hyperbolic motion, they should determine its area, too. This leads to the following remarkable relation:

GAUSS-BONNET FORMULA (FOR HYPERBOLIC TRIANGLES). *The area of the hyperbolic triangle Δ with angles (α, β, γ) is equal to the "angle defficiency":*

$$\operatorname{area} \Delta = \pi - (\alpha + \beta + \gamma).$$

PROOF. For the ideal triangle $\Delta(\alpha, \beta, 0)$ depicted on Figure 2.2, it can be checked by a direct calculation. The general case follows by representing an arbitrary triangle as a difference of two ideal ones, as shown on Figure 2.3. $\square$

Of course, the angles should determine the lengths of the edges of the trinagle as well. It follows from the following Sine Theorem (combined with the Gauss-Bonnet):

HYPERBOLIC SINE THEOREM. *The lengths a, b, c of the edges of the triangle $\Delta(\alpha, \beta, \gamma)$ are related to its angles as follows:*

$$\frac{\sinh a}{\sin \alpha} = \frac{\sinh b}{\sin \beta} = \frac{\sinh c}{\sin \gamma}.$$

PROJECT 2.30. *Study a proof of the Hyperbolic Sine Theorem, as well as other basic aspects of the Hyperbolic Trigonometry, see* [**Be1**, §7].

Similarly, we can consider *hyperbolic polygons* P. As for triangles, we allow some of their vertices to be ideal, representing *cusps* of P. In fact, will also need to consider more general ideal polygones with *ideal sides* (i.e., arcs of $\mathbb{T}$).

EXERCISE 2.31. *Show that if P is a hyperbolic n-gon without ideal sides then*

$$\operatorname{area} P = \pi(n-2) - \sum \alpha_i,$$

where α_i are its angles.

An important feature of the hyperbolic geometry is that projections to geodesics are exponentially contracting:

EXERCISE 2.32. (i) *For any geodesic γ in $\mathbb{H}$ and any point $z \in \mathbb{H}$, there is a unique point $\pi_\gamma(z) \in \gamma$ which is closest to z on γ. Moreover, the geodesic connscting z to $\pi_\gamma(z)$ is orthogonal to γ.*

(ii) *For any two points $z, \zeta \in \mathbb{H}$ on distance $\geq R$ from γ, we have:*

$$\operatorname{dist}_{\mathrm{hyp}}(\pi_\gamma(z), \pi_\gamma(\zeta)) \leq e^{-R} \operatorname{dist}_{\mathrm{hyp}}(z, \zeta).$$

2.4.8. *Fuchsian groups: dynamical structure.* A *Fuchsian group* Γ is a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R})$ acting on the hyperbolic plane. Let us start with a structural theorem for these actions.

THEOREM 2.33. *Let Γ be a Fuchsian group acting on $(\hat{\mathbb{C}}, \overline{\mathbb{D}}, \mathbb{T})$. Then there is a closed Γ-invariant set $\Lambda = \Lambda(\Gamma) \subset \mathbb{T}$ with the following properties:*

(i) *Λ is the limit set of any* $\operatorname{orb}(z)$, *$z \in \hat{\mathbb{C}}$. In prticular, the action of Γ on Λ is minimal: any orbit is dense in Λ.*

(ii) *The action of Γ on the complementary set, $\Omega(\Gamma) := \hat{\mathbb{C}} \smallsetminus \Lambda(\Gamma)$, is properly discontinuous.*

(iii) *Λ is the closure of hyperbolic fixed points.*

(iv) *If Γ contains a parabolic map, then Λ is the closure of parabolic fixed points.*

(v) *If Γ contains an elliptic map then it has finite order, it has a unique fixed point α in $\overline{\mathbb{D}}$, and this point belongs to $\mathbb{D}$ (in particular, it does not belong to Λ).*

(vi) *Λ is either the whole circle $\mathbb{T}$, or a Cantor set, or esle $|\Lambda| \leq 2$. In the latter case, there are four options:*

a) *Γ is a finite cyclic group of elliptic rotations around a fixed point $\alpha \in \mathbb{D}$, and $\Lambda = \emptyset$. The normal form for Γ is $z \mapsto e(k/\mathfrak{q})\, z$, $k \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$.*

b) *Γ is an infinite cyclic group of parabolic translations with a common fixed point $\alpha \in \mathbb{T}$, and $\Lambda = \{\alpha\}$. The $\mathbb{H}$-normal form*[4] *for Γ is*

$$z \mapsto z + n, \quad n \in \mathbb{Z}.$$

c) *Γ is an infinite cyclic group of hyperbolic maps with common fixed points $\alpha, \beta \in \mathbb{T}$, and $\Lambda = \{\alpha, \beta\}$. The $\mathbb{H}$-normal form for Γ is*

$$\Gamma_\rho = \{z \mapsto \rho^n z, \quad n \in \mathbb{Z}\} \text{ for some } \rho \in (0,1). \tag{2.10}$$

d) *Γ is an infinite diahedral group*[5] *that has an orbit $\{\alpha, \beta\} \subset \mathbb{T}$ of size two. Moreover, $\Lambda = \{\alpha, \beta\}$. The $\mathbb{H}$-normal form for Γ is generated by the above cyclic group Γ_ρ (for some $\rho \in (0,1)$) and the involution $z \mapsto -1/z$.*

PROOF. Assume first that Γ is infinite and that there are no points fixed under the whole group. Since Γ is discrete, it is countable, $G = \{g_n\}_{n \in \mathbb{N}}$ where $g_0 = \mathrm{id}$, and moreover $g_n \to \infty$ in $\mathrm{PSL}(2, \mathbb{R})$ (meaning that the g_n eventually escape any compact subset of $\mathrm{PSL}(2, \mathbb{R})$).

Let $\omega(z)$ stand for the limit set of $\operatorname{orb} z$ (i.e., it consists of the limits of all convergent subsequences $(g_{n(k)} z)$ as $n_k \to \infty$). It is a non-empty closed subset of $\hat{\mathbb{C}}$. By Proposition 2.11, if $a \in \omega(z)$ then $a \in \omega(\zeta)$ for all points $\zeta \in \hat{\mathbb{C}}$ except at most one point b. But if b is exceptional then so is gb for any $g \in \Gamma$ (since $\omega(gb) = \omega(b)$), implying that the exceptional point must be fixed under the whole group. As we assume that such fixed points do not exist, there are no exceptional points either, so $a \in \omega(z)$ for any $z \in \hat{\mathbb{C}}$. Consequently, the limit set $\omega(z)$ is independent of z, and we can call it $\Lambda \equiv \Lambda(\Gamma)$.

---

[4]i.e., the normal form in the upper-half plane model $\mathbb{H}$ for the hyperbolic plane.

[5]i.e., the semidirect product of $\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$, with the latter acting on $\mathbb{Z}$ by reflection: $n \mapsto -n$.

We can also make the following conclusions:

- Λ is c Γ-invariant (since for any $z \in \hat{\mathbb{C}}$, we have: $g(\Lambda) = \omega(gz) = \omega(z) = \Lambda$).
- $\Lambda \subset \mathbb{T}$ (since $\omega(z) \subset \mathbb{T}$ for $z \in \mathbb{T}$).

This confirms (i) (under our current assumptions).

Proposition 2.11 also implies that any sequence of distinct elements $g_n \subset \Gamma$ contains a sunbsequence g_{n_k} such that for some points $a, b \in \Lambda$, we have: $g_{n_k} z \to a$ and $g_{n_k}^{-1} z \to b$ uniformly on compact subsets of $\Omega = \hat{\mathbb{C}} \smallsetminus \Lambda$.[6] This implies that the action of Γ on Ω is properly discontinuous, confirming (ii).

Let $\mathrm{Fix}(\Gamma)$ be the union of the sets of fixed points of all maps $g \in \Gamma$, $g \neq \mathrm{id}$. Notice that if α is a fixed point for some $f \in \Gamma$ and $g \in \Gamma$ then $g\alpha$ is a fixed point for $g \circ f \circ g^{-1}$. It follows that $\mathrm{Fix}(\Gamma$ is Γ-invariant, and moreover, the action of Γ on $\mathrm{Fix}(\Gamma)$ preserves type of points (hyperbolic/parabolic/eilliptic).

By minimality of the action, if Λ contains one fixed point of certain type then fixed points of that type are dense in Λ. Obviously, any hyperbolic or parabolic fixed point (if exists) belongs to Λ. This confirms (iv). To complete the proof of (iii), we need to show that hyperbolic points always exist (under our assumptions). It follows from the following assertion[:

Exercise 2.34. *Let A and B be two elements of* $\mathrm{PSL}(2, \mathbb{R})$ *that do not share fixed points. If both of them are parabolic, or else if one of them is elliptic, then the commutator $\{A, B\} := ABA^{-1}B^{-1}$ is hyperbolic.*

Remark 2.35. Another approach to this issue is to find a non-peripheral closed geodesic (not necessarily simple) on the quotient Riemann surface $S = \mathbb{D}/\Gamma$ (or rather, on the quotient orbifold). Its lift to $\mathbb{D}$ is an axis for a hyperbolic deck transformation (compare Prop. 2.53 below).

We leave to the reader to verify (v), and pass to (vi). If Λ has non-empty interior in $\mathbb{T}$, then by (iii), $\operatorname{int}\Lambda$ contains a fixed point α of some hyperbolic transformation $g \in \Gamma$. But then the orbit $\displaystyle\bigcup_{n \in \mathbb{Z}} g^n(\Lambda)$ can omit at most one point on $\mathbb{T}$. Since Λ is invariant and closed, it must coincide with the whole circle $\mathbb{T}$.

Thus, if $\Lambda \neq \mathbb{T}$ then it is nowhere dense. Let us show that it is perfect. Notice that if $g \in \Gamma$ is a hyperbolic map with fixed points α and β, then $\{\alpha, \beta\} \subset \omega(z)$ for any $z \in \mathbb{T} \smallsetminus \{\alpha, \beta\}$. It follows that if $|\Lambda| > 2$ then hyperbolic fixed points are not isolated in Λ. Now (iii) implies that there are no isolated points in Λ.

Let us now deal with the remaining special cases.

Case 1) *Assume there is a point α fixed under the whole group.* If $\alpha \in \mathbb{D}$ then Γ is a group of elliptic rotations around α. Since it is discrete, it must be finite, leading to the option (vi-a)

Let now α lie on the absolute $\mathbb{T}$. Let us pass to the upper half-plane model $\mathbb{H}$ putting α at infinity. Then Γ becomes a subgroup of affine transformations $z \mapsto \rho z + c$ with $\rho \in \mathbb{R}^*$, $c \in \mathbb{R}$. If $\rho = 1$ for all $g \in \Gamma$ then Γ is a discrete subgroup of the one-parameter parabolic group $z \mapsto z + t$. It follows that Γ is cyclic and $\Lambda = \{\alpha\}$, yielding option (vi-b).

---

[6] Compare with the notion of normality and the Montel Theorems below (§4).

Otherwise Γ contains a hyperbolic element g, which can be brought to the form $g : z \mapsto \rho z$. If there is some $h \in \Gamma$ that does not fix 0 then $h(z) = \lambda z + c$ with $c \neq 0$, and $g^n \circ h \circ g^{-n}$ is equal to $z \mapsto \lambda z + \mu^n c$, $n \in \mathbb{Z}$, which is not discrete.

It follows that G is a discrete subgroup of the one-parameter hyperbolic group $z \mapsto \rho z$, $\rho \in \mathbb{R}^*$. Hence G is cyclic again, and $\Lambda = \{0, \infty\}$ (in the normal form), which amounts to case (vi-c)

Case 2) Assume $|\Lambda| \leq 2$.

Assume $\Lambda = \emptyset$, which is equivalent for Γ to be finite. Then all elments of Γ must have finite order, so they must be elliptic. Then by Exercise 2.34, they must share a fixed point, bringing us back to Case 1).

If $|\Lambda| = 1$ then $\Lambda = \{\alpha\}$, where α is fixed under the whole Γ, bringing us back to Case 1) again.

If $|\Lambda| = 2$ then we can normalize it so that $\Lambda = \{0, \infty\}$. Then Γ contains a normal subgroup Γ_ρ (2.10) fixing these points, so $[\Gamma : \Gamma_\rho] \leq 2$. Thus, either $\Gamma = \Gamma_\rho$, bringing us back to Case 1) once again, or Γ is a semidirect product of Γ_ρ and an involution permuting 0 and ∞. Such an involution has a form $z \mapsto -\lambda/z$ with $\lambda > 0$, which is conjugate to $\zeta \mapsto -1/\zeta$ by rescaling $z = \sqrt{\lambda}\,\zeta$. We arrive at case (vi-d). □

The set $\Lambda(\Gamma)$ is naturally called the *limit set* of Γ, while the complementary set $\Omega(\Gamma)$ is called the *set of discontinuity*. One says that Γ is a Fuchsian group of *first kind* if $\Lambda(\Gamma) = \mathbb{T}$. If $\Lambda(\Gamma)$ is a Cantor set then Γ is called a Fuchsian group of *second kind*. The special groups listed in item (vi) are called *elementary*. We see that $|\Lambda| \leq 2$ for such a group, while Λ is uncountable for all others.

2.4.9. *Hyperbolic Riemann surfaces.* Since Γ acts properly discontinuous on Ω, the quotient space $S := \mathbb{D}/\Gamma$ is Hausdorff. Moreover, if Γ acts freely on $\mathbb{D}$, then the complex structure and the hyperbolic metric naturally descend from $\mathbb{D}$ to S, and we obtain a *hyperbolic Riemann surface*.

Under the above circumstances, one says that S is *uniformized by a Fuchsian group*. The Geometric Uniformization Theorem (see Appendix 2 below) yields:

THEOREM 2.36. *Any complete hyperbolic*[7] *Riemann surface can be uniformized by a Fuchsian group.*

2.4.10. *Cusp and annulus.* Let us take a look at elementary Fuchsian groups listed in items (vi-b) and (vi-c) of Theorem 2.33. Let us start with parabolic cyclic groups, Case (vi-b).

The quotient of a horoball $\mathbb{H}_h$ by a discrete cyclic group of parabolic transformations $\mathbb{Z} =< z \mapsto z + n >$ is called a *cusp*. Conformally it is the punctured disk $\mathbb{D}^*$, hyperbolically it is the *pseudosphere*. Simple closed curves $\mathbb{L}_t/\mathbb{Z} \subset \mathbb{H}_h/\mathbb{Z}$, $t > h$ (see §2.4.4) are also called *horocycles* (in the cusp).

EXERCISE 2.37. *Any cusp $\mathbb{H}_h/\mathbb{Z}$ has infinite hyperbolic diameter but a finite hyperbolic area. The hyperbolic length of the horocycle $\mathbb{L}_t/\mathbb{Z}$ goes to zero as $t \to \infty$.*

[7] In this statement, S is assumed to be *endowed* with a metric of constant negative curvature, while the general Uniformization Theorem does not assume this. (A similar remark applies to Theorems 2.4 and 2.12 above.)

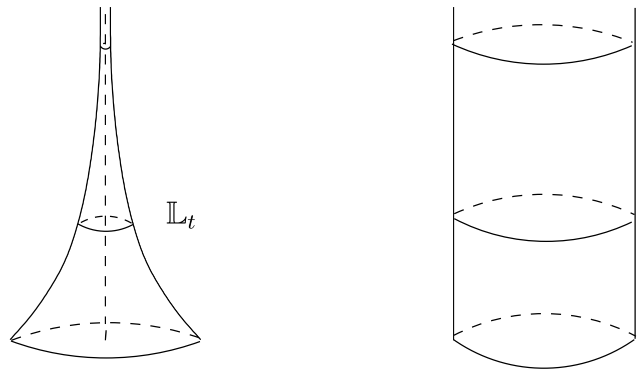

FIGURE 2.4. The hyperbolic cusp (on the left) vs the flat half-infinite cylider (on the right). Metrically they are quite different, though conformally they are equivalent.

EXERCISE 2.38. *Write down explicitly the universal covering* $\mathbb{H} \to \mathbb{D}^*$ *and show that the hyperbolic metric on* $\mathbb{D}^*$ *is equal to* $ds = \dfrac{|dz|}{|z \log |z|\,|}$.

EXERCISE 2.39. *Let* S *be a Riemann surface, and let* $f : S \to \mathbb{D}^*$ *be a holomorphic covering. If* $\deg f < \infty$ *then* S *is isomorphic to* $\mathbb{D}^*$*; otherwise* S *is isomorphic to* $\mathbb{H}$.

Let us now pass to the hyperbolic cyclic groups: Case (vi-c) of Theorem 2.33.

EXERCISE 2.40. (i) *The quotient* $\mathbb{H}/\Gamma_\rho$ *is conformally equivalent to an annulus* $\mathbb{A} := \mathbb{A}(1, R)$ *with some* $R > 1$ (*which one?*).

(ii) *The circle* $\gamma_{\mathbb{A}} := \mathbb{T}_{\sqrt{R}}$ *is a simple closed hyperbolic geodesic in* $\mathbb{A}$. *Calculate its hyperbolic length.*

(iii) *There is an anti-holomorphic involution* $\sigma : \mathbb{A} \to \mathbb{A}$ *such that* $\gamma_{\mathbb{A}} = \mathrm{Fix}(\sigma)$.

This hyperbolic geodesic $\gamma_{\mathbb{A}}$ is called the *equator* of A.

Let us now consider any conformal annulus A. By definition, it is conformally equivalent to a round annulus $\mathbb{A}(r, R)$ with some $0 < r < R < \infty$. (Of course, it can be normalized so that $r = 1$ as above.) Such annuli can be easily conformally classified:

EXERCISE 2.41. *Two round annuli are conformally equivalent if and only if* $R/r = R'/r'$. *The only conformal isomorphisms* $\mathbb{A}(r, R) \to \mathbb{A}(r', R')$ *are complex rescaling*

$$z \mapsto \lambda\, e(\theta), \quad \lambda = \frac{R'}{r'} : \frac{R}{r}, \quad \theta \in \mathbb{R}/\mathbb{Z}.$$

Hence the *modulus of* A,

$$\operatorname{mod} A := \frac{1}{2\pi} \log \frac{R}{r} \in (0, \infty),$$

is well defined and is the complete conformal invariant of A.

REMARK 2.42. In §6.3.1 we will give an intrinsic definition of the modulus, in terms of the *extremal length* of some path family, which is one of the main tools in Holomorphic Dynamics.

Moreover, Exercise 2.40(ii) implies that any conformal annulus A has an *equator* (a simpe closed hyperbolic geodesic whose homotopy class generates $\pi_1(A)$) and yields the following expression for its hyperbolic length:

$$l_{\text{hyp}}(\gamma_A) = \frac{\pi}{\operatorname{mod} A}. \tag{2.11}$$

(Note that in the round model $\mathbb{A} \equiv \mathbb{A}(r, R)$, The equator becomes the "geometric mean circle", $\gamma_{\mathbb{A}} = \{|z| = \sqrt{Rr}\}$.)

REMARK 2.43. Later on we will: a) see that the equator is the *only* simple closed geodesic in a conformal annulus (see §2.4.16); b) discuss another useful model for a conformal annulus, a *flat cylinder* (see §2.6.1).

If A is a conformal annulus with boundary then $\operatorname{mod} A$ is defined as the modulus of $\operatorname{int} A$.

The punctured disk $\mathbb{D}^* \approx \mathbb{H}/\mathbb{Z}$ can be viewed as the annulus $\mathbb{A}(0, 1)$, so it is natural to let $\operatorname{mod} \mathbb{D}^* = \infty$. Similarly, $\mathbb{C}^*$ can be viewed as the annulus $\mathbb{A}(0, \infty)$. All the more, we let $\operatorname{mod} \mathbb{C}^* = \infty$.

EXERCISE 2.44. *Show that:*

(i) $\mathbb{D}^*$ *is not conformally equivalent to any annulus* $\mathbb{A}(r, R)$*,* $0 < r < R < \infty$*;*

(ii) $\mathbb{C}^*$ *is not conformally equivalent to any of the above surfaces.*

We see that a Riemann surface under considerarion "knows" wherther it has a *puncture* or an *ideal circle* at infinity of its end, which contrasts with the topological situation discussed above (see §1.7.8).

REMARK 2.45. Later on (§5) we will see that any Riemann surface which is a *topological* annulus is conformally equivalent to one of the models from the last exercise. Moreover, the type of an embedded annulus for $A \subset \hat{\mathbb{C}}$ can be recognized by checking whether the components of $\hat{\mathbb{C}} \smallsetminus A$ are singletons or continua (see §6.3.3).

2.4.11. *Hyperbolic orbifolds.* Assume now the action of a Fuchsian group Γ on $\mathbb{D}$ is not free, so it has fixed points $z_i \in \mathbb{D}$. Since the action of Γ on $\mathbb{D}$ is properly discontinuous,

- these points are isolated, and so, there are at most countably many of them;
- each $\operatorname{Stab}(z_i)$ is finite, and hence is a cyclic group $\mathbb{Z}/\mathfrak{q}_i\mathbb{Z}$.

Each z_i projects to a singular point α_i of the quotient $S := \mathbb{D}/\Gamma$. Thus, we obtain an orbifol $\mathcal{O}$ with signature $(S; \{\mathfrak{q}_i)\}$. Moreover, the hyperbolic metric on $\mathbb{D}$ descends to a hyperbolic metric on S with cone singularities of angles $2\pi/\mathfrak{q}_i$ at the α_i.

Elementary Fuchsian groups of this kind are covered by Cases (vi-a) and (vi-d) of Theorem 2.33. In the former case, we obtain the cone with angle $2\pi/\mathfrak{q}$ (which has signature $(\mathbb{D}; \mathfrak{q})$).

EXERCISE 2.46. *Describe the orbifold that appears in* Case (vi-d).

2.4.12. *Modular group, surface, and function J.* The *modular group* is the group $\Gamma := \mathrm{PSL}(2,\mathbb{Z})$ naturally acting on the closed upper half-plane $\mathrm{cl}\,\mathbb{H}^+ = \mathbb{H}^+ \cup \mathbb{R}$ by Möbius transformations.

EXERCISE 2.47. (i) *The modular group Γ is generated by the parabolic map $\gamma : z \mapsto z+1$ and the order two elliptic map $\delta : z \mapsto -1/z$.*
(ii) *Letting $\Pi^+ := \{z \in \mathbb{H}^+ : \ |\operatorname{Re} z| \le 1\}$, the set $\Delta := \Pi^+ \smallsetminus \mathbb{D}$ is a fundamental donain for Γ.*
(iii) *The fundamental domain Δ contains three fixed points (on its boundary), $i = e(1/4)$ is the order two fixed point (for δ) while $e(1/6)$, $e(1/3)$ are order three fixed points (for $\delta\gamma^{-1}$ and $\delta\gamma$ respectively).*
(iv) *The quotient $\mathbb{H}/\Gamma$ is a hyperbolic orbifold $\mathfrak{M}$ supported on $\mathbb{C}$ with one cusp (at infinity of $\mathbb{C}$) and two cone points, of order two and three.*
(v) *The limit set for Γ is the whole circle $\hat{\mathbb{R}} \equiv \mathbb{R} \cup \{\infty\}$. (So Γ is a Fuchsian group of the first kind.)*
(v) *The modular group is isomorphic to the free product of two cyclic groups:*

$$\Gamma \approx (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/3\mathbb{Z}).$$

The quotient $\mathfrak{M} := \mathbb{H}/\Gamma$ is called the *modular surface* and the corresponding projection $J : \mathbb{H} \to \mathfrak{M} \approx \mathbb{C}$ is called the *modular function*.

EXERCISE 2.48. *Let $\hat{\Delta} := \Delta \cap \{\operatorname{Re} z \ge 0\}$. It is a hyperbolic triangle with vertices $i, e(1/6), \infty$. Consider the group $\hat{\Gamma}$ generated by reflections with respect to the sides of $\hat{\Delta}$.*
(i) *Show that the translations of $\hat{\Delta}$ under $\hat{\Gamma}$ tessellate $\mathbb{H}$.*
(ii) *Show that the modular group Γ is the unique index two normal subgoup of $\hat{\Gamma}$.*

The modular group will naturally appear in the discussion of the moduli space of complex tori and four-times-punctured spheres (see §§2.6.3, 2.6.4).

2.4.13. *Thrice-punctured sphere, ideal triangle group, and modular function λ.* Let us now consider the thrice-punctured sphere[8] $\mathbb{C} \smallsetminus \{0,1\}$. For this domain, there is a simple explicit construction of its uniformization by a Fuchsian group. Namely, let us consider an ideal triangle Δ in the hyperbolic plane, that is, the geodesic triangle with vertices on the absolute[9] (see Figure 2.4.13). By the Riemann Mapping Theorem, it can be conformally mapped onto the upper half-plane $\mathbb{H}_+$ so that its vertices go to the points $0, 1$ and ∞. By the Schwarz Reflection Principle, this conformal map can be extended to the three symmetric ideal triangles obtained by reflection of Δ in its edges. Each of these symmetric rectangles will be mapped onto the lower half-plane $\mathbb{H}_-$. Then we can extend this map further to the six symmetric rectangles each of which will be mapped onto $\mathbb{H}_+$ again, etc. Proceeding in this way, we obtain the desired universal covering $\lambda : \mathbb{D} \to \mathbb{U}$ called a *modular function*.

EXERCISE 2.49. *Verify the following properties:*
(i) *The union of the above triangles tile the whole disk $\mathbb{D}$;*
(ii) *The modular function λ provides us with the Universal covering $\mathbb{D} \to \mathbb{C} \smallsetminus \{0,1\}$;*

[8]Note that all thrice-punctured spheres are equivalent under the action of the Möbius group $\text{Möb}(\hat{\mathbb{C}})$.
[9]Note that all these triangles are equivalent under the action of $\mathrm{PSL}(2,\mathbb{R})$.

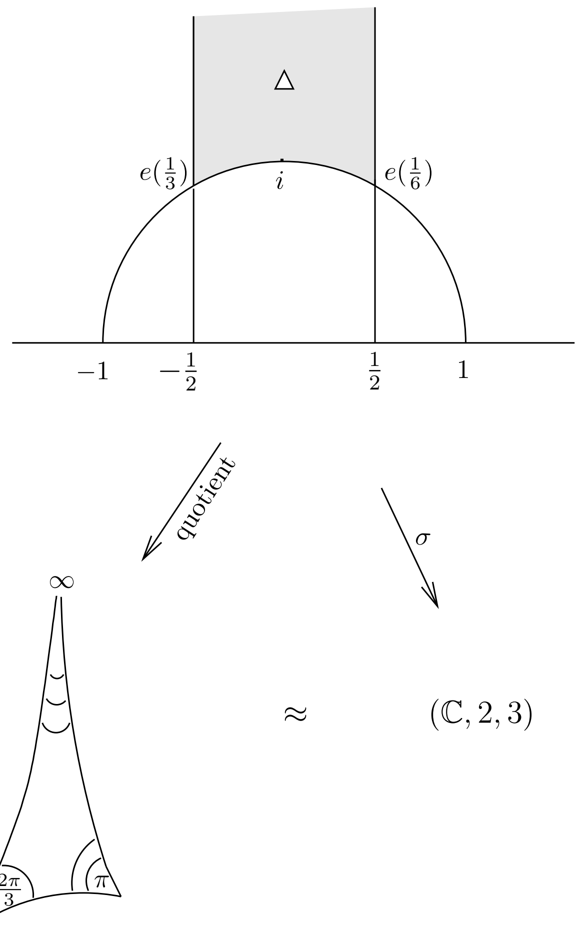

FIGURE 2.5. Modular group, surface, and function J.

(iii)) *Its group of deck transformations is the index 6 normal subgroup of the modular group* $\mathrm{PSL}(2,\mathbb{Z})$ *consisting of matrices congruent to* $I \mod 2$

(iv) *There is a natural Galois orbifold covering* $\mathbb{C} \smallsetminus \{0,1\} \to \mathfrak{M}$ *ofr degree 6. Describe its group of deck transformations.*

The above Fuchsian group is called the *congruent group* Γ_2, or the *ideal triangle group.*

Pushing the hyperbolic metric down by λ from $\mathbb{D}$ to $\mathbb{C} \smallsetminus \{0,1\}$, we endow $\mathbb{C} \smallsetminus \{0,1\}$ with a complete hyperbolic metric (with three cusps corresponding to $0, 1, \infty$).

2.4.14. *Hyperbolic triangle groups.* We have seen above a short list of examples of spherical and parabolic orbifolds with three singular points corresponding to chess-board triangle tessellations of S^2 and $\mathbb{R}^2$ (see Exercises 2.13, 2.6). We have also seen examples of the modular groups corresponding to the chess-board tessellations of the hyperbolic plane $\mathbb{H}$ by triangles with angles $\{0, \pi/3, \pi/2\}$ and by ideal triangles with all angles 0 (respectively). It turns out that there exists a plenty

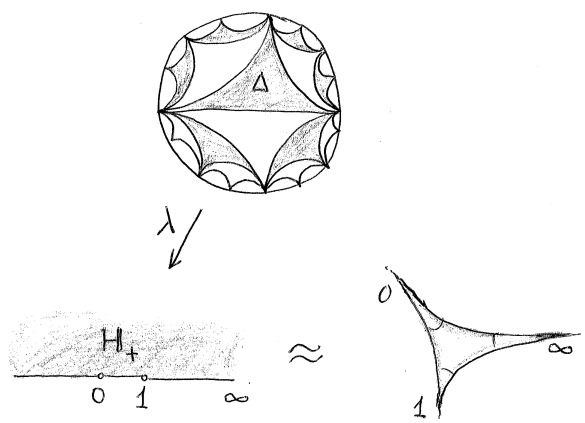

FIGURE 2.6. Modular function λ as the universal covering over the thrice-punctured sphere, which is presented in the flat and hyperbolic models.

of such hyperbolic examples. Indeed, by Theorem 2.29, for any triple $\{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\}$ of numbers in $\{2, 3, \dots\} \cup \{\infty\}$ with $\frac{1}{\mathfrak{p}} + \frac{1}{\mathfrak{q}} + \frac{1}{\mathfrak{r}} < 1$ there exists a hyperbolic triangle Δ with angles $\pi/\mathfrak{p}$, $\pi/\mathfrak{q}$, and $\pi/\mathfrak{r}$. Let $\sigma_{\mathfrak{p}}$, $\sigma_{\mathfrak{q}}$, and $\sigma_{\mathfrak{r}}$ be the reflections with respect to the edges of these triangles, and let $\widehat{\Gamma}$ be a group generated by these reflections.

PROBLEM 2.50. (i) $\widehat{\Gamma}$ *is a discrete group generating a checker-board tesselletation of* $\mathbb{H}$ *by the transformations of* Δ*, with* Δ *serving as a fundamental domain.*

(ii) $\widehat{\Gamma}$ *contains the index two subgroup* $\Gamma \subset \mathrm{PSL}(2, \mathbb{R})$ *whose fundamental domain is a rectangle consisting of the union of two triangles, black and white.*

(iii) *The orbifold* $\mathcal{O} \equiv \mathcal{O}_{\{\mathfrak{p},\mathfrak{q},\mathfrak{r}\}}$ *corresponding to* Γ *has signature* $(S^2; \{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\})$ *and Euler characteristic*

$$\chi(\mathcal{O}) = \frac{1}{\mathfrak{p}} + \frac{1}{\mathfrak{q}} + \frac{1}{\mathfrak{r}} - 1 < 0.$$

2.4.15. *Fundamental domain.* For a finitely generated Fuchsian group Γ, let us consider its Dirichlet fundamental domain Δ (see §1.5.1). Its construction implies that it is a hyperbolic polygon, possibly ideal. Its ideal vertices correspond to the cusps of the quotient Riemann surface $S := \mathbb{H}/\Gamma$, while ideal sides correspond to the ideal boundary components for S. Moreover, the sides of Δ are paired: for any side e, there is an element $\gamma_e \in \Gamma$ such that $\gamma_e(e)$ is a side e' of Δ, so $\gamma(\Delta)$ is attached to Δ along e'. So, § is obtained from Δ by gluing e to e' by means of γ_e.

It may also happen that for a vertex $v \in \mathbb{H}$ of Δ, there is a composition γ_v of some γ_e's that fixes v. If v is non-ideal, then γ_v is elliptic with a rational rotation number $1/\mathfrak{q}$; it produces an orbifold singularity on S of index $\mathfrak{q}$. Otherwise, γ_v is parabolic producing a cusp for S.

Vice versa, according to the *Poinaré Theorem*, any hyperboplic polygon with above properties generates a Fuchsian group.

Project 2.51. *Work out details of the above description.*

2.4.16. *Simple closed curves on Riemann surfaces.* Let us now consider a Fuchsian group Γ and the corresponding hyperbolic Riemann surface $S = \mathbb{D}/\Gamma$. Hyperbolic geodesics on S are (obviously) projections of the hyperbolic geodesics on $\mathbb{D}$; horocycles on S are (by definition) projections of the horocycles on $\mathbb{D}$. (A *simple* horocycle is a horocycle without self-intersections.)

Let γ be a non-trivial simple closed curve on S, and let $[\gamma]$ be the class of simple closed curves freely homotopic to γ. To this class corresponds a conjugacy class $A(\gamma)$ of deck transformations (see Corollary 1.60 and Exercise 1.61).

Exercise 2.52. *Show that elements* $\delta \in A(\gamma)$ *are primitive.*

Since deck transformations cannot be elliptic, the elements of $A(\gamma)$ are either all hyperbolic or all parabolic. Accordingly, we say that the class $[\gamma]$ itself is either hyperbolic or parabolic.

Proposition 2.53. (i) *If the class* $[\gamma]$ *is hyperbolic then it is represented by a unique closed hyperbolic geodesic* $\delta \in [\gamma]$. *This geodesic minimizes the hyperbolic length of the closed curves in* $[\gamma]$.

(ii) *If the class* $[\gamma]$ *is parabolic then* S *contains a neighborhood* U *isometric to a cusp, and* $[\gamma]$ *is represented by any horocycle in it. In this case, the class contains arbitrary short curves.*

Proof. Let us consider a lift $\tilde{\gamma}$ of γ, and let $G = <\phi^n>_{n\in\mathbb{Z}}$ be its stabilizer.

(i) If ϕ is hyperbolic then it has two fixed points, x_- and x_+, on the absolute, and then the closure of $\tilde{\gamma}$ in $\overline{\mathbb{D}}$ is a topological interval with endpoints x_1 and x_+. Let us consider the hyperbolic geodesic $\tilde{\delta}$ in $\mathbb{D}$ with endpoints $x_\pm$. It is invariant under the action of the cyclic group G. In fact, it is completely invariant. Indeed, if $\psi(\tilde{\delta}) \cap \tilde{\delta} \neq \emptyset$ for some $\psi \in \Gamma \smallsetminus G$, then $\psi(\tilde{\gamma}) \cap \tilde{\gamma} \neq \emptyset$ as well, which is impossible since γ does not have self-intersections. Hence the projection of $\tilde{\delta}$ to S is equal to $\tilde{\delta}/G$, which is the desired simple closed geodesic representing $[\gamma]$.

FIGURE 2.7. Tightening of a loop to a geodesic or a horocycle.

(ii) If ϕ is parabolic then it has a single fixed point x on the absolute, and the closure of $\tilde{\gamma}$ in $\overline{\mathbb{D}}$ is a topological circle touching $\mathbb{T}$ at x (a "topological horocycle centered at x").

Let $\tilde{U}$ be the corresponding topological horoball bounded by $\tilde{\gamma}$. Let us show that it is completely invariant under G. Indeed, for $\psi \in \Gamma \smallsetminus G$, $\psi(\tilde{U})$ is a topological horoball centered at $\beta(x) \neq x$. But since γ is a simple curve, $\psi(\tilde{\gamma}) \cap \tilde{\gamma} = \emptyset$ for any $\beta \in \Gamma \smallsetminus G$. Since two topological horoballs with disjoint boundaries are disjoint, $\psi(\tilde{U}) \cap \tilde{U} = \emptyset$.

It follows that $\tilde{U}/G$ is is isometrically embedded into $\mathbb{D}/\Gamma = S$. But $\tilde{U}/G$ is a conformal punctured disk containing some standard cusp $\mathbb{H}_h/\mathbb{Z}$. Thus, this cusp isometrically embeds into S as well, and its horocycles give us desired representatives of $[\gamma]$. $\square$

In case (i) of the above statement we say that the class $[\gamma]$ (or, the curve γ itself) is *represented by a geodesic*. In case (ii) it is *represented by a horocycle*.

A free homotopy of a simple loop to a representing geodesic or a horocycle will be called *tightening*. It can be done without increasing the length:

EXERCISE 2.54. *A simple closed ε-loop can be tightened to a simple closed geodesic or a simple closed horocycle through a family of ε-loops.* (*An ε-loop is a loop of length at most ε.*)

2.4.17. *Ideal boundary.* Let us now consider a tame end E of S, and let $F \subset S$ be a topological cylinder representing E. It is bounded by a peripheral curve γ.

If γ is trivial then S is simply connected, and hence $S = \mathbb{D}$. In this case, the ideal compactification of S is naturally defined as the closed disk, $\mathrm{cl}^I\, \mathbb{D} = \overline{\mathbb{D}}$, with the absolute being its ideal boundary, $\partial^I \mathbb{D} = \mathbb{T}$.

Otherwise, γ is either parabolic or hyperbolic, so it is represented either by a horocycle or by a geodesic (Proposition 2.53).

LEMMA 2.55. *If γ is parabolic, then two options can occur:*

(i) *The end E is a cusp; it can be completed by adding one point, ∞_E, and the completed surface has a natural conformal structure;*

(ii) *$S \approx \mathbb{D}^*$ and the end E corresponds to annuli $\mathbb{A}(r, 1) \subset \mathbb{D}^*$, $r \in (0, 1)$; it can be completed by adding $\mathbb{T}$ to $\mathbb{D}^*$.*

In the latter case, we let $\partial^I \mathbb{D}^* \equiv \partial^I E = \mathbb{T}$ and $\mathrm{cl}^I\, \mathbb{D}^* = \mathbb{D}^* \cup \partial^I \mathbb{D}^* = \overline{\mathbb{D}} \smallsetminus \{0\}$.

PROOF. By Proposition 2.53, γ is homotopic to a horocycle in a cusp $U \subset S$, which is conformally isomorphic to $\mathbb{D}^*$. The isomorphism $\phi : U \to \mathbb{D}^*$ provides us with an embedding of U into $\mathbb{D}$. By glueing S and $\mathbb{D}$ by means of ϕ, we obtain a desired Riemann surface

$$\mathrm{cl}^I_E\, S := S \cup_\phi \mathbb{D}.$$

completing S at infinity by adding one ideal point.

If the cusp U represents the same end as F, we are in case (i). If U and F represent different ends of S, then γ partitions S into two topological cylinders, so S is a topological cylinder as well, with the fundamental group generated by $[\gamma]$. Then $S \approx \mathbb{H}/ < A >= \mathbb{D}^*$, where $A : z \mapsto z+1$ is the parabolic deck transformation (appropriately normalized) corresponding to $[\gamma]$. $\square$

Let us now assume that a peripheral curve γ on S is hyperbolic, so it is represented by a geodesic δ.

LEMMA 2.56. *The cylinder F lifts to a topological bigon $\tilde{F}$ on the universal covering $\mathbb{D}$ bounded by a lift $\tilde{\gamma}$ of γ and an interval $I \subset \mathbb{T}$. The stabilizer of F and of I in the covering Fuchsian group Γ is the cyclic group $G =< A >$ generated by a hyperbolic transformation A corresponding to γ. The quotient I/G is a circle completing the cylinder F at infinity. The geodesic δ lifts to a geodesic $\tilde{\delta}$ sharing endpoints with $\tilde{\gamma}$. The bigon $\Delta = \Delta_I$ bounded by $\tilde{\delta}$ and $\bar{I}$ covers a cylinder Δ/G in S representing the same end E.*

The circle I/G is called the *ideal boundary* of the end E. We denote it $\partial^I E \equiv \partial^I F$, and we let $\mathrm{cl}^I_E\, S := S \cup \partial^I E$.

PROOF. Since $\pi_1(F)$ is the cyclic group generated by $[\gamma]$, the cylinder F lifts to a domain $\tilde{F} \subset \mathbb{D}$ with a cyclic stabilizer $G =< A >$ generated by the hyperbolic deck transformation A. Moreover, F is completely invariant under G. Since $\partial F = \gamma$, the boundary of F in $\mathbb{D}$ is equal to $\tilde{\gamma}$, so $\tilde{F}$ is one of the two components of $\mathbb{D} \smallsetminus \tilde{\gamma}$. In $\overline{\mathbb{D}}$, this component is bounded by $\tilde{\gamma} \cup \bar{I}$, where I is an open arc of $\mathbb{T}$ sharing the endpoints with $\tilde{\gamma}$. It follows that I is also completely invariant under the cyclic group G. Since G acts on I totally discontinuously, we have $I \subset \mathbb{T} \smallsetminus \Lambda$. As $\partial I = \partial \tilde{\gamma} \subset \Lambda$, we conclude that I is a gap in Λ.

The quotient $(\tilde{F} \cup I)/G$ is a semi-closed cylinder with the circle I/G attached to the open cylinder F. So, it provides a completion of F at infinity.

Let us check that the bigon Δ is completely invariant under G. Indeed, $A(\Delta_I) = \Delta_{A(I)} = \Delta_I$ for any $A \in G$. On the other hand, if $A \in \Gamma \smallsetminus G$ then $A(I) \cap I = \emptyset$. Hence

$$A(\Delta_I) \cap \Delta_I = \Delta_{A(I)} \cap \Delta_I = \emptyset.$$

We conclude that $F' := \Delta/G$ is a cylinder in S bounded by δ. Moreover, $F' \cap F$ contains a cylinder $\tilde{U}/G$, where $\tilde{U} \subset \mathbb{D}$ is a small G-invariant neighborhood of I. It follows that F' and F represent the same end of S. $\square$

For surfaces of finite topological type, the above assertion can be reversed: any gap in Λ corresponds to some ideal boundary circle of S:

PROPOSITION 2.57. *Assume that Γ is finitely generated, but not a parabolic cyclic group. Let us consider a gap I in Λ. Then the stabilizer of I in Γ is a cyclic group G generated by a hyperbolic transformation A. The quotient I/G is an ideal boundary circle of S completing some non-cuspidal end.*

PROOF. Let us consider a curve $\tilde{\sigma} : [0, 1) \to \mathbb{D}$ landing at some point $b \in I$, i.e., $\tilde{\sigma}(t) \to b$ as $t \to 1$. It projects to a curve σ on S converging to some end E. Since all the ends of S are tame, σ is eventually trapped in a cylinder F representing E. Let $\tilde{F}$ be its lift to $\mathbb{D}$ that contains a tail of $\tilde{\sigma}$.

If E were cusp then the closure of $\tilde{F}$ would touch $\mathbb{T}$ at a single parabolic point, so b is this point. But this is impossible since parabolic points belong to Λ.

Thus, E is a non-cuspidal end. If γ were parabolic then by Lemma 2.55 (ii), $S \approx \mathbb{D}^*$, and Γ would be a cyclic parabolic group, which is ruled out by the assumption.

Hence γ is hyperbolic, and as such, is represented by a geodesic. Then the boundary of $\tilde{F}$ on the absolute is a gap $\tilde{I} \subset \mathbb{T} \setminus \Lambda$ corresponding to the end E (see Lemma 2.56). Since the gaps I and $\tilde{I}$ overlap, they coincide. So, I is stabilized by the cyclic group G generated by A. We are done. $\square$

Thus, the simultaneous ideal completion of all non-cuspidal ends of S can be obtained by taking the quotient of $\overline{\mathbb{D}} \setminus \Lambda$ by the action of the Fuchsian group Γ. We call it $\mathrm{cl}^I S$, while the full ideal boundary will be called $\partial^I S$.

Remark 2.58. Note that the above result shows that *a finitely generated Fuchsian group* Γ *does not have wandering intervals* , i.e., there are no gaps I in Λ such that $A(I) \cap I = \emptyset$ for all $A \in \Gamma \setminus \{\mathrm{id}\}$.

Let us summarize our discussion:

Theorem 2.59. *Let* Γ *be a finitely generated Fuchsian group, and let* S *be the corresponding hyperbolic Riemann surface* $\mathbb{H}/\Gamma$. *Any cuspidal end* E *of* S *can be completed by an ideal puncture* ∞_E*, with complex structure extended through* ∞_E. *Any non-cuspidal end* E *can be completed by an ideal circle* $\partial^I E$ *at infinity.*

This completion produces a compact Riemann surface $\mathbf{S}$, the full *ideal compactification of* S . This compactification is *conformally natural*:

Proposition 2.60. *Any conformal isomorphism* $\phi : S \to S'$ *between hyperbolic Riemann surfaces extends to a conformal isomorphism* $\mathbf{\Phi} : \mathbf{S} \to \mathbf{S}'$ *between their ideal compactifications.*

Proof. The isomorphism ϕ lifts to an equivariant isomorphism $\hat{\phi} : \mathbb{D} \to \mathbb{D}$ between the universal coverings. Being Möbius, it extends to $\overline{\mathbb{D}}$. Being equivariant, ϕ maps the limit set Λ for S to the limit set Λ' for S'. On the complement, we obtain an equivariant isomorphism $\overline{\mathbb{D}} \setminus \Lambda \to \overline{\mathbb{D}} \setminus \Lambda$, which descends to an isomorphism $\Phi : \partial^I S \to \partial^I S'$. By the Removability of isolated singularities, Φ extends through the ideal punctures to a conformal isomorphism $\mathbf{\Phi} : \mathbf{S} \to \mathbf{S}'$. $\square$

More generally, we have:

Exercise 2.61. *Any holomorphic covering* $f : S \to S'$ *of finite degree between hyperbolic Riemann surfaces extends continuously to a covering* $\mathbf{f} : \mathbf{S} \to \mathbf{S}'$ *between their ideal compactifications (holomorphic on* $\mathrm{int}\,\mathbf{S}$*).*

2.4.18. *Convex core.* Let $\Lambda = \Lambda(\Gamma)$ be the limit set of a Fuchsian group Γ of second kind, and let $\pi : \mathbb{D} \to S$ be the projection onto the quotient Riemann surface. Since Λ is invariant under Γ, the convex hull $\hat{\Lambda}$ is Γ-invariant as well. Hence it covers a Riemann surface $C = C_S$ with boundary called the *convex core* of S.

Proposition 2.62. *The natural embedding* $C \to S$ *is a homotopy equivalence.*

Proposition 2.63. *The group* Γ *is convex co-compact if and only if the convex core* C *is compact.*

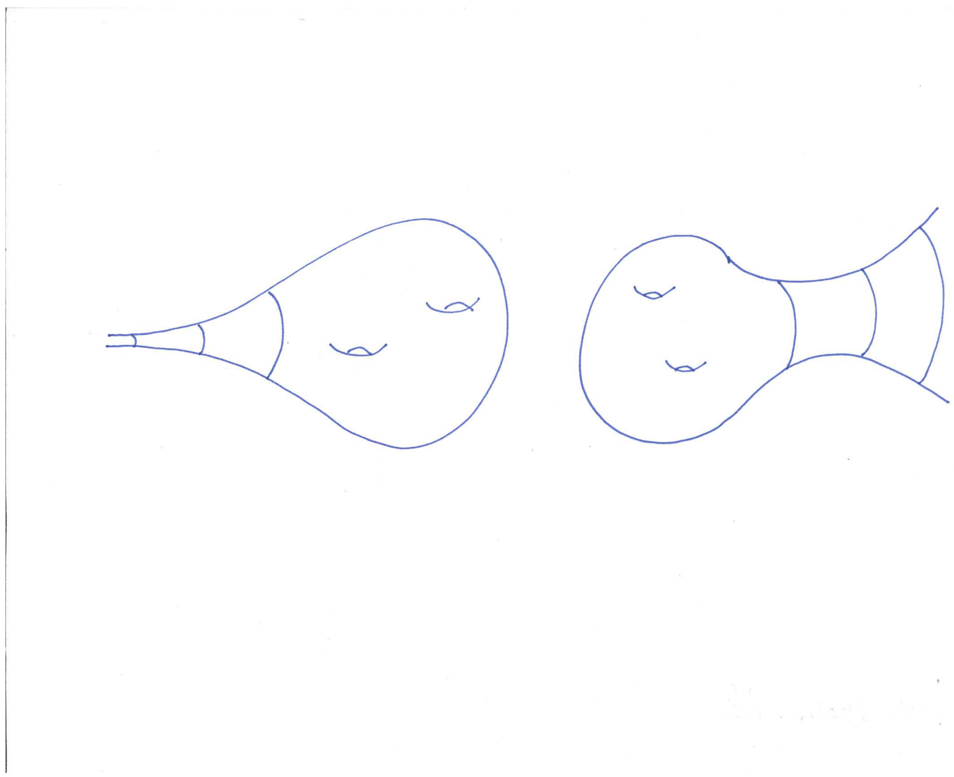

FIGURE 2.8. This picture illustrates a geometric difference between two types of tame hyperbolic ends: a cusp associated with a horocycle and and an "open end" associated with a peripheral geodesic.

2.4.19. *Linking.* Let us consider a configuration of two pairs of points on a topological circle S^1, $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$, where all four points are assumed to be distinct. There are two possible relative positions of these pairs: they can be *linked* or *unlinked.* Linking means that the points alternate when one goes around the circle, i.e., both intervals with endpoints x_1, x_2 contain a y-point. Otherwise, Y is contained in one of these intervals. These properties are intrinsic for S^1, but in case when S^1 is the boundary of a 2-disk, they can be nicely recognized from the inside:

EXERCISE 2.64. *For two pairs of points X and Y on the unit circle $\mathbb{T} = \partial\mathbb{D}$, the following properties are equivalent:*

(i) *X and Y are linked;*

(ii) *The hyperbolic geodesics* $[x_1, x_2]$ *and* $[y_1, y_2]$ *cross in* $\mathbb{D}$*;*

(iii) *Any continua* X' *and* Y' *in* $\overline{\mathbb{D}}$ *such that* $X \subset X' \subset \overline{\mathbb{D}} \smallsetminus Y$ *and* $Y \subset Y' \subset \overline{\mathbb{D}} \smallsetminus X$ *intersect.*

More generally, let X and Y be two disjoint closed non-singleton subsets of a topological circle S^1. We say that X and Y are *unlinked* if any two pairs of distinct points, $\{x_1, x_2\} \subset X$ and $\{y_1, y_2\} \subset Y$, are unlinked. The previous Exercise yields:

Exercise 2.65. *For sets* X *and* Y *as above the following properties are equivalent:*

(i) X *and* Y *are unlinked;*

(ii) X *is contained in a single gap of* Y *(and the other way around);*

(iii *The hyperbolic convex hulls* $\hat{X}$ *and* $\hat{Y}$ *are disjoint;*

Two closed curves γ_1 and γ_2 on a surface S intersect (or "cross") *essentially* if any two curves γ_1' and γ_2', respectively homotopic to γ_1 and γ_2, intersect (so, the intersection cannot be removed by deforming the curves).

Exercise 2.66. *Let* S *be a hyperbolic Riemann surface and let* $\pi : \mathbb{D} \to S$ *be its universal covering. Two closed curves,* γ_1 *and* γ_2*, on* S *intersect essentially iff they admit lifts,* $\tilde{\gamma}_1$ *and* $\tilde{\gamma}_2$*, to* $\mathbb{D}$ *that converge to linked pairs of points,* $\{a_1, b_1\}$ *and* $\{a_2, b_2\}$*, in* $\mathbb{T}$*.*

2.5. Geodesic laminations.

2.5.1. *Glossary.* A *geodesic lamination* $\mathcal{L}$ in $\mathbb{D}$ is a closed subset of $\mathbb{D}$ (called $\operatorname{supp}\mathcal{L}$) partitioned into complete hyperbolic geodesics (*leaves* of $\mathcal{L}$). In other words, there is a *complete* geodesic $\gamma_z \subset \operatorname{supp}\mathcal{L}$ passing through any point $z \in \operatorname{supp}\mathcal{L}$, and these geodesics are either equal or disjoint. The leaves of the geodesic lamination vary uniformly continuously, together with derivatives, in the following sense:

Exercise 2.67. *Let* (γ_n, z_n) *be a sequence of disjoint pointed geodesics in* $\mathbb{D}$ *such that* $z_n \to z \in \mathbb{D}$*. Then the naturally parametrized* $\bar{\gamma}_n$ *converge to a geodesic* $\bar{\gamma}$ *through* z *in the* C^1*-topology on the space of paths* $[-\infty, +\infty] \to \overline{\mathbb{D}}$*.*

A *gap* Q in the geodesic lamination is a component of $\mathbb{D} \smallsetminus \operatorname{supp}\mathcal{L}$.

Exercise 2.68. *Show that any gap* Q *in* $\mathcal{L}$ *is hyperbolically convex. Moreover, the closure* $\overline{Q}$ *is the convex hull of its ideal boundary* $\partial^I Q \subset \mathbb{T}$*.*

In particular, if $\partial^I Q$ is finite then Q is a hyperbolic polygon. More generally, we say that a gap Q of a geodesic lamination is of *countable type* if $\partial^I Q$ is countable.

We say that a geodesic lamination is *clean* if no two gaps of countable type share an edge. Any geodesic lamination can be *cleaned* by removing common edges of gaps of countable type.

Given a clean lamination, let us blacken all gaps of countable type and possibly some other gaps observing the condition that *no two black gaps are adjacent* (i.e., they do not share a leaf of the lamination). We call such a lamination *colored*.

Three disjoint geodesics in $\mathbb{D}$ can be in two combinatorial positions: one of them can separate the other two, or not. In the non-separating case, the geodesics "bound" an m-gon in $\overline{\mathbb{D}}$ with $3 \le m \le 6$ (with $m-3$ ideal sides and $6-m$ cuspidal vertices). More generally, if we have $n \ge 3$ disjoint non-separating geodesics (i.e., none of them separates any other two) then they "bound" an m-gon in $\overline{\mathbb{D}}$ with $n \le m \le 2n$.

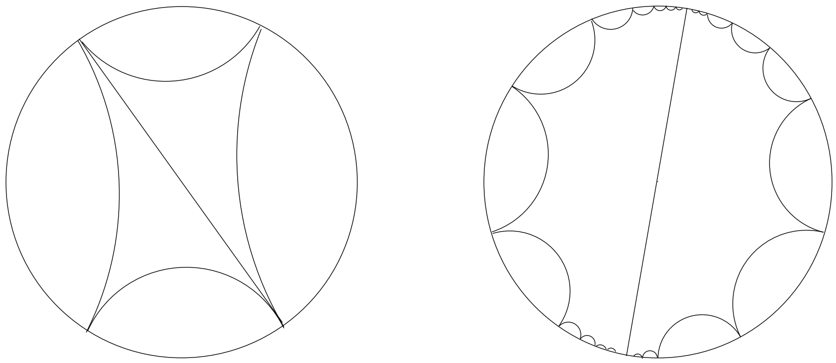

FIGURE 2.9. Examples of unclean laminations: a rectangle and an ∞-gon with a diagonal.

EXERCISE 2.69. *Let $\mathcal{L}$ be a geodesic lamination, and let $\{\gamma_i\}_{i=1}^n$ be a non-separating family of $n \geq 3$ leaves of $\mathcal{L}$. Then the polygon bounded by the γ_i contains a gap of $\mathcal{L}$.*

A simple example of a geodesic lamination is the one with geodesics

$$(e(\theta, e(-\theta))_\#, \quad \theta \in (0, 1/2),$$

connecting $\mathbb{R}$-symmetric points of $\mathbb{T}$.[10] Its support is the whole disk $\mathbb{D}$ (so it is actually a foliation). We will refer to it as the *Chebyshev lamination* (or *foliation)* $\mathcal{L}_{\text{Ч}}$ (for the reason that will become clear later, §32.5.1).

EXERCISE 2.70. (i) *Assume that* $\operatorname{supp} \mathcal{L}$ *contains a domain* Π. *Then there exist two disjoint* (*open*) *ideal intervals $I, J \subset \mathbb{T}$ and an orientation reversing homeomorphism $h : I \to J$ such that*

$$\Pi = \bigcup_{x \in I} (x, h(x))_\#.$$

(ii) *If* $\operatorname{supp} \mathcal{L} = \mathbb{D}$ *then the lamination is topologically equivalent to the Chebyshev foliation.*

In case (i), we refer to $\mathcal{L}$ as a *vertical geodesic foliation* in the rectangle Π. A particular case of this situation is when the intervals I and J share an endpoint a. Then the rectangle Π degenerates to a topological *sector* based on the ideal interval $I \cup \{a\} \cup J$. Moreover, the foliation $\mathcal{L}$ is not topologically transverse to $\mathbb{T}$ at a, so we call the latter an *ideal singular point* for $\mathcal{L}$. For instance, the Chebyshev foliation has two ideal singular points, ± 1.

2.5.2. *Geodesic laminations, equivalence relations, and lc hulls.* We say that a geodesic lamination $\mathcal{L}$ is *polygonal* if all the gaps in $\mathcal{L}$ are polygons. Notice that if a polygonal lamination is clean then it is *maximal* among clean laminations.

Geodesic laminations provide us with a visualization of equivalence relations $\sim$ of $\mathbb{T}$. Such an equivalence relation is called *unlinked* if all the equivalence classes are pairwise unlinked.

---

[10] Recall from §50.2 that $(a, b)_\#$ stands for the interval with endpoints a and b disregarding their order.

To any closed unlinked equivalence relation $\sim$ on $\mathbb{T}$ we can associate a colored geodesic lamination $\mathcal{L}(\sim)$ in $\mathbb{D}$ as follows:

1) Take the hyperbolic convex hull $H(X)$ of each equivalence class X;

2) Consider the boundary components (in $\mathbb{D}$) of these convex hulls as the leaves of $\mathcal{L}(\sim)$.

3) Blacken all the gaps of type $H(X)$.

Exercise 2.71. *Check that $\mathcal{L}(\sim)$ is indeed a colored (and in particularly, clean) geodesic lamination.*

Let us define an equivalence relation $\approx$ on $\mathbb{C}$ by declaring that its equivalence classes are the above convex hulls $H(X)$ or singletons.

Proposition 2.72. *For any colored lamination $\mathcal{L}(\sim)$, we have:* (i) *The quotients $K := \overline{\mathbb{D}}/\approx$ is a lc compact space;*

(ii) *The quotient $\mathbb{C}/\approx$ is a topological plane $\mathbb{R}^2$;*

(iii) *K is a hull in $\mathbb{R}^2$.*

Proof. (ii) Though we could use Moore's Theorem, let us sketch a direct argument.

Let $\pi : \overline{\mathbb{D}} \to K$ be the natural projection. The lamination $\mathcal{L}$ has two types of gaps: a gap of first kind represents a single equivalence class (which collapses to a point under π), while a gap of second kind is partitioned into singletons (so, π is injective on such a gap G). In the latter case, we say that $\pi(G)$ is a "component of interior" of K (which is just a term at the moment).

Let us start with a couple of special cases:

a) In case of a finite lamination, the quotient $(\mathbb{C}, \mathbb{D})/\approx$ is homeomorphic to $(\mathbb{R}^2, K)$, where K is a tree T of bubbles $\overline{D}_i$ (i.e. piecewise smooth closed Jordan disks) attached one to another. Let us realize T as an actual tree embedded into K (e.g., by marking a "center" a_i in each bubble D_i and connecting them through K by piecewise smooth arcs).

b) If $\mathcal{L}$ consists of boundary leaves L_i of some gap G of first kind, then K is a point $x = \pi(\overline{G})$ with countably many bubbles $\overline{D}_i$ attached to it (corresponding to the components of $\mathbb{D} \smallsetminus L_i$ disjoint from G).

c) If $\mathcal{L}$ consists of boundary leaves L_i of some gap G of second type, then K is a closed Jordan disk $\overline{D} = \pi(\overline{G})$ (see Exercise 1.8) with countably many bubbles $\overline{D}_i$ attached to it at points $\pi(L_i)$.

In general, let us consider a sequence of finite laminations $\mathcal{L}^n$ converging to $\mathcal{L}$ such that $\mathcal{L}^{n+1}$ is obtained from $\mathcal{L}^n$ by adding either a single leave or all the peripheral leaves bounding one gap G. Moreover, assume any peripheral leaf belongs to some $\mathcal{L}^n$. Let us construct inductively the corresponding sequence of quotient bubble trees $(\mathbb{R}^2, K^n = \bigcup \overline{D}_i^n, T^n)$ so that $K^{n+1} \subset K^n$, $\partial K^n \cap \partial K^{n+1}$ is a finite set of points of T^n, and $T^{n+1} \supset T^n$. Let $X = \bigcap K^n$.

For a bubble $\overline{D}_i^n$ centered at a_i^n, the set $T_i^n := T^n \cap \overline{D}_i^n$ is a star rooted at a_i^n. For the sake of this discussion, let us call its edges the *internal radii* of D_i^n.

If $\mathcal{L}^{n+1}$ is obtained from $\mathcal{L}^n$ by adding a single leaf L, then K^{n+1} is obtained from K^n by a simple pinching of some bubble $\overline{D}_i^n$ producing two new bubbles $\overline{D}_j^{n+1}$ and $\overline{D}_{j+1}^{n+1}$ touching at the point $x = \pi(L)$. This procedure can be realized in $\mathbb{R}^2$

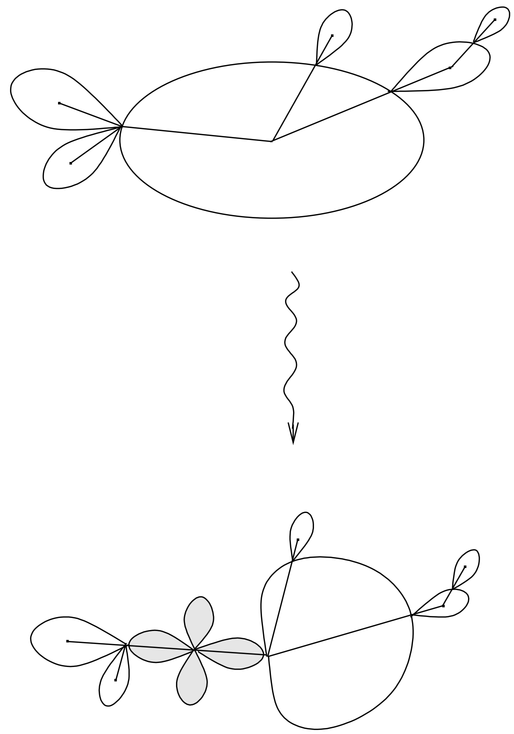

FIGURE 2.10. Adding a rectangular gap to a lamination amounts to tuning one of the bubbles by a flower with four petals (shadowed).

by putting x at the middle (with respect to the arc length) of the appropriate internal radius and putting the bubbles $\overline{D}_j^{n+1}$, $\overline{D}_{j+1}^{n+1}$ into narrow neighborhoods of the corresponding half-radii.

If $\mathcal{L}^{n+1}$ is obtained from $\mathcal{L}^n$ by adding boundary leaves of a gap G of first kind, then the construction is similar. Again, we can put the point $x = \pi(G)$ at the middle of the appropriate internal radius, add to T_i^n several new short edges sticking out of x, and attach to x a bouquet of bubbles contained in small neighborhoods of the corresponding half-radii and new edges. In the course of this procedure, the bubble D_i^n is replaced with a smaller bubble, while all other bubbles D_k^n remain untouched, see Figure 2.10.

Finally, if $\mathcal{L}^{n+1}$ is obtained from $\mathcal{L}^n$ by adding boundary leaves of a gap G of second kind, then we realize $\pi(\overline{G})$ as a small Jordan disk centered in the middle of the appropriate inner radius of D_i^n, with attached bubbles which are contained in small neighborhoods of the corresponding half-radii or new short edges.

In this way we ensure that the diameters of the bubbles and interior components (that appear in the above construction) go to 0 as $n \to \infty$, which implies that X is a hull. Moreover, there is a natural homeomorphism $(\mathbb{C}, K) \to (\mathbb{R}^2, X)$.

EXERCISE 2.73. *Supply missing pieces and details of the proof.*

$\square$

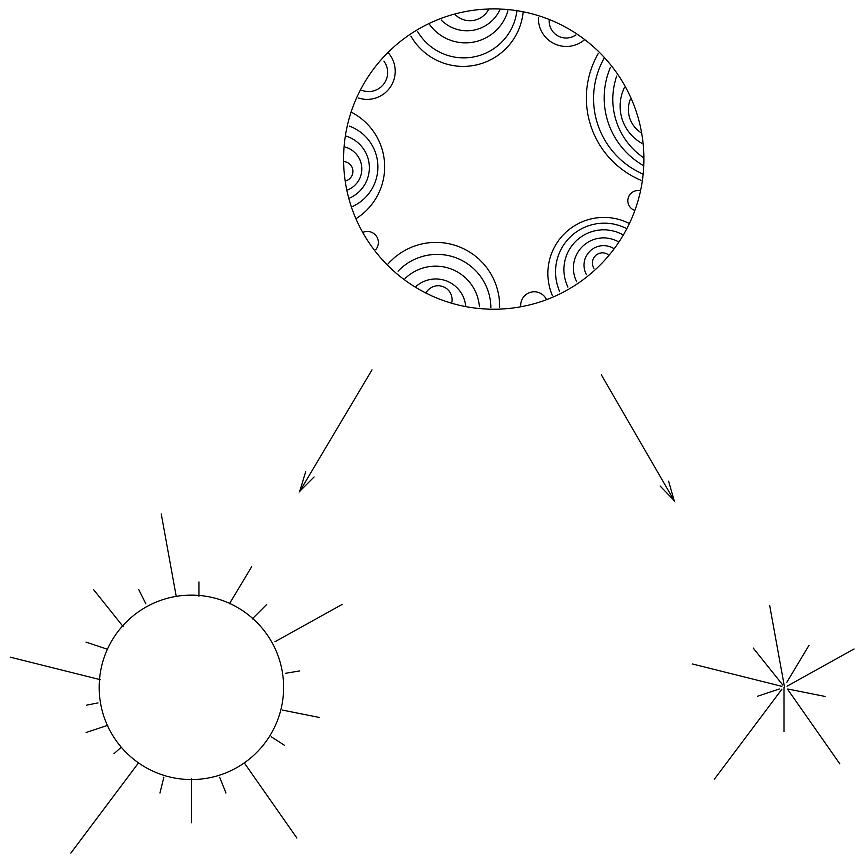

FIGURE 2.11. Fat and thin hedgehogs.

Somewhat informally, we will also refer to K as the *quotient of* $\overline{\mathbb{D}}$ *mod the colored lamination* $\mathcal{L}$, and will use notation $K_{\max} = \overline{\mathbb{D}}/\mathcal{L}$ for it. Similarly, we let $\mathbb{C}/\mathcal{L}$ be the corresponding quotient of the whole plane. Such a representation of a hull $K \subset \mathbb{C}$ is called *the pinched disk model* for K.

The minimal way of coloring a clean lamination $\mathcal{L}$ is to blacken only gaps of countable type. In this way we obtain the *maximal quotient* $K_{\max}$ associated with this lamination (as any other quotient K is the quotient of $K_{\max}$).

EXAMPLE 2.74 (Fat and thin hedgehogs). Consider a Cantor set $X \subset \mathbb{T}$ and it convex hull Q. Filling in every component of $\mathbb{D} \smallsetminus Q$ with a geodesic foliation, we obtain a geodesic lamination $\mathcal{L}$ on $\mathbb{D}$. The hull $K_{\max} = \overline{\mathbb{D}}/\mathcal{L}$ is a *hedgehog* (a closed Jordan disk $\overline{D}$ with needles attached densely to its boundary). In this case, there is also the minimal hull $K_{\min}$ obtained from $K_{\max}$ by collapsing $\overline{D}$ to a point It is another version of a *hedgehog* whose needles are attached to a single point. We will distinguish these hedgehogs as *fat* and *thin* respectively. See Figure 2.11.

Sometimes we will need to consider non-closed subsets $X \subset \mathbb{D}$ partitioned into complete geodesics. We refer to these objects as *geodesic pre-laminations* supported on X. There is a natural partial order on the space of geodesic pre-laminations: $\mathcal{L}' \succ \mathcal{L}$ if $\mathcal{L}'$ is an extension of $\mathcal{L}$ to a bigger support.

We say that two geodesic pre-laminations $\mathcal{L}$ and $\mathcal{L}'$ are *unlinked*, $\mathcal{L} \parallel \mathcal{L}'$, if any two leaves $\gamma \in \mathcal{L}$ and $\gamma' \in \mathcal{L}'$ either coincide or disjoint.

EXERCISE 2.75. (i) *Any geodesic pre-lamination supported on X extends to a geodesic lamination supported on $\overline{X}$ (we will refer to this extension as the* closure *of $\mathcal{L}$).*
(ii) *If two geodesic pre-laminations are unlinked then so are their closures.*

2.6. Cylinders, rectangles, tori, and four-times-punctured-spheres.

2.6.1. *Flat cylinder.* Let us consider a flat cylinder $\mathrm{Cyl} \equiv \mathrm{Cyl}_l^h = \mathbb{S}_h/(l \cdot \mathbb{Z})$ (obtained by taking the quotient of the strip of height h by the cylic translation group with generator $z \mapsto z + l$). Endow it with the *flat metric* $ds = |dz|$ induced from the strip. The exponential map

$$\mathbb{S}_h \to \mathbb{A} \equiv \mathbb{A}(1, \exp(2\pi h/l)), \quad z \mapsto e(-z/l)$$

induces a conformal isomorphism between Cyl and the annulus $\mathbb{A}$, so Cyl is a conformal annulus with modulus

$$\operatorname{mod} \mathrm{Cyl} = \operatorname{mod} \mathbb{A} = h/l. \tag{2.12}$$

By (2.4.2) the hyperbolic metric on Cyl_h^l is equal to

$$ds_{\rho_{\mathbb{A}}} = \frac{\pi}{h} \frac{|dz|}{\sin(\pi y/h)}. \tag{2.13}$$

In this model, the equator $\gamma \equiv \gamma_{\mathrm{Cyl}}$ becomes the round circle on the middle height, $h/2$. Its hyperbolic length is equal to

$$l_{\mathrm{hyp}}(\gamma) = \frac{\pi l}{h} \tag{2.14}$$

(which matches with (2.11)).

Let finally note that the punctured disk $\mathbb{D}^*$ is conformally equivalent, via the exponential map $e : \mathbb{H} \to \mathbb{D}^*$, to the *half-infinite (flat) cylinder* $\mathbb{T} \times \mathbb{R}_+ \approx \mathbb{H}/\mathbb{Z}$. (Compare with the *bi-infinite* cylinder model for $\mathbb{C}^*$, §2.2.)

2.6.2. *Modulus of a rectangle.* A *flat rectangle* Π is a standard Euclidean rectangle in $\mathbb{R}^2$ (the one we learn at the elementary school), for instance, a *standard rectangle* $\Pi_l^h = [0, l] \times [0, h]$. A *marking* of a flat rectangle is a choice of two opposite sides declared to be *horizontal* (while the other two are declared to be *vertical*); of course, a standard rectangle Π_l^h is naturally marked. The *modulus* $\operatorname{mod} \Pi$ of a marked rectangle is the ratio of the lengths of its vertical and horizontal sides, so $\operatorname{mod} \Pi_l^h = h/l$. Change of marking replaces the modulus to the inverse one.

Two marked rectangles Π and Π' are called *affinely/conformally/etc equivalent* if there is an affine/conformal/etc isomorphism $\Pi \to \Pi'$ that maps the horizontal sides of Π to the horizontal sides of Π'. Obviously, marked Π and Π' are affinely equivalent if and only if $\operatorname{mod} \Pi = \operatorname{mod} \Pi'$. It is still true, albeit less obvious, for conformal equivalence:

EXERCISE 2.76. *Two marked flat rectangles Π and Π' are conformally equivalent if and only if $\operatorname{mod} \Pi = \operatorname{mod} \Pi'$.*

For three real points, $a < b < c$, let $\mathbb{H}(a,b,c) \equiv \mathbb{H}(a,b,c,\infty)$ be the upper half plane with these points marked on its ideal boundary (and ∞ marked by default). By an affine automorphism of $\mathbb{H}$, two of these points can be normalized, e.g. the triple can be brought to the form $(a,0,1)$ with some $a<0$.

EXERCISE 2.77. *The elliptic integral*

$$E(z) = \int_1^z \frac{dz}{\sqrt{z(z-1)(z-a)}}$$

induces a conformal isomorphism between $\mathbb{H}$ *and some standard rectangle* $\Pi(a) = \Pi(l, h(a))$. *Moreover,* $\operatorname{mod}\Pi(a)$ *depends on a continuously and monotonically, and*

$$\operatorname{mod}\Pi(a) \to 0 \text{ as } a \to -\infty, \ \operatorname{mod}\Pi(a) \to \infty \text{ as } a \to 0.$$

Thus, $\operatorname{mod}\Pi(a)$ *assumes all possible values.*

2.6.3. *Modulus of the torus.* Let us take a closer look at the actions of the group $\Gamma \approx \mathbb{Z}^2$ on the (oriented) affine plane $P \approx \mathbb{C}$ by translations (see §2.2). We would like to classify these actions up to affine conjugacy, i.e., two actions T and S are considered to be equivalent if there is an (orientation preserving) affine automorphism $A : P \to P$ and an algebraic automorphism $i : \Gamma \to \Gamma$ such that for any $\gamma \in \Gamma$ the following diagram is commutative:

$$\begin{array}{ccc} P & \underset{T^\gamma}{\longrightarrow} & P \\ A\downarrow & & \downarrow A \\ P & \underset{S^{i(\gamma)}}{\longrightarrow} & P \end{array} \tag{2.15}$$

This is equivalent to classifying the quotient tori P/T^Γ up to conformal equivalence (since a conformal isomorphism between the quotient tori lifts to an affine isomorphism between the universal covering spaces conjugating the actions of the covering groups).

The conjugacy A in the above definition will also be called *equivariant* with respect to the corresponding group actions.

The problem becomes easier if to require first that $i = \mathrm{id}$ in (2.15). Fix an *uncolored pair* of generators α and β of Γ. Since T acts by translations and since P is affine, the ratio

$$\tau = \tau(T) = \frac{T^\beta(z) - z}{T^\alpha(z) - z}$$

makes sense and is independent of $z \in P$. Moreover, by switching the generators α and β we replace τ with $1/\tau$. Thus, we can color the generators in such a way that $\operatorname{Im}\tau > 0$. (With this choice, the basis of P corresponding to the generators $\{\alpha, \beta\}$ is positively oriented.)

EXERCISE 2.78. *Show that two actions* T *and* S *of* $\Gamma =< \alpha, \beta >$ *are affinely equivalent with* $i = \mathrm{id}$ *if and only if* $\tau(T) = \tau(\tilde{T})$.

According to the discussion in §1.7.15, the choice of generators of Γ means *(uncolored) marking* of the corresponding torus. Thus, the marked tori are classified by a single complex modulus $\tau \in \mathbb{H}$.

Forgetting the marking amounts to replacement one basis $\{\alpha, \beta\}$ in Γ by another, $\{\tilde\alpha, \tilde\beta\}$. If both bases are positively oriented then there exists a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$$

such that $\tilde\alpha = a\,\alpha + b\,\beta$, $\tilde\beta = c\,\alpha + d\,\beta$. Hence

$$\tilde\tau = \frac{a\tau + b}{c\tau + d}.$$

Thus, the unmarked tori are parametrized by a point $\tau \in \mathbb{H}$ modulo the action of the modular group $\mathrm{PSL}(2, \mathbb{Z})$ on $\mathbb{H}$ (see §2.4.12).

Remark. Passing from $\mathrm{SL}(2, \mathbb{Z})$ to $\mathrm{PSL}(2, \mathbb{Z})$ has an underlying geometric reason. All tori $\mathbb{C}/\Gamma$ have a conformal symmetry $z \mapsto -z$. It change marking $\{\alpha, \beta\}$ by $-I\{\alpha, \beta\}$. Thus, remarking by $-I$ acts trivially on the space of marked tori.

We also know from Exercise 2.47 that the modular surface $\mathfrak{M} = \mathbb{H}/\mathrm{PSL}(2, \mathbb{Z})$ is an orbifold supported by $\mathbb{C}$ with two cone points, of order two and three. Thus, the unmarked tori are parametrized by a single *modulus* $\mu \in \mathfrak{M}(\mathbb{T}^2) \equiv \mathfrak{M} \approx \mathbb{C}$.

Exercise 2.79. *What is the special property of the tori corresponding to the cone points?*

In the dynamical context we will encounter tori $\mathbb{T}^2_\rho$ obtained by taking the quotients of $\mathbb{C}^*$ by cyclic groups generated by complex scalings $L_\rho : z \mapsto \rho z$, where $|\rho| \neq 1$. Assume for definiteness that $|\rho| < 1$.

Note that $\mathbb{T}^2_\rho$ can be explicitly "cooked" by taking a *fundamental annulus* $A_\rho := \{|\rho| \le |\zeta| \le 1]$ and gluing its boundary components by the scaling relation (identifying a point $\zeta \in \mathbb{T}$ with $\rho\,\zeta \in \mathbb{T}_{|\rho|}$).

This torus has a *marked generator* $\alpha \in \Gamma := \pi_1(\mathbb{T}^2) \approx \mathbb{Z}^2$ represented by the equator in A_ρ (which is the generator of the subgroup of Γ associated with the covering $\mathbb{C}^* \to \mathbb{T}^2_\rho$).

The second generator β of Γ is represented by a proper arc in the fundamental annulus A_ρ connecting two related boundary points, e.g., connecting 1 to ρ by an arc of a logarithmic spiral given in the polar coordinates (r, θ) as

$$r = |\rho|^t, \ \theta = t \arg \rho, \quad 0 \le t \le 1.$$

The $(2\pi i)$-ambiguity in the choice of $\arg \rho$ corresponds to the "twist" ambiguity in the choice of β, i.e., replacement of β with $\beta + n\alpha$.

Let us consider the universal covering of $\mathbb{C}^*$ explicitly given by $\exp : \mathbb{C} \to \mathbb{C}^*$. Then the complementary generator β lifts to the deck translation $z \mapsto z + \log \rho$, with the branch of $\log \rho$ corresponding to the above choice of $\arg \rho$. For a given choice of $\log \rho$, the torus becomes fully marked, with the modulus

$$\tau = \frac{\log \rho}{2\pi i} \in \mathbb{H}.$$

Making different choices of $\log \rho$ amounts to taking the quotient of $\mathbb{H}$ by the cyclic group $\tau \mapsto \tau + n$, $n \in \mathbb{Z}$. In this way, the moduli space of partially marked tori $\mathbb{T}^2_\rho$, $\rho \in \mathbb{D}^*$, gets naturally identified with $\mathbb{H}/\mathbb{Z}$, where $\mathbb{H}$ represents the moduli space of the marked tori.

2.6.4. *Cross-ratios and four-times-punctured spheres.* Let us mark a set of four distinct points, $\{a, b, c, d\}$, on the Riemann sphere (or, equivalently, puncture these points out of $\hat{\mathbb{C}}$). Then we can form six cross-ratios, obtained from the basic one

$$\lambda = [a, b, c, d] := \frac{(c-a)(d-b)}{(b-a)(d-c)}$$

by permuting the variables. The basic property of cross-ratios is that they are preserved under the Möbious transformations,

$$[a, b, c, d] = [M(a), M(b), M(c), M(d)] \quad \text{for any} \quad M \in \text{Möb}(\hat{\mathbb{C}}),$$

and vice versa, if $[a, b, c, d] = [\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}]$ then there exist a Möbius transformation M such that $\tilde{a} = M(a)$ etc.

EXERCISE 2.80. (i) *Various cross ratios are obtained one from another by the action of the cross-ratio group* Γ_{cr} *comprising six Möbius transformations*

$$\lambda \mapsto \left\{ \lambda; \quad 1-\lambda, \quad \frac{1}{\lambda}, \quad \frac{\lambda}{\lambda-1}; \quad \frac{\lambda-1}{\lambda}, \quad \frac{1}{1-\lambda} \right\}, \tag{2.16}$$

which is isomorphic to the symmetric group S_3.

(ii) *Two four-times-punctures spheres,* $\hat{\mathbb{C}} \smallsetminus (a, b, c, d)$ *and* $\hat{\mathbb{C}} \smallsetminus (\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})$*, with colored points, are conformally equivalent iff the corresponding cross-ratios are equal:* $[a, b, c, d] = [\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}]$.

(iii) *Two four-times-punctures spheres,* $\hat{\mathbb{C}} \smallsetminus \{a, b, c, d\}$ *and* $\hat{\mathbb{C}} \smallsetminus \{\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}\}$*, with uncolored points, are conformally equivalent iff the corresponding cross-ratios are related by the gross-ratio group:* $[a, b, c, d] = \gamma([\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}])$ *for some* $\gamma \in \Gamma_{\text{cr}}$.

By moving the triple (a, b, d) to $(0, 1, \infty)$ by a Möbius transformation, we turn the basic cross-ratio $[a, b, c, d]$ into $c \in \mathbb{C} \smallsetminus \{0, 1\}$. Thus, the space of colored four-times-punctured spheres is isomorphic to the thrice-punctured sphere $\mathbb{C} \smallsetminus \{0, 1\}$. Its quotient by the Möbius action of the cross-ratio group, $\mathfrak{M}(\hat{\mathbb{C}} \smallsetminus \{a, b, c, d\}) = (\mathbb{C} \smallsetminus \{0, 1\})/\Gamma_{\text{cr}}$ is the space of uncolored four-times-punctured spheres. From Exercise 2.49 we know that this space is the modular surface $\mathfrak{M} \approx \mathbb{C}$ (or rather: the modular orbifold).

EXERCISE 2.81. *Write down explicitely the covering* $\mathbb{C} \smallsetminus \{0, 1\} \to \mathfrak{M}$.

The reader has certainly noticed that the modular surface $\mathfrak{M}$ has appeared on two occasions: as the space of conformal tori and as the space of four-times-punctures spheres. There is a good reason for this. Indeed, to any torus $\mathbb{T}^2_\mu = \mathbb{C}/\mathbb{L}$, $\mu \in \mathfrak{M}(\mathbb{T}^2)$, we can associate a four–times-punctured sphere by taking its quotient by the involution $\sigma : z \mapsto -z$. Vice versa, giving a four-times-punctured sphere $\hat{\mathbb{C}} \smallsetminus X$, where $X = \{a, b, c, d\}$, parametrized by the cross-ratio $\nu = [a, b, c, d] \in \mathfrak{M}(\hat{\mathbb{C}} \smallsetminus X)$, we can construct a double branched covering $T_\nu \to \hat{\mathbb{C}} \smallsetminus X$ (for some $\nu \in \mathfrak{M}(\hat{\mathbb{C}} \smallsetminus X)$) branched over X, by gluing two copies of $\hat{\mathbb{C}}$ along two slits pairing the pucntures (which is in the most classical way of constructing Riemann surfaces).

EXERCISE 2.82. *Show that two tori are conformally equivalent iff the corresponding four-times-punctured spheres are.*

Thus, we obtain an intrinsic isomorphism between $\mathfrak{M}(\mathbb{T}^2)$ and $\mathfrak{M}(\hat{\mathbb{C}} \smallsetminus X)$.

Let us note that the above construction requires the Uniformization Theorem for the conformal tori T_ν doubly covering the four-times-punctured sphere $\mathbb{C} \smallsetminus$

$\{0, 1, \lambda\}$. However, on this occasion we do not need its full power as it can be explicitly constructed as follows:

• Realize the torus in question as the algebraic curve

$$\Gamma := \{(z, w) \in \mathbb{C}^2 : \ w^2 = z(z-1)(z-\nu)\},$$

and let $\hat{\Gamma}$ be its Universal covering.

• Consider the Abelian differential $\omega := dz/w$ on Γ.

• For any base point $p_\circ \in \hat{\Gamma}$, consider the Abelian integral

$$I : \hat{\Gamma} \to \mathbb{C}, \quad p \mapsto \int_{p_\circ}^{p} \omega, \quad p \in \hat{\Gamma}.$$

Let $\mathbb{L}$ be its lattice of *periods*, i.e. of the image of the fiber containing $p_\circ$.

• The map I descends to the desired uniformization $T_\nu \to \mathbb{C}/\mathbb{L}$.

Project 2.83. *Work out details of this construction.*

2.7. Flat structures and geometry of quadratic differentials.

2.7.1. *Flat structures with cone singularities and boundary corners.* Recall that a Euclidean, or flat, structure on a surface S is an atlas of local charts related by Euclidean motions. However, for topological reasons, many surfaces do not admit any flat structure: the Gauss-Bonnet Theorem bans such a structure on any compact surface except the torus (see below). On the other hand, if we allow some simple singularities, then these obstruction disappears.

Everybody is familiar with a Euclidean cone of angle $\alpha \in (0, 2\pi)$. To give a formal definition, just take a standard Euclidean wedge of angle α and glue its sides by the isometry. It is harder to define (and even harder to visualize) a cone of angle $\alpha > 2\pi$. One possible way is to partition α into several angles $\alpha_i \in (0, 2\pi)$, $i = 0, 1, \dots n-1$, to take wedges W_i of angles α_i, and paste W_i to W_{i+1} by gluing the sides isometrically (where i is taken mod n) (and then to check, by taking a "common subdivision", that the result is independent of the particular choice of the angles α_i).

But there is a more natural way. Consider a smooth universal covering $e : \mathbb{H} \to \mathbb{D}^*$ over the punctured disk, and endow $\mathbb{H}$ with the pullback of the Euclidean metric, $e^{-y}|dz|$. Let us define the wedge $W = W(\alpha)$ of angle α as the strip $\{z : 0 \leq \operatorname{Re} z \leq \alpha\}$ completed with one point at $\operatorname{Im} z = +\infty$. If we isometrically glue the sides of this wedge, we obtain the cone $C = C(\alpha)$ of angle α. (We can also define $C(\alpha)$ as the one-point completion at $+\infty$ of the quotient $\mathbb{H}/\alpha\mathbb{Z}$.)

Exercise 2.84. *Let γ be a little circle around a cone singularity of angle α. Check that the tangent vector γ' rotates by angle α as we go once around γ.*

According to the discussion in Appendix 2.12.2, *a cone singularity x with angle $\alpha = \alpha(x)$ carries curvature $2\pi - \alpha$.*

Let us now consider a compact flat surface S with boundary. Assume that the boundary is piecewise linear with corners. It means that near any boundary point, S is isometric to a wedge $W(\alpha)$ with some $\alpha > 0$. Points where $\alpha \neq \pi$ are called *corners of angle* α (as the corners are isolated, there are only finitely many of them). The *rotation* $\rho(x)$ at a corner $x \in \partial S$ of angle $\alpha = \alpha(x)$ is defined as $\pi - \alpha$ (see Appendix 2.12.2).

2.7.2. *Gauss-Bonnet Formula (for flat surfaces).*

THEOREM 2.85. *If S is a compact flat surface with cone singularities and piecewise linear boundary with corners then*

$$\sum K(x) + \sum \rho(y) = 2\pi\chi(S),$$

where the first sum is taken over the cone singularities while the second sum is taken over the boundary corners.

This is certainly a particular case of the general Gauss-Bonnet formula (2.35) from Appendix 2.12.2, but in our special case we will give a direct combinatorial proof of it.

PROOF. Let us triangulate S by Euclidean triangles in such a way that all cone singularities and all boundary corners are contained in the set of vertices. Let α_i be the list of the angles of all triangles. Summing these angles over the triangles, we obtain:

$$\sum \alpha_i = \pi(\#\text{ triangles}).$$

On the other hand, summation over the vertices gives:

$$\sum \alpha_i = 2\pi(\#\text{ regular vertices}) + \sum_{cones} \alpha(x) + \sum_{corners} \alpha(y)$$

$$= 2\pi\,(\#\text{ vertices}) - \sum_{cones} K(x) - \sum_{corners} \rho(y) + \pi(\#\text{ corners}).$$

Hence

$$\sum K(x) + \sum \rho(y) = \pi\,(2(\#\text{ vertices}) + (\#\text{ corners}) - (\#\text{ triangles})) = 2\pi\chi(S),$$

where the last equality follows from

$$3(\#\text{ triangles}) = 2(\#\text{ edges}) + (\#\text{ corners}).$$

□

2.7.3. *Geodesics.* Let S be a flat surface with cone singularities. A piecewise smooth curve $\gamma(t)$ in S is called a *geodesic* if it is locally shortest, i.e., for any $x = \gamma(t)$ there exists an $\varepsilon > 0$ such that for any $t_1, t_2 \in [t-\varepsilon, t+\varepsilon]$, $\gamma : [t_1, t_2] \to S$ is the shortest path connecting $\gamma(t_1)$ to $\gamma(t_2)$.

Obviously, any geodesic is piecewise linear: a concatenation of straight Euclidean intervals meeting at cone points. Moreover, both angles between two consecutive intervals in a geodesic must be at least π (in particular, the intervals cannot meet at a cone point with angle $< 2\pi$).

EXERCISE 2.86. *Verify these assertions by exploring geodesics on a cone $C(\alpha)$.*

THEOREM 2.87. *Let S be a closed flat surface with only negatively curved cone singularities. Then for any path $\gamma : [0,1] \to S$, there is a unique geodesic $\delta : [0,1] \to S$ homotopic to γ rel the endpoints.*

PROOF. *Existence.* Let L be the infimum of the lengths of smooth paths homotopic to γ rel the endpoints. We can select a minimizing sequence of piecewise linear paths with the intervals of definite length. Such paths form a precompact sequence in S, so we can select a subsequence converging to a path δ in S of length L. Obviously, δ is a local minimizer, and hence is a geodesic.

Uniqueness. Let γ and δ be two geodesics on S homotopic rel the endpoints. They can be lifted to the universal covering $\hat{S}$ to geodesics $\hat{\gamma}$ and $\hat{\delta}$ with common endpoints. We can assume without loss of generality that the endpoints a and b are the only intersection points of these geodesics (replacing them if needed by the arcs $\hat{\gamma}'$ and $\hat{\delta}'$ bounded by two consecutive intersection points). Then $\hat{\gamma}$ and $\hat{\delta}$ bound a polygon Π with vertices at a and b and some corner points x_i. Let y_j be the cone points in $\operatorname{int}\Pi$. By the Gauss-Bonnet formula,

$$(\pi - \rho(a)) + (\pi - \rho(b)) + \sum(\pi - \rho(x_i)) + \sum K(y_j) = 2\pi.$$

But the first two terms in the left-hand side are less than π while the others are negative – contradiction. □

2.7.4. Euc(2)- *and* Euc(1)*-structures.* Let S^* stand for a flat surface S with its cone singularities punctured out.

A *parallel line field* on S is a family of tangent lines $l(z) \in \mathrm{T}_z S$, $z \in S^*$, that are parallel in any local chart of S.

Let $j : \mathrm{Euc}(\mathbb{C}) \to U(2)$ be the natural projection that associates to a Euclidean motion its rotational part. Let $\mathrm{Euc}(n)$ stand for the j-preimage of the cyclic group of order n in $U(2)$. In other words, motions $A \in \mathrm{Euc}(n)$ are compositions of rotations by $2\pi k/n$ and translations. (So, the complex coordinate, they assume the form $A : z \mapsto e(k/n)z + c$.)

Lemma 2.88. *A flat surface S admits a parallel line field if and only if its Euclidean structure can be refined to a* Euc(2)*-structure.*

Proof. Let S be Euc(2)-surface and let $\theta \in \mathbb{R}/\operatorname{mod}\pi\mathbb{Z}$. Given a local chart, we can consider the parallel line field in the θ-direction. Since the θ-direction is preserved $(\operatorname{mod}\pi)$ by the group Euc(2), we obtain a well defined parallel line field on S^*.

Vice versa, assume we have a parallel line field on S^*. Then we can rotate the local charts so that this line field becomes horizontal. The transit maps for this atlas are Euclidean motions preserving the horizontal direction, i.e., elements of Euc(2). □

Lemma 2.89. *S admits a parallel line field if and only if all cone angles are multiples of π.*

Proof. Any tangent line can be parallelly transported along any path in S^*. Since S is flat, the result is independent of the choice of a path within a certain homotopy class. S admits a parallel line field if and only if the holonomy of this parallel transport around any cone singularity is trivial, i.e., it rotates the line by a multiple of π. But the holonomy around a cone singularity of angle α rotates the line by angle α. □

2.7.5. *Abelian & quadratic differentials vs translation surfaces.* Next, we will relate flat geometry to complex geometry. Namely, any flat surface S is naturally a Riemann surface. Indeed, since Euclidean motions are conformal, the flat structure induces complex structure on S^*. To extend it through cone singularities, consider a conformal isomorphism $\phi : \mathbb{H}/\alpha\mathbb{Z} \to \mathbb{D}^*$, $z \mapsto e(z/\alpha)$. It extend to a homeomorphism $C(\alpha) \to \mathbb{D}$ that serves as a local chart on the cone $C(\alpha)$.

EXERCISE 2.90. *Show that the extension of the conformal structures from S^* to S is unique.*

PROPOSITION 2.91. *Let S be a a smooth surface.*
(i) *Prescribing a* Euc(1)*-structure on S (with cone singularitires) is equivalent to precribing an Abelian differential ω on a Riemann surface supported on S.*
(ii) *Prescribing a* Euc(2)*-structure on S (with cone singularitires) is equivalent to precribing a quadratic differential q on a Riemann surface supported on S.*

For instance, if we have a holomorphic function $\phi : S \to \mathbb{C}$ then we can pull back the standard Euclidean structure on $\mathbb{C}$ to obtain a Euc(1)-structure on S with cone singularities at zeros of ϕ. The corresponding Abelian differential is $\omega = d\phi$.

More generally, we can consider a holomorphic function on the Universal covering, $\Phi : \hat{S} \to \mathbb{C}^*$, which is transformed multiplicatively under the action of the group Γ of deck transformations:

$$\Phi(\gamma z) = c(\gamma) \cdot \phi(z), \quad \gamma \in \Gamma,$$

where $c : \Gamma \to \mathbb{C}^*$ is a multiplicative homomorphism. The pullback of the standard Euclidean structure on $\mathbb{C}$ to S by $\log \Phi : \hat{S} \to \mathbb{C}$ descends to a Euc(1)-structure on $S^* := S \smallsetminus$ zeros of Φ. The correspomnding Abeliean differential is the log-derivative of Φ, $\omega = d\Phi/\Phi$.

2.8. More on orbifolds.

2.8.1. *Triangle groups: summary.* Putting together Exercises 2.13, 2.6, and 2.50, we obtain:

THEOREM 2.92. *For any triple $\{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\}$ of numbers in $\{2, 3, \dots\} \cup \{\infty\}$, there exists an orbifold $\mathcal{O} \equiv \mathcal{O}_{\{\mathfrak{p},\mathfrak{q},\mathfrak{r}\}}$ with signature $(S^2, \{\mathfrak{p}, \mathfrak{q}, \mathfrak{r}\})$. It is uniformized by a spherical, Euclidean, or hyperbolic triangle group, depending on whether the Euler characteristic*

$$\chi(\mathcal{O}) = \frac{1}{\mathfrak{p}} + \frac{1}{\mathfrak{q}} + \frac{1}{\mathfrak{r}} - 1$$

positive, zero, or negative.

2.8.2. *Almost all 2D orbifolds are good.*

THEOREM 2.93. *There only two bad* (2D) *orbifolds:*[11] $(S^2; \mathfrak{p})$ *and* $(S^2; \{\mathfrak{p}, \mathfrak{q}\})$, *with $\mathfrak{p} \neq \mathfrak{q}$.*

PROOF. Let $\mathcal{U}$ be the Universal covering of an orbifold $\mathcal{O}$, and let M be its underlying surface. Since M is a simply connected surface, it is the topological disk or the topological sphere. Assume the singular set $X \subset M$ contains more than two points. Then take a Jordan disk $\Delta \subset M$ containing three if these points. Adding a "cup" to this disk, we obtain a topological sphere S^2 with three singular points. By Theorem 2.92, the corresponding orbifold can be uniformized by a triangle group. It induces a non-trivial orbifold covering of $\mathcal{U}$ corresponding to the disk Δ with three singular points (as in §1.7.13). Hence $\mathcal{U}$ cannot be universal in this case.

So, we are reduced to the case of the disk or the sphere with at most two singular points. The disk with two singular points is topologically equivalent to the sphere with three singular points (one of which has infinite index), which is

[11] Notice that both of them have positive Euler characteristic.

uniformized by a triangular group once again. The sphere with two two singular points of the same index $\mathfrak{q}$ can be realized as $M = \hat{\mathbb{C}}$ with $X = \{0, \infty\}$, which is uniformized by the map $f_{\mathfrak{q}} : \hat{\mathbb{C}} \to hC$, $z \mapsto z^{\mathfrak{q}}$. The remaining cases correspond to the excdeptional bad orbifolds: see Example 1.115. □

2.8.3. *Orbifolds of finite conformal type.* We say that a Riemann orbifold $\mathcal{O}$ has a *finite conformal type* if it has a compact underlying surface with finitely many singular points (maybe, of infinite index). Such an orbifold has a finitely generated fundamental group and a finite Euler characteristic. We have already encountered a number of examples: Euclidean orbifolds listed in Exercise 2.6, spherical orbifolds listed in Exercise 2.13, and hyperbolic orbifolds uniformized by finitely generated Fuchsian groups of first kind. Remarkably, these examples exhaust the full list of good orbifolds of finite conformal type! See Theorem 5.9 below.

2.9. Schwarzian derivative and projective structures.

2.9.1. *Definition.* The fastest way to define the Schwarzian derivative Sf is by means of a mysterious formula:

$$Sf = \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2. \tag{2.17}$$

However, there is a bit longer but better motivated approach.

Let us try to measure how a function f at a non-critical point z deviates from a Möbius transformation. Möbius transformations depend on three complex parameters. So, one expects to find a unique Möbius transformation A_z that coincides with f to the second order. Then

$$f(\zeta) - A_z(\zeta) \sim \frac{b}{6}\,(\zeta - z)^3$$

near z, and we let $Sf(z) = b/f'(z)$.

Remark 2.94. Division by $f'(z)$ ensures scaling invariance of the Schwarzian derivative: $S(\lambda f) = Sf$. Coefficient $1/6$ provides a convenient normalization suggested by the Taylor formula: it makes $Sf(z) = f'''(z)$ for a normalized map $f(\zeta) = \zeta + O(|\zeta - z|^3)$.

Exercise 2.95. *Show that by postcomposing with a Möbius transformation, $f \mapsto A \circ f$, any univalent map f near z can be brought to the just mentioned normal form* (*with the* 2*-jet equal to* id).

The best Möbius approximation to f is easy to write down explicitly. Let $f(\zeta) = a_0 + a_1(\zeta - z) + a_2(\zeta - z)^2 + \dots$ near z with $a_1 = f'(z) \neq 0$. Then

$$A_z(\zeta) = a_0 + \frac{a_1(\zeta - z)}{1 - \beta(\zeta - z)} \quad \text{with} \quad \beta = \frac{a_2}{a_1},$$

the 3d Taylor coefficient of $f - A_z$ is $(a_3 - a_2^2/a_1)$, and (2.17) follows.

2.9.2. *Chain rule.*

LEMMA 2.96. *Let f be a holomorphic function on a domain U. Then $Sf \equiv 0$ on U if and only if f is a restriction of a Möbius map to U.*

PROOF. Sufficiency is obvious from the definition: If f is a Möbius map then $A_z = f$ at any point z, and $Sf(z) = 0$.

Vice versa, assume $Sf \equiv 0$ on U. Then f is a solution of a 3d order analytic ODE

$$f''' = \frac{3}{2}\frac{(f'')^2}{f'} = 0$$

on $U \setminus C_f$, where C_f is the critical set of f. Such a solution is uniquely determined by its 2-jet[12] at any point $z \in U \setminus C_f$. Hence $f = A_z$. □

Similarly, one can prove:

EXERCISE 2.97. *Let f and g be two holomorphic functions on a domain U. Then $Sf \equiv Sg$ on U if and only if $f = A \circ g$ for some Möbius map A.*

LEMMA 2.98 (Chain Rule).

$$S(f \circ g) = (Sf \circ g) \cdot (g')^2 + Sg. \tag{2.18}$$

PROOF. Since the Schwarzian derivative is translationally invariant on both sides (i.e., $S(T_1 \circ f \circ T_2) = Sf$ for any translations T_1 and T_2), it is sufficient to check (2.18) at the origin and to assume that $g(0) = f(0) = 0$. Furthermore, by Exercise 2.97, postcomposition of f with a Möbius transformation would not change either side of (2.18). In this way, we can bring f to a normalized form:

$$f(\zeta) = \zeta + \frac{Sf(0)}{6}\,\zeta^3 + \dots \tag{2.19}$$

and then painlessly check (2.18) by composing (2.19) with the 3-get of g. □

In particular, for a Möbius transformation A, we have:

$$S(f \circ A) = (Sf \circ A) \cdot (A')^2, \tag{2.20}$$

which coincides with the transformation rule for quadratic differentials. It suggests that the Schwarzian should be viewed not as a function but rather as a quadratic differential $Sf(z)\,dz^2$. This point of view is not quite right on *Riemann* surfaces, but it becomes exactly correct on *projective* surfaces.

2.9.3. *Projective surfaces.* A *projective structure* on a Riemann surface S is an atlas of holomorphic local charts with Möbius transit maps. A surface endowed with a projective structure is called a *projective surface*. Projective morphisms are defined naturally, so that we can refer to isomorphic projective surfaces.

Of course, the Riemann sphere $\hat{\mathbb{C}}$ has a natural projective structure, and any domain $U \subset \hat{\mathbb{C}}$ inherits it. If we have a group Γ of Möbius transformations acting properly discontinuously and freely on U then the quotient Riemann surface $V = U/\Gamma$ inherits a unique projective structure that makes the quotient map $\pi : U \to V$ projective. In particular, any hyperbolic Riemann surface V is endowed with the *Fuchsian projective structure* coming from the uniformization $\pi : \mathbb{H} \to V$.

[12]Recall that a n-jet of a function f at z is its Taylor approximant of order n at z.

Given a meromorphic function f on a projective surface V, the Chain Rule (2.20) tells us that the local expressions $Sf(z)\,dz^2$ determine a global quadratic differential on V.

EXERCISE 2.99. *Check carefully this assertion.*

More generally, let us consider two projective structures f and g on a Riemann surface V given by atlases $\{f_\alpha\}$ and $\{g_\beta\}$ respectively. Then the Chain Rule (more specifically, Exercises 2.97 and 2.99) tell us that the local expressions $S(f_\alpha \circ g_\beta^{-1})(z)\,dz^2$ determine a global quadratic differential on V endowed with the g-structure. This differential is denoted $S\{f, g\}$. It measures the distance between f and g.

In particular, given a holomorphic map $f : V \to W$ between two projective surfaces, we obtain a quadratic differential $S\{f^*(W), V\}$ on[13] V. Writing f in projective local coordinates ($\zeta = f(z)$), we obtain the familiar expression, $Sf(z)\,dz^2$, for this differential. It measures the deviation of f from being projective.

2.10. Appendix 0: Weierstrass $\mathcal{P}$-functions. An meromorphic fuction $f : \mathbb{C} \to \hat{\mathbb{C}}$ is called *periodic* if $f(z + w) = f(z)$ for some $w \in \mathbb{C}^*$. Then the set of all periods (with 0 added to them) form a *lattice* $\mathbb{L} \subset \mathbb{C}$, i.e., a discrete subgroup of $\mathbb{C}$. Such a lattice is either ismorphic to $\mathbb{Z}$ (rank 1 case) or to $\mathbb{Z}^2$ (rank 2 case).

2.10.1. *Trigonometric functions.* Let us say that an entire function $f : \mathbb{C} \to \mathbb{C}$ is *triginometric* if it is periodic with rank 1 group of periods:

$$f(z + an) = f(z), \quad n \in \mathbb{Z}.$$

After an affine change of variable, we can make $a = 1$. Familiar examples are $e(z) \equiv e^{2\pi i z}$ and $\cos z$. The former provides us with the universal covering $\mathbb{C} \to \mathbb{C}^*$, while the latter provides us with a Galois branched covering $\mathbb{C} \to \mathbb{C}$, branched over $\{\pm 1\}$ (which can be also viewed as the universal covering of the orbifold with signature $(\mathbb{C}, \{2, 2, \})$. These coverings are nicely visualized by means of checkerboard tilings.

EXERCISE 2.100. *Any trignomentric function (normalized so that $a = 1$) can be expanded into a Fourier series*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n e(nz).$$

EXERCISE 2.101. (i) *For any $n \in \mathbb{Z}$ there is a polynomial $Ч_n$ such that*

$$\cos(nz) = Ч_n(\cos z). \tag{2.21}$$

For instance, $Ч_2(z) = 2z^2 - 1$.

(ii) *All critical points of these polynomials are simple and lie on the interval $(-1, 1)$; the only critical values are ± 1 (only 1 for $n = 2$).*

These polynomials are called *Chebyshev*.

[13]Here we notationally identify surfaces with their projective structures

2.10.2. *Weierestrass $\mathcal{P}$-function.* A *elliptic function* is a doubly periodic meromorpohic function $f : \mathbb{C} \to \hat{\mathbb{C}}$. So, there is a rank two lattice[14] $L \subset \mathbb{C}$ such that $f(z+w) = f(z)$ for any $w \in L$. Equivalently, it can be viewed as a holomorphic branched covering from the torus $\mathbb{C}/L$ to the Riemann sphere $\hat{\mathbb{C}}$. This branched covering has a degree d that we call the (*associated*) *degree* of f. (Note that formally speaking, f itself has infinite degree).

Basic examples are provided by the *Weierstrass $\mathcal{P}$-functions* of associated degree two. Such a function can be explicitely represented by the following series:

$$\mathcal{P}(z) = \frac{1}{z^2} + \sum_{w \in L} \left[\frac{1}{(z-w)^2} - \frac{1}{w^2} \right] \tag{2.22}$$

PROJECT 2.102. *Justify that expression* (2.22) *represents indeed a degree two elliptic function.*

By a complex rescaling $z \mapsto \lambda z$, $\lambda \in \mathbb{C}^*$, any lattice $\mathbb{L} \subset \mathbb{C}$ can be brought to a form $\{m + n\tau\}_{(m,n)\in\mathbb{Z}^2}$ with $\operatorname{Im}\tau > 0$. Considering Weierstrass $\mathcal{P}$-functions up to rescalings $\mathcal{P}(\lambda z)$, we obtain a complex one parameter family of them, $\{\mathcal{P}_\tau(z)\}_{\tau\in\mathbb{H}}$.

EXERCISE 2.103. (i) *For any $n \in \mathbb{Z}$ there is a rational function R_n of degree n^2 such that*

$$\mathcal{P}(nz) = \mathcal{L}_n(\mathcal{P}_n(z)). \tag{2.23}$$

(ii) *All critical values of the $\mathcal{L}_n$ are contained in the 4 critical values of $\mathcal{P}$.*

(iii) *A similar function $\mathcal{L}_\rho$ is associiated to any complex multiplication $z \mapsto \rho z$ for a rank 2 lattice $\mathbb{L}$.*

These rational functions are called *Lattès maps*.

2.11. Appendix 1 : Tensor calculus in complex dimension one.

2.11.1. *General notion.* For $(n,m) \in \mathbb{Z}^2$, an (n,m)-*tensor* on a Riemann surface S is an object τ that can be locally written as a differential form

$$\tau(z)\, dz^n d\bar{z}^m. \tag{2.24}$$

Formally speaking, to any local chart $z = \gamma(x)$ on S corresponds a function $\tau_\gamma(z)$, and this family of functions satisfy the transformation rule: if $\zeta = \delta(x)$ is another local chart and $z = \phi(\zeta)$ is the transit map, then

$$\tau_\delta(\zeta) = \tau_\gamma(\phi(\zeta))\, \phi'(\zeta)^n \overline{\phi'(\zeta)}^m. \tag{2.25}$$

The regularity of the tensor (e.g., τ can be measurable, smooth or holomorphic) is determined by the regularity of all its local representative τ_γ.

Even when dealing with globally defined tensor, we will often use local notation (2.24), and we will usually use the same notation for a tensor and the representing local function.

Disregarding the regularity issue, tensors form a bigraded commutative semigroup: if τ and τ' are respectively (m,n)- and (m',n')-tensors, then $\tau\tau'$ is an $(m+m', n+n')$-tensor.

[14] i.e., a discrete subgroup generated by some basis $w_1, w_2 \in \mathbb{R}^2$

A holomorphic $(1,0)$-tensor $\omega(z)dz$ is called an *Abelian differential*; a holomorphic $(2,0)$-tensor $q(z)dz^2$ is called a *quadratic differential*. More generally, we can consider *meromorphic* $(n,0)$-tensors, e.g., meromorphic quadratic differentials.

A $(-1,1)$-tensor $\mu(z)d\bar{z}/dz$ is called a *Beltrami differential*. Notice that the absolute value of a Beltrami differential, $|\mu|$, is a global function on S. (*In this book all Beltrami differentials under consideration are assumed measurable and bounded.*)

A $(1,1)$-tensor $\rho = \rho(z)\, dzd\bar{z}$ with $\rho \geq 0$ is a *conformal Riemannian metric* $\rho(z)|dz|^2$ on S. Its area form

$$\frac{i}{2}\,\rho(z)\, dz \wedge d\bar{z} = \rho(z)\, dx \wedge dy$$

is a tensor of the same type (both are transformed by the factor $|\phi'(\zeta)|^2$). This allows us to integrate $(1,1)$-tensors:

$$\int \rho = \frac{i}{2}\int \rho(z)\, dz \wedge d\bar{z}.$$

For instance, if q is a quadratic differential then $|q|$ is a $(1,1)$-form, so that we can evaluate $\int |q|$ (at least locally). If q is a quadratic differential and μ is a Beltrami differential, then $q\mu$ is again a $(1,1)$-form, so the local integral $\int q\mu$ makes sense.

A $(-1,0)$-tensor $\dfrac{v(z)}{dz}$ has the same type as a vector field. Indeed, in this case the tensor rule (2.25) assumes the form $v_\gamma(\phi(\zeta)) = \phi'(\zeta)\, v_\delta(\zeta)$ that coincides with the transformation rule for vector fields.

EXERCISE 2.104. (i) *Let $v = v(z)/dz$ be a C^1-smooth vector field near ∞ on $\hat{\mathbb{C}}$. Show that $v(z) = az^2 + bz + O(1)$. Moreover, $v(\infty) = 0$ iff $a = 0$.*

(ii) *A vector field $v(z)/dz$ is holomorphic on the whole sphere $\hat{\mathbb{C}}$ iff*

$$v(z) = az^2 + bz + c.$$

EXERCISE 2.105. (i) *Let $q = q(z)\, dz^2$ be a meromorphic quadratic differential near ∞ on $\hat{\mathbb{C}}$. If $q(z) \asymp z^{-n}$, $n \leq 3$, then q has a pole of order $4-n$ at ∞. In particular, q has at most a simple pole at ∞ iff $q(z)$ vanishes to the third order at ∞, i.e., $q(z) = O(|z|^{-3})$.*

(ii) *$q \in \mathcal{Q}^1(\hat{\mathbb{C}})$ iff $q(z)$ is a rational function with simple poles in $\mathbb{C}$ that vanishes to the third order at ∞.*

(iii) *# poles $-$# zeros of q is equal to 4.*

2.11.2. *∂ and $\bar{\partial}$.* The differential of a function $\tau(z)$ can be expressed in $(z,\bar{z})$-coordinates as follows:

$$d\tau = \partial_x \tau\, dx + \partial_y \tau\, dy = \partial_z \tau\, dz + \partial_{\bar{z}} \tau\, d\bar{z},$$

where

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y). \tag{2.26}$$

This suggests to introduce differential operators ∂ and $\bar{\partial}$ (acting from functions to $(1,0)$- and $(0,1)$-forms respectively):

$$\partial\tau = \partial_z \tau\, dz, \quad \bar{\partial}\tau = \partial_{\bar{z}}\tau\, d\bar{z}, \quad \text{so } d = \partial + \bar{\partial}.$$

EXERCISE 2.106. *Check that* $\partial\tau$ *and* $\bar\partial\tau$ *are correctly defined* $(1,0)$*- and* $(0,1)$*-forms respectively. .*

We will sometimes use notation ∂ and $\bar\partial$ for the partial derivatives (2.26) as well, unless it can lead to a confusion.

Using the semigroup structure, we can extend these differential operators to arbitrary tensors:

$$\partial(\tau\, dz^n d\bar z^m) = \partial\tau\, dz^n d\bar z^m = \partial_z \tau\, dz^{n+1} d\bar z^m,$$

$$\bar\partial(\tau\, dz^n d\bar z^m) = \bar\partial\tau\, dz^n d\bar z^m = \partial_{\bar z} \tau\, dz^n d\bar z^{m+1}.$$

These operators increase the grading by $(1,0)$ and $(0,1)$ respectively.

For instance, if v is a vector field viewed as a $(-1,0)$ tensor, then $\bar\partial v$ is a Beltrami differential.

REMARK 2.107. The above *commutative tensor* operators should not be confused with their *anti-commutative exterior* counterparts acting on differential forms. For instance, if $\omega = \omega(z)\, dz$ is a holomorphic $(1,0)$-form then in the tensor sense $\partial\omega = \omega'(z)\, dz^2$, while $\partial\,\omega = 0$ in the exterior sense .

2.11.3. *Pullback and push-forward.* Let $f : S \to T$ be a holomorphic map between two Riemann surfaces. Then any (n,m)-form τ on T can be pulled back to an (n,m)-form $f^*\tau$ on S, which in is given in local coordinates by the expression

$$f^*(\tau(w)\, dw^m d\bar w^m) = \tau(f(z)) f'(z)^n \bar f'(z)^m\, dz^n d\bar z^m.$$

Moreover, if τ is a holomorphic/meromorphic $(n,0)$-form then so is $f^*(\tau)$.

If f is invertible then of course forms can be also pushed forward. For $\tau = \tau(z)\, dz^n d\bar z^m$, it looks as follows:

$$f_*\tau \equiv (f^{-1})^*(\tau) = \frac{\tau(z)}{f'(z)^n \bar f'(z)^m}\, dw^n d\bar w^m \quad \text{substituting } z = f^{-1}(w).$$

It is less standard that tensors can be also pushed forward by non-invertible maps (at least, by branched coverings of finite degree) by summing up the local push-forwards over the preimages:

$$f_*\tau = \sum (f_i)_*(\tau) = \sum_{z_i \in f^{-1}(w)} \frac{\tau(f(z_i))}{f'(z_i)^n \bar f'(z_i)^m}\, dw^n d\bar w^m \text{ substituting } z_i = f_i^{-1}(w).$$

where f_i is the local branch of f near $z_i \in f^{-1}(w)$. This expression is well defined outside the set V of critical values of f.

Moreover, if τ is a meromorphic $(n,0)$-form with the polar set P then $f_*\tau$ is also meromorphic, with the polar set contained in $f(P) \cup V$. Indeed, outside $f(P)\cup V$, the push-forward $f_*(\tau)$ is a holomorphic $(n,0)$-form with at most power growth near $f(P)\cup V$.

This discussion applies directly to the case of meromorphic quadratic differentials $q = q(z)\, dz^2$, which will be the main case of our interest:

$$f_* q = \sum (f_i)_* q = \sum_{z_i \in f^{-1}(w)} \frac{q(z_i)}{f'(z_i)^2}.$$

In the case of an area form $\rho\, dz \wedge d\bar{z}$, the push-forward operation is actually standard as it corresponds to the push-forward of the measure with density ρ:

$$f_*(\rho\, dz \wedge d\bar{z}) = \sum_{z \in f^{-1}(w)} \frac{\rho(z)}{|f'(z)|^2}.$$

Since the area is conserved under invertible changes of variable, we have:

$$\int f^*\rho = d \int \rho, \quad \int f_*\rho = \int \rho \tag{2.27}$$

(assuming ρ has a finite total mass).

2.11.4. *Push-forward is a contraction in $\mathcal{Q}^1$. Integrability* of a meromorphic quadratic differential q on a Riemann surface S means integrability of the corresponding area form $|q|$. Let $\mathcal{Q}^1(S)$ stand for the space of integrable meromorphic quadratic differential on S, and $\mathcal{Q}^1_{\rm loc}(S)$ stand for the space of *locally integrable* ones. Note that $q \in \mathcal{Q}^1_{\rm loc}$ if and only if it has only simple poles.

For $q \in \mathcal{Q}^1(S)$, transformation rules (2.27) (together with the triangle inequality) imply:

$$\int |f_*q| \leq \int f_*|q| = \int |q| \tag{2.28}$$

Thus, *the push-forward operator is contracting in the space of integrable holomorphic quadratic differentials.* This property plays a key role in the Thurston theory, see §39.

Exercise 2.108. *Consider a holomorphic quadratic differential $q = q(z)dz^2$ on the whole Riemann sphere $\hat{\mathbb{C}}$, so $q(z)$ is a rational function.*

(i) *What is the condition that q has zero/pole at ∞. If so, what is its order?*

(ii) *q is integrable if and only if all its poles (including at ∞) are simple;*

(iii) *For $f : z \mapsto z^d$ and $q = z^n dz^2$, calculate f^*q and f_*q.*

Lemma 2.109. *Let $f : S \to T$ be a holomorphic covering of degree d, and let q be an integrable quadratic differential on S. Then*

$$\int |f_*q| = \int |q| \tag{2.29}$$

if and only if $f^*(f_*q) = d\, q$.

Proof. Equality (2.29) is equivalent to attaining equality in (2.28). Since both q and f_*q are continuous outside a finite set and $|f_*q| \leq f_*|q|$ everywhere, integral equality in (2.28) is equivalent to pointwise equality $|f_*q| = f_*|q|$. But equality in the triangle inequality is attained if and only if all the terms have the same phase, so

$$f_*q = c_i\, (f_i)_*q, \quad c_i > 0.$$

Being positive and holomorphic in z, the factors c_i must be constants. Applying the pullback f_i^* to the last equation, we obtain:

$$f^*(f_*q) = c_i\, q \quad \text{near } z_i \in f^{-1}z.$$

But the ratio $f^*(f_*q)/q$ is a global meromorphic function: if it is locally constant, it must be globally constant, so $f^*(f_*q) = c\, q$. Finally, by (2.27)

$$\int |f^*(f_*q)| = d \int |f_*q| = d \int |q|,$$

so $c = d$. □

2.11.5. *Duality.*

LEMMA 2.110. *Let $f : S \to T$ be a holomorphic covering of degree d. Consider a meromorphic quadratic differential $q \in \mathcal{Q}^1(S)$ and a measurable essentially bounded Beltrami differential μ on T. Then*

$$\int_S q \cdot f^*\mu = \int_T f_* q \cdot \mu.$$

PROOF. It is sufficient to check that

$$\int_U q \cdot f^*\mu = \int_V f_* q \cdot \mu$$

for a base of neighborhoods V on T and $U = f^{-1}(V)$. Since f is covering, we can choose the V so that

$$U = \bigsqcup_{i=1}^{d} U_i,$$

where the restrictions $f_i = (f : U_i \to V)$ are biholomorphic. Then

$$\int_U q \cdot f^*\mu = \sum \int_{U_i} q \cdot f^*\mu = \sum \int_{U_i} f^*((f_i)_* q \cdot \mu)$$
$$= \sum \int_{U_i} (f_i)^* q \cdot \mu = \int_U (f_i)^* q \cdot \mu.$$

□

REMARK 2.111. All the above statements concerning covering maps extend immediately to maps $f : S \to T$ that are coverings over $T \smallsetminus A$ where A is a discrete subset. This includes branched coverings (see §3).

2.12. Appendix 2: Bits of 2D Riemannian geometry.

2.12.1. *Classification of Riemannian surfaces of constant curvature.* The notion of *Riemannian surface* is not the same as of *Riemann surface*. The former is a smooth surface endowed with a Riemannian metric

$$ds^2 = E\,dx^2 + 2F\,dx\,dy + G\,dy^2 \quad \text{(a local expression)}.$$

Our main models are *homogeneous surfaces* that have a constant curvature K: the standard sphere $S^2 \subset \mathbb{R}^3$ ($K > 0$), the Euclidean plane $\mathbb{R}^2$ ($K = 0$), and the hyperbolic plane $\mathbb{H}$ ($K < 0$). For definiteness, we will always normalize the curvature (by rescaling the metric) so that $K \in \{-1, 0, +1\}$.

PROPOSITION 2.112. *A Riemannian surface S of constant curvature K is endowed with the associated geometric structure: spherical (for $K > 0$), flat (for $K = 0$), or hyperbolic (for $K < 0$). Vice versa: such a geometric structure induces a metric of constant curvature on S.*

PROOF. The latter assertion is obvious as the metric of constant curvature on S can be obtained by pulling back the corresponding homogenious metric by the local charts ϕ_i of the given geometric structture (spherical, flat, or hyperbolic).

The direct assertion follows from the fact that any Riemannian surface of constant curvature is locally isometric to one of three homogeneous surfaces. It can be accomplished by analysing the *Liouville equation*

$$\Delta\phi = K\, e^{\phi}$$

for the metric $ds = e^{-\phi}\,|dz|$ of constant curvature K (see [**DNF**, §13, Thm 5]). $\square$

Geometric Uniformization Theorem. *Any simply connected complete Riemannian surface of constant curvature K is either the standard sphere $S^2 \subset \mathbb{R}^3$ ($K > 0$), or the Euclidean plane $\mathbb{R}^2$ ($K = 0$), or the hyperbolic plane $\mathbb{H}$ ($K < 0$).*

Proof. Take a base point $x \in S$ and consider a local isometry $h : (U, x) \to (V, 0)$ to the model homogeneous surface $S_\circ$ (S^2, $\mathbb{R}^2$ or $\mathbb{H}$). Assume first $K \le 0$. Then by the Hadamard Theorem (see [**DoC**, §5-6]), h extends to a global diffeomorphism $\hat{h} : S \to S_\circ$ which is isometric on the geodesic rays emanating from x. Then the local isometries along any such ray $\mathcal{R}$ coinciding with $\hat{h}$ on $\mathcal{R}$ glue into an isometry near $\mathcal{R}$. It is easy to see that they must match for nearby rays, implying that $\hat{h}$ is a global isometry.

This construction of promoting local isometries (or other geometric structures) to a global one is called *developing.*

For $K > 0$, S must be compact (see [**DoC**, §5-9, Exercise 1]). By the Gauss-Bonnet Formula, $\chi(S) > 0$, so S must be a topological sphere. Let $S_\circ$ be the standard sphere. Then the developing map $S \mapsto S_\circ$ is a locally isometric covering. As S_0 is simply connected, it is a global isometry. $\square$

By taking the Universal covering, we obtain the *full classification* of complete Riemannian surfaces of constant curvature:

Corollary 2.113. *Any complete Riemannian surface of constant curvature K is isometric to one of the following surfaces:*

(i) *Spherical case: the standard sphere $S^2 \subset \mathbb{R}^3$ ($K > 0$);*

(ii) *Flat case: the Eucliuidean plane $\mathbb{R}^2$, or the flat cylinder $\mathbb{T} \times \mathbb{R}$, or the torus $\mathbb{T}^2$ ($K = 0$);*

(iii) *Hyperbolic case: The quotient of the hyperbolic plane $\mathbb{H}^2$ modulo a Fuchsian group ($K < 0$).*

2.12.2. *Gauss-Bonnet formula for variable metrics.* Formally speaking, we can skip a discussion of this general version of the Gauss-Bonnet formula as we have verified it directly in all special cases that we need. However, it does give a deeper insight into the matter. The reader can consult, e.g., [**DoC**] for a proof.

Let S be a compact smooth Riemannian surface, maybe with boundary. Let $K(x)$ be the Gaussian curvature at $x \in S$, and let $\kappa(x)$ be the geodesic curvature at $x \in \partial S$. The Gauss-Bonnet formula related these geometric quantities to topology of S:

$$\int_S K d\sigma + \int_{\partial S} \kappa ds = 2\pi\chi(S), \tag{2.30}$$

where $d\sigma$ and ds are the area and length elements respectively.

In particular, if S is closed then

$$\int_S K d\sigma = 2\pi\chi(S), \tag{2.31}$$

which, in particular, implies that there are no flat structures on a closed surface of genus $g \neq 0$. Also, for a surface S of constant curvature $K \in \{\pm 1\}$, we obtain:

$$\operatorname{area} S = \pm 2\pi\chi(S). \tag{2.32}$$

The boundary term in (2.30) admits a nice interpretation. Let us parametrize a closed boundary curve γ with the length parameter, so that $\gamma'(t)$ is the unit tangent vector to γ. Then for nearby points $\gamma(t)$ and $\gamma(\tau)$, where $\tau = t + \Delta t > t$, let $v(t,\tau)$ be the tangent vector $\gamma'(\tau)$ parallelly transported from $\gamma(\tau)$ back to $\gamma(t)$. Then let $\theta(t,\tau)$ be the angle between $\gamma'(t)$ and $v(t,\tau)$ (taking with positive sign if v points "into S". Summing these angles up over a partition of γ into small intervals, we obtain the *rotation number* of the tangent vector. It coincides with $\displaystyle\int_\gamma \kappa ds$.

Note that if ∂S consists of geodesics, the boundary term in (2.30) disappears, and it assumes the same form (2.31) as in the closed case.

If we allow the Riemannian metric to have an isolated singularity at some point $x \in S$ then using the Gauss-Bonnet formula for a small disk around x, we can assign the Gaussian curvature to x:

$$K(x) = 2\pi - \lim_{\gamma \to x} \int_\gamma \kappa ds, \tag{2.33}$$

provided the limit exists. (Here γ is a small circle around x, and $K(x)$ is assumed to be integrable.)

In particular, if x is a cone singularity of angle $\theta \in (0, +\infty)$, then it support curvature

$$K(x) = 2\pi - \theta. \tag{2.34}$$

If we allow a corner of angle $\alpha \in (0, \infty)$ at a boundary point $y \in \partial S$ (see §2.7.1), we can assign the rotation number $\rho(y) = \pi - \alpha \in (\pi, -\infty)$ to it as the angle between the incoming and outgoing tangent vectors.

Then the Gauss-Bonnet formula is still valid for surfaces with singularities and boundary corners, assuming the following form:

$$\int_S K d\sigma + \sum_{sing} K(x) + \int_{\partial S} \kappa ds + \sum_{corners} \rho(y) = 2\pi\chi(S). \tag{2.35}$$

With these definitions, the Gauss-Bonnet Formula immediately extends to orbifolds. For instance, let $\mathcal{O}$ be an orbifold of finite conformal type with signature $(S; \{\mathfrak{q}_i\})$, where S is a closed underlying surface. Endow it with an orbifold Riemannian metric with curvature $K(x)$ and the area form $d\sigma$. As the curvature form $K\, d\sigma$ of this metric is calculated in the orbifold local charts, it does not account to the curvatures K_i of the cone singularities. By (2.35) and (2.33), we have:

$$2\pi\,\chi(S) = \int K\, d\sigma + \sum K_i = \int K\, d\sigma + \sum \left(2\pi - \frac{2\pi}{\mathfrak{q}_i}\right),$$

yielding

$$\int K\, d\sigma = 2\pi\,\chi(\mathcal{O}). \tag{2.36}$$

(So, the orbifold Euler characteristic amounts to the *regular* portion of the total curvature.)

3. Holomorphic proper maps and branched coverings

3.1. First observations.

EXERCISE 3.1. *Show than any non-constant holomorphic map between two Riemann surfaces is topologically holomorphic.*

The following assertion generalizes Exercise 2.61 to branched coverings:

EXERCISE 3.2. *Any holomorphic branched covering $f : S \to S'$ of finite degree between hyperbolic Riemann surfaces extends continuously to a branched covering $\mathbf{f} : \mathbf{S} \to \mathbf{S}'$ between their ideal compactifications (holomorphic on* $\operatorname{int} \mathbf{S}$*).*

EXERCISE 3.3. (i) *A holomorphic branched covering $f : \mathbb{D} \to \mathbb{D}$ of degree d is a Blaschke product*

$$f(z) = \lambda \prod_{k=0}^{d} \frac{z - a_k}{1 - \bar{a}_k z}, \quad \text{where } |\lambda| = 1,\ a_k \in \mathbb{D}.$$

(ii) *A holomorphic branched covering $f : \mathbb{H}_+ \to \mathbb{H}_+$ of degree d with $f(\infty) = \infty$ has a form*

$$f(z) = \lambda_0 z + a_0 - \sum_{k=1}^{d-1} \frac{\lambda_k}{z - a_k}, \quad \text{where } \lambda_k > 0,\ a_k \in \mathbb{R},\ k = 0, 1, \dots, d - 1.$$

3.2. Riemann-Hurwitz formula. This formula gives us a beautiful relation between topology of the surfaces S and T, and branching properties of f.

RIEMANN - HURWITZ FORMULA. Let $f : S \to T$ be a branched covering of degree d between two topological surfaces of finite type. Let C be the set of branched points of f. Then

$$\chi(S) = d \cdot \chi(T) - \sum_{a \in C} (\deg_a f - 1).$$

Let us define the *multiplicity* of $a \in C$ as a critical point to be equal to $\deg_a f - 1$ (in the holomorphic case it is the multiplicity of a as the root of the equation $f'(a) = 0$). Then the sum in the right-hand side of the Riemann-Hurwitz formula is equal to the *number of critical points of f counted with multiplicities.*

PROOF. Let us first assume that S and T are closed Riemann surfaces.

Let us consider a triangulation $\mathcal{T}$ of T such that all critical values of f are vertices of $\mathcal{T}$. By the Euler formula,

$$\chi(T) = v(\mathcal{T}) - e(\mathcal{T}) + t(\mathcal{T}),$$

where v,e and t stand for the number of vertices, edges and faces (triangles) of $\mathcal{T}$. Let $\mathcal{S}$ be the lift of this triangulation to S. Then

$$t(\mathcal{S}) = d \cdot t(\mathcal{T}), \quad e(\mathcal{S}) = d \cdot e(\mathcal{T}), \quad v(\mathcal{S}) = d \cdot v(\mathcal{T}) - \sum_{a \in C} (\deg_a f - 1),$$

and the conclusion follows.

To deal with non-closed case, consider the one-point-per-end compactifications $\hat{S}$ and $\hat{T}$ of our surfaces. If S and T are of finite type then these surfaces are closed. Since f is proper, it continuously extends to a map $\hat{f} : \hat{S} \to \hat{T}$. This map is certainly proper. By Exercise 1.103, it is topologically holomorphic. Thus, it is a branched covering (of the same degree d). As in the above calculation, we have

$$|\mathcal{E}(S)| = d \cdot |\mathcal{E}(T)| - \sum_{e \in \mathcal{E}(S)} (\deg_e \hat{f} - 1).$$

Putting this together with the Riemann-Hurwitz formula for $\hat{f}$ implies the desired. □

Remark 3.4. One could also define $\chi(S)$ for an open surface S using *ideal* triangulations, with some vertices being at infinity, ∞_e (see §1.7.5). Then the proofs for closed and open cases become identical.

Remark 3.5. The formula also applies to surfaces with boundary, with the same proof (or by removing the boundary, which does not change the Euler characteristic). Or else, one can use triangulations of the bordered surfaces.

Corollary 3.6. *Under the above circumstances, assume that T is a topological disk. Then S is a topological disc as well if and only if there are $d-1$ critical points in S (counted with multiplicities).*

Proof. A surface S is a topological disk if and only if $\chi(S) = 1$. □

3.3. Topological Argument Principle. Consider the punctured plane $\mathbb{R}^2 \smallsetminus \{b\}$. If $\gamma : S^1 \to \mathbb{R}^2 \smallsetminus \{b\}$ is a smooth oriented Jordan curve then one can define the winding number of γ around b as

$$w_b(\gamma) = \int_\gamma d(\arg(x-b)).$$

Since the 1-form $d(\arg(x-b))$ is closed, the winding number is the same for homotopic curves. Hence we can define the *winding number $w_b(\gamma)$ for any continuous Jordan curve $\gamma : S^1 \to \mathbb{R}^2 \smallsetminus \{b\}$* by approximating it with a smooth Jordan curves.

Furthermore, the winding number can be linearly extended to any simplicial 1-cycle in $\mathbb{R}^2 \smallsetminus \{b\}$ with integer coefficients (i.e., a formal combination of oriented Jordan curves in $\mathbb{R}^2 \smallsetminus \{b\}$) and then factored to the first homology group. It gives an isomorphism

$$w : H_1(\mathbb{R}^2 \smallsetminus \{b\}) \to \mathbb{Z}, \quad [\gamma] \mapsto w_b(\gamma). \tag{3.1}$$

Exercise 3.7. *Prove the last statement.*

Proposition 3.8. *Let D be a Jordan disc and let $f : \bar{D} \to \mathbb{R}^2$ be a continuous map that does not assume some value $b \in \mathbb{R}^2$ on ∂D. If $w_b(f|\,\partial\mathbb{D}) \neq 0$ then f assumes the value b in D.*

Proof. Obviously, the curve $\gamma = (f : \partial D \to \mathbb{R}^2)$ is contractible in $f(\bar{D})$. If $b \notin f(D)$ then γ would be contractible in $\mathbb{R}^2 \smallsetminus \{b\}$, so it would have zero winding number around b. □

Let $x \in D$ be an isolated preimage of $b = fx$. Then one can define the $\operatorname{ind}_x(f)$ as follows. Take a disk $V \subset D$ around x that does not contain other preimages of $b = fx$. Take a positively oriented Jordan loop $\gamma \subset V \smallsetminus \{x\}$ around x whose image does not pass through b, and calculate the winding number of the curve $f : \gamma \to \mathbb{R}^2 \smallsetminus \{b\}$ around b:

$$\operatorname{ind}_x(f) = w_{fx}(f \circ \gamma).$$

Clearly it does not depend on the loop γ, since the curves corresponding to different loops are homotopic without crossing b.

PROPOSITION 3.9. *Let $D \subset \mathbb{R}^2$ be a domain bounded by a Jordan curve Γ, and let $f : \bar{D} \to \mathbb{R}^2$ be a continuous map such that the curve $f \circ \Gamma$ does not pass through some point $b \in \mathbb{R}^2$. Assume that the preimage of this point $f^{-1}b$ is discrete in D. Then*

$$\sum_{x \in f^{-1}b} \operatorname{ind}_x(f) = w_b(f \circ \Gamma),$$

provided Γ is positively oriented.

PROOF. Note first that since $f^{-1}b$ is a discrete subset of a compact set $\bar{D}$, $f^{-1}x$ is actually finite, so that the above sum makes sense.

Select now small Jordan loops γ_i around points $x_i \in f^{-1}b$, and orient them anticlockwise. Since Γ and these loops bound a 2-cell, $[\Gamma] = \sum [\gamma_i]$ in $H_1(\bar{D} \smallsetminus f^{-1}b)$. Hence $f_*[\Gamma] = \sum f_*[\gamma_i]$ in $H_1(\mathbb{R}^2 \smallsetminus \{b\})$. Applying the isomorphism (3.1), we obtain the desired formula. $\square$

EXERCISE 3.10. *Let $f : D \to \mathbb{R}^2$ be a continuous map, and let $a \in D$ be an isolated point in the fiber $f^{-1}b$, where $b = f(a)$. Assume that $\operatorname{ind}_a(f) \neq 0$. Then f is locally surjective near a, i.e., for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $f(\mathbb{D}_\varepsilon(a)) \supset \mathbb{D}_\delta(b)$.*

Hint: For a small ε-circle γ around a, the curve $f \circ \gamma$ stays some positive distance δ from b. Then for any $b' \in \mathbb{D}_\delta(b)$ we have: $\operatorname{ind}_b(f \circ \gamma) = \operatorname{ind}_b(f \circ \gamma) \neq 0$. But if $b' \notin f(\mathbb{D}_\varepsilon(a))$ then the curve $f \circ \gamma$ could be shrunk to b without crossing b'.

3.3.1. *Degree of proper maps.*

3.4. Lifts.

LEMMA 3.11. *Let $f : (S, a) \to (T, b)$ and $\tilde{f} : (\tilde{S}, \tilde{a}) \to \tilde{T}, \tilde{b})$ be two double branched between topological disks (with or without boundary) coverings branched at points a and $\tilde{a}$ respectively. Then any homeomorphism $h : (T, b) \to (\tilde{T}, \tilde{b})$ lifts to a homeomorphism $H : (S, a) \to (\tilde{S}, \tilde{a})$ which makes the diagram*

$$\begin{array}{ccc} (S,a) & \xrightarrow{H} & (\tilde{S},\tilde{a}) \\ f\downarrow & & \downarrow \tilde{f} \\ (T,b) & \xrightarrow[h]{} & (\tilde{T},\tilde{b}) \end{array}$$

commutative. Moreover, the lift H is uniquely determined by its value at any unbranched point $z \neq a$. Hence there exists exactly two lifts.

If the above surfaces are Riemann and the map h is holomorphic then then the lifts H are holomorphic as well.

PROOF. Puncturing all the surfaces at their preferred points, we obtain four topological annuli. The maps f and $\tilde{f}$ restrict to the unbranched double coverings between respective annuli, while h restricts to a homeomorphism. The image of the fundamental group $\pi_1(S \smallsetminus \{a\})$ under f consist of homotopy classes of curves with winding number 2 around b, and similar statement holds for $\tilde{f}$. Since the winding number is preserved under homeomorphisms,

$$h_*(f_*(\pi_1(S \smallsetminus \{a\})) = \tilde{f}_*(\pi_1(\tilde{S} \smallsetminus \{\tilde{a}\})). \tag{3.2}$$

By the general theory of covering maps, h admits a lift

$$H : S \smallsetminus \{a\} \to \tilde{S} \smallsetminus \{\tilde{a}\}$$

which makes the "punctured" diagram (3.2) commutative. Moreover, this lift is uniquely determined by the value of H at any point $z \in S \smallsetminus \{a\}$.

Extend now H at the branched point by letting $H(a) = \tilde{a}$. It is clear from the local structure of branched coverings that this extension is continuous (as well as the inverse one), so that it provides us with the desired lift.

If all the given maps are holomorphic then the lift H is also holomorphic on the punctured disk $S \smallsetminus \{a\}$. Since isolated singularities are removable for bounded holomorphic maps, the extension of H to the whole disk is also holomorphic. □

EXERCISE 3.12. *Similar statement holds for branched coverings f and $\tilde{f}$ with a single branched point (of any degree). Analyze the situation with two branched points.*

EXERCISE 3.13. *Assume that all the topological disks in the above lemma are $\mathbb{R}$-symmetric and that all the maps commute with the reflection σ with respect to $\mathbb{R}$. Assume also that $h(f(T \cap \mathbb{R})) = \tilde{f}(\tilde{T} \cap \mathbb{R})$. Then both lifts H also commute with σ (in particular, they preserve the real line).*

3.5. Galois branched coverings.

EXERCISE 3.14. *If a Galois branched covering $f : S \to T$ is holomorphic then the deck transformations $\gamma \in \Gamma$ are holomorphic automorphisms of S. Vice versa, if in the previous Exercise the deck transformations are holomorphic automorphisms then the natural projection $S \to S/\Gamma$ is holomorphic.*

The following statement shows that any branched covering can be "symmetrized" in a controlled way:

PROPOSITION 3.15. *Let $f : S \to T$ be a holomorphic branched cover of Riemann surfaces of degree d. Then there is a Galois branched cover $g : \Sigma \to T$ of degree at most $d!$ that factors as $g = f \circ h$ for some $h : \Sigma \to S$. Moreover, g is ramified only over critical values of f.*

PROOF. . Let $\mathcal{V}$ be the set of critical values of f and let $T^* = T \smallsetminus \mathcal{V}$. Let $\mathcal{C} = f^{-1}(\mathcal{V})$ and let $S^* = S \smallsetminus \mathcal{C}$. Then $f : S^* \to T^*$ is an unbranched covering of degree d. By Exercise 1.58, there is a Galois covering $g : \Sigma^* \to T^*$ of degree at most $d!$ that factors through some covering $h : \Sigma^* \to S^*$. Moreover, we can endow Σ^* with the pullback complex structure to make both g and h holomorphic.

Let us now complete these coverings to obtain branched coverings. To this end, it is enough to treat h. Let us take any $c \in \mathcal{C}$ and a little disk D around it. Let $D^* = D \smallsetminus \{c\}$, and let us consider a component U^* of $h^{-1}(D^*)$ Then $f : U^* \to D^*$ is a finite degree covering. Hence U^* is isomorphic to a punctured disk $\mathbb{D}^*$ (see Exercise 2.39), so it represents a cusp end e of Σ^*. Let us complete it to a conformal disk $U = U^* \cup \{\infty_e\} \approx \mathbb{D}$ and extend h holomorphically to the completion (see Proposition 2.59 and a remark after it).

Doing this with all punctures c and all punctured disks U^*, we complete Σ^* to a Riemann surface Σ such that h admits a holomorphic extension $\Sigma \to S$. Obviously, this extension is proper, so it is a branched covering. □

Note in conclusion that the notion of a holomorphic Galois branched covering is equivalent to the notion of a covering between Riemann orbifolds (compare Exercise 1.113).

3.6. Telescope. For dynamical purposes, let us prepare a simple but very useful lemma.

TELESCOPING LEMMA. *Let $U_i \subset \mathbb{C}$, $i = 0, 1, \dots, l$, be a family of open topological disks, and let $\phi_i : U_i \to \mathbb{C}$, $i = 0, 1, \dots, l-1$, be a family univalent maps such that $\phi_i(U_i) \supset U_{i+1}$. Let $\Phi := \phi_{n-1} \circ \cdots \circ \phi_0$. Then* $\operatorname{Dom} \Phi$ *is a topological disk $D \subset U_0$ univalently mapped onto U_l. Furthermore, if $\phi_i(U_i) \Supset U_{i+1}$ then $D \Subset U_0$.*

COROLLARY 3.16. *Under the above circumstances, there is a nest of topological disks*

$$D \equiv D^0 \subset D^1 \subset \cdots \subset D^l \equiv U_0$$

such that D^k is mapped univalently under $\Phi_k := \phi_k \circ \cdots \circ \phi_0$ onto U_l. Furthermore, if $\phi_i(U_i) \Supset U_{i+1}$ then $D^k \Subset D^{k+1}$, $k = 0, 1, \dots, l-1$.

4. Riemann, Montel, Koebe

4.1. Little Montel Theorem.

THEOREM 4.1 (Little Montel). *Any bounded family of holomorphic functions is normal.*

PROOF. It is because the derivative of a holomorphic function can be estimated via the function itself. Indeed by the Cauchy formula

$$|\phi'(z)| \le \frac{\max_{\zeta \in U} |\phi(\zeta)|}{\operatorname{dist}(z, \partial U)^2}.$$

Thus, if a family of holomorphic functions ϕ_n is uniformly bounded, their derivatives are uniformly bounded on compact subsets of U. By the Arzela-Ascoli Criterion, this family is precompact in the space $C(U)$ of continuous functions. Since $\mathcal{M}(U)$ is closed in $C(U)$, we see that the original family is precompact in the space $\mathcal{M}(U)$. □

4.2. Riemann Mapping Theorem. For dynamical applications, we will not need the full strength of the Uniformization Theorem: only uniformization of plane domains will be relevant. Let us start with the most classical case:

RIEMANN MAPPING THEOREM. *Any simply connected domain $D \subset \hat{\mathbb{C}}$ whose complement contains more than one point is conformally equivalent to the unit disk $\mathbb{D}$. The conformal isomorphism $\phi : \mathbb{D} \to D$ is unique up to pre-composition with a Möbius transformation $M \in \operatorname{Aut}(\mathbb{D})$.*[15]

[15]For instance, it is uniquely determined by its value at 0 and the image of the tangent vector $1 \in \mathrm{T}_0\mathbb{D}$ under $D\phi(0)$.

Proof. The uniqueness part is obvious, so let us focus on the existence.

First, notice that D can be conformally mapped onto a bounded domain in $\mathbb{C}$. Indeed, since $\hat{\mathbb{C}} \setminus D$ contains more than one point and D is simply connected, $\hat{\mathbb{C}} \setminus D$ is in fact a continuum. Let us take two points $a_1, a_2 \in \hat{\mathbb{C}} \setminus D$, and move them to $0, \infty$ by a Möbius transformation. This turns D into a domain in $\mathbb{C}^*$.[16] Since D is simply connected, the square root map $Q : z \mapsto \sqrt{z}$ has a single-valued branch on D. Applying it, we obtain a domain whose complement has non-empty interior (the image of the other branch of Q). Moving ∞ to this complement by a Möbius transformation, we make D a bounded domain in $\mathbb{C}$.

Let us now take a point $a \in D$, and consider the space $\mathcal{C}$ of conformal embeddings $\psi : D \to \mathbb{D}$ normalized so that $\psi(a) = 0$. Note that $\mathcal{C} \neq \emptyset$ since D can be embedded into $\mathbb{D}$ by an affine map. By the Little Montel Theorem, $\mathcal{C}$ is normal. Hence we can find a conformal map $\psi_0 \in \mathcal{C}$ that maximizes the derivative $|\psi'(a)|$ over the class $\mathcal{C}$.

We claim that ψ_0 conformally maps D onto $\mathbb{D}$. The only issue is surjectivity. Assume there is a point $a \in \mathbb{D} \setminus \psi_0(D)$. Let $B : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ be a double branched covering with critical point at a.

Exercise 4.2. *Write down B explicitly.*

Since $\psi_0(D)$ is simply connected, there is a single-valued branch $B^{-1} : \psi_0(D) \to \mathbb{D}$. By the Schwarz Lemma, $|B'(0)| < 1$, and hence the embedding $B^{-1} \circ \psi_0 : (D, a) \to (\mathbb{D}, 0)$ has a bigger derivative at a than ψ – contradiction. □

In particular, if we mark a point $a \in D$, then the uniformization $\phi : (\mathbb{D}, 0) \to (D, a)$ is unique up to rotations of $\mathbb{D}$. Since rotations preserve the foliations $\mathbb{D}$ by the radii and circles centered at 0, their images under ϕ are well defined. In this way we obtain two orthogonal analytic foliations of $D \setminus \{a\}$, by (*Green*) *rays* $\mathcal{R}^\theta := \{\phi(re(\theta)) : \ r \in (0, 1)\}$ and *equipotential* $\mathcal{E}^r := \{\phi(re(\theta)) : \ \theta \in \mathbb{R}/\mathbb{Z}\}$.

Note that this definition is consistent with the one given in §10.9 since the Green function G_a is equal to $-\log|\phi^{-1}(z)|$.

4.3. Normal families and Big Montel Theorem. Let U be a Riemann surface, and let $\mathcal{M}(U)$ be the space of meromorphic functions $\phi : U \to \bar{\mathbb{C}}$. Supply the *target* Riemann sphere $\bar{\mathbb{C}}$ with the spherical metric d_s and the space $\mathcal{M}(U)$ with the topology of uniform convergence on compact subsets of U. Thus $\phi_n \to \phi$ if for any compact subset $K \subset U$, $d_s(\phi_n(z), \phi(z)) \to 0$ uniformly on K. Since locally uniform limits of holomorphic functions are holomorphic, $\mathcal{M}(U)$ is *closed* in the space $C(U)$ of continuous functions $\phi : U \to \bar{\mathbb{C}}$ (endowed with the topology of uniform convergence on compact subsets of U).

Exercise 4.3. *Endow $\mathcal{M}(U)$ with a metric compatible with the above convergence that makes $\mathcal{M}(U)$ a complete metric space.*

It is important to remember that the target should be supplied with the *spherical* rather than *Euclidean* metric even if the original family consists of *holomorphic* functions. In the limit we can still obtain a meromorphic function, though of a very special kind:

[16] We will keep notation D for various domains conformally equivalent to D.

Exercise 4.4. *Let* $\phi_n : U \to \mathbb{C}$ *be a sequence of holomorphic functions converging to a meromorphic function* $\phi : U \to \bar{\mathbb{C}}$ *such that* $\phi(z) = \infty$ *for some* $z \in U$. *Then* $\phi(z) \equiv \infty$, *and thus* $\phi_n(z) \to \infty$ *uniformly on compact subsets of* U.

A family of meromorphic functions on U is called *normal* if it is precompact in $\mathcal{M}(U)$.

Exercise 4.5. *Show that normality is the local property: If a family is normal near each point* $z \in U$, *then it is normal on* U.

Exercise 4.6. *If a domain* $U \subset \mathbb{C}$ *is supplied with the Euclidean metric* $|dz|$ *while the target* $\bar{\mathbb{C}}$ *is supplied with the spherical metric* $|dz|/(1+|z|^2)$, *then the corresponding "ES norm" of the differential* $D\phi(z)$ *is equal to* $|\phi'(z)|/(1+|\phi(z)|^2)$, $z \in U$. *Show that a family of meromorphic functions* $\phi_n : U \to \bar{\mathbb{C}}$ *is normal if and only if the ES norms* $\|D\phi_n(z)\|$ *are uniformly bounded on compact subsets of* U.

Exercise 4.7. *A sequence of holomorphic functions is normal if and only if one can extract from any subsequence a further subsequence which is either locally bounded or divergent (locally uniformly) to* ∞.

Theorem 4.8 (Montel). *If a family of meromorphic functions* $\phi_n : U \to \bar{\mathbb{C}}$ *does not assume three values then it is normal.*

Proof. Since normality is a local property, we can assume that U is a disk. Let us endow it with the hyperbolic metric ρ. Let a, b, c be omitted values on $\bar{\mathbb{C}}$, and let ρ' be the hyperbolic metric on the thrice punctured sphere $\bar{\mathbb{C}} \smallsetminus \{a, b, c\}$.

By the Schwarz Lemma, all the functions ϕ_n are contractions with respect to these hyperbolic metrics. By Proposition 7.5 (iii), the spherical metric is dominated by ρ', so the ϕ_n are uniformly Lipschitz from metric ρ to the spherical metric. Normality follows. □

Theorem 4.9 (Refined Montel). *Let* $\{\phi_n : U \to \bar{\mathbb{C}}\}$ *be a family of meromorphic functions. Assume that there exist three meromorphic functions* $\psi_i : U \to \bar{\mathbb{C}}$ *such that for any* $z \in U$ *and* $i \neq j$ *we have:* $\psi_i(z) \neq \psi_j(z)$ *and* $\phi_n(z) \neq \psi_i(z)$. *Then the family* $\{\phi_n\}$ *is normal.*

Proof. Let us consider the holomorphic family of Möbius transformations $h_z : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ depending on $z \in U$ as a parameter such that

$$h_z : (\psi_1(z), \psi_2(z), \psi_3(z)) \mapsto (0, 1, \infty).$$

Then the family of functions $\Phi_n(z) = h_z(\phi_n(z))$ omits value $0, 1, \infty$, and hence is normal by Theorem 4.8. It follows that the original family is normal as well. □

Exercise 4.10. *Show that the theorem is still valid if the functions* ψ_j *are different but* $\psi_i(z) = \psi_j(z)$ *is allowed for some* $z \in U$.

Given a family $\{\phi_n\}$ of meromorphic functions on U, we can define its *set of normality* as the maximal open set $F \subset U$ on which this family is normal.

4.4. Koebe Distortion Theorem. We will now discuss one of the most beautiful and important theorems of the classical geometric functions theory.

The *inner radius* $r_D(a) \equiv \operatorname{dist}(a, \partial D)$ of a pointed disk (D, a) is the biggest round disk $\overline{\mathbb{D}}(a, \rho)$ contained in $\overline{D}$. The *outer radius* $R_D(a) \equiv \operatorname{dist}_{\mathrm{H}}(a, \partial D)$ is the radius of the smallest disk $\overline{\mathbb{D}}(a, \rho)$ containing $\overline{D}$. (If $a = 0$, we will simply write r_D and R_D.) The *shape* (or *dilatation*) of a disk D around a is the ratio $R_D(a)/r_D(a)$.

THEOREM 4.11 (Koebe Distortion). *Let $\phi : (\mathbb{D}, 0) \to (D, a)$ be a conformal isomorphism, and let $k \in (0, 1)$, $D_k = \phi(\mathbb{D}_k)$. Then there exist constants $C = C(k)$ and $L = L(k)$ (independent of a particular ϕ!) such that*

$$\frac{|\phi'(z)|}{|\phi'(\zeta)|} \leq C(k) \text{ for all } z, \zeta \in \mathbb{D}_k \tag{4.1}$$

and

$$L(k)^{-1}|\phi'(0)| \leq r_{D_k,a} \leq R_{D_k}(a) \leq L(k)\,|\phi'(0)|. \tag{4.2}$$

In particular, the inner radius of the image $\phi(\mathbb{D})$ around a is bounded from below by an absolute constant times the derivative at the origin:

$$r_{\phi(D)}(a) \geq \rho|\phi'(0)| > 0. \tag{4.3}$$

The expression in the left-hand side of (4.1) is called the *distortion* of ϕ. Thus, estimate (4.1) tells us that the function ϕ restricted to $\mathbb{D}_k$ has a *uniformly bounded distortion* (depending on κ only). Estimate (4.2) tells that the shape of the domain D_k around a is uniformly bounded. This shape is also called the *dilatation* of h on $\mathbb{D}_\kappa$. So, univalent functions have *uniformly bounded dilatation* on any disk $\mathbb{D}_\kappa$. Note that since *any* proper topological disk in $\mathbb{C}$ can be uniformized by $\mathbb{D}$, there could be no possible bounds on the distortion and dilatation of ϕ in the whole unit disk $\mathbb{D}$. However, once the disk is slightly shrunk, the bounds appear!

The Koebe Distortion Theorem is equivalent to the normality of the space of normalized univalent functions:

THEOREM 4.12. *The space $\mathcal{U}$ of univalent functions $\phi : (\mathbb{D}, 0) \to (\mathbb{C}, 0)$ with $|\phi'(0)| = 1$ is compact* (*in the topology of uniform convergence on compact subsets of* $\mathbb{D}$).

Let us make a simple but important observation:

LEMMA 4.13. *Let $\phi : (\mathbb{D}, 0) \to (\mathbb{C}, 0)$ be a univalent function normalized so that $|\phi'(0)| = 1$. Then the image $\phi(\mathbb{D})$ cannot contain the whole unit circle $\mathbb{T}$.*

PROOF. Otherwise the inverse map ϕ^{-1} would be well defined on some disk $\mathbb{D}_r$ with $r > 1$, and the Schwarz Lemma would imply $|D\phi^{-1}(0)| \leq 1/r < 1$, contrary to the normalization assumption. □

Proof of Theorem 4.12. By Lemma 4.13, for any $\phi \in \mathcal{U}$ there is a $\theta \in \mathbb{R}$ such that the rotated function $e^{i\theta}\phi$ does not assume value 1. Since the group of rotation is compact, it is enough to prove that the space $\mathcal{U}_0 \subset \mathcal{U}$ of univalent functions $\phi \in \mathcal{U}$ which do not assume value 1 is compact.

Let us puncture $\mathbb{D}$ at the origin, and restrict all the functions $\phi \in \mathcal{U}_0$ to the punctured disk $\mathbb{D}^*$. Since all the ϕ are univalent, they do not assume value 0 in $\mathbb{D}^*$. By the Montel Theorem, the family $\mathcal{U}_0$ is normal on $\mathbb{D}^*$.

Let us show that it is normal at the origin as well. Take a Jordan curve $\gamma \subset \mathbb{D}^*$ around 0, and let Δ be the disk bounded by γ. Restrict all the functions $\phi \in \mathcal{U}_0$ to γ. By normality in $\mathbb{D}^*$, the family $\mathcal{U}_0$ is either uniformly bounded on γ, or admits a sequence which is uniformly going to ∞. But the latter is impossible since all the curves $\phi_n(\gamma)$ intersect the interval $[0, 1]$ (as they go once around 0 and do not go around 1). Thus, the family $\mathcal{U}_0$ is uniformly bounded on γ. By the Maximum Principle, it is is uniformly bounded, and hence normal, on Δ as well.

Thus, the family $\mathcal{U}_0$ is precompact. What is left, is to check that it contains all limiting functions. By the Argument Principle, limits of univalent functions can be either univalent or constant. But the latter is not possible in our situation because of normalization $|\phi'(0)| = 1$. □

EXERCISE 4.14. (i) *Show that a family $\mathcal{F}$ of univalent functions $\phi : \mathbb{D} \to \mathbb{C}$ is precompact in the space of all univalent functions if and only if there exists a constant $C > 0$ such that*

$$|\phi(0)| \le C \quad \text{and } C^{-1} \le |\phi'(0)| \le C \text{ for all } \phi \in \mathcal{F}.$$

(ii) *Let (Ω, a) be a pointed domain in $\mathbb{C}$ and let $C > 0$. Consider a family $\mathcal{F}$ of univalent functions $\phi : \Omega \to \mathbb{C}$ such that $|\phi(a)| \le C$. Show that this family is normal if and only if there exists $\rho > 0$ such that each function $\phi \in \mathcal{F}$ omits some value ζ with $|\zeta| < \rho$.*

Proof of the Koebe Distortion Theorem. Compactness of the family $\mathcal{U}$ immediately yields that functions $\phi \in \mathcal{U}$ and their derivatives are uniformly bounded on any smaller disk $\mathbb{D}_k$, $k \in (0,1)$. Combined with the fact that all functions of $\mathcal{U}$ are univalent, compactness also implies a lower bound on the inner radius $r_{\phi(D_k)}$ and on the derivative $\phi'(z)$ in $\mathbb{D}_k$. These imply estimates (4.1) and (4.2) on the distortion and shape by normalizing a univalent function $\phi : \mathbb{D} \to \mathbb{C}$, i.e., considering

$$\tilde{\phi}(z) = \frac{\phi(z) - a}{\phi'(0)} \in \mathcal{U}.$$

(Note that this normalization does not change either distortion of the function or its dilatation.)

Estimate (4.3) is an obvious consequence of the left-hand side of (4.2). □

We have given a qualitative version of the Koebe Distortion Theorem, which will be sufficient for all our purposes. The quantitative version provides sharp constants $C(k)$, $L(k)$, and ρ, all attained for a remarkable extremal *Koebe function* $f(z) = z/(1-z)^2 \in \mathcal{U}$. The sharp value of the constant ρ is particularly remarkable:

KOEBE 1/4−THEOREM. *Let $\phi : (\mathbb{D}, 0) \to (\mathbb{C}, 0)$ be a univalent function with $\phi'(0) = 1$. Then $\phi(\mathbb{D}) \supset \mathbb{D}_{1/4}$, and this estimate is attained for the Koebe function.*

We will sometimes refer to the Koebe 1/4−Theorem rather than its qualitative version (4.3), though as we have mentioned, the sharp constants never matter for us.

EXERCISE 4.15. *Find the image of the unit disk under the Koebe function.*

Let us finish with an invariant form of the Koebe Distortion Theorem:

THEOREM 4.16. *Consider a pair of conformal disks $\Delta \Subset D$. Let*

$$\operatorname{mod}(D \smallsetminus \Delta) \ge \mu > 0.$$

Then any univalent function $\phi : D \to \mathbb{C}$ has a bounded (in terms of μ) distortion on Δ:

$$\frac{|\phi'(z)|}{|\phi'(\zeta)|} \le C(\mu) \text{ for all } z, \zeta \in \Delta.$$

The proof will make use of one important property of the modulus of an annulus: if an annulus is getting pinched, then its modulus is vanishing:

LEMMA 4.17. *Let* $0 \in K \subset \mathbb{D}$, *where* K *is compact. If*

$$\operatorname{mod}(\mathbb{D} \setminus K) \geq \mu > 0$$

then $K \subset \mathbb{D}_k$ *where the radius* $k = k(\mu) < 1$ *depends only on* μ.

PROOF. Assume there exists a sequence of compact sets K_i satisfying the assumptions but such that $R_i \to 1$, where R_i is the outer radius of K_i around 0. Let us uniformize $D \setminus K_i$ by a round annulus, $h_i : \mathbb{A}(\rho_i, 1) \to \mathbb{D} \setminus K_i$. Then $\rho_i \leq \rho \equiv e^{-\mu} < 1$. Thus, the maps h_i are well-defined on a common annulus $A = \mathbb{A}(\rho, 1)$. By the Little Montel Theorem, they form a normal family on A, so that we can select a converging subsequence $h_{i_n} \to h$.

Let $\gamma \subset A$ be the equator of A. Then $h(\gamma)$ is a Jordan curve in $\mathbb{D}$ which separates the sets K_{i_n} (with sufficiently big n) from the unit circle - contradiction. $\square$

Remark. The extremal compact sets in the above lemma (minimizing k for a given μ) are the straight intervals $[0, ke^{i\theta}]$.

Proof of Theorem 4.16. Let us uniformize D by the unit disk, $h : \mathbb{D} \to D$, in such a way that $h(0) \in \Delta$. Let $\tilde{\Delta} = h^{-1}\Delta$ and $\tilde{\phi} = \phi \circ h$. By Lemma 4.17, $\tilde{\Delta} \subset \mathbb{D}_k$, where $k = k(\mu) < 1$. By the Koebe Theorem, the distortion of the functions h and $\tilde{\phi}$ on $\tilde{\Delta}$ is bounded by some constant $C = C(k)$. Hence the distortion of ϕ is bounded by C^2. $\square$

We will often use the following informal formulation of Theorem 4.16: "If $\phi : D \to \mathbb{C}$ is a univalent function and $\Delta \subset D$ is *well inside* D, then ϕ has *a bounded distortion* on Δ". Or else: "If a univalent function $\phi : \Delta \to \mathbb{C}$ has a *definite space* around Δ, then it has a bounded distortion on Δ".

Let us summarize some of the above results in a very useful comparison of the derivative of a univalent function with the inner radius of its image:

COROLLARY 4.18. *For any univalent function* $\phi : (\mathbb{D}, 0) \to (D, a)$, *we have:*

$$r_D(A) \leq |\phi'(0)| \leq 4\, r_D(a).$$

PROOF. The left-hand side estimate follows from Lemma 4.13 by normalizing ϕ. The Koebe 1/4-Theorem implies the right-hand side one: $r_D(a) \geq \dfrac{1}{4}|\phi'(0)|$. $\square$

5. Uniformization Theorem

5.1. Statement. The following theorem of Riemann and Koebe is the most fundamental result of complex analysis:

THEOREM 5.1. *Any simply connected Riemann surface is conformally equivalent to either the Riemann sphere* $\hat{\mathbb{C}}$, *or to the complex plane* $\mathbb{C}$, *or the unit disk* $\mathbb{D}$.

We also say that *any simply connected Riemann surface as a conformal sphere, or a conformal plane, or a conformal disk.*

5.2. Classification of Riemann surfaces. Consider now any Riemann surface S. Let $\pi : \hat{S} \to S$ be its universal covering. Then the complex structure on S naturally lifts to $\hat{S}$ turning S into a simply connected Riemann surface which holomorphically covers S. Thus, we come up with the following classification of Riemann surfaces:

THEOREM 5.2. *Any Riemann surface S is conformally equivalent to one of the following surfaces:*

- *The Riemann sphere $\hat{\mathbb{C}}$ (spherical case);*
- *The complex plane $\mathbb{C}$, or the punctured plane $\mathbb{C}^*$, or a torus $\mathbb{T}^2_\tau$, $\tau \in \mathbb{H}$ (parabolic case);*
- *The quotient of the hyperbolic plane $\mathbb{H}^2$ modulo a discrete group of isometries (hyperbolic case).*

Thus, *any* Riemann surface comes endowed with one of the three geometries: projective (§2.3), affine (§2.2), or hyperbolic (§2.4). In particular, *any hyperbolic Riemann surface S is endowed with the canonical hyperbolic metric*, the push-forward of the hyperbolic metric from $\mathbb{H}^2$ to S.

5.3. Smooth annuli. We will now pass to domains on Riemann surfaces beginning with annuli:

PROPOSITION 5.3. *Let $A \Subset S$ be a topological annulus in a Riemann surface S with piecewise smooth boundary. Then A is conformally equivalent to a standard annulus $\mathbb{A}(r, R)$.*

PROOF. Let us call one of the boundary components of A "inner", $\partial^i A$, and the other one "outer", $\partial^o A$ (compare §1.7.12). Let us consider the "harmonic measure" of the outer component, i.e. a harmonic function $u(z)$ on A vanishing on $\partial^i A$ and $\equiv 1$ on $\partial^o A$ (see §10.8). Let u^* be its harmonic conjugate, This function is not single valued, but rather gets changed by the period

$$p = \int_\gamma *du$$

under the monodromy along a non-trivial cycle γ in A (see §10.1). Hence the holomorphic function

$$f = \exp \frac{2\pi}{p}(u + iu^*)$$

is single valued. Moreover, it *properly* maps A onto the round annulus $\mathbb{A}(1, e^{1/p})$ and has degree one (since f homeomorphically maps the equipotentials of A onto the round circles. The conclusion follows. □

5.4. Simply connected domains.

PROPOSITION 5.4. *Let $D \Subset S$ be a simply connected domain on a Riemann surface S with piecewise smooth boundary. Then D is conformally equivalent to the unit disk $\mathbb{D}$.*

PROOF. Take a base point $z_0 \in D$, and let D_ε be a coordinate disk of radius $\varepsilon > 0$ centered at z_0. Then $U \smallsetminus D_\varepsilon$ is a topological annulus with piecewise smooth boundary, so by Proposition 5.3 there is a conformal map $\phi_\varepsilon : D \smallsetminus D_\varepsilon \to \mathbb{A}(r(\varepsilon), 1)$ onto a round annulus. By the Little Montel Theorem, the family of maps ϕ_ε is normal on $D \setminus \{z_0\}$.[17] Let us select a converging subsequence $\phi_{\varepsilon_k} \to \phi$ as $\varepsilon_k \to 0$, where $\phi : D \setminus \{z_0\} \to \mathbb{D}$ is a holomorphic map. By Removability of isolated singularities, ϕ holomorphically extends through z_0.

Let us show that $\phi : D \to \mathbb{D}$ is proper. It is sufficient to check that for any $r \in (0,1)$, the preimage $\phi^{-1}(\mathbb{D}_r)$ is compactly contained in D. Indeed, take any $R \in (r,1)$. By invariance of the modulus,

$$\operatorname{mod}(\phi_\varepsilon^{-1}(\mathbb{A}(R,1)) = \frac{1}{2\pi} \log \frac{1}{R} > 0, \quad \text{for any } \varepsilon > 0 \text{ sufficiently small.}$$

By Lemma 4.17, $\operatorname{dist}(\phi_\varepsilon^{-1}(\mathbb{T}_R), \partial D) \geq \rho > 0$ for some $\rho = \rho(R) > 0$. Letting $\varepsilon \to 0$, we conclude that $\operatorname{dist}(\phi^{-1}(\mathbb{T}_r), \partial D) \geq \rho > 0$, and properness of ϕ follows.

So, ϕ has a well defined degree. Since degree is stable under perturbations, $\deg \phi = \deg \phi_{\varepsilon_k}$ for all k sufficiently large. Thus $\deg \phi = 1$, and hence ϕ is a conformal isomorphism. $\square$

5.5. Simply connected Riemann surfaces. We are now ready to prove the Uniformization Theorem: it is covered by the following two results.

THEOREM 5.5. *Any simply connected open Riemann surface S is isomorphic to either the disk $\mathbb{D}$ or to the complex plane $\mathbb{C}$.*

PROOF. Fix a base point $p \in S$ and some reference local chart ψ near it. By Lemma 1.80, S can be exhausted by a nest of (open) topological disks D_n with piecewise smooth boundary:

$$p \in D_0 \Subset D_1 \Subset \dots, \quad \bigcup D_n = S.$$

By Proposition 5.4, for each n, there is a conformal map $\phi_n : (D_n, p) \to (\mathbb{D}_{R_n}, 0)$ normalized so that $\phi_n'(p) = 1$, where the derivative is calculated with respect to the reference local chart ψ.

The Schwarz Lemma implies that $R_0 < R_1 < \dots$. Let $R := \lim R_n \in (0, \infty]$. By the Koebe Theorem, for any $m \in \mathbb{N}$, the sequence $(\phi_n)_{n=m}^\infty$, being restricted to D_m, is precompact. By the diagonal procedure, we can select a subsequence ϕ_{n_k} converging on each D_m. The limits patch together into the desired conformal map $\Phi : S \to \mathbb{D}_R$ (where $\mathbb{D}_\infty \equiv \mathbb{C}$). $\square$

THEOREM 5.6. *Any simply connected closed Riemann surface S is isomorphic to the Riemann sphere $\hat{\mathbb{C}}$.*

PROOF. By the Fundamental Theorem of 2D Topology, S is a topological sphere. So, if we puncture out a point p from S, we obtain a topological disk. By Theorem 5.5, $S \setminus \{p\}$ is isomorphic to either $\mathbb{D}$ or $\mathbb{C}$.

Assume the former, and let $\phi : S \setminus \{p\} \to \mathbb{D}$ be an isomorphism. By Removability of isolated singularities for conformal maps, ϕ extends to a holomorphic map $\hat{\phi} : S \to D$. But such maps do not exist for variety of reasons (e.g., the image $\hat{\phi}(S)$ must be simultaneously compact and open; or by the Maximum Principle).

So, there is an isomorphism $\phi : S \setminus \{p\} \to \mathbb{C}$. Using the Removability of isolated singularities once again, we extend it to an isomorphism $\hat{\phi} : S \to \hat{\mathbb{C}}$. $\square$

[17] On normality with varying domains of definition, see §7.7.

5.6. Extension to ends revisited. Let us now formulate a semi-local version of Exercise 3.2:

LEMMA 5.7. *Let $f : S \to S'$ be a holomorphic map between two Riemann surfaces. Let e and e' be tame ends of S and S' respectively such that f induces a proper map $e \mapsto e'$.*

(i) *If E is a cusp then f extends holomorphically through the ideal punctures at ∞.*

(ii) *If E is non-cuspidal then f extends continuously to a covering between the ideal circles at infinity of the ends.*

PROOF. Let us consider a topological cylinder $F' \subset S'$ representing the end E', and let $F \subset S$ be the component of $f^{-1}(F')$ that has e as one of its ends. Such a component exists as f properly maps E to E'. Moreover, if F' is sufficiently small then $f : F \to F'$ is proper, and hence is a branched covering of finite degree. Then it has finitely many critical points, and F' can be further shrunk so that $f : F \to F'$ is unbranched. By 1.100, F is a topological cylinder, and hence it represents the end e. By Exercise 2.61, $f : F \to F'$ extends to a covering between the ideal completions of these cylinders, implying the assertion. $\square$

COROLLARY 5.8. *Let $i : S \hookrightarrow S'$ be an embedding between two Riemann surfaces. Let E and E' be tame ends of S and S' respectively such that i properly maps E to E'. Then i continuously extends to a homeomorphism $\partial^I_E S \to \partial^I_{E'} S'$ between the ideal circles at infinity of the ends.*

Under these circumstances, we identify E and E' by means of the extended homeomorphism i.

5.7. Uniformization of orbifolds.

THEOREM 5.9. *Let $\mathcal{O}$ be a Riemann orbifold of finite conformall type with non-exceptional signature (i.e., different from $(S^2; \mathfrak{p})$ and $(S^2, \{\mathfrak{p}.\mathfrak{q}\},\ \mathfrak{p} \neq \mathfrak{q})$. Then $\mathcal{O}$ falls into one of the following three types:*

(i) *Parabolic type: $\chi(\mathcal{O}) = 0$. In this case $\mathcal{O}$ is uniformized by a discrete group of Euclidean motions acting on $\mathbb{C}$ (listed in Exercise 2.6), and possesses the canonical flat structure (up to scaling).*

(ii) *Elliptic type: $\chi(\mathcal{O}) > 0$. In this case $\mathcal{O}$ is uniformized by a finite group of rotations of the round sphere (listed in Exercise 2.13), and possesses the canonical spherical structure.*

(iii) *Hyperbolic type: $\chi(\mathcal{O}) < 0$. In this case $\mathcal{O}$ is uniformized by a finitely generated Fuchsian group of first kind, and possesses the canonical hyperbolic structure.*

PROOF. By Theorem 2.93, our orbifold is good, so it can be uniformized by a discrete group Γ of autoimorphisms of a simply connected Riemann surface S. By the Uniformization Theorem, S is either $\mathbb{C}$, or $\hat{\mathbb{C}}$ or $\mathbb{H}$, endowed with the canonical geometric structure (flat, spherical, or hyperbolic respectively), making Γ a group of motions. This endows $\mathcal{O}$ with the corresponding geometric structure. By the Gauss-Bonnet Formula for orbifolds (2.36), the Euler characteristic $\chi(\mathcal{O})$ has the same sign as the curvature of the structure. $\square$

Accordingly Riemann orbifolds are classified as *elliptic, parabolic, or hyperbolic* (similarly to Riemann surfaces). Moreover, almost all Riemann orbifolds are hyperbiolic: the short list of parabolic and elliptic ones is provided in Exercises 2.6 and 2.13.

6. Extremal length and width

6.1. Definitions. Let us now introduce one of the most powerful tools of Conformal Geometry. Given a path family Γ in a Riemann surface U, we will define a conformal invariant $\mathcal{L}(\Gamma)$ called the extremal length of Γ. Consider a measurable conformal metric $\rho|dz|$ on $\mathbb{C}$ with finite total mass,

$$m_\rho(U) = \int \rho^2 dx \wedge dy < \infty, \quad \text{where } \rho : U \to [0, \infty]$$

(such metrics are called *admissible*). Let

$$l_\rho(\gamma) = \int_\gamma \rho|dz|,$$

stand for the length of $\gamma \in \Gamma$ in this metric (with the convention $l_\rho(\gamma) = \infty$ if γ is non-rectifiable, or $\rho|\,\gamma$ is not measurable, or else it is not integrable[18]). Define the *ρ-length of* Γ as

$$l_\rho(\Gamma) = \inf_{\gamma \in \Gamma} l_\rho(\gamma).$$

Normalize it in the scaling invariant way:

$$\mathcal{L}_\rho(\Gamma) = \frac{l_\rho(\Gamma)^2}{m_\rho(U)},$$

and define the *extremal length* of Γ as

$$\mathcal{L}(\Gamma) = \sup_\rho \mathcal{L}_\rho(\Gamma),$$

where the supremum is taken over all admissible metrics.

A metric ρ on which this supremum is attained (if exists) is called *extremal*.

Exercise 6.1. *Show that the value of $\mathcal{L}(\Gamma)$ does not change if one uses only smooth admissible metrics ρ.*

Let us summarize immediate consequences of the definition:

Exercise 6.2. • Extension of the family: *If a path family Γ' contains a family Γ, then $\mathcal{L}(\Gamma') \le \mathcal{L}(\Gamma)$.*

• Overflowing: *If Γ overflows Γ' (i.e., each path of Γ contains some path of Γ'), then $\mathcal{L}(\Gamma) \ge \mathcal{L}(\Gamma')$.*

• Independence of the ambient surface: *If $U \subset U'$ and Γ is a path family in U then $\mathcal{L}(\Gamma) = \mathcal{L}(\Gamma')$.* (This justifies skipping of "U" in the notation.)

• Disregarding small subfamilies: *If Γ is a smooth foliation of some domain and $\Gamma' \subset \Gamma$ comprises almost all curves of Γ then $\mathcal{L}(\Gamma') = \mathcal{L}(\Gamma)$.*

[18] For this to make sense, we should think of ρ as an actual *function* rather than a class of functions up to modification on null-sets. It is also convenient to assume that ρ is defined *everywhere*.

The *extremal width* of the family Γ is defined as the inverse to its length:

$$\mathcal{W}(\Gamma) = \mathcal{L}(\Gamma)^{-1}.$$

One can also conveniently define it as follows:

EXERCISE 6.3. $\mathcal{W}(\Gamma) = \inf m_\rho(U)$, *where the infimum is taken over all metrics with* $\rho(\gamma) \geq 1$ *for all paths* $\gamma \in \Gamma$. (In this context, such metrics are called *admissible*. Sometimes, to distinguish it from the previous admissibility requirement, we call these metrics $\mathcal{W}$-*admissible*.)

REMARK 6.4. One should think that a family is "big" if it is "wide", i.e., it has big extremal width. So, big families are short.

The extremal length and width are *conformal invariants*:
If $\phi : U \to U'$ *is a conformal isomorphism between two Riemann surfaces such that* $\phi(\Gamma) = \Gamma'$, *then* $\mathcal{L}(\Gamma) = \mathcal{L}(\Gamma')$. This immediately follows from the observation that ϕ transfers the family of admissible metrics on U to the family of admissible metrics on U' preserving all the quantities in question.

6.2. Electric circuits laws. We will now formulate two crucial properties of the extremal length and width that show that they behave respectively like the resistance and the conductance in electric circuits.

Let Γ_1, Γ_2 and Γ be three path families on U. We say that Γ *disjointly overflows* Γ_1 and Γ_2 if any path $\gamma \in \Gamma$ contains a pair of disjoint paths $\gamma_1 \in \Gamma_1$ and $\gamma_2 \in \Gamma_2$.

SERIES LAW. *Assume that a family* Γ *disjointly overflows families* Γ_1 *and* Γ_2. *Then*

$$\mathcal{L}(\Gamma) \geq \mathcal{L}(\Gamma_1) + \mathcal{L}(\Gamma_2),$$

or equivalently,

$$\mathcal{W}(\Gamma) \leq \mathcal{W}(\Gamma_1) \oplus \mathcal{W}(\Gamma_2).$$

Here $x \oplus y = (1/x + 1/y)^{-1}$ is the *harmonic sum* of x and y (which is conjugate to the usual sum by means of $z \mapsto x^{-1}$).

PROOF. Let ρ_1 and ρ_2 be arbitrary admissible metrics. By appropriate rescalings, we can normalize them so that

$$l_{\rho_i}(\Gamma_i) = m_{\rho_i}(U) = \mathcal{L}_{\rho_i}(\Gamma_i), \quad i = 1, 2.$$

Let $l_\rho = \max(\rho_1, \rho_2)$. Since any $\gamma \in \Gamma$ contains two disjoint paths $\gamma_i \in \Gamma_i$, we have:

$$l_\rho(\gamma) \geq l_{\rho_1}(\gamma_1) + l_{\rho_2}(\gamma_2) \geq l_{\rho_1}(\Gamma_1) + l_{\rho_2}(\Gamma_2) = \mathcal{L}_{\rho_1}(\Gamma_1) + \mathcal{L}_{\rho_2}(\Gamma_2).$$

Taking the infimum over all $\gamma \in \Gamma$, we obtain:

$$l_\rho(\Gamma) \geq \mathcal{L}_{\rho_1}(\Gamma_1) + \mathcal{L}_{\rho_2}(\Gamma_2).$$

On the other hand, since $\rho \leq \rho_1 + \rho_2$, we have:

$$m_\rho(U) \leq m_{\rho_1}(U) + m_{\rho_2}(U) = \mathcal{L}_{\rho_1}(\Gamma_1) + \mathcal{L}_{\rho_2}(\Gamma_2).$$

Hence

$$\mathcal{L}_\rho(\Gamma) \geq \mathcal{L}_{\rho_1}(\Gamma_1) + \mathcal{L}_{\rho_2}(\Gamma_2).$$

Taking the supremum over all normalized metrics ρ_1 and ρ_2, we obtain the desired inequality. $\square$

We say that two path families, Γ_1 and Γ_2, are *disjoint* if they are contained in disjoint measurable sets.

PARALLEL LAW. *Let* $\Gamma = \Gamma_1 \cup \Gamma_2$. *Then*

$$\mathcal{W}(\Gamma) \le \mathcal{W}(\Gamma_1) + \mathcal{W}(\Gamma_2).$$

Moreover, if Γ_1 *and* Γ_2 *are disjoint then*

$$\mathcal{W}(\Gamma) = \mathcal{W}(\Gamma_1) + \mathcal{W}(\Gamma_2).$$

PROOF. This time, let us consider metrics ρ_1 and ρ_2 that are $\mathcal{W}$-admissible in the sense of Exercise 6.3, so that $l_{\rho_i}(\Gamma_i) \ge 1$. Let again $\rho = \max(\rho_1, \rho_2)$. Then $l_\rho(\Gamma) \ge 1$ as well, and hence

$$\mathcal{W}(\Gamma) \le m_\rho(U) \le m_{\rho_1}(U) + m_{\rho_2}(U).$$

Taking the infimum over the metrics ρ_i, we obtain the desired inequality.

Assume now that Γ_1 and Γ_2 are disjoint. Let X_1 and X_2 be two disjoint measurable sets supporting the respective families. Take any admissible metric ρ with $\rho(\Gamma) \ge 1$, and let $\rho_i = \rho | X_i$. Then $\rho_i(\Gamma_i) \ge 1$ as well, and hence

$$m_\rho(U) \ge m_{\rho_1}(U) + m_{\rho_2}(U) \ge \mathcal{W}(\Gamma_1) + \mathcal{W}(\Gamma_2).$$

Taking the infimum over admissible ρ, we obtain the opposite inequality. □

REMARK 6.5. Both laws extend immediately to the case of n families $\Gamma_1, \dots, \Gamma_n$.

6.3. Annulus, rectangle and torus revisited.

6.3.1. *Modulus as the extremal length.* We will now calculate the modulus of an annulus (see §2.6.1) in terms of the extremal length. Consider an open flat cylinder $\mathrm{Cyl} \equiv \mathrm{Cyl}^l_h$ with circumference l and height h. Proper curves $(0,1) \to \mathrm{Cyl}$ going from the top to the bottom of Cyl will be called *vertical*. [19] Among these curves there are *genuinely vertical*, that is, straight intervals perpendicular to the top and the bottom. *Horizontal curves* in Cyl are closed curves homotopic to the top and the bottom of Cyl. Among them there are *genuinely horizontal*, that is, the circles parallel to the top and the bottom. Genuinely vertical and horizontal curves form the *vertical* and *horizontal foliations* respectively.

If A is a conformal annulus, then it is isomorphic to a flat cylinder, and will be freely identified with it. Curves in A corresponding to (genuinely) vertical/horizontal curves in the cylinder will be called in the same way. In particular, the vertical and horizontal foliations in $\mathbb{A}(r, R)$ are respectively comprised of radial intervals and co-centric circles.

PROPOSITION 6.6. *Let* Γ *be a family of vertical curves in a conformal annulus* $A \approx \mathrm{Cyl}^l_h$ *containing almost all genuinely vertical ones. Then*

$$\mathcal{L}(\Gamma) = \operatorname{mod} A = h/l.$$

Moreover, the Euclidean metric on the cylinder Cyl *is extremal.*

PROOF. We will identify A with a flat cylinder $\mathrm{Cyl} \equiv \mathrm{Cyl}^l_h$. Take first the flat metric e on the cylinder. Then $l_e(\gamma) \ge h$ for any $\gamma \in \Gamma$, so that, $l_e(\Gamma) = h$. On the other hand, $m_e(\Gamma) = lh$. Hence

$$\mathcal{L}(\Gamma) \ge \mathcal{L}_e(\Gamma) = h^2/lh = \operatorname{mod} \mathrm{Cyl}.$$

Take now any admissible metric ρ on Cyl. Let $\gamma_\theta \in \Gamma$ be the genuinely vertical curve through $\theta \in \mathbb{R}/l\mathbb{Z}$. Then $l_\rho(\Gamma) \le l_\rho(\gamma_\theta)$ for any $\theta \in \mathbb{R}/l\mathbb{Z}$. Integrating this

[19] In this discussison, we do not require that vertical curves land at any points of ∂A.

over $\mathbb{R}/l\mathbb{Z}$ (using that $\gamma_\theta \in \Gamma$ for a.e. $\theta \in \mathbb{R}/l\mathbb{Z}$)) and applying the Fubini Theorem and the Cauchy-Schwarz inequality, we obtain:

$$(l \cdot l_\rho(\Gamma))^2 \le \left(\int_{\mathbb{R}/lZ} l_\rho(\gamma_\theta)\, d\theta\right)^2 = \left(\int_{\mathrm{Cyl}} \rho\, dm_e\right)^2 \le lh\, m_\rho(\mathrm{Cyl}). \tag{6.1}$$

Hence $\mathcal{L}_\rho(\mathrm{Cyl}) \le h/l = \operatorname{mod} \mathrm{Cyl}$, and the conclusion follows. □

EXERCISE 6.7. *Show that the Euclidean metric is the unique extremal metric on* Cyl.

There is also the "dual" way to evaluate the same modulus:

EXERCISE 6.8. *Let Γ be a family of horizontal curves in* Cyl *containing almost all genuinely horizontal curves. Then*

$$\operatorname{mod} \mathrm{Cyl} = \mathcal{W}(\Gamma).$$

6.3.2. *Gröztsch Inequality.* The following inequality plays an outstanding role in holomorphic dynamics (the name we use for it is widely adopted in the dynamical literature):

PROPOSITION 6.9. *Consider a conformal annulus A containing n disjoint conformal annuli $A_1, \dots A_n$ homotopically equivalent to A. Then*

$$\operatorname{mod} A \ge \sum \operatorname{mod} A_k.$$

PROOF. Let Γ_k be the horizontal family of A_k and Γ be the horizontal family in A. By the Parallel Law, $\mathcal{W}(\Gamma) \ge \sum \mathcal{W}(\Gamma_k)$, and the conclusion follows from Exercise 6.8. (Dually, one can apply the Series Law to the extremal length of the vertical families.)

□

6.3.3. *Euclidean geometry of an annulus.* The length-area method allows one to relate $\operatorname{mod} A$ to the Euclidean geometry of A. As a simple illustration, let us show that $\operatorname{mod} A$ *is bounded by the distance between the inner and the outer complements of A rel the size of the inner complement*:

LEMMA 6.10. *Consider a topological annulus $A \subset \mathbb{C}$. Let K and Q stand for its inner and outer complements*[20] *respectively. Then*

$$\operatorname{mod} A \le C(1 + \operatorname{dist}(K, Q)/ \operatorname{diam} K).$$

PROOF. Let Γ be the family of horizontal curves in A. According to Exercise 6.8, we need to bound $\mathcal{W}(\Gamma)$.

Take points $a \in K$ and $c \in Q$ on minimal distance $\operatorname{dist}(K, Q)$, and then select a point $b \in K$ such that $\operatorname{dist}(a, b) > \operatorname{diam} K/2$. Consider a family Δ of closed Jordan curves $\gamma \subset \mathbb{C} \smallsetminus \{a, b, c\}$ with winding number 1 around a and b and winding number 0 around c. Since $\Gamma \subset \Delta$, $\mathcal{W}(\Gamma) \le \mathcal{W}(\Delta)$.

Let us estimate $\mathcal{L}(\Delta)$ from below. Rescale the configuration $\{a, b, c\}$ (without changing notations) so that $|a - b| = 1$ and $|a - c| = d$, where

$$\operatorname{dist}(K, Q)/ \operatorname{diam} K \le d \le 2 \operatorname{dist}(K, Q)/ \operatorname{diam} K.$$

[20]For an annulus with comlicated boundary, they are defined as follows. Take a homotopically non-trivial Jordan curve $\gamma \subset A$, and let the inner/outer complement be the union of components of $\mathbb{C} \smallsetminus A$ lying in the inner/outer component of $\mathbb{C} \smallsetminus \gamma$, respectively.

Take now the unit neighborhood B of the union $[a,b] \cup [a,c]$ of two intervals, and endow it with the Euclidean metric e (extended by 0 outside B). Then $l_e(\Delta) \geq 1$ so this family is $\mathcal{W}$-admissible. Moreover, $m_e(B) \leq C(1+d)$, and hence $\mathcal{W}(\Delta) \leq C(1+d)$ as well. ☐

Corollary 6.11. *If neither K nor Q is a singleton then* $\operatorname{mod} A < \infty$.

Corollary 6.12. *If $\mu := \operatorname{mod} A \geq \mu > C$ (with C from* Lemma 6.10) *then a homotopically non-trivial round annulus of modulus $\geq \log \mu$ can be inscribed into A.*

Proof. A round annulus of outer radius $\operatorname{dist}(Q,K)$ and inner radius $\operatorname{diam} K$ can be inscribed into A. ☐

Exercise 6.13. *Show that under the above circumstances,*

(i) *If* $\operatorname{mod} A \geq \mu > C$ *then* $\operatorname{diam} K \leq M \exp(-\alpha\mu)$, *with absolute* $M > 0$ *and* $\alpha > 0$.

(ii) *There is a lower bound:* $\operatorname{mod} A \geq \mu(\operatorname{dist}(K,Q)/\operatorname{diam}(K)) > 0$.

An annulus $A \subset \hat{\mathbb{C}}$ is called *ε-pinched* if

$$\operatorname{dist}(K,Q) \leq \varepsilon \min(\operatorname{diam} K, \operatorname{diam} Q),$$

where K and Q are the inner and outer complementary components of A.

Proposition 6.14. *The modulus of an annulus $A \subset \hat{\mathbb{C}}$ is δ-small iff A is ε-pinched (quantitatively).*

Problem 6.15. *Let $U = A \cup K$. Show that*

$$\frac{\operatorname{area} U}{\operatorname{area} K} \geq 1 + 4\pi \operatorname{mod} A.$$

6.3.4. *Divergence Property.* The Gröztsch Inequality proves an effecient tool to recognize cusps as ends of Riemann surfaces.

Proposition 6.16. *Let $(A_n)_{n=0}^{\infty}$ be a nest of disjoint annuli in $\mathbb{C}$, and let $K_0 \supset K_1 \supset \dots$ be the corresponding nest of their inner components A_n.*

If $\sum \operatorname{mod} A_n = \infty$ then $K := \bigcup K_n$ is a singleton.

Proof. Let us take a horizontal curve γ in A_0 and consider the annulus A bounded by γ and K. By the Gröztsch Inequality,

$$\operatorname{mod} A \geq \sum_{n=1}^{\infty} \operatorname{mod} A_n = \infty.$$

Corollary 6.11 concludes the argument. ☐

6.3.5. $\mathbb{R}$-*symmetric case.* Let us consider a pair of nested intervals, $L \Subset \operatorname{int} I$, and let $H^{\pm}$ be the components of $I \smallsetminus L$. Let us consider the affiliated conformal annulus

$$A(I,L) := (\mathbb{C} \smallsetminus \mathbb{R}) \cup (H^+ \cup H^-).$$

and let

$$\operatorname{mod}(I:L) := \operatorname{mod} A(I,L), \quad \operatorname{mod}_{\mathbb{R}}(I:L) := \min_{\pm} \frac{|H^{\pm}|}{|L|}. \tag{6.2}$$

LEMMA 6.17. *We have:*

$$\operatorname{mod}(I:L) \geq \varepsilon > 0 \Leftrightarrow \operatorname{mod}_{\mathbb{R}}(I:L) \geq \delta > 0.$$

PROOF. To obtain implication $\Leftarrow$, consider the round annulus A whose inner circle is based upon L as a diameter, and the outer circle is based upon the scaled interval $(1+\delta)\cdot I$.

The other implication follows from Lemma 6.10. □

COROLLARY 6.18. *Let (I, L) and (I', L') be two pairs of real intervals, and let $\phi : (\mathbb{C}, I, L) \to (\mathbb{C}, I', L')$ be a conformal embedding that restricts to a diffeomorphism between the interval pairs. Then*

$$\operatorname{mod}_{\mathbb{R}}(I:L) \geq \delta > 0 \Rightarrow \operatorname{mod}_{\mathbb{R}}(I':L') \geq \varepsilon > 0.$$

PROOF. By conformal invariance and monotonicity of the modulus, we have:

$$\operatorname{mod}(I':L') \geq \operatorname{mod} \phi(A(I,L)) = \operatorname{mod} A(I,L) \equiv \operatorname{mod}(I:L),$$

and the conclusion follows from the lemma. □

EXERCISE 6.19. *Derive the above Corollary from the Koebe Distortion Theorem.*

6.3.6. *Shrinking nests of annuli.* For a set $X \subset \mathbb{C}$, let us say that a sequence of disjoint annuli $A_n \subset \mathbb{C}$ is *nested around* X if for any any n, A_n separates both A_{n+1} and X from ∞. (We will also call it a "nest of annuli around X".)

COROLLARY 6.20. *Consider a nest of annuli A_n around X. If $\sum \operatorname{mod} A_n = \infty$ then X is a single point.*

PROOF. Only the first annulus, A_1, can be unbounded in $\mathbb{C}$. Take some disk $D = \mathbb{D}_R$ containing A_2, and consider the annulus $D \smallsetminus X$. By the Gröztsch Inequality,

$$\operatorname{mod}(D \smallsetminus X) \geq \sum_{n \geq 2} \operatorname{mod} A_n = \infty.$$

Hence X is a single point. □

6.3.7. *Quadrilaterals* (*rectangles*). This discussion is parallel to the above discussion of annuli, so we will be brief. Let us consider a standard marked rectangle $\Pi = \Pi[l,h] = [0,l] \times [0,h]$ (see §2.6.2). As in the case of an annulus, we can naturally define (*topologically*) *vertical and horizontal* paths in Π, as well *genuinely* vertical and horizontal ones. The latter form the vertical and horizontal *foliations*.

EXERCISE 6.21. (i) *Let Γ be a vertical path family in $\Pi[l,h]$ that contains almost all genuinely vertical paths. Then $\mathcal{L}(\Gamma) = \operatorname{mod}(\Pi)$.*

(ii) *More generally, let Γ be a genuinely vertical lamination in $\Pi[l,h]$ (i.e., a family of genuinely vertical paths) supported on a measurable set Λ. Let κ be the horizontal projection of Λ. Then $\mathcal{L}(\Gamma) = h/\kappa$.*

A *quadrilateral* or a *conformal rectangle* $\mathrm{Q}(a,b,c,d)$ is a conformal disk Q with four marked points a,b,c,d on its ideal boundary. We will often let $Q = Q(a,b,c,d)$ so that there is no notational difference between the quadrilateral and the underlying disk. (If the underlying disk is called, say, D then the corresponding quadrilateral is denoted accordingly, $D = D(a,b,c,d)$.)

A quadrilateral has four ideal boundary sides. As in the case of a rectangle *marking* of a quadrilateral is a choice of pair of opposite sides called *horizontal*, while the other pair is called *vertical*.

As an important example, let us consider the quadrilateral

$$\mathbb{H}(a,0,1) \equiv \mathbb{H}(a,0,1,\infty), \quad a<0,$$

based on the upper half-plane $\mathbb{H}$, marked so that $[1,\infty]$ is a horizontal side.

PROPOSITION 6.22. *Any marked quadrilateral Q is conformally equivalent to a unique (up to scaling) standard marked rectangle.*

PROOF. By the Riemann Mapping Theorem, Q can be conformally mapped (as a marked quadrilateral) onto a marked quadrilateral $\mathbb{H}(a,0,1,\infty)$ with some $a<0$. By Exercise 2.77, this quadrilateral is conformally equivalent to a standard rectangle.

Uniqueness follows from Exercise 2.76. □

At this point, we can define various conformal notions and objects ($\operatorname{mod} Q$, genuinely vertical foliation, etc) for any marked quadrilateral Q by transferring them from a standard rectangle Π conformally equivalent to Q. Assertions of Exercises 2.76 and 6.21 immediately extend to general marked quadrilaterals.

As in the annulus case, the length-area method allows one to relate conformal and Euclidean quantities:

EXERCISE 6.23. *Show that for $R>1$*

$$\frac{1}{4\pi}\log R \le \operatorname{mod}(\mathbb{H}(0,1,R)) \le -\frac{4\pi}{\log(1-1/R)}.$$

(*Here the left-hand estimate is good for big R, while the right-hand one is good for $R\approx 1$.*)

6.3.8. *Tori.* Let us now consider a flat torus $\mathbb{T}^2$. Given a non-zero homology class $\alpha\in H_1(\mathbb{T}^2)$, we let Γ_α be the family of closed curves on $\mathbb{T}^2$ representing α (we call them α-*curves*). Among these curves, there are closed geodesics, α-*geodesics* (they lift to straight lines in the universal covering $\mathbb{R}^2$) . They form a foliation. All these geodesics have the same length, l_α.

EXERCISE 6.24. *Let Γ be a family of α-curves containing all α-geodesics. Then*

$$\mathcal{W}(\Gamma) = \frac{\operatorname{area}\mathbb{T}^2}{l_\alpha^2}.$$

An annulus A embedded into $\mathbb{T}^2$ is called an α-*annulus* if its horizontal curves represent the class α. (In this case, we also say that A *itself represents the class* α.) The following estimate finds interesting applications in dynamics and geometry (see §24.6):

PROPOSITION 6.25. *Let $A_1,\dots,A_n$ be a family of disjoint α-annuli. Then*

$$\sum \operatorname{mod} A_i \le \frac{\operatorname{area}\mathbb{T}^2}{l_\alpha^2}.$$

PROOF. Let Γ_i be the family of horizontal curves of the annulus A_i. Then by the Parallel Law, $\sum\mathcal{W}(\Gamma_i)\le\mathcal{W}(\Gamma_\alpha)$, and the result follows from Exercises 6.8 and 6.24. □

6.4. Dirichlet integral.

6.4.1. *Definition.* Consider a Riemann surface S endowed with a smooth conformal metric ρ. The *Dirichlet integral (D.I.)* of a function $\chi : S \to \mathbb{C}$ is defined as

$$D(\chi) = \int \|\nabla\chi\|_\rho \, dm_\rho,$$

where the norm of the gradient and the area form are evaluated with respect to ρ. However:

EXERCISE 6.26. *The Dirichlet integral is independent of the choice of the conformal metric ρ. In particular, it is invariant under conformal changes of variable.*

In the local coordinates, the Dirichlet integral is expressed as follows:

$$D(h) = \int (|h_x|^2 + |h_y|^2) dm = \int (|\partial h|^2 + |\bar{\partial} h|^2) dm.$$

In particular, for a conformal map $h : U \hookrightarrow \mathbb{C}$ we have the area formula:

$$D(h) = \int |h'(z)|^2 dm = \operatorname{area} h(U).$$

6.4.2. *D.I. of a harmonic function.*

EXERCISE 6.27. *Consider a flat cylinder $A = S^1 \times (0, h)$ with the unit circumference. Let $\chi : A \to (0, 1)$ be the projection to the second coordinate (the "height" function) divided by h. Then $D(\chi) = 1/h$.*

Note that the function χ in the exercise is a harmonic function with boundary values 0 and 1 on the boundary components of the cylinder (i.e., the solution of the Dirichlet problem with such boundary values).

EXERCISE 6.28. *Such a harmonic function is unique up to switching the boundary components of A, which leads to replacement of χ by $1 - \chi$.*

Due to the conformal invariance of the Dirichlet integral (as well as the modulus of an annulus and harmonicity of a function), these trivial remarks immediately yield a non-trivial formula:

PROPOSITION 6.29. *Let us consider a conformal annulus A. Then there exist exactly two proper harmonic functions $\chi_i : A \to (0, 1)$ (such that $\chi_1 + \chi_2 = 1$) and $D(\chi_i) = 1/\operatorname{mod}(A)$.*

6.4.3. *Multi-connected case.* Let S be a compact Riemann surface with boundary. Let $\partial S = (\partial S)_0 \sqcup (\partial S)_1$, where each $(\partial S)_i \neq \emptyset$ is the union of several boundary components of ∂S. Let us consider two families of curves: the "vertical family" Γ^v consisting of arcs joining $(\partial S)_0$ to $(\partial S)_1$, and the "horizontal family" Γ^h consisting of Jordan multi-curves separating $(\partial S)_0$ from $(\partial S)_1$. (A *multicurve* is a finite union of Jordan curves.)

Let $\chi : S \to [0, 1]$ be the solution of the Dirichlet problem equal to 0 on $(\partial S)_0$ and equal to 1 on $(\partial S)_1$.

THEOREM 6.30.

$$\mathcal{L}(\Gamma^v) = \mathcal{W}(\Gamma^h) = \frac{1}{D(h)}.$$

The *modulus of S rel the boundaries* $(\partial S)_0$ and $(\partial S)_1$ is defined as the above extremal length:

$$\operatorname{mod}((\partial S)_0, (\partial S)_1) = \mathcal{L}(\Gamma^v).$$

Remark. Physically, we can think of the pair $(\partial S)_0$ and $(\partial S)_1$ in S as an electric condensator. The harmonic function χ represents the potential of the electric field created by the uniformly distributed charge on $(\partial S)_1$. The Dirichlet integral $D(\chi)$ is the *energy* of this field. Thus, $\operatorname{mod}((\partial S)_0, (\partial S)_1) = 1/D(\chi)$ is equal to the ratio of the charge to the energy, that is, to the *capacity* of the condensator.

6.5. Non-Crossing Principle. Let us say that two path families Γ and Δ *cross* if every path of Γ crosses every path of Δ. The Non-Crossing Principle asserts that two big path families do not cross:

Non-Crossing Principle. *Let Γ be a horizontal foliation in a quadrilateral or an annulus, and let Δ be another path family. If $\mathcal{W}(\Gamma)\cdot\mathcal{W}(\Delta) > 1$ then Γ and Δ do not cross.*

Proof. Let Π be the quadrilateral or the annulus supporting Γ, and let $\Gamma^\perp$ be the vertical path family in Π. If Δ crosses Γ then it overflows $\Gamma^\perp$, and hence

$$\mathcal{L}(\Delta) \geq \mathcal{L}(\Gamma^\perp) = \mathcal{W}(\Gamma).$$

□

Exercise 6.31. *Assume Δ is also a horizontal foliation in a quadrilateral or an annulus. If $\mathcal{W}(\Gamma)\cdot\mathcal{W}(\Delta) = 1$, then Γ and Δ do not cross unless they are supported on the same quadrilateral and Δ is equal to $\Gamma^\perp$.*

The Non-Crossing Principle can be sharpened to an assertion that *two wide path families have a relatively small overlap* (which will be used only in vol. III).

Let us consider a genuinely vertical lamination in a quadrilateral Q (see Exercise 6.21). After uniformizing it by a standard rectangle $\Pi = \Pi[l, h]$, its projection to the horizontal side induces a transverse measure ν on Λ (defined up to scaling). If Q is embedded into a Riemann surface S and γ is a path on S, we say that γ *intersects less than ε-portion of the total width of* Λ if

$$\nu\{\lambda \in \Lambda : \lambda \cap \gamma \neq \emptyset\} < \varepsilon\nu(\Lambda)$$

(note that this condition does not depend on the normalization of ν). The same discussion applies to the case of annulus.

Small Overlapping Principle. *Let $\kappa \geq 1$. Let us consider a genuinely vertical lamination Λ on some conformal annulus or quadrilateral $Q \subset S$, and let Γ be another path family on S. If $\mathcal{W}(\Lambda) > \kappa$ and $\mathcal{W}(\Gamma) \geq \kappa$, then there exists a path $\gamma \in \Gamma$ that intersects less than $1/\kappa$-portion of the total width of Λ.*

Proof. Assume for definiteness that Q is a quadrilateral. Let $\phi : \Pi[a, h] \to R$ be the uniformization of Q by a standard rectangle normalized so that the horizontal projection of $\phi^*\Lambda$ (which is a genuinely vertical lamination in Π) has length κ. By Exercise 6.21,

$$\mathcal{W}(\Lambda) = \mathcal{W}(\phi^*(\Lambda)) = \frac{\kappa}{h}.$$

Since $\mathcal{W}(\Lambda) > \kappa$, we conclude that $h < 1$, and thus

$$\operatorname{area}(\phi^*\Lambda) = h \cdot \mathcal{W}(\phi^*(\Lambda)) < \kappa.$$

To bound $\mathcal{W}(\Gamma)$, let us the push-forward the Euclidean metric e on Π to the quadrilateral Q, i.e., let $\rho = \phi_*(e\,|\Lambda)$. If a curve $\gamma \in$
$Gamma$ intersects at least $1/\kappa$-portion of the total width of Λ, then the transverse length of γ is at least 1, and hence

$$l_\rho(\gamma) = \nu(\gamma) \geq 1.$$

If this happened for every $\gamma \in \Gamma$ then we would have

$$\mathcal{W}(\Gamma) \leq \text{area}_\rho(\Lambda) = \text{area}(\phi^*\Lambda) < \kappa,$$

contradicting the assumption. □

6.6. Transformation rules.

6.6.1. *General rules.* As we know, both extremal length and extremal width are conformal invariants. More generally, we have:

LEMMA 6.32. *Let $f : U \to V$ be a holomorphic map between two Riemann surfaces, and let $\mathcal{G}$ be a family of curves on U. Then*

$$\mathcal{L}(f(\Gamma)) \geq \mathcal{L}(\Gamma).$$

Moreover, if f is at most $d - to - 1$, then

$$\mathcal{L}(f(\Gamma)) \leq d \cdot \mathcal{L}(\Gamma).$$

PROOF. Let ρ be a conformal metric on U. Let us push-forward the area form m_ρ by f. We obtain the area form $m_\tau = f_*(m_\rho)$ of some conformal metric τ on V. Then $\text{area}_\tau(V) = \text{area}_\rho(U)$ and $f^*(\tau) \geq \rho$. It follows that

$$\mathcal{L}_\rho(\Gamma) \leq \mathcal{L}_\tau(f(\Gamma)) \leq \mathcal{L}(f(\Gamma)).$$

Taking the supremum over ρ completes the proof of the first assertion.

For the second assertion, let us consider a conformal metric τ on V and pull it back to U, $\rho = f^*\tau$. Then $l_\rho(\gamma) = l_\tau(f(\gamma))$ for any $\gamma \in \Gamma$, while $m_\rho(U) \leq d \cdot \text{area}_\tau(V)$. Hence

$$\mathcal{L}(\Gamma) \geq \mathcal{L}_\rho(\Gamma) \geq \frac{1}{d}\mathcal{L}_\tau(f(\Gamma)),$$

and taking the supremum over τ completes the proof. □

COROLLARY 6.33. *Under the circumstances of the previous lemma, let Δ be a family of curves in V satisfying the following lifting property: any curve $\gamma \in \Delta$ contains an arc that lifts to some curve in Γ. Then $\mathcal{L}(\Delta) \geq \mathcal{L}(\Gamma)$.*

PROOF. The lifting property means that the family Δ overflows the family $f(\Gamma)$. Hence $\mathcal{L}(\Delta) \geq \mathcal{L}(f(\Gamma))$, and the conclusion follows. □

6.6.2. *Coverings of an annulus.* Let us start with a particular case which is most important for dynamical applications.

PROPOSITION 6.34. *Let U and U' be two conformal disks, and let $f : U \to U'$ be a holomorphic branched covering of degree D. Let $B' \Subset U'$ be a Jordan disk and let $B \Subset U$ be a component of $f^{-1}(B')$. Let $d = \deg(f : B \to B')$. Then*

$$d \cdot \mathcal{W}(U' \smallsetminus B') \leq \mathcal{W}(U \smallsetminus B) \leq D \cdot \mathcal{W}(U' \smallsetminus B')$$

Proof. Let Γ and Γ' be the vertical path families on the annuli $U \smallsetminus B$ and $U' \smallsetminus B'$ respectively. Take an arbitrary $\mathcal{W}$-admissible (in the sense of Exercise 6.3) conformal metric on $U' \smallsetminus B'$, so $l_{\rho'}(\gamma') \geq 1$ for any $\gamma' \in \Gamma'$. Let $\rho = f^*(\rho')$ be its pullback to $U \smallsetminus B$.

Take any path $\gamma \in \Gamma$ and orient it from the outer boundary. The intersection $\gamma \cap f^{-1}(\overline{B}')$ is closed in γ, so we can take the first intersection point b. Let $\gamma_0 \subset \gamma$ be the initial piece of γ that ends at b. Then the image $\gamma_0' := f(\gamma_0)$ begins on $\partial U'$, ends at $f(b) \in B'$, and except for the endpoint, is contained in $U' \smallsetminus \overline{B}'$. Thus, $\gamma_0' \in \Gamma'$ and γ_0 is a lift of γ. Hence

$$l_\rho(\gamma) \geq l_\rho(\gamma_0) = l_{\rho'}(\gamma_0') \geq 1.$$

Thus, metric ρ is $\mathcal{W}$-admissible for Γ.

By the definition of extremal width given in Exercise 6.3,

$$\mathcal{W}(U \smallsetminus B) = \mathcal{W}(\Gamma) \leq \text{area}_\rho(U \smallsetminus B) = D \cdot \text{area}_{\rho'}(U' \smallsetminus B').$$

Taking the infimum over all $\mathcal{W}$-admissible ρ' we obtain the desired right-hand side inequality.

To prove the left-hand side inequality, let us consider the genuinely vertical foliation $\mathcal{F}'$ on the annulus $U' \smallsetminus B'$. This time, let us orient it from B' to $\partial U'$. Let us cut $U' \smallsetminus B'$ along the critical leaves of $\mathcal{F}'$, i.e., the leaves passing through the critical values of f. If there are no such leaves, let us cut $U' \smallsetminus B'$ along one leaf of $\mathcal{F}$.

We obtain a tiling of $U' \smallsetminus B'$ by rectangles Π_i' such that

$$\sum \mathcal{W}(\Pi_i') = \mathcal{W}(U' \smallsetminus B').$$

Each Π_i' lifts to d disjoint rectangles Π_{ij} in $U \smallsetminus B$ (with horizontal sides on ∂B and ∂U) each of which is conformally equivalent to Π_i'. By Monotonicity of the width, the Parallel Law, and conformal invriance of the width, we obtain:

$$\mathcal{W}(U \smallsetminus B) \geq \mathcal{W}\left(\bigcup \Pi_{ij}\right) = \sum \mathcal{W}(\Pi_{ij}) = d \cdot \sum \mathcal{W}(\Pi_i') = d \cdot \mathcal{W}(U' \smallsetminus B').$$

$\square$

Exercise 6.35. *Generalize* Proposition 6.34 *to the case when B and B' are finite unions of Jordan disks and the restriction $f : B \to B'$ is a branched covering of degree d. Conclude that if $f : U \smallsetminus B \to U' \smallsetminus B'$ is a covering of degree d then*

$$\text{mod}(U' \smallsetminus B') = d \,\text{mod}(U \smallsetminus B),$$

where $\text{mod}(U \smallsetminus B)$ *stands for the extremal length of the path family connecting ∂B to ∂U in $U \smallsetminus B$.*

6.7. Maximal, canonical, and covering annuli.

Lemma 6.36. *Let $S \subset \hat{\mathbb{C}}$ be a domain on the Riemann sphere. Then in the homotopy class of any non-trivial simple closed curve $\gamma \subset S$, there exists an embedded open annulus $\mathbf{A} \equiv \mathbf{A}_\gamma$ of maximal modulus. This modulus is infinite if and only if γ is a peripheral curve around a puncture (i.e., an isolated point of $\hat{\mathbb{C}} \smallsetminus S$).*

PROOF. The statement is vacuous for $\hat{\mathbb{C}}$ and $\mathbb{C}$, and obvious for $\mathbb{C}^*$. So, we can assume that $|\hat{\mathbb{C}} \setminus S| \geq 3$, and thus S is hyperbolic. Let us consider the family Φ of all conformal embeddings

$$\phi : \mathbb{A}(1, r) \to S,$$

where the radius $r > 1$ is variable. Let $r_{\max}$ be the sup of all possible radii. Select a monotonic sequence of radii $r_n \in (1, r_{\max})$ converging to $r_{\max}$, and corresponding functions $\phi_n \in \Phi$, $\phi_n : \mathbb{A}(1, r_n) \to S$. Since S is hyperbolic, Montel's Theorem implies that for any m, the restricted sequence $\phi_n | \mathbb{A}(1, r_m)$, $n = m, m+1, \dots$ admits a convergent subsequence. By the diagonal procedure, we can select a subsequence $\phi_{n(k)} \in \Phi$ that converges on each annulus A_m. Its limit provides us with an extremal embedding $\mathbb{A}(1, r_{\max}) \to S$ whose image $\mathbf{A}_\gamma$ is a desired annulus of maximal modulus,

$$\operatorname{mod} \mathbf{A}_\gamma = \frac{1}{2\pi} \log r_{\max},$$

in the given homotopy class.

If γ is a peripheral curve around a puncture, then $\operatorname{mod} \mathbf{A}_\gamma = \infty$ since $\operatorname{mod} \mathbb{D}^* = \infty$. The inverse statement follows e.g., from Exercise 6.13. $\square$

We call such annuli $\mathbf{A}_\gamma$ *maximal*, and we let $\operatorname{mod}[\gamma] \equiv \operatorname{mod} \mathbf{A}_\gamma$ be their moduli. Of course, the above discussion applies to bordered Riemann surfaces S as well. In this case, we allow γ to be a component of ∂S, letting $\operatorname{mod}[\gamma]$ be the modulus of the corresponding maximal peripheral annulus.

EXERCISE 6.37. *Generalize the above result to an arbitrary Riemann surface S (including tori).*

THEOREM 6.38. *Let S be a domain in $\hat{\mathbb{C}}$. Then in each non-trivial homotopy class $[\gamma]$ of simple closed curves there is a unique maximal annulus $\mathbf{A}_\gamma$.*

The uniqueness part of this statement is much deeper than the existence result proved above, and its proof provides us with a beautiful insight into the geometry of the maximal annuli. Namely, these annuli are horizontal annuli of so called *Strebel quadratic differentials*, i.e, quadratic differentials whose non-singular horizontal trajectories are all circles. Any maximal annulus is obtained by cutting the sphere along the separatrices of such a differential (see [**GaL**, §11]). We will leave this picture without a proof and in our future discussion will refrain from using it, but will keep it in mind as a good intuition.

Fix now some $M > 1$. Let us consider a maximal annulus $\mathbf{A}$ in S with $\operatorname{mod} \mathbf{A} > 2M$ that does not represent a puncture (but can represent a boundary curve). Let us uniformize it by a Euclidean cylinder $C = S^1 \times (0, h)$, where S^1 has length 1 (so $h = \operatorname{mod} A > 2M$). Round cylinders $S^1 \times (0, M]$ and $S^1 \times [h - M, h)$ are called *buffers* in C. Note that they are disjoint since $h > 2M$. *Buffers* B^o and B^i in $\mathbf{A}$ are the corresponding annuli in $\mathbf{A}$ (where labels "o" and "i" stand for the "outer" and "inner" respectively[21]). Removing the buffers, we obtain an M*-canonical annulus*

$$A = \mathbf{A} \setminus (B^o \cup B^i).$$

LEMMA 6.39. *For any $M > 1$, any two M-canonical annuli are disjoint.*

[21] On the Riemann sphere, this labeling is arbitrary, but it makes a clear sense when $\mathbf{A} \subset \mathbb{C}$.

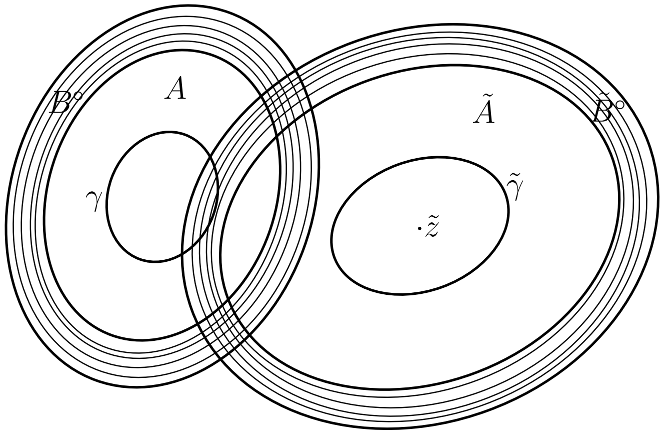

FIGURE 6.1. Intersection of the annuli forces the buffers cross.

PROOF. Let γ, $\tilde{\gamma}$ be two non-trivial Jordan curves in S, and let $A \subset \mathbf{A}$, $\tilde{A} \subset \tilde{\mathbf{A}}$ be the corresponding canonical and maximal annuli. If γ and $\tilde{\gamma}$ essentially cross then any horizontal curve in $\mathbf{A}$ crosses any horizontal curve in $\tilde{\mathbf{A}}$ contradicting the Non-Crossing Principle.

So, the curves γ and $\tilde{\gamma}$ are essentially disjoint. Replacing them with homotopic ones, we can assume that γ and $\tilde{\gamma}$ are disjoint in the first place. Then one of the Jordan disks in $\hat{\mathbb{C}}$ bounded by γ contains $\tilde{\gamma}$. Let us call it the *outer* side of γ. Similarly we can define the outer side of $\tilde{\gamma}$. Since the curves are not homotopic, the intersection of their outer complementary components contains a point of $\hat{\mathbb{C}} \smallsetminus S$. This point can be placed at ∞, making the inner complementary components bounded.

Since our curves are non-trivial, we can select a point $\tilde{z} \in \hat{\mathbb{C}} \smallsetminus S$ lying "inside" $\tilde{\gamma}$ (i.e., in the inner complementary component of $\tilde{\gamma}$) and hence "outside" γ.

Let the buffer B^o and $\tilde{B}^o$ lie on the outer sides of A and $\tilde{A}$, respectively. If $A \cap \tilde{A} \neq \emptyset$ then any horizontal curve $\tilde{\delta}$ in $\tilde{B}^o$ is forced to enter the annulus A (see Figure 6.1). At the same time, $\tilde{\delta}$ must "go around" $\tilde{z}$ forcing it to go outside $\mathbf{A}$. It follows that $\tilde{\delta}$ must cross the whole buffer B^o.

We conclude that the horizontal path families in the buffers B^o and $\tilde{B}^o$ cross each other. Since both have width > 1, we arrive at a contradiction with the Non-Crossing Principle. ☐

Thus, we obtain the *canonical multicurve* on S comprising the equators of all the canonical annuli. The corresponding homotopy classes are also called canonical.

Along with the maximal annuli $\mathbf{A}_\gamma$, we can consider the covering annuli $\mathbb{A}_\gamma$ from §1.7.13.

LEMMA 6.40. *For any domain S in $\hat{\mathbb{C}}$ and any non-trivial Jordan curve $\gamma \subset S$, we have:* $\operatorname{mod} \mathbb{A}_\gamma - 2 < \operatorname{mod} \mathbf{A}_\gamma < \operatorname{mod} \mathbb{A}_\gamma$.

PROOF. Under the covering $q : \mathbb{A}_\gamma \to S$, the fundamental group $\pi_1(\mathbb{A}_\gamma)$ projects to the cyclic group Γ generated by $[\gamma]$. Since $\Gamma = \pi_1(\mathbf{A}_\gamma)$, the annulus $\mathbf{A}_\gamma$ lifts to an annulus $\hat{\mathbf{A}}_\gamma \subset \mathbb{A}_\gamma$ that conformally projects onto $\mathbf{A}_\gamma$. The upper estimate for $\operatorname{mod} \mathbf{A}_\gamma$ follows.

The lower estimate follows from Lemma 7.15 as mod $\mathbf{A}_\gamma$ is bounded from below by the modulus of the geometric collar $\mathcal{N}_\eta(\gamma)$, while the latter is at least $\pi/l - 2 = \mathrm{mod}\,\mathbb{A}_\gamma - 2$. □

6.8. Canonical weighted arc diagram. Let S be a hyperbolic Riemann surface of finite topological type, and let $\hat{S}^I$ be its ideal compactification (see §2.4.17). We assume that $\partial^I S \neq \emptyset$. Let α be a non-trivial proper arc on S landing on boundary circles of $\hat{S}^I$ (perhaps, on the same one). To the homotopy class of α we will now associate a weight $\mathcal{W}(\alpha) \geq 0$ as follows.[22]

Let $\pi : \mathbb{D} \to S$ be the universal covering of S, let $\Gamma \approx \pi_1(S)$ be the Fuchsian group of deck transformations acting on $\mathbb{D}$, and let $\Lambda \subset \mathbb{T}$ be its limit set. It is a Cantor set, and the covering π extends continuously to a covering $\pi : \mathbb{T} \smallsetminus \Lambda \to \partial^I S$ (keeping the same notation). So, for any complementary interval ("gap") $\hat{J} \subset \mathbb{T} \smallsetminus \Lambda$, the projection $\pi|\,\hat{J}$ is the universal covering over some component J of the ideal boundary. Moreover, if $\gamma \in \Gamma$ is a deck transformation corresponding to the loop J, then γ is a hyperbolic Möbius map keeping invariant one of the gaps $\hat{J}$ over J, and the boundary of $\hat{J}$ consists of the fixed points of γ.

Let J and J' be the boundary components of S connected by the arc α. Then α lifts to an arc $\hat{\alpha}$ on $\mathbb{D}$ connecting some gaps $\hat{J}$ and $\hat{J}'$ that cover J and J' respectively. If α is non-trivial then the intervals $\mathrm{cl}\,\hat{J}$ and $\mathrm{cl}\,\hat{J}'$ are disjoint, and hence they can be viewed as the horizontal sides of a quadrilateral $\mathbf{\Pi} \equiv \mathbf{\Pi}(\hat{\alpha})$ supported on $\overline{\mathbb{D}}$. Define

$$\overline{\mathcal{W}}(\alpha) \equiv \hat{\mathcal{W}}(J, J') := \mathcal{W}(\mathbf{\Pi})$$

as the width of this quadrilateral. In other words, let us uniformize $\mathbf{\Pi}$ by s standard rectangle $\mathbf{P} = [0, a] \times [0, 1]$ so that J and J' correspond to the horizontal sides of $\mathbf{P}$. Then $\overline{\mathcal{W}}(\alpha) = a$.

If $\overline{\mathcal{W}}(\alpha) > 2$, let us define the *square buffers* of $\mathbf{\Pi}$ as the quadrilateral corresponding to the lateral squares $[0, 1) \times [0, 1]$ and $(a-1, a] \times [0, 1]$ in $\mathbf{P}$. Removing the square buffers from $\mathbf{\Pi}$ we obtain a quadrilateral $\hat{\Pi} = \Pi(\hat{\alpha})$. The *weight* of α is defined as the width of $\hat{\Pi}$:

$$\mathcal{W}(\alpha) := \mathcal{W}(\hat{\Pi}) = \overline{\mathcal{W}}(\alpha) - 2 = a - 2.$$

Note that this width is independent of the choice of the lift $\hat{\alpha}$ since the corresponding quadrilaterals are related by Möbius transformations of $\mathbb{D}$.

In case when $\overline{\mathcal{W}}(\alpha) \leq 2$, we let $\mathcal{W}(\alpha) = 0$.

The family of arcs α with $\mathcal{W}(\alpha) > 0$ is called the *canonical arc diagram* $\mathcal{A} = \mathcal{A}(S)$ of S.

LEMMA 6.41. *Quadrilaterals $\hat{\Pi}$, $\hat{\Pi}'$ corresponding to different lifts $\hat{\alpha}$, $\hat{\alpha}'$ of an arc $\alpha \in \mathcal{A}$ are disjoint.*

PROOF. Otherwise these rectangles would have intersecting vertical sides, $L \subset \hat{\Pi}$ and $L' \subset \hat{\Pi}'$, $L \cap L' \neq \emptyset$. But then their buffers B and B' attached to these sides would *cross* each other in the sense the corresponding vertical path families cross. Since the buffers have width 1, this would contradict to the Non-Crossing Principle (accompanied with Exercise 6.31) . □

[22]We will not notationally distinguish an arc and its homotopy class.

REMARK 6.42. In fact, we could define arcs allowing self-intersections. What the above argument shows is that the canonical arcs automatically avoid ones.

COROLLARY 6.43. *The projection* $\pi : \hat{\Pi} \to S$ *is an embedding.*

Thus, the projection $\Pi = \Pi(\alpha) := \pi(\hat{\Pi}(\hat{\alpha}))$ is an embedded rectangle in S (obviously independent on the choice of the lift $\hat{\alpha}$). We call it the *canonical rectangle* corresponding to the arc α.

LEMMA 6.44. *Any two canonical rectangles,* $\Pi(\alpha)$ *and* $\Pi(\beta)$*, are disjoint.*

PROOF. Otherwise, some of their lifts, $\hat{\Pi}(\hat{\alpha})$ and $\hat{\Pi}(\hat{\beta})$, would intersect. But this can be ruled out by the same argument as in Lemma 6.41. $\square$

Together with Proposition 1.94, this implies:

COROLLARY 6.45. *The canonical arc diagram* $\mathcal{A}_S$ *contains at most* $-3\chi(S)$ *arcs. In particular, for a disk with* n *holes (which is the only case needed for the dynamical applications) we obtain at most* $3(n-1)$ *arcs.*

Thus, we have at most $-3\chi(S)$ disjoint canonical rectangles $\Pi(\alpha)$, $\alpha \in \mathcal{A}$, on S. Putting together the vertical foliations on these rectangles, we obtain the *canonical foliation* on S.

7. Hyperbolic metric and Schwarz Lemma

7.1. Schwarz Lemma. In terms of the hyperbolic metric, the elementary Schwarz Lemma can be brought to a conformally invariant form that plays an outstanding role in holomorphic dynamics:

SCHWARZ LEMMA. *Let* $\phi : S \to S'$ *be a holomorphic map between two hyperbolic Riemann surfaces. Then*
• *either* ϕ *is a strict contraction, i.e.,* $\|D\phi(z)\| < 1$ *for any* $z \in S$*, where the norm of the differential is evaluated with respect to the hyperbolic metrics of* S *and* S'*;*
• *or else,* ϕ *is a covering map, and then it is a local isometry:* $\|D\phi(z)\| = 1$ *for any* $z \in S$.

PROOF. Given a point $z \in S$, let $\pi : (\mathbb{D}, 0) \to (S, z)$ and $\pi' : (\mathbb{D}, 0) \to (S', \phi(z))$ be the universal coverings of the Riemann surfaces S and S' respectively. Then ϕ can be lifted to a holomorphic map $\tilde{\phi} : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$. By the elementary Schwarz Lemma, $|\tilde{\phi}'(0)| < 1$ or else $\tilde{\phi}$ is a conformal automorphism of $\mathbb{D}$ (in fact, rotation). This yields the desired dichotomy for ϕ. $\square$

In particular, if $S \subset S'$ then $\rho_S \geq \rho_{S'}$ (a smaller Riemann surface is "more hyperbolic"). Moreover, if $S \neq S'$ then $d\rho_S(z) > d\rho_{S'}(z)$ for any $z \in S$.

COROLLARY 7.1. *Let* $S \subset S'$ *be a nest of two hyperbolic Riemann surfaces,* $S \neq S'$*, and let* $f : S \to \Sigma$ *be an (unramified) covering map to a Riemann surface* Σ*. Then for any* $z \in S$ *we have*

$$\|Df(z)\|_{S',\Sigma} > 1,$$

where the norm of $Df(z)$ *is evaluated from the hyperbolic metric of* S' *to that of* Σ.

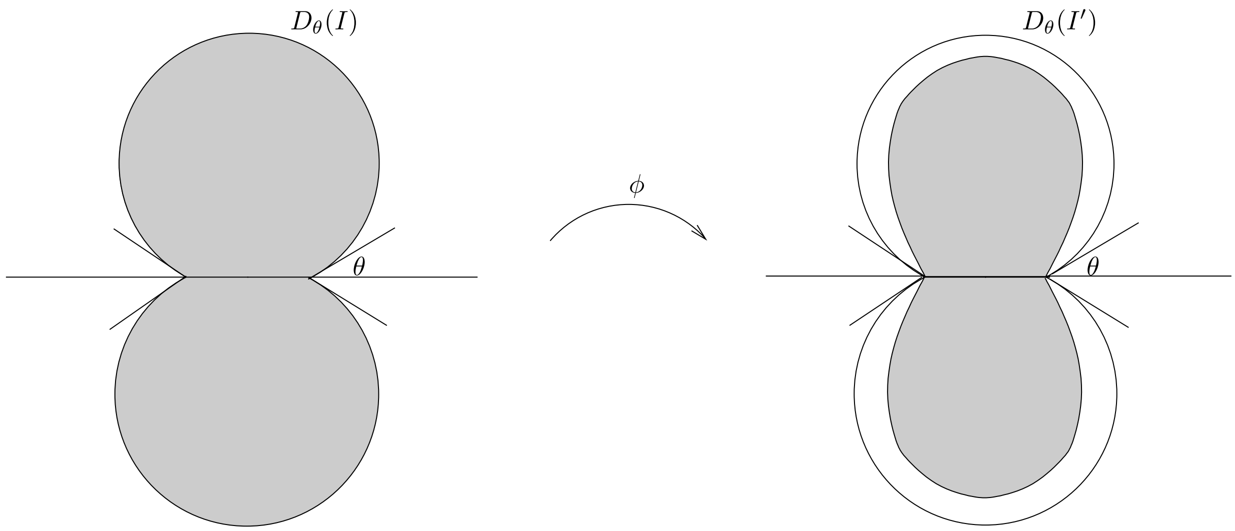

FIGURE 7.1. Symmetric Schwarz Lemma.

PROOF. Since f is a local isometry from the hyperbolic metric of S to that of Σ, we have

$$\|Df(z)\|_{S',\Sigma} = \frac{d\rho_S}{d\rho_{S'}}(z), \tag{7.1}$$

and the desired estimate follows from the remark preceding this Corollary. □

7.2. Symmetric Schwarz Lemma.

7.2.1. *Formulation.* Let us formulate an $\mathbb{R}$-symmetric version of the Schwarz Lemma (with the notation introduced in §2.4.5): Let $\mathfrak{U}_\mathbb{R}$ be the class of $\mathbb{R}$-symmetric univalent maps $\phi : \mathbb{C}(I) \to \mathbb{C}(I')$ between slit planes that restrict to diffeomorphisms $I \to I'$ between open intervals $I, I' \subset \mathbb{R}$.

SYMMETRIC SCHWARZ LEMMA. *Let $\phi : \mathbb{C}(I) \to \mathbb{C}(I')$ be a map of class $\mathfrak{U}$. Then for any $\theta \in (0, \pi)$, we have $\phi(\mathbb{D}_\theta(I)) \subset \mathbb{D}_\theta(I')$.*

PROOF. Follows from the Schwarz Lemma and Exercise 2.23 (iii) □

Any open interval $I = (a, b) \subset \mathbb{R}$ can be considered as a model of the *hyperbolic line* endowed with the hyperbolic metric

$$ds = \frac{2(b-a)\,dx}{(x-a)(b-x)}.$$

This metric can be also viewed as induced from the disk $\mathbb{D}(I) \approx \mathbb{H}^2$.

COROLLARY 7.2. *Let $\phi : \mathbb{C}(I) \to \mathbb{C}(I')$ be a univalent map of class $\mathfrak{U}_\mathbb{R}$. If ϕ is not Möbius then it strictly contracts the hyperbolic metric:*

$$\|D\phi(x)\|_{\mathrm{hyp}} < 1 \quad \forall\, x \in I. \tag{7.2}$$

If $L \Subset \operatorname{int} I$ is a smaller interval with $\operatorname{mod}_{\mathbb{R}}(I : L) \geq \delta > 0$ *then the hyperbolic norm in* (7.2) *is bounded by $\rho(\delta) < 1$, and the map ϕ has a bounded distortion on L:*

$$\left| \frac{D\phi(x)}{D\phi(y)} \right| \leq C(\delta) \quad \forall\, x, y \in L.$$

Proof. By the Symmetric Schwarz Lemma, $\phi(\mathbb{D}(I)) \subset \mathbb{D}(I')$. Contracting property (7.2) follows from the Schwarz Lemma (for the maps $\mathbb{D} \to \mathbb{D}$). Consequently, the hyperbolic length of $\phi(L)$ is bounded by that of L, which is bounded in terms of δ. Hence the contraction is uniform on L by the Definite Schwarz Lemma below (or just by a normality argument). Finally, the distortion bounds follow from the Koebe Theorem. □

7.2.2. *Lipschitz control.* Given an interval $I = (a, b)$ and a point $z \in \mathbb{C}(I)$, let

$$\operatorname{ang}(z, I) = \min\{|\arg(z - b)|,\ |\arg(a - z)|\},$$

where the argument is selected in the range $[0, \pi]$. In words, $\operatorname{ang}(z, I)$ is the smallest of the angles between the intervals $[a, z]$, $[b, z]$ and the corresponding real rays $(a, -\infty]$, $[b, +\infty)$ of the real line. We let

$$\mathbb{C}_\theta(I) := \{z \in \mathbb{C}(I) :\ \operatorname{ang}(z, I) > \theta\}. \tag{7.3}$$

Lemma 7.3. *Under the circumstances of the Symmetric Schwarz Lemma, let us consider a point $z \in \mathbb{C}_\theta(I)$ with* $\operatorname{dist}(z, I) \geq |I|$. *Then*

$$\frac{\operatorname{dist}(\phi(z), I')}{|I'|} \leq C \frac{\operatorname{dist}(z, I)}{|I|} \quad \text{with } C = C(\theta).$$

Proof. Let $\overline{\mathbb{D}}_\eta(I)$ be the smallest (closed) geodesic neighborhood of I containing z, and let $\mathbb{D}^+_\eta(I)$ be its upper half. Assume for definiteness that $z \in S := \partial\, \mathbb{D}^+_\eta(I)$ and $|z - b| \leq |z - a|$. Let $\Gamma \subset S$ be the circle arc connecting b to z; its angular size γ is less than π.

Since $|z - b| \geq |I|$, the angle between Γ and $z - b$ at b is at least $\theta/2$, so $\gamma \geq \theta$. Hence

$$\operatorname{dist}(z, I) = |z - b| \geq C^{-1} \operatorname{diam} \mathbb{D}_\eta(I), \quad \text{with } C = C(\theta).$$

By the Symmetric Schwarz Lemma, $\phi(z) \in \mathbb{D}_\theta(I'))$. Hence

$$\frac{\operatorname{dist}(\phi(z), I')}{|I'|} \leq \frac{\operatorname{diam}(\mathbb{D}_\eta(I'))}{|I'|} = \frac{\operatorname{diam}(\mathbb{D}_\eta(I))}{|I|} \leq C \frac{\operatorname{dist}(z, I)}{|I|},$$

as asserted. □

7.3. Hyperbolic metric blows up near the boundary. For a domain $U \subset \hat{\mathbb{C}}$, let $d(z)$ stand for the *spherical distance* from $z \in U$ to ∂U.

Lemma 7.4. *Let $\mathbf{S}$ be a Riemann surface, $x \in \mathbf{S}$, and assume that the punctured surface $S = \mathbf{S} \smallsetminus \{x\}$ is hyperbolic with the hyperbolic metric ρ. Then*

$$d\rho(z) \asymp -\frac{|dz|}{|z| \log |z|},$$

where z is a local coordinate on $\mathbf{S}$ with $z(x) = 0$.

Proof. By Proposition 2.53, a standard cusp $\mathbb{H}_h/\mathbb{Z}$ is isometrically embedded into S so that its puncture corresponds to x. On the other hand, by means of the exponential maps $\mathbb{H} \to \mathbb{D}^*$, $z \mapsto e^{2\pi i z}$, the cusp $\mathbb{H}_h/\mathbb{Z}$ is isometric to the punctured disk $\mathbb{D}_r^*$, $r = e^{-2\pi h}$, in the hyperbolic metric of $\mathbb{D}^*$. By Exercise 2.38, the latter has the desired form in the plane coordinate of $\mathbb{D}_r^*$ (which extends to a local coordinate on $\mathbf{S}$ near x). Hence it has the desired form in any other local coordinate on $\mathbf{S}$ near x. $\square$

Proposition 7.5. *For any hyperbolic plane domain $U \subset \hat{\mathbb{C}}$, there exists $\kappa = \kappa(U) > 0$ such that:*

$$\frac{d\rho_U}{d\rho_{\mathrm{sph}}}(z) \geq -\frac{\kappa}{d(z) \log d(z)}, \quad z \in U,$$

where σ is the spherical metric.

Proof. Take some point $z \in U$, and find the closest to it point $a \in \partial U$. Since ∂U consists of at least three points, we can find two more points, $b, c \in \partial U$, such that the points a, b, c are ε-separated on $\overline{\mathbb{C}}$, where $\varepsilon > 0$ depends only on U. Let us consider the Möbius transformation ϕ that moves (a, b, c) to $(0, 1, \infty)$. By Exercise 2.10, these transformations are uniformly bi-Lipschitz in the spherical metric, which reduces the problem to the case when $(a, b, c) = (0, 1, \infty)$. But in this case, $\rho_U(z)$ dominates the hyperbolic metric on $\mathbb{U} = \mathbb{C} \smallsetminus \{0, 1\}$, and the desired estimate follows from Lemma 7.4. $\square$

Exercise 7.6. *More generally, let $\mathbf{S}$ be a Riemann surface endowed with a conformal Riemannian metric σ, and let K be a compact subset of $\mathbf{S}$ such that $\mathbf{S} \smallsetminus K$ is a hyperbolic Riemann surface with hyperbolic metric ρ. Then there exists a $\kappa = \kappa(\mathbf{S}, K) > 0$ such that*

$$\frac{d\rho}{d\sigma}(z) \geq -\frac{\kappa}{d(z) \log d(z)}, \qquad z \in \mathbf{S} \smallsetminus K,$$

where $d(z) = \mathrm{dist}(z, K)$.

7.4. Hyperbolic metric on simply connected domains. For simply connected plane domains, the hyperbolic metric can be very well controlled:

Lemma 7.7. *Let $D \subset \mathbb{C}$ be a conformal disk endowed with the hyperbolic metric ρ_D. Then*

$$\frac{1}{4}\, \frac{|dz|}{\mathrm{dist}(z, \partial D)} \leq d\rho_D(z) \leq \frac{|dz|}{\mathrm{dist}(z, \partial D)}\,.$$

Remark. Of course, particular constants in the above estimates will not matter for us.

Proof. Let $r = \mathrm{dist}(z, \partial D)$; then $\mathbb{D}(z, r) \subset D$. Consider a linear map $h : \mathbb{D} \to \mathbb{D}(z, r)$ as a map from $\mathbb{D}$ into D. By the Schwarz Lemma, it contracts the hyperbolic metric. Hence

$$d\rho_D(z) \leq h_*(d\rho_{\mathbb{D}}(0)) = h_*(|d\zeta|) = |dz|/r.$$

To obtain the opposite inequality, consider the Riemann mapping $\psi : (\mathbb{D}, 0) \to (D, z)$. By definition of the hyperbolic metric,

$$d\rho_D(z) = \psi_*(d\rho_{\mathbb{D}}(0)) = \psi_*(|d\zeta|) = \frac{|dz|}{|\psi'(0)|}\,.$$

But by the Koebe 1/4-Theorem, $r \le |\psi'(0)|/4$, so that $d\rho_D(z) \ge |dz|/4r$. □

The $1/d$-*metric* on a plane domain U is a continuous Riemannian metric with the length element $|dz|/d(z)$. The previous lemma tells us that the hyperbolic metric on a simply connected domain is equivalent to the $1/d$-metric.

7.5. Definite Schwarz Lemma. Montel's compactness allows one to turn the Schwarz Lemma into a definitive form. Let us begin with an elementary version:

LEMMA 7.8. *Let $\phi : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ be a holomorphic map that omits a point z with $|z| \le r < 1$. Then $|\phi'(0)| \le \sigma(r) < 1$.*

PROOF. By the Little Montel Theorem and the Hurwitz Theorem, the space of maps in question is compact (for a given $\rho < 1$). Hence the Schwarz Lemma becomes definite on this space. □

Now the Uniformization Theorem immediately turns this elementary fact into an invariant geometric property:

LEMMA 7.9. *Let $\phi : (S, a) \to (S', a')$ be a holomorphic map between hyperbolic Riemann surfaces. If $\rho_{S'}(a', \partial(\phi S)) \le r$ then $\|D\phi(a)\| \le \sigma(r) < 1$, where the norm is evaluated with respect to the hyperbolic metrics.*

PROOF. Following the proof of the Schwarz Lemma given in §7.1, lift ϕ to a holomorphic map $\tilde{\phi} : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$. By assumption, there is a point $z \in \partial(\phi S)$ such that $\rho_{S'}(a', z) \le r$. Then $\tilde{\phi}$ omits a point $\tilde{z}$ such that

$$\rho_{\mathbb{D}}(\tilde{z}, 0) = \operatorname{dist}_{S'}(z, a') \le r.$$

By Lemma 7.8,

$$\|D\phi(a)\| = |\tilde{\phi}'(0)| \le \sigma(r) < 1.$$

□

COROLLARY 7.10. *For a nest of two hyperbolic Riemann surfaces $S \subset S'$ and any $z \in S$ such that $\rho_{S'}(z, \partial S) \le r$ we have:*

$$\frac{d\rho'}{d\rho}(z) \le \sigma(r) < 1.$$

COROLLARY 7.11. *Let $S \subset S'$ be a nest of two hyperbolic Riemann surfaces, and let $f : S \to S'$ be an (unramified) covering map. Then for $z \in S$ we have*

$$\|Df(z)\|_{S'} \ge \lambda(r) > 1, \quad \text{provided } \operatorname{dist}_{S'}(z, \partial S) \le r.$$

PROOF. It follows from (7.1) and Corollary 7.10. □

Thus, if $S \ne S'$ then a covering map $f : S \to S'$ as above is *locally strictly expanding* in the hyperbolic metric of S'. (If $S = S'$ then it is a local isometry.)

EXERCISE 7.12. *Let $A' \supset A \supset \mathbb{T}$ be a nest of two annuli symmetric with respect to the unit circle $\mathbb{T}$ such that*

$$0 < \mu' \le \operatorname{mod} A' \le 1/\mu', \quad 0 < \mu \le \operatorname{mod} A \le 1/\nu.$$

Then for any $z \in \mathbb{T}$ we have:

$$\frac{d\rho'}{d\rho}(z) \le \sigma(\mu, \mu') < 1.$$

Moreover, if $g : A \to A'$ is a holomorphic double covering then

$$\|Dg(z)\|_{A'} \geq \lambda(\mu') > 1.$$

Let us conclude with the inverse of Corollary 7.10:

EXERCISE 7.13. *For a nest of two hyperbolic Riemann surfaces $S \subset S'$ and any $z \in S$ such that $\rho_{S'}(z, \partial S) \geq r > 0$ we have:*

$$\frac{d\rho'}{d\rho}(z) \geq \xi(r) > 0, \quad \text{where } \xi(r) \to 1 \text{ as } r \to \infty.$$

Thus, the hyperbolic metrics on two Riemann surfaces $S \subset S'$ are comparable in terms of $\rho_{S'}(z, \partial S')$.

7.6. Thin-thick decomposition.

7.6.1. *Definite cusp neighborhoods.*

LEMMA 7.14. *There is a universal $\xi > 0$ such that any cusp on any hyperbolic Riemann surface S has a neighborhood bounded by a horocycle of length ξ.*

PROOF. Let Γ be the Fuchsian group covering S. Assume it contains the translation $T : z \mapsto z + 1$ covering a cusp of interest. Let us consider horocycles $\mathbb{L}_h \equiv \mathbb{L}_h(\infty)$ in $\mathbb{H}$ centered at ∞ and horocycles

$$\mathbb{L}_r(a) = \{z \in \mathbb{H} : |z - (a + ir/2)| = r/2\}$$

centered at $a \in \mathbb{R}$, and let $\mathbb{H}_h \equiv \mathbb{H}_h(\infty)$, $\mathbb{H}_r(a)$ be the corresponding horoballs (see §2.4.4). Note that T translates points of L_1 by distance 1 in the intrinsic horocyclic metric.

Let us show that if $\gamma(\mathbb{L}_1(\infty)) = \mathbb{L}_r(a)$ for some $\gamma \in \Gamma$ and $a \in \mathbb{R}$ then $r < e^2$. [This will impy the desired assertion since under these circumstances the horoball $\mathbb{H}_{e^2}$ will project to a cusp neighborhood on S as a cyclic covering.]

Without loss of generality we can assume that $a = 0$. Let us consider the horocycle $\mathbb{L}_e \subset \mathbb{H}_1$ on hyperbolic distance 1 from $\mathbb{L}_1$. Then $\gamma(\mathbb{L}_e) = \mathbb{L}_{r/e}$ (which in the horocylce in $\mathbb{H}_r(0)$ on hyperbolic distance 1 from $\mathbb{L}_r$). Let us consider the strip $S := \mathrm{cl}(\mathbb{H}_1 \smallsetminus \mathbb{H}_e)$ and the crescent $C := \mathrm{cl}(\mathbb{H}_r(0) \smallsetminus \mathbb{H}_{r/e}(0)) = \gamma(S)$.

Let $\delta := \gamma \circ T \circ \gamma^{-1}$. It is a parabolic map fixing 0 and translatng $\mathbb{L}_r(0)$ by 1 in the intrinsic horocyclic metric. Assume $r \geq e^2$. Then each horocycle $\mathbb{L}_\rho$ with$\rho \in [r/e, r]$ contains a fundamental interval for δ that fits into S.

Let us take a sequence $n_k \to \infty$. Take a point $z^0 \in S$ and consider its translate $z^0 + n_0 \in S$ and let $\zeta^0 := \gamma(z^0 + n_0) \in C$. Then $z^1 := \delta^{m_0}(\zeta^0) \in S$ for some m_0. Now repeat the procedure: let $\zeta^1 := \gamma(z^1 + n_1)$ and $z^2 := \delta^{m_1}(\zeta_1) \in S$, and so on.

Proceeding this way, we will construct a non-escaping infinite Γ-orbit. □

7.6.2. *Geometric collars.* Given a simple closed geodesic γ on a hyperbolic Riemann surface $S = \mathbb{H}/\Gamma$, let

$$\mathcal{N}_\eta(\gamma) := \{z \in S : \ \mathrm{dist}_{\mathrm{hyp}}(z, \gamma) < \eta\}.$$

For instance, let us consider the strip model $\mathbb{S}$ for the hyperbolic plane (see §2.4.2) and a standard cylinder $\mathbb{A}_l = \mathbb{S}/l\mathbb{Z}$. Its equator $\boldsymbol{\gamma}_l := (\mathbb{R} + i\pi/2)/l\mathbb{Z}$ is a hyperbolic geodesic of length l. In this case, $\mathcal{N}_\eta(\boldsymbol{\gamma}_l)$ is also a standard cylinder

$$\mathcal{N}_\eta^{\mathrm{st}}(\boldsymbol{\gamma}_l) := \{z : \ 1/2 - h < \mathrm{Im}\, z < 1/2 + h\}/l\mathbb{Z} \quad \text{where} \quad \int_{1/2}^{1/2+h} \frac{dy}{\sin y} = \eta.$$

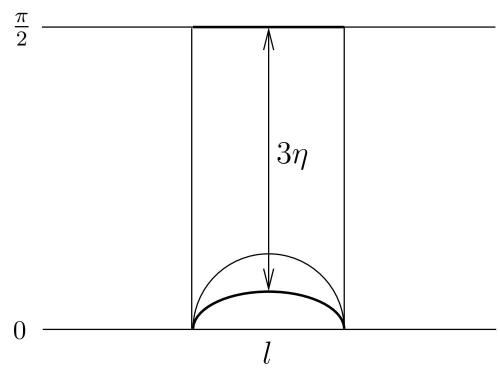

FIGURE 7.2. Geometric collar of a geodesic.

PROPOSITION 7.15. *Ther exists an $\eta = \eta(l)$ such that: for any simple closed geodesic $\gamma \subset S$ of length l, its η-neighborhood* $\operatorname{cl}\mathcal{N}_\eta(\gamma)$ *is isometric to the standard cylinder* $\operatorname{cl}\mathcal{N}_\eta^{\mathrm{st}}(\boldsymbol{\gamma}_l)$*; Moreover,*

$$\eta(l) = \log\frac{1}{l} + O(1) \quad \text{and} \quad \operatorname{mod}\frac{\pi}{l} - 2 < \mathcal{N}_\eta(\gamma) < \frac{\pi}{l} \quad \text{as } l \to 0.$$

The neighborhood $\mathcal{N}_\eta(\gamma)$ will be called the *geometric collar* of γ.

LEMMA 7.16. *Let us consider an interval $I \subset \mathbb{R}$ of Euclidean length $l < \pi$. Let $\boldsymbol{\delta} \subset \mathbb{S}$ and $\boldsymbol{\beta} \subset \mathbb{S}$ be respectively the Euclidean semi-circle and the hyperbolic geodesic in $\mathbb{S}$ sharing the endpoints with I. Then $\boldsymbol{\beta}$ lies under $\boldsymbol{\delta}$.*

PROOF. Let $\chi_I : \partial\mathbb{S} \to \{0, 1,\}$ be the characteristic function of I and let $h_{\boldsymbol{\delta}} : \mathbb{S} \to \mathbb{R}_+$ be its harmonic extension to $\mathbb{S}$. Similarly, let $h_{\boldsymbol{\beta}} : \mathbb{H} \to \mathbb{R}_+$ be the harmonic extension of $\chi_I \,|\, \mathbb{R}$ to $\mathbb{H}$. Then $h_{\boldsymbol{\delta}} \,|\, \partial\mathbb{S} \geq h_{\boldsymbol{\beta}} \,|\, \partial\mathbb{S}$ (with a strict inequality on $\mathbb{R} + i\pi$). By the Maximum Principle (see §10.5), $h_{\boldsymbol{\delta}} \,|\, \mathbb{S} > h_{\boldsymbol{\beta}}$. But

$$\boldsymbol{\delta} = \{z \in \mathbb{H} :\ h_{\boldsymbol{\delta}}(z) = 1/2\}, \quad \boldsymbol{\beta} = \{z \in \mathbb{S} :\ h_{\boldsymbol{\beta}}(z) = 1/2\}.$$

Hence $h_{\boldsymbol{\delta}} \,|\, \boldsymbol{\beta} > 1/2 = h_{\boldsymbol{\delta}} \,|\, \delta$, so $\boldsymbol{\delta}$ is separated from I by $\boldsymbol{\beta}$. $\square$

Proof of Prop. 7.15. (i) Let us realize the universal covering of S as the strip $\mathbb{S}$ so that γ is lifted to the horizontal geodesic $\mathbb{R} + \pi i/2$. The corresponding deck transformation T is the translation by l. Let us consider any deck transformation R that does not belong to the cyclic group of T, and let $\boldsymbol{\beta} := R(\boldsymbol{\gamma})$. Since the geodesic $\boldsymbol{\gamma}$ is simple, $\boldsymbol{\beta}$ does not intersect it, so both ends of $\boldsymbol{\beta}$ lie either on $\mathbb{R}$ or on $\mathbb{R} + \pi i$. Assume for definiteness that the former holds.

Let $I \subset \mathbb{R}$ be the ideal interval sharing the endpoints with $\boldsymbol{\beta}$ (the "shadow" of β). Since $T(\boldsymbol{\beta})$ is disjoint from $\boldsymbol{\beta}$, $|I| \leq l$. Hence $\boldsymbol{\beta}$ lies under the geodesic $\boldsymbol{\beta}' \subset \mathbb{R}$ whose shadow is an interval $I' \supset I$ of length l. The distance from I' to $\mathbb{R} + i\pi/2$ depends only on l; call it $3\eta(l)$. Then the closed η-neighborhood of $\mathbb{R} + i\pi/2$ is disjoint from all of its translates by deck transformations. The conclusion follows.

By Lemma 7.16, $\boldsymbol{\beta}$ lies under the semi-circle $\boldsymbol{\delta}$ based upon I. The Euclidean distance between $\mathbb{R}+i\pi/2$ and $\boldsymbol{\delta}$ is equal $(1-l)/2$, implying the desired asymptotics by elementary estimates. □

Exercise 7.17. (i) *A simple closed geodesic γ of length $\le l$ is $3\eta(l)$-separated from any disjoint simple closed geodesic (where $\eta(l)$ is defined in the above proof).*

(ii) *The geometric collars of disjoint simple closed geodesics are disjoint.*

7.6.3. *Collars around geodesics.* Recall from §6.7 that given a simple closed curve γ on a surface S, $\mathbf{A}_\gamma$ stands for the maximal embedded annulus in the homotopy class of γ.

Collar Lemma. *Let S be a hyperbolic Riemann surface. Then there exist absolute $\bar{l} > 0$ and M such that if for some simple closed geodesic, $l_{\text{hyp}}(\gamma) < \bar{l}$ then*

$$|\operatorname{mod} \mathbf{A}_\gamma - \frac{\pi}{l_{\text{hyp}}(\gamma)}| \le M.$$

Proof. Let us represent the universal covering of S a the strip $\mathbb{S}(\pi)$ so that the geodesic γ lifts to the horizontal line $\{\operatorname{Im} z = \pi/2\}$. The corresponding deck transformation covering γ is the translation $T : z \mapsto z + l$ with $l = l_{\text{hyp}}(\gamma)$. Any other lift of γ is a curve contained in a disk $\mathbb{D}(a, l/2)$ centered at $\partial\mathbb{S}(\pi)$. If $\bar{l} < \pi$, then this lift lies outside the strip $\{|\operatorname{Im} z - \pi/2|\} < (\pi - \bar{l})/2$. Hence this strip projects cyclically to S, and the conclusion follows. □

7.6.4. *Local weights.* Let us take a component J of the ideal boundary $\partial^I S$. It represents a loop on S. Let $\pi_J : \mathbb{A}_J \to S$ be the annulus covering of S corresponding to this loop. (For simplicity, we will often skip label J in the notation, so $\mathbb{A} = \mathbb{A}_J$, $\pi = \pi_J$.) The *local weight* of J is defined as the width of this annulus:

$$\mathcal{W}_{\text{loc}}(J) := \mathcal{W}(\mathbb{A}_J).$$

The following lemma makes a connection between conformal and hyperbolic geometries of S:

Lemma 7.18. *Let γ be the peripheral hyperbolic geodesic on S homotopic to J. Then*

$$\mathcal{W}_{\text{loc}}(J) = \frac{1}{\pi}\, l_{\text{hyp}}(\gamma).$$

Proof. By definition of the covering π, the geodesic γ lifts to a simple closed curve $\hat{\gamma} \subset \mathbb{A}$ such that $\pi : \hat{\gamma} \to \gamma$ is a homeomorphism. Since π is a local isometry, $\hat{\gamma}$ is a closed geodesic, and $l_{\text{hyp}}(\hat{\gamma}) = l_{\text{hyp}}(\gamma)$. But there is only one simple closed geodesic on $\mathbb{A}$, and by (2.12) we have:

$$\mathcal{W}(\mathbb{A}) = \frac{1}{\pi}\, l_{\text{hyp}}(\hat{\gamma}).$$

Putting the above ingredients together, we obtain the desired formula. □

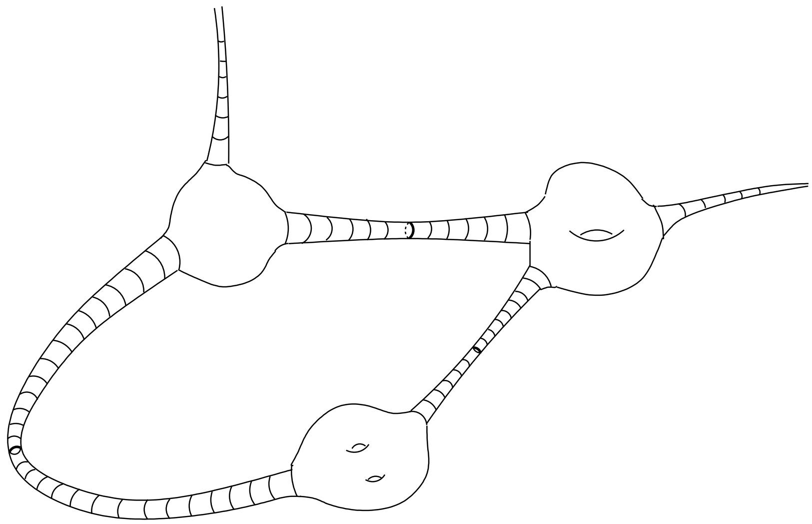

Figure 7.3. Thin-thick decomposition.

7.6.5. *Thin-thick decomposition (without border).* The ε-thick part of a Riemann surface is the set of points with injectivity radius $\geq \varepsilon$.

Theorem 7.19. *There exists an absolute $\varepsilon > 0$ with the following property. Let S be a hyperbolic Riemann surface of finite type without ideal circles at infinity. Then S is represented as the union of finitely many definite cusps, finitely many canonical annuli and the ε-thick part.*

7.6.6. *Comparison of hyperbolic distances.*

Exercise 7.20. *Let γ_1 and γ_2 be two geodesic segments in the hyperbolic plane such that their endpoints stay distance at most d apart. Then*

$$|l_{\text{hyp}}(\gamma_1) - l_{\text{hyp}}(\gamma_2)| \leq 2d.$$

Lemma 7.21. *Let $U \subset V$ be two hyperbolic Riemann surfaces. For $x \in U$, let $\rho(x) \geq 1$ be the conformal density of the hyperbolic metric of U with respect to the hyperbolic metric of V. Let $L(x) = \text{dist}_V(x, \partial U)$, where dist_V stands for the hyperbolic distance in V. Then*

$$\rho(x) = 1 + O(e^{-L}).$$

Proof. Let us consider the universal covering $\pi : (\mathbb{D}, 0)) \to (V, x)$, and let $\hat{U}$ be the component of $\pi^{-1}(U)$ containing 0. Since π is a local isometry, $\text{dist}_{\mathbb{D}}(0, \partial\hat{U}) = L$. Hence $\tilde{U} \supset \mathbb{D}_r$ with $r = 1 - O(e^{-L})$.

The hyperbolic metric in $\mathbb{D}_r$ is obtained from that in $\mathbb{D}$ by scaling by $1/r$. By the Schwarz Lemma, the hyperbolic metric in $\hat{U}$ is dominated by that in $\mathbb{D}_r$. Hence $\hat{\rho}(0) \leq 1/r = 1 + O(e^{-L})$, where $\hat{\rho}$ is the density of the hyperbolic metric in $\hat{U}$ with respect to that in $\mathbb{D}$. Since $\pi : \hat{U} \to U$ is covering, it is a local isometry as well, so $\hat{\rho}(0) = \rho(x)$, and we are done. $\square$

7.6.7. *Hyperbolic geometry of a quadrilateral.*

LEMMA 7.22. *Let us consider an arc α connecting ideal boundary components J and J', and let γ and γ' be the peripheral geodesics homotopic to J and J'. Then*

$$l_{\text{hyp}}(\gamma \cap \Pi(\alpha)) = l_{\text{hyp}}(\gamma' \cap \Pi(\alpha)) = \pi\, \mathcal{W}(\alpha) + O(1).$$

PROOF. Let us consider a lift $\hat{\alpha}$ of α to the universal covering $\mathbb{D}$. It connects some ideal intervals $\hat{J}$ and $\hat{J}'$ covering J and J' respectively. The geodesics γ and γ' lift to geodesics $\hat{\gamma}$ and $\hat{\gamma}'$ sharing the endpoints with $\hat{J}$ and $\hat{J}'$ respectively. Let $\hat{\Pi}$ be the lift of Π to $\mathbb{D}$.

The quadrilateral in $\mathbb{D}$ with horizontal sides $\hat{J}$ and $\hat{J}'$,

$$\mathbf{\Pi}(\hat{\alpha}) \approx \mathbf{\Pi}^a \equiv [0,a] \times [0,1],$$

is symmetric with respect to the geodesic δ corresponding to the horizontal axis in $\mathbf{\Pi}^a$ on mid-height. This symmetry interchanges the geodesic segments $\hat{\gamma} \cap \hat{\Pi}$ and $\hat{\gamma}' \cap \hat{\Pi}$ implying that they have the same length. So, we can focus on one of these geodesics, say $\hat{\gamma} \equiv \hat{\gamma}_a$ in the $\mathbf{\Pi}^a$-model (with all notations preserved).

Let $\hat{\gamma}(x)$ be the point on $\hat{\gamma}$ with the horizontal coordinate $x \in [0,a]$. Then $\operatorname{Im} \hat{\gamma}_a(1)$ decreases as a decreases (the geodesic $\hat{\gamma}_a$ goes "down" as a decreases, which can be seen using the comparison principle for the harmonic measure, as $\hat{\gamma}$ is the 1/2-level set for the harmonic function h on $\mathbf{\Pi}^a$ which is equal to 1 on on $\hat{J}$ and vanishes on the rest of the boundary of $\mathbf{\Pi}^a$). It follows that for $a \geq 2$, $1/2 - \operatorname{Im} \gamma_a(1)$ is bounded by $1/2 - \operatorname{Im} \gamma_2(1)$. For the same reason (or by symmetry), $1/2 - \operatorname{Im} \hat{\gamma}_a(a-1)$ is bounded as well.

We have two geodesics segments, $\hat{\gamma} \cap \hat{\Pi}$ and $\delta \cap \hat{\Pi}$ with the endpoints staying bounded hyperbolic distance apart. By Exercise 7.20,

$$l_{\text{hyp}}(\hat{\gamma} \cap \hat{\Pi}) = l_{\text{hyp}}(\delta \cap \hat{\Pi}) + O(1).$$

Finally, we should compare the hyperbolic length of $\delta \cap \hat{\Pi}$ in $\mathbf{\Pi}^a$ and in the infinite strip $\mathbb{S} = \{0 < \operatorname{Im} z < 1\}$ (as we know, the latter is equal to $\pi\mathcal{W}(\alpha)$). We claim that they differ by a bounded amount. Indeed, let $d(x) = \min(x, a-x)$ be the distance (both Euclidean and hyperbolic in $\mathbb{S}$) from $(x, 1/2) \in \hat{\Pi}$ to the vertical boundary of $\hat{\Pi}$. By Lemma 7.21, the ratio between the hyperbolic metrics in question at that point is $1 + O(e^{-d(x)})$, which implies the conclusion. □

Let $\mathcal{A}_J$ stand for the family of canonical arcs landing on J. For the corresponding peripheral geodesic $\gamma = \gamma_J$, we let

$$\gamma_{\text{thick}} \equiv \gamma_{J,\text{thick}} := \gamma \smallsetminus \bigcup_{\alpha \in \mathcal{A}_J} \Pi(\alpha). \tag{7.4}$$

LEMMA 7.23. *There exists an absolute $\varepsilon > 0$ such that for any two peripheral geodesics γ and γ', we have:* $\operatorname{dist}_{\text{hyp}}(\gamma_{\text{thick}}, \gamma'_{\text{thick}}) > \varepsilon$.

PROOF. For a small $\varepsilon > 0$, let us consider two peripheral geodesics, surrounding boundary curves J and J' respectively, with $\operatorname{dist}_{\text{hyp}}(\gamma_{\text{thick}}, \gamma'_{\text{thick}}) < \varepsilon$. Let a and a' be the closest points on these geodesics, and let α_0 be the the shortest geodesic connecting a and a' (the common perpendicular to our peripheral geodesics).

Let us lift this configuration to the universal covering $\mathbb{D}$ so that a and a' go to imaginary symmetric points, $i\hat{a} = -i\hat{a}'$. (We mark the lifted objects in $\mathbb{D}$ with "*hat*".) Then $\hat{\alpha}_0 = [\hat{a}', \hat{a}]$ is contained in the full geodesic $\hat{\alpha} := (-i, i)$, where α represents an arc in S connecting J to J'.

Moreover, $\hat{\gamma}$ and $\hat{\gamma}'$ are uniformly close almost horizontal (in the Euclidean sense) geodesics chopping off in $\mathbb{D}$ a geometrically thin quadrilateral Q symmetric with respect to both axes. The complement $\mathbb{T} \smallsetminus \bar{Q}$ consists of two symmetic arcs λ and λ' of size close to $1/2$ covering J and J' respectively. These arcs are horizontal sides of a rectangle $\mathbf{\Pi}$. Since they are very long compared wih the complementary arcs of $\mathbb{T}$, the rectangle $\mathbf{\Pi}$ is wide. Hence it produces the canonical rectangle $\Pi(\alpha)$ whose lift $\hat{\Pi}(\alpha)$ is obtained by chopping off two square buffers from $\mathbf{\Pi}$. Moreover, $\hat{\alpha}$ is the genuinely vertical geodesic in $\mathbf{\Pi}$, which belongs to $\hat{\Pi}$. Thus, the points a and a' belong to $\Pi(\alpha)$, which is in the thin part of our surface S. The conclusion follows. □

Corollary 7.24. *Any two disjoint closed peripheral geodesics of length $\leq l$ are at least $\varepsilon(l) > 0$ separated.*

Proof. On the thin part, the geodesics are $\varepsilon(l)$-separated since the canonical rectangles $\Pi(\alpha)$ from (7.4) intersecting γ or γ' have a bounded width (by Lemma 7.22). On the thick part, they are ε-separated with a uniform $\varepsilon > 0$ by Lemma 7.23. □

7.6.8. *Thin-thick decomposition* (*bordered case*).

Theorem 7.25. *For a bordered hyperbolic Riemann surface S of finite type, we have:*

$$\sum_{J \in \partial^i S} \mathcal{W}_{\mathrm{loc}}(J) = 2 \sum_{\alpha \in \mathrm{WAD}(S)} \mathcal{W}(\alpha) + O(|\chi(S)|).$$

Proof. Let us consider all the peripheral geodesics $\gamma = \gamma_J$, $J \in \partial^I S$, and their slices $\gamma \cap \Pi(\alpha)$ by the canonical rectangles, $\alpha \in \mathrm{WAD}(S)$. By Lemma 7.22, we have:

$$2 \sum_{\alpha} \mathcal{W}(\Pi(\alpha) = \frac{1}{\pi} \sum_{J,\alpha} l_{\mathrm{hyp}}(\gamma_J \cap \Pi(\alpha)) + O(|\chi(S)),$$

where the factor "2" appears because each arc α meets two geodesics γ_J. Also, by Lemma 7.18,

$$\sum \mathcal{W}_{\mathrm{loc}}(J) = \frac{1}{\pi} l_{\mathrm{hyp}}(\gamma_J).$$

Comparing these inequalities, we see that the desired one boils down to

$$\sum_J l_{\mathrm{hyp}}(\gamma_{J,\mathrm{thick}}) = O(|\chi(S)|).$$

Let $\varepsilon > 0$ be from Lemma 7.23. For each peripheral geodesic γ, consider a maximal 2ε-separated net of points x_i^J on $\gamma_{J,\mathrm{thick}}$. Let n_J be the number of these points. Then

$$\varepsilon\, n_J \asymp l_{\mathrm{hyp}}(\gamma_{J,\mathrm{thick}}).$$

By the choice of ε, the hyperbolic half-disks $B(x_i^J, \varepsilon)$ (lying on the interior side of the corresponding geodesics) are pairwise disjoint. Hence

$$\sum_J \frac{\pi \varepsilon^2}{2} n_J \leq \mathrm{area}_{\mathrm{hyp}}(S_{\mathrm{conv}}),$$

where S_{conv} is the *convex hull* of S, i.e., the subsurface bounded by all peripheral geodesics. By the *Gauss-Bonnet formula*, the latter is equal to $\pi\chi(S)$. The conclusion follows. □

7.7. Carathéodory convergence.

7.7.1. *Convergence of domains.* Let us consider the space $\mathcal{D}$ of all pointed conformal disks (D, a) in the complex plane. This space can be endowed with a natural topology called *Carathéodory.* We will describe it it terms of convergence:

Definition 7.26. A sequence of pointed disks $(D_n, a_n) \in \mathcal{D}$ converges to a disk $(D, a) \in \mathcal{D}$ if:

(i) $a_n \to a$;

(ii) Any compact subset $K \subset D$ is eventually contained in all disks D_n:

$$\exists\, N: \;\; K \subset D_n \;\forall\, n \geq N;$$

(iii) If U is a topological disk contained in infinitely many domains D_n then U is contained in D.

Note that this definition allows one to pinch out big bubbles from the domains D_n.

Exercise 7.27. *a) Define a topology on $\mathcal{D}$ that generates the Carathéodory convergence.*

b) Show that if ∂D_n converges to ∂D in the Hausdorff metric then the disks D_n converge to D in the Carathéodory sense.

The above purely geometric definition can be reformulated in terms of the uniformizations of the disks under consideration. Let us uniformize any pointed disk $(D, a) \in \mathcal{D}$ by a conformal map $\phi : \mathbb{D} \to D$ *positively normalized* so that $\phi(0) = a$ and $\phi'(0) > 0$.

Proposition 7.28. *A sequence of pointed disks $(D_n, a) \in \mathcal{D}$ converges to a pointed disk $(D, a) \in \mathcal{D}$ if the corresponding sequence of normalized uniformizations $\phi_n : \mathbb{D} \to D_n$ converges to the positively normalized uniformization $\phi : \mathbb{D} \to D$ uniformly on compact subsets of $\mathbb{D}$.*

Proof. Assuming $\phi_n \to \phi$, let us check properties (i)-(iii) of Definition 7.26. The first one is obvious. To verify (ii), take a compact subset K of D. Then $\phi(\mathbb{D}_r) \supset K$ for some $r < 1$. Hence $\text{dist}(\phi(\mathbb{T}_r), K) > 0$ and the curve $\phi : \mathbb{T}_r \to \mathbb{C}$ has winding number 1 around any point of K. Since $\phi_n \to \phi$ uniformly on $\mathbb{T}_r$, eventually all the curves $\phi_n : \mathbb{T}_r \to \mathbb{C}$ have winding number 1 around all points of K. Then $\phi_n(\mathbb{D}_r) \supset K$.

Let us now verify (iii). It is enough to check that any disk $V \Subset U$ is contained in D. For such a disk, we have: $\text{mod}(D_n, V) \geq \mu > 0$ for all n. Let $W_n = \phi_n^{-1}(V_n)$. By the conformal invariance, $\text{mod}(\mathbb{D}, W_n) \geq \mu$ as well. Hence $W_n \subset \mathbb{D}_{1-2\varepsilon}$ for some $\varepsilon > 0$ (by Lemma 4.17 or 6.10). Using conformal invariance of moduli and Lemma 6.10 once again, we conclude that $\text{dist}(\phi_n(\mathbb{T}_{1-\varepsilon}), V) \geq \rho > 0$. Since eventually $|\phi(z) - \phi_n(z)| < \rho/2$ on $\mathbb{T}_{1-\varepsilon}$, the curve $\phi : \mathbb{T}_{1-\varepsilon} \to \mathbb{C}$ has the same winding number around any point of V as $\phi_n : \mathbb{T}_{1-\varepsilon} \to \mathbb{C}$, and the latter is equal to 1 (for n sufficiently big). Hence $\phi(\mathbb{D}_{1-\varepsilon}) \supset V$, as required.

Vice versa, assume $(D_n, a_n) \to (D, a)$ in the Carathéodory topology. By Property (ii) of Definition 7.26, the domains D_n eventually contain the disc $\mathbb{D}(a, r_D(a)/2)$ (where $r_D(a)$ stands for the inner radius of the domain D with respect to $a \in D$, see §4.4). By Corollary 4.18, $|\phi_n(0)| \geq r_D(a)/2$.

On the other hand, by Property (iii), the domains D_n do not eventually contain the disc $\mathbb{D}(a, 2r_D(a))$. By Corollary 4.18, $|\phi_n'(0)| \leq 8r_D(a)$.

Thus, $|\phi_n(0)| \asymp 1$. By the Koebe Distortion Theorem (see Exercise 4.14), the family $\{\phi_n\}$ is precompact in the space of univalent functions. But by the first part of this lemma, any limit function $\phi = \lim \phi_{n(k)}$ is the positively normalized uniformization of (D, a) by $(\mathbb{D}, 0)$. It follows that the ϕ_n converge to this uniformization. $\square$

For $r \in (0, 1)$, let $\mathcal{D}_r$ stand for the family of pointed disks $(D, a) \in \mathcal{D}$ with

$$r \leq r_D(a) \leq 1/r.$$

COROLLARY 7.29. *The space $\mathcal{D}_r$ is compact.*

PROOF. Let $\phi_D : (\mathbb{D}, 0) \to (D, a)$ be the positively normalized uniformization of D. By Corollary 4.18, $r \leq \phi_D'(0) \leq 4/r$ By the Koebe Distortion Theorem (see Exercise 4.14), the family of univalent functions ϕ_D, $D \in \mathcal{D}_r$, is compact. By Proposition 7.28, the space $\mathcal{D}_r$ is compact as well. $\square$

7.7.2. *Convergence of maps.* With these notions in hands, we can define *convergence of a sequence of functions $\psi_n : (D_n, a_n) \to (\mathbb{C}, b_n)$ on varying domains.* Namely, the functions ψ_n converge to a function $\psi : (D, a) \to (\mathbb{C}, b)$ if the pointed domains (D_n, a_n) converge to (D, a), and $\psi_n \to \psi$ uniformly on compact subsets of D. (This makes sense since for any $K \Subset D$, all but finitely many functions ψ_n are well defined on K.)

REMARK 7.30. We will often suppress mentioning of the base points a_n, as long as it would not lead to a confusion.

We can now naturally define *normality* of a family of functions $\psi_n : D_n \to \mathbb{C}$ with varying domains of definition. In case when the D_n converge to some domain D, we also say that "the family $\{\psi_n\}$ is normal on D".

The statement of the Montel Theorem admits an obvious adjustment in this setting: If the family of domains D_n is Carathéodory precompact and the functions $\psi_n : D_n \to \hat{\mathbb{C}}$ omit three values on the Riemann sphere, then the family $\{\phi\}_n$ is normal.

7.7.3. *Space of annuli maps.* In conclusion, let us consider the space $\mathcal{A}$ of pointed conformal annuli (A, a) whose equator E contains a and separates 0 from ∞.

EXERCISE 7.31. (i) *Carry the above discussion for the space $\mathcal{A}$.*
(ii) *Show that for any $\mu \in (0, 1), R > 1$, the subspace*

$$\mathcal{A}_{\mu,R} := \{A : \ \mu \leq \operatorname{mod} A \leq 1/\mu, \ R^{-1} < |a| < R\}$$

is Carathéodory compact.

Let us consider the space $\mathfrak{C}_d$ of annuli coverings $f : (A, a) \to (A', a')$ (with $A, A' \in \mathcal{A}$) of degree d. Let $\mathfrak{C}_d(\mu, R)$ be the subspace of maps $f \in \mathfrak{C}_d$ with $A \in \mathcal{A}_{\mu,R}$.

EXERCISE 7.32. *For anty $d \in \mathbb{Z}_+$, $\mu \in (0, 1)$ and $R > 1$. the space $\mathfrak{C}_d(\mu, R)$ is compact.*

8. Carathéodory boundary

8.1. Prime ends. As we know from §2.4.17, any non-cuspidal tame end E of a Riemann surface S can be completed by attaching to it ideal circle $\partial^I e$. It turns out that in case when $S \Subset \hat{S}$ for some ambient Riemann surface $\hat{S}$, this completion can be described in terms of the ambient geometry. This is the goal of this section.

We will follow the general strategy outlined in §1.7.6: to define a notion of an "end", we need a notion of a "fjord" and a notion of "escaping nest" of fjords.

8.1.1. *Nested fjords.* We will focus on the case of a conformal disk in the Riemann sphere. So, let us consider a pointed conformal disk $(D, a) \subset (\hat{\mathbb{C}}, a)$. A *(genuine) cross-cut* in D is an arc $\sigma : [0, 1] \to \overline{D} \smallsetminus \{a\}$ such that $\operatorname{int} \sigma = \sigma(0, 1) \subset D$ while $\partial\sigma := \sigma\{0, 1\} \subset \partial D$. A *generalized cross-cut* in D is a proper arc $\sigma : (0, 1) \to D$ not passing through a.

Exercise 8.1. *Any generalized cross-cut σ divides D into two domains,*

The component of $D \smallsetminus \sigma$ that does not contain a is called a *prime fjord F*. A prime fjord is specified as genuine or generalized according to the quality of its cross-cut.

Exercise 8.2. *Show that for any $\varepsilon > 0$, a generalized cross-cut $\sigma : (0, 1) \to D$ can be ε-approximated (in the Hausdorff metric) with a genuine cross-cut σ' : $[0, 1] \to \overline{D}$ coinciding with σ on $[\delta, 1 - \delta]$ (for some $\delta > 0$) and such that the ends $\sigma'[0, \delta]$ and $\sigma'[1 - \delta, 1]$ have length $< \varepsilon$.*

Now we can define *equivalent nests* of prime fjords as in §1.7.5.

Remark 8.3. Sometimes it is convenient to consider *continuous* nests of prime fjords F_t, $t \in (0, \varepsilon)$ (where $F_\tau \subset F_t$ for $\tau < t$).

Let $\psi : D \to \mathbb{D}$ be the Riemann mapping. A nest of prime fjords F_n is *escaping* if $\operatorname{diam} \psi(F_n) \to 0$. For instance, a nest of standard fjords F_n is escaping if and only if $|\theta_n^+ - \theta_n^-| \to 0$ and $t_n \to 0$, where $\theta_n^\pm$, t_n are respectively the angles and the level of the rays and the equipotential bounding the F_n.

Following the general strategy, we now define a *prime end E* of D as an equivalence class of escaping nests of (generalized) prime fjords. The definition is designed so that there is a natural one-to-one correspondence between prime ends and points of the ideal boundary $\partial^I D \approx \mathbb{T}$. In these terms, topology on the ideal compactification $\partial^I D \approx \overline{\mathbb{D}}$ can be described as follows (compare §1.7.5). Given a prime fjord F, let $\mathcal{U}_n(F)$ be the union of F and all the prime-ends that are subordinated to F. The base of topology of $\operatorname{cl}^I D$ comprises all the sets $\mathcal{U}(F)$, together with all open sets of D.

According to a general definition from §1.7.6, the *impression* of the prime end E represented by a nest of prime fjords (F_n) is defined as

$$I(E) = \bigcap_n \overline{F}_n.$$

8.1.2. *Shrinking cross-cuts.* The above discussion is quite tautological as the notion of "escaping fjords" is defined in terms of the Riemann mapping. What will make it useful is that prime ends can be characterized in terms of the ambient geometry of the domain D.

A nest of cross-cuts is *shrinking* if $\operatorname{length}(\sigma_n) \to 0$ (in the spherical metric).

LEMMA 8.4. *Any prime end is represented by a nest of (genuine) prime fjords with shrinking cross-cut.*

PROOF. Let $b \in \mathbb{T}$ and $e = e(b)$ be the corresponding prime end of $\mathbb{D}$. It is represented by the (continuous) nest of circular cross-cuts $\sigma_r := \mathbb{T}(b, r) \cap \mathbb{D}$ around b.

The images $\phi(\sigma_r)$ form a continuous nest of (generalized) cross-cuts of D. These cross-cuts do not necessarily shrink, but as we will see in a moment, some of them do.

It will be slightly more convenient to replace $\mathbb{D}$ with the upper-half plane $\mathbb{H}$ and to put b at the origin. Let us consider half-circles $S(r) = \mathbb{T}_r \cap \mathbb{H}$ around 0. We will show that there is a sequence of good radii $r_i \to 0$ such that the cross-cuts $\phi(S(r_i))$ of D shrink. To this end, let us consider half-annuli $\Pi_r = \mathbb{A}(r/2, r) \cap \mathbb{H}$ viewed as rectangles whose horizontal sides are the semi-circles. Let $\mathcal{F}_r$ be the horizontal foliation of Π_r by the half-circles S_ρ, $r/2 < \rho < r$. The extremal length of this foliation is equal to $1/\operatorname{mod} \Pi_r = \pi/\log 2$. By the conformal invariance, the foliation $\phi(\mathcal{F}_r)$ has the same extremal length.

Let l_r be the minimal spherical length of the curves of $\phi(S_\rho)$, $r/2 < \rho < r$. By definition of the extremal length,

$$\frac{l_r^2}{\operatorname{area}(\phi(\Pi_r))} \le \mathcal{L}(\phi(\mathcal{F}_r)) = \frac{\pi}{\log 2}, \tag{8.1}$$

where the "area" stands for the spherical area. Since $\operatorname{area}(\phi(A_r)) \to 0$ as $r \to 0$, we conclude that $l_r \to 0$ as well, which gives us the desired nest with shrinking cross-cuts. $\square$

REMARK 8.5. a) Notice that the good radii r_i constructed above have the property that $r_{i+1} \ge r_i/2$.

b) Note also that the rate of shrinking of the above cross-cuts is uniform with respect to the choice of $b \in \mathbb{T}$. Indeed, all the rectangles Π_r are contained in the annulus $\mathbb{A}[1 - r, 1)$ (in the disk model), and $\operatorname{area} \phi(\mathbb{A}[1 - r, 1)) \to 0$. This makes estimate (8.1) uniform.

To reverse the above lemma, we will need the following useful fact that can be called *uniform continuity of the Riemann mapping on continua*:

LEMMA 8.6. *Let $D \subset \hat{\mathbb{C}}$ be a conformal disk, and let $\psi : D \to \mathbb{D}$ be the Riemann mapping. Then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that for any continuum $\gamma \subset D$ with* $\operatorname{diam} \gamma < \delta$ *(in the spherical metric) we have* $\operatorname{diam}(\psi(\gamma)) < \varepsilon$.

PROOF. Assume this is not the case. Then we can find a sequence of continua $\gamma_n \subset D$ such that $\gamma_n \to b \in \partial D$, while the images $\psi(\gamma_n)$ converge (in the Hausdorff metric) to some closed interval $\omega \subset \mathbb{T}$ (which *a priori* could even coincide with the whole circle $\mathbb{T}$). Let ω' be a closed sub-interval of $\operatorname{int} \omega$.

As in Lemma 8.4, let us pass again to the half-plane model, so that $\omega' \subset \mathbb{R}$. Let us also put b at the origin. Also, let $\phi = \psi^{-1} : \mathbb{D} \to D$.

Lemma 8.4 implies that for any $i \in \mathbb{N}$ there is a finite family of half-circles $S_{r_i}(a_i) = \{|z - a_i| = r_i,\ \operatorname{Im} z > 0\}$ such that:

$$[a_i - r_i, a_i + r_i] \subset \omega, \quad \omega' \subset \bigcup (a_i - r_i, a_i + r_i)$$

and

$$r_i < 1/i, \quad l(\phi(S_{r_i}(a_i)) < 1/i.$$

Since each of these half-circles intersects $\psi(\gamma_n)$, and $\gamma_n \to 0$, we conclude that $\phi(S_{r_i}(a_i)) \to 0$ (in the Hausdorff metric).

Concatenating arcs of the above circles, we obtain a genuine cross-cut σ_i in $\mathbb{D}$ with a base interval T_i such that $\omega' \subset \operatorname{int} T_i \subset T_i \subset \omega$. Moreover, we can arrange the construction so that the base intervals T_i form an increasing nest. Then for any i and $j > i$ sufficiently big, we can find a Jordan curve $\Gamma_{i,j}$ composed by an arc of σ_i and an arc of σ_j. It bounds a Jordan disk Δ_{ij}. By the Maximum Principle,

$$\max_{z\in\Delta_{ij}} |\phi(z)| \to 0 \quad \text{as } i, j \to \infty.$$

Letting $j \to \infty$, we conclude that

$$\limsup |\phi(z)| \leq \varepsilon_i \quad \text{as } z \to \omega',$$

where $\varepsilon_i \to 0$. Hence $\phi(z) \to 0$ uniformly as $z \to \omega'$, so ϕ admits a continuous extension to ω' by letting $\phi|\,\omega' \equiv 0$. But this is impossible. $\square$

EXERCISE 8.7. *Can you justify this assertion?*

COROLLARY 8.8. *Let $D \subset \hat{\mathbb{C}}$ be a conformal disk, and let $\psi : D \to \mathbb{D}$ be the Riemann mapping. If a curve $\gamma : [0, 1) \to D$ lands at some boundary point $b \in \partial D$ as $t \to 1$, then its image $\psi(\gamma)$ lands at some point $e(\theta)$ of the circle $\mathbb{T}$.*

Finally, we can invert Lemma 8.4:

COROLLARY 8.9. *Let $D \subset \hat{\mathbb{C}}$ be a conformal disk, and let $\psi : D \to \mathbb{D}$ be the Riemann mapping. If (F_n) is a nest of genuine prime fjords in D with shrinking cross-cuts, then $(\psi(F_n))$ is a nest of fjords in $\mathbb{D}$ of the same quality. Hence the latter shrinks to some point of of $\mathbb{T}$.*

Lemma 8.4 and Corollary 8.9 show that for a conformal disk $D \subset \hat{\mathbb{C}}$, any prime end of D is represented by a nest of (genuine) prime fjords with shrinking cross-cuts, and vice versa: any such a nest represents some ideal end. This brings us to the standard definition of a prime end as *an equivalence class of nests of (genuine) prime fjords with shrinking cross-cuts*. We see that a prime end gives a view of an ideal boundary point in terms of the spherical geometry (for a conformal disk in $\hat{\mathbb{C}}$). With this understanding, we will also refer to the ideal boundary $\partial^I D$ as the *Carathéodory boundary* and will use notation $\partial^C D$ for it. Accordingly, the ideal compactification $\operatorname{cl}^I D$ will also be called the *Carathéodory compactification* $\operatorname{cl}^C D$.

We are ready to formulate a fundamental result of the classical boundary values theory:

CARATHÉODORY BOUNDARY THEOREM. *The Riemann mapping $\phi : \mathbb{D} \to D$ extends to a homeomorphism $\hat{\phi} : \overline{\mathbb{D}} \to \operatorname{cl}^C D$.*

EXERCISE 8.10. *Let us consider two conformal disks $D, D' \subset \hat{\mathbb{C}}$ and a homeomorphism $h : \overline{D} \to \overline{D}'$. Then $h : D \to D'$ continuously extends to a homeomorphism $\hat{h} : \operatorname{cl}^C D \to \operatorname{cl}^C D'$. Moreover, if $D = D'$ and $h|\,\partial D = \mathrm{id}$ then $\hat{h}|\,\partial^C D = \mathrm{id}$.*

COROLLARY 8.11. *Let $D \subset \hat{\mathbb{C}}$ be a conformal disk, and let $\psi : D \to \mathbb{D}$ be the Riemann mapping. Let $h : \overline{D} \to \overline{D}$ be a homeomorphism. Then the conjugate homeomorphism $H = \phi \circ h \circ \phi^{-1} : \mathbb{D} \to \mathbb{D}$ admits a continuous extension to $\overline{\mathbb{D}}$. Moreover, if $h|\,D = \mathrm{id}$ then $H|\,\mathbb{T} = \mathrm{id}$.*

In conclusion, let us consider an arbitrary domain D in $\hat{\mathbb{C}}$ with a tame end E corresponding to a connected component K of $\hat{\mathbb{C}} \setminus D$. Then K has an annular collar $A \subset D$ representing E. Uniformize A by a round annulus $\mathbb{A}(r, 1)$ so that E corresponds to $\mathbb{T}$. Then one can develop the theory of prime ends for E in the same way as above identifying the ideal boundary $\partial^I_E D \approx \mathbb{T}$ with its Carathéodory boundary, $\partial^C_E D$.

Exercise 8.12. *Go through details of this construction.*

8.2. Local connectivity and Conformal Schönflies Theorem.

Exercise 8.13. *Show that the inverse Riemann map $\phi : \mathbb{D} \to D$ extends continuously to a point $a \in \partial\mathbb{D}$ if and only if the corresponding impression $I(\hat{\phi}(a))$ is a singleton.*

Exercise 8.14. *Let $\Gamma \subset \mathbb{C}$ be an immersed smooth closed curve with transverse self-intersections, and let D be the unbounded component of $\mathbb{C} \setminus \Gamma$. Then the uniformization $\phi : \mathbb{C} \setminus \overline{\mathbb{D}} \to D$ admits a continuous extension to a map $\mathbb{T} \to \Gamma$. Moreover, for any $a \in \Gamma$, $\operatorname{card}(\phi^{-1}(a))$ is equal to the number of components of intersection $\mathbb{D}(a, \varepsilon) \cap D$, where $\varepsilon > 0$ is sufficiently small.*

The next classical theorem will motivate some central problems of holomorphic dynamics:

Carathéodory-Torhorst Theorem. *The following properties are equivalent:*

(i) The inverse Riemann mapping $\phi : \mathbb{D} \to D$ extends to a continuous map $\overline{\mathbb{D}} \to \overline{D}$;

(ii) ∂D is locally connected;

(iii) $\hat{\mathbb{C}} \setminus D$ is locally connected.

Proof. (i) $\implies$ (ii) by Exercise 1.13.

(ii) $\implies$ (iii) by Exercise 1.16.

(iii) $\implies$ (i). Assume ϕ does not admit a continuous extension to $\overline{\mathbb{D}}$. Then there is a point $a \in \partial\mathbb{D}$ such that the corresponding prime end $\hat{\phi}(a)$ has a non-singleton impression $I = I(\hat{\phi}(a))$. Let us consider a nest of semi-circles δ_n shrinking to a whose images $\gamma_n := \phi(\delta_n)$ form a nest $\bar{\gamma}$ of cross-cuts representing the prime end $\hat{\phi}(a)$ (see the proof of the Carathéodory Boundary Theorem). By selecting a subsequence, we can assume that the cross-cuts γ_n shrink to some point $y \in \partial D$.

Since I is not a singleton, $\operatorname{diam} D^+_n(\bar{\gamma}) \not\to 0$. Hence there exist $\varepsilon > 0$ and a sequence of points $\zeta_n = \phi(z_n) \in D^+_n(\bar{\gamma})$ such that $\operatorname{dist}(\zeta_n, \gamma_n) > \varepsilon$. Let us connect z_n to 0 by the straight interval $[0, z_n]$; it crosses δ_n at some point b_n. As the distance $d(\phi(0), \phi(b_n))$ stays away from 0, we can assume it is bigger than ε as well.

Thus, both arcs, $\phi[0, b_n]$ and $\phi[b_n, z_n]$ must intersect the circle of radius $\varepsilon/2$ around y (for n sufficiently big). Then there is a subarc

$$\omega_n \subset \mathbb{D}(y, \varepsilon/2) \cap \phi[0, z_n] \subset D$$

with endpoint on this circle that crosses γ_n at a single point $\phi(b_n)$. This arc separates the endpoints of γ_n in $\mathbb{D}(y, \varepsilon/2) \setminus D$, contradicting local connectivity of $\hat{\mathbb{C}} \setminus D$ at y. $\square$

As a consequence, we obtain:

Conformal Schönflies Theorem. *Let $\gamma \subset \hat{\mathbb{C}}$ be a Jordan curve and D be a component of $\hat{\mathbb{C}} \smallsetminus \gamma$. Then the Riemann uniformization $\phi : \mathbb{D} \to D$ extends to a homeomorphism $\overline{\mathbb{D}} \to \overline{D}$.*

8.3. Accessibility and landing rays. Let $K \subset \mathbb{C}$ be a hull and let $J = \partial K$. Then $D := \hat{\mathbb{C}} \smallsetminus K$ is a hyperbolic disk, so $D \smallsetminus \{\infty\}$ supports two orthogonal analytic foliations, by rays and equipotentials centered at ∞ (see §§4.2, 10.9). In this situation, they are also called *external rays* and *external equipotentials* for K.

Let $B : \mathbb{C} \smallsetminus K \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ be the Riemann mapping normalized so that $B(z) \sim z$ as $z \to \infty$. By definition, the *external rays* and *equipotentials* for K are the pullbacks of the straight rays emanated from ∞ and round circles centered at ∞ (compare §4.2). Let $\mathcal{R}^\theta$ be the pullback of the straight ray $\{re(\theta) : \ r > 1\}$, and let $\mathcal{R}^\theta(r) = B^{-1}(re(\theta))$.

Let $\omega(\mathcal{R}^\theta)$ stand for the *limit set* of the ray $\mathcal{R}^\theta$, i.e., the set of all subsequential limits

$$a = \lim_{r_n \to 1} \mathcal{R}^\theta(r_n) \in J.$$

We say that a ray $\mathcal{R}^\theta$ *lands* at some point $a \in J$ if $\mathcal{R}^\theta(r) \to a$ as $r \to 1$ (or equivalently, $\omega(\mathcal{R}^\theta) = \{a\}$). In this case, $\overline{\mathcal{R}}^\theta = \mathcal{R}^\theta \cup \{a\}$ is called a *closed external ray.*

By the Carathéodory-Torhorst Theorem, we have:

Corollary 8.15. *If $K \subset \mathbb{C}$ is a locally connected hull then every external ray $\mathcal{R}^\theta$ lands at some point $z(\theta) \in J$.*

Lindelöf Theorem. Let $D \subset \hat{\mathbb{C}}$ be a conformal disk, and let $\phi : \mathbb{D} \to D$ be the Riemann mapping. Assume there is a curve $\delta : [0, 1) \to \mathbb{D}$ landing at $e^{2\pi i\theta} \in \mathbb{T}$ whose image $\gamma(t) := \phi(\delta(t))$ lands at some point $\zeta \in D$. Then the ray $\mathcal{R}^\theta$ also lands at ζ, with the same access as γ.

Proof. Let us replace $\mathbb{D}$ with the half-plane $\mathbb{H}$ so that $0 \in \partial\mathbb{H}$ corresponds to $e^{2\pi i\theta}$ (keeping the notation for ϕ). Then the ray $\mathcal{R}$ in question is equal to $\phi(i \cdot \mathbb{R}_+)$.

In the course of the proof of the Carathéodory Boundary Theorem we constructed a shrinking nest of cross-cuts $\sigma_i = \phi(S(r_i))$, where $S(r) = \mathbb{T}_r \cap \mathbb{H}$ and $r_{i+1} \geq r_i/2$. It follows that the ray $\mathcal{R}$ intersects cross-cuts σ_i at points z_i such that $l_{\text{hyp}}(\mathcal{R}[z_i, z_{i+1}]) \leq \log 2$, where $\mathcal{R}[z_i, z_{i+1}]$ is the arc of $\mathcal{R}$ in between points z_i and z_{i+1} and $\text{length}_{\text{hyp}}$ stands for the hyperbolic length in D.

Since near ∂D, the hyperbolic metric blows up compared with the spherical one (Proposition 7.5) we conclude that $\text{length}_{\text{sph}}\, \mathcal{R}([z_{n+1}, z_i]) \to 0$ as $i \to \infty$. Since the cross-cuts σ_i uniformly converge to ζ, so do the arcs $\mathcal{R}[z_i, z_{i+1}]$, and the conclusion follows. □

Exercise 8.16. *Let D be a conformal disk and let $\phi : \mathbb{D} \to D$ be the Riemann mapping. If some ray $\mathcal{R}^\theta$, $\theta \in \mathbb{R}/\mathbb{Z}$, lands at some point $\zeta \in \partial D$ then $\phi(z) \to \zeta$ uniformly on any Stolz sector centered at $e(\theta)$.*

The Lindelöf Theorem, together with Corollary 8.8, imply:

Corollary 8.17. *Let $D \subset \hat{\mathbb{C}}$ be a conformal disk. If a curve γ in D lands at some boundary point $b \in \partial D$, then there is a ray $\mathcal{R}^\theta$ landing at b with the same access as γ.*

LEMMA 8.18. *Let D be a conformal disk and let $\psi : D \to \mathbb{D}$ be the Riemann mapping. Let $\gamma, \gamma' : [0, 1) \to D$ be two curves landing at some points $b, b' \in \partial D$ respectively, and let $e(\theta), e(\theta') \in \mathbb{T}$ be the landing points of their images $\psi(\gamma), \psi(\gamma')$ (existing by* Corollary 8.8). *Then $\theta' = \theta$ iff $b' = b$ and the curves γ, γ' represent the same access to b. Moreover, $\mathcal{R}^\theta$ is the only ray landing at b with the same access.*

PROOF. Let $\gamma : [0, 1) \to D$ and $\gamma(t) \to b$ as $t \to 1$. For each small $\varepsilon > 0$, let us define a cross-cut σ_ε as the component of $\partial\mathbb{D}(b, \varepsilon) \cap D$ that separates $\gamma(0)$ from the $\gamma(t)$ with t sufficiently close to 1. These cross-cuts form a (continuous) nest with shrinking cross-cuts that represents a prime end $p \in \partial^C D$. Moreover, $\gamma(t) \to p$ as $t \to 1$ in the Carathéodory compactification $\mathrm{cl}^C D$ of D. By the Carathéodory Boundary Theorem, $\delta := \psi(\gamma(t))$ converges to the corresponding point $a = e^{2\pi i\theta} := \hat{\psi}(p)$ of $\mathbb{T}$.

If γ' lands at b with the same access, then the prime end p constructed above is the same for these two curves. Hence $\psi(\gamma')$ converges to the same point $a \in \mathbb{T}$. In particular, $\mathcal{R}^\theta$ is the only ray that can land at b with the same access as γ.

Vice versa, assume that the images δ and δ' of γ and γ' converge to the same point $e(\theta)$ of $\mathbb{T}$. By the Lindelöf Theorem, the ray $\mathcal{R} \equiv \mathcal{R}^\theta$ lands at the same point of ∂D as each curve γ and γ'. So, γ and γ' land at the same point b.

Let us show that γ and γ' represent the same access to b as $\mathcal{R}$. It is sufficient to deal with one of them, say γ. We can also assume without loss of generality that γ is smooth and transverse to $\mathcal{R}$. If $\delta = \psi(\gamma)$ intersects the interval $I := [0, a)$ at two consecutive points $\delta(t)$ and $\delta(\tau)$, $t < \tau$, then we can pull the arc $\delta[t-\varepsilon, \tau+\varepsilon]$ off so that it becomes disjoint from I, without changing the access of γ to b. Performing this to all intersections one by one, we replace δ with a curve landing at a and disjoint from I, without changing the access of γ to b. Furthermore, by means of the loop-erasing procedure (see Lemma 1.9), we can turn γ into an arc landing at the same point b.

Thus, we can assume without loss of generality that δ is an arc disjoint from I. Let us connect in $\mathbb{D}$ the beginning points ($\delta(0)$ and 0) of δ and I to obtain a Jordan curve $\Gamma \subset \mathbb{D} \cup \{a\}$. Let $\Delta \subset \mathbb{D}$ be the open Jordan disk bounded by Γ. By the Carathéodory Boundary Theorem, the conformal mapping $\phi \equiv \psi^{-1} : \Delta \to D$ extends continuously to $\overline{\Delta}$. Hence any homotopy $\delta_t \subset \Delta$ between δ and I rel a induces a homotopy $\gamma_t := \phi(\delta_t) \subset D$ between γ and $\mathcal{R}$ rel b. $\square$

We see that landing of a ray at a certain boundary point $b \in \partial D$ and the number of landing rays can be detected purely topologically by looking at accesses to this point. As the rays naturally correspond to prime ends, in the locally connected case there is one-to-one correspondence between all accesses to all boundary points and all prime ends.

A cross-cut, and the corresponding fjord, is called *Green* if it is composed of two arcs of Green rays and an arc of an equipotential (centered at a). Lemma 8.18 implies that any genuine cross-cut can be replaced with a Green cross-cut with the same accesses. It follows that any escaping nest of prime fjords is equivalent to a nest of Green prime fjords. So, prime ends can be defined in terms of Green prime ends as well.

8.4. Appendix: Radial limits for bounded functions. Local connectivity implies accessibilty of all points of a hull. Remarkably, "almost all" points can be accessed for an arbitrary hull.

Let $\phi : \mathbb{D} \to \mathbb{C}$ be a holomorphic function. We say that it has a *radial limit* for $\theta \in \mathbb{R}/\mathbb{Z}$ if there exists a limit

$$\phi(e(\theta)) := \lim_{r \nearrow 1} \phi(re(\theta)).$$

We say that ϕ has *almost all radial limits* if the radial limits exist for Lebesgue almost all $\theta \in \mathbb{R}/\mathbb{Z}$.

The following classical result by Fatou (1906) might be the first application of the Lebesgue Intergation Theory to Complex Analysis that inuagurated the Boundary Value Theory for holomorphic functions:

Fatou Theorem. *Any bounded holomorphic function $\phi : \mathbb{D} \to \mathbb{C}$ has almost all radial limits.*

See [**GaM**, Ch.I,§2] for the proof.

Exercise 8.19. *Show that for any hull $K \subset \mathbb{C}$, external rays $\mathcal{R}^\theta$ land for a.e. $\theta \in \mathbb{R}/\mathbb{Z}$. (It can be also formulated by saying that almost all point with respect to the harmonic measure on K are accessible.)*

Uniqueness Theorem. *Let $\phi : \mathbb{D} \to \mathbb{C}$ be a bounded holomorphic function. If there is a set $\Theta \subset \mathbb{R}/\mathbb{Z}$ of positive length such that $\phi(\theta) \equiv c$ on Θ, then $\phi \equiv c$ on $\mathbb{D}$.*

This classical result is attributed to F. & M. Riesz or Privalov. See [**GaM**, Ch. VI, §2] for the further discussion.

Remark 8.20. These results are important for the Ergodic Theory of polynomials. However, they do not play much role in the theory developed in the 2nd volume of this book (§23.2 notwithstanding).

9. Puzzle and pinched disk models

9.1. Cut-curves and puzzle pieces for hulls.

9.1.1. *Terminology.* Let us develop some terminology concerning intersections of a curve γ with a hull $K \subset \mathbb{C}$.

Assume first that γ is an arc intersecting K for a single parameter, which we can place at 0. Let $a = \gamma(0)$. We say that γ *touches* K at a if the local branches $\gamma_- : (-\varepsilon, 0] \to \mathbb{C}$ and $\gamma_+ : [0, \varepsilon) \to \mathbb{C}$ represent the same access to a from $D := \mathbb{C} \smallsetminus K$. Otherwise we say that γ *cuts* K at a.

An embedded curve γ (which can be closed) is called a *cut-curve* for K if γ *cuts* K at every intersection point $a_i = \gamma(t_i) \in K \cap \gamma$. In particular, we can talk about:

- a *Jordan cut-curve*;
- a *cut-line*, i.e., a properly embedded line $\mathbb{R} \to \mathbb{C}$ which is a cut-curve. (It can also be viewed as a Jordan cut-curve in $\hat{\mathbb{C}}$.)

Note that in both cases, the intersection $\gamma \cap K$ is finite.

Let $L : \mathbb{R} \to \mathbb{C}$ be a cut-line crossing K at points $a_i = L(t_i)$, where $t_1 < \cdots < t_n$. It is a concatenation of two topological rays

$$L_0 : (-\infty, t_1] \to D \cup \{a_1\}, \quad L_n : [t_n, +\infty) \to \{a_n\} \cup D,$$

and $n-1$ arcs $L_i : [t_i, t_{i+1}] \to \{a_i\} \cup D \cup \{a_{i+1}\}$ whose interiors lie in D. The Lindelöf Theorem implies that

• L_0 is homotopic in D rel a_1 to an external ray in D landing at a_1 (and similarly, for L_n);

• Each L_i, $i = 1, \dots, n-1$, is homotopic in D rel $\{a_i, a_{i+1}\}$ to a concatenation of two arcs of Riemann rays in D and an equipotential arc.

So, the whole cut-line L is homotopic in $\hat{\mathbb{C}}$ rel $(L \cap K) \cup \{\infty\}$ to a line concatenated of arcs of external rays and equipotentials in D (in a "minimal" way). The latter cut-lines are called *Green*. The ray part of such a line is called *vertical*, while the equipotential part is called *horizontal*.[23]

We say that a cut-line is *simple* if it crosses K at a single point a. If such a line is Green, it comprises two closed external rays, $\overline{\mathcal{R}}^{\theta_1}$ and $\overline{\mathcal{R}}^{\theta_2}$, landing at a.

A point $a \in K$ is called a *cut-point* if there is a simple cut-line L through this point. In this case, the components of $\mathbb{C} \smallsetminus L$ are called (open[24]) *sectors* bounded by L (*rooted at* a). For a Green sector S bounded by rays $\mathcal{R}^{\theta_1}$ and $\mathcal{R}^{\theta_2}$, we call $\mathrm{Sh}(S) := (\theta_1, \theta_2) \subset \mathbb{R}/\mathbb{Z}$ the *shadow* of S at infinity (where the arc (θ_1, θ_2) is selected so that the rays $\mathcal{R}^\theta$ with $\theta \in (\theta_1, \theta_2)$ are contained in S).

A sector S rooted at a cut-point a is called *perfect* if the intersection $T^* := (K \cap S) \smallsetminus \{a\}$ is connected. In this case, the closure $T := \mathrm{cl}\, T^* = T^* \cup \{a\}$ is called a *branch* of K at a, while T^* itself is called an *unrooted branch*. Under these circumstances, there are no non-peripheral accesses to a from $S \smallsetminus K$. (By a *peripheral access* to a we mean an access given by a boundary curve of ∂S.) For a perfect Green sector, it follows that no external ray $\mathcal{R}^\theta \subset S$ lands at a.

A *puzzle piece* (for K) is a closed Jordan disk P bounded by a Jordan cut-curve γ. The cut-points $K \cap \partial P$ are called *vertices* of P. Given a vertex $a \in \partial P$, the local sector S of ∂P rooted at a and contained in $\mathrm{int}\, P$ is called the *corner* of P at a. A puzzle piece is called *perfect* if all its corners are such.

9.1.2. *Chopping off subhulls.* More generally than above, we say that two hulls, K_1 and K_2, *touch* at a point $a \in K_1 \cap K_2$ if the set $K := K_1 \cup K_2$ admits a cut-line through a locally separating $K_1 \smallsetminus \{a\}$ from $K_2 \smallsetminus \{a\}$.

The following lemma shows that a simple cut-line cuts a hull into two subhulls touching at the cut-point.

LEMMA 9.1. *Let $K \subset \mathbb{C}$ be a hull, L be a simple cut-line through a cut-point $a \in \partial K$, and let S be a corresponding global sector rooted at a. Then the intersection $S \cap K$ is non-empty, and $\overline{S} \cap K = \mathrm{cl}(S \cap K) = (S \cap K) \cup \{a\}$ is a subhull of K.*

PROOF. The intersection $S \cap K$ is non-empty since L is a cut-line. Furthermore,

$$\overline{S} \cap K = (S \cap K) \cup (\partial S \cap K) = (S \cap K) \cup (L \cap K) = (S \cap K) \cup \{a\}.$$

Let us show that $\overline{S} \cap K$ does not contain relatively clopen subsets $X \subset S$. Assume otherwise. Since $\overline{S} \cap K$ is closed, X is closed in K as well. On the other hand, as $X \subset S \cap K$ and X is relatively open in $\overline{S} \cap K$, it is relatively open in $S \cap K$. But since S is open, $S \cap K$ is relatively open in K. By "transitivity", X

---

[23]This terminology will naturally be extended to various objects below that involve cut-lines (sectors, puzzlle pieces, corners).

[24]We will also allow closed and semi-open sectors, with obvious adjustments.

is relatively open in K. Thus, X is clopen in K contradicting connectivity of the latter.

Now, if $\overline{S} \cap K$ is disconnected then it is decomposed into a disjoint union of relatively clopen subsets, $\overline{S} \cap K = X_1 \sqcup X_2$. Since only one of them can contain a, which is the only point on $\partial S \cap K$, one of the sets X_i is contained in S, contradicting the above assertion.

Thus, $\overline{S} \cap K$ is a continuum. Let us show that it is full. Indeed, if U is a bounded component of its complement, then $\partial U \subset K$. Since K is full, $U \subset \operatorname{int} K$. Since U is in the complement of $\overline{S} \cap K$, it intersects $\operatorname{int} K \smallsetminus \overline{S}$. But then it is contained in $\operatorname{int} K \smallsetminus \overline{S}$ (since any connected subset of K intersecting both S and $\mathbb{C} \smallsetminus \overline{S}$ must contain a, which is on the boundary of K). Then $\partial U \subset K \smallsetminus S$, while by its definition $\partial U \subset \overline{S} \cap K$. It follows that $\partial U \subset \{a\}$, which is of course an absurd.

To complete the proof, we need to check that $\operatorname{cl}(S \cap K) = (S \cap K) \cup \{a\}$. Since we already know that the latter set is closed (equal to $\overline{S} \cap K$), it is left to notice that $a \in \operatorname{cl}(S \cap K)$, for otherwise $S \cap K$ would be a clopen subset of $\overline{S} \cap K$. □

We can now proceed inductively to show:

Corollary 9.2. *Let $K \subset \mathbb{C}$ be a hull.*

(i) *Let L_i be a finite family of disjoint simple cut-lines with $L_i \cap K = \{a_i\}$. Then any component S of $\mathbb{C} \smallsetminus \bigcup L_i$ intersects K, and*

$$\overline{S} \cap K = \operatorname{cl}(S \cap K) = (S \cap K) \cup \{a_j \in \partial S\}$$

is a subhull of K.

(ii) *For any puzzle piece P, the intersection $P \cap K$ is a subhull of K.*

Remark 9.3. Notice that any two components S_1 and S_2 as above are separated by a cut-line L, and hence their closures can only touch at a *single* point $a = L \cap K$.

In fact, Corollary 9.2 is still valid for infinitely many cut-lines:

Corollary 9.4. *Let $K \subset \mathbb{C}$ be a hull, and let L_i be a countable family of disjoint simple cut-lines with $L_i \cap K = \{a_i\}$. Assume $z \in K$ is distinct from all the a_i, and let $I \equiv I(z) \subset K$ be the set of points in K that are not separated from z by the cut-lines L_i, together with those a_i that are not separated from z by other cut-lines L_j, $j \neq i$. Then:*

(i) *I is either a subhull of K or a singleton.*

(ii) *If none of the a_i is separated from z by another cut line L_j ($j \neq i$), then $I(z)$ is a hull.*

(iii) *For any nest of puzzle pieces $P_1 \supset P_2 \supset \dots$, the intersection $K \cap \bigcap P_n$ is either a subhull of K or a singleton.*

Proof. (i) Consider subhulls $I_n \equiv I_n(z)$ associated with the first n cut-lines $L_1, \dots, L_n$. By the previous corollary, they form a nest of subhulls. As $\bigcap I_n = I$, the conclusion follows.

(ii) Under these circumstances, all the a_i belong to I, so I is not a singleton (provided there is more than one point a_i; otherwise the assertion is directly covered by Lemma 9.1).

Part (iii) is a particular case of (i). □

Letting S_i be the sector bounded by L_i that does not contain z, we can also describe $I(z)$ as follows:

$$I(z) = \mathbb{C} \setminus \bigcup S_i.$$

So, $I(z)$ is obtained from K by *chopping off infinitely many sectors* S_i. Moreover, in case (ii) these sectors are pairwise disjoint.

Note that in general the relation between accesses to a hull K and a subhull Q is quite loose. Indeed, different accesses to K can represent the same access to Q, while on the other hand, an access to Q may not represent any access to K. We can only say that the set of acceses to K at z is naturally mapped to the set of accesses to Q at z. [This map is obtained by viewing a curve $\gamma \subset \mathbb{C} \setminus K$ landing at z as a curve in $S \setminus Q$.] However, under the above circumstances, we can say more:

LEMMA 9.5. *Under the circumstances of* Corollary 9.4 (ii) *, the set of acceses to K at z is naturally injected to the set of accesses to $I \equiv I(z)$ at z.*

PROOF. Let us consider two curves $\gamma_1, \gamma_2 \subset \mathbb{C} \setminus K$ landing at z representing two different accesses to K at z. Together, they form a simple cut-line L through K at z, which bounds two sectors $S_\pm$. By Lemma 9.1, the sets $K_\pm := (K \cap S_\pm) \cup \{z\}$ are two subhulls of K touching at z; so, they are not singletons. Accordingly, the set A of cut-points a_i is partitioned into two disjoint subsets, $A_\pm := \{a_i \in S_\pm\} = \{a_i \in K_\pm\}$. Let

$$I_\pm := I \cap K_\pm = (I \cap S_\pm) \cup \{z\}.$$

Then I_+ is obtained from K_+ by chopping off sectors S_i rooted at $a_i \in K_+$ and the sector S_-, all disjoint. By Corollary 9.2 (ii), I_+ is a not a singleton; similar, neither is I_-. It follows that $\gamma_\pm$ represent different accesses to I at z. □

9.1.3. *Puzzle ends, impressions, and rigidity.* Let K be a hull, and let $D := \mathbb{C} \setminus K$. Following the general framework of §1.7.6, the family $\mathcal{P}$ of puzzle pieces allows us to compactify the domain D. In this setting, we define *puzzle fjords* as the intersections $F(P) := \operatorname{int}(P \cap D)$ for various $P \in \mathcal{P}$. An *escaping nest* of fjords $F_0 \supset F_1 \supset \dots$ (or, the corresponding puzzle pieces P_n) is defined by the property that any puzzle piece $P \in \mathcal{P}$ either contains $P_\infty := \bigcap P_n$ or $\operatorname{int} P$ is disjoint from P_∞.

EXERCISE 9.6. *Show that the last property is equivalent to saying that P_∞ cannot be cut into two pieces by a cut-line of K.*

With these in hands, we have a notion of *puzzle ends* $E \equiv E_{\mathcal{P}}$, *puzzle boundary* $\partial^{\mathcal{P}} D$, and *puzzle compactifications* $\operatorname{cl}^{\mathcal{P}} D$. We also have a notion of the *puzzle impression* $P_\infty(E)$ of a puzzle end E defined as $\bigcap P_n$ for any escaping nest (P_n) representing E.

For $a \in K$, we define the puzzle end $E(a)$ as the end represented by escaping puzzle nests (P_n) such that $a \in \operatorname{int} P_n$. The corresponding puzzle impression $P_\infty(a)$ is the set of points $\zeta \in K$ that cannot be separated from a by a cut-line (i.e., there is no cut-line L such that a and ζ lie in different components of $\mathbb{C} \setminus L$). Equivalently, $P_\infty(a)$ is the intersection of all puzzle pieces containing a in its interior.

Corollary 9.4 implies:

COROLLARY 9.7. (i) *Let K be a hull. Then any puzzle impression $P_\infty(a)$ is either a subhull of K or a singleton.*

(ii) *There is a natural pojection* $K \to \partial^{\mathcal{P}} D$ *whose fibers are puzzle impressions.*

(iii) *If all puzzle impressions are singletons then* K *is homeomorphic to* $\partial^{\mathcal{P}} D$.

If the puzzle impression $P_\infty(a)$ is a singleton then a is called *rigid*. Equivalently, a point $a \in J$ is *rigid* if there is a shrinking nest of puzzle pieces P_n containing a in their interiors. If these pieces can be selected perfect, then we say that a is *perfectly rigid*.

PROPOSITION 9.8. (i) *If a point* $a \in K$ *is rigid then* K *is weakly locally connected at* a.

(ii) *If* $a \in K$ *is perfectly rigid then* K *is locally connected at* a.

(iii) *If all points of* K *are rigid then* K *is locally connected at* a.

PROOF. (i) In this case, puzzle pieces P_n containing a in their interior provide a base of closed neighborhoods of a whose intersections with K are connected.

(ii) In this case, the $\operatorname{int} P_n$ form a base of open neighborhoods of a whose intersections with K are connected.

(iii) follows from Exercise 1.11. □

This provides us with a very useful condition for local connectivity that will be applied numerous times in the dynamical context. Let us formulate it in a user friendly way:

COROLLARY 9.9. *Let* K *be a hull. Assume that for some point* $a \in K$, *there exists a nest of puzzle pieces* P_n *such that* $a \in \operatorname{int} P_n$ *and* $\operatorname{diam} P_n \to 0$. *Then* K *is rigid, and hence weakly locally connected, at* a. *If this happens for all* $a \in K$ *then* K *is locally connected.*

Of course, under the circumstances of the last assertion, K is nowhere dense.

9.1.4. *Puzzle and rays.* By definitions, a ray $\mathcal{R}^\theta$ *converges to a puzzle end* E if for *some* (and then for *any*) escaping nest (P_n) representing E, we have:

$$\forall\ n \in \mathbb{N}\ \exists\ t_n > 0 \text{ such that } \mathcal{R}^\theta(t) \in \operatorname{int} P_n \text{ for } t \in (0, t_n).$$

In particular, for $a \in K$, a ray $\mathcal{R}^\theta$ converges to a puzzle end $E(a)$ iff for any puzzle piece P containing a in its interior, the ray $\mathcal{R}^\theta$ is eventually trapped in $\operatorname{int} P$ (i.e., there exists $t_0 > 0$ such that $\mathcal{R}^\theta(t) \in \operatorname{int} P$ for all $t \in (0, t_0)$).

EXERCISE 9.10. *Let* K *be a hull.*

(i) *If* σ *is a cross-cut contained in a puzzle piece* P *then the corresponding prime end fjord* F *is contained in* $\operatorname{int} P$.

(ii) *If a prime end impression* $I(E)$ *intersects a puzzle piece* P, *then* $I(E) \subset P$. *Hence, if the limit set* $\omega(\mathcal{R})$ *of an external ray* $\mathcal{R}$ *intersects* P *then* $\omega(\mathcal{R}) \subset P$.

(iii) *For any puzzle end* $E_{\mathcal{P}}$, *there is at least one prime end* E_C *subordinated to* $E_{\mathcal{P}}$. *Hence there is at least one external ray* $\mathcal{R}^\theta$ *converging to* $E_{\mathcal{P}}$.

(iv) *There is a natural continuous surjective projection* $\pi : \operatorname{cl}^C D \to \operatorname{cl}^{\mathcal{P}} D$ *extending the identical map* $D \to D$. *Hence* $\partial^{\mathcal{P}} D$ *is locally connected.*

COROLLARY 9.11. *Assume a point* $a \in K$ *is rigid. If* a *belongs to the impression* $I(E_C)$ *of some prime end* E_C *then* $I(E_C) = \{a\}$, *and hence the corresponding external ray lands at* a.

Thus, a rigid point a has a well defined full preimage $Q \equiv \Phi^{-1}(a) \subset \mathbb{T}$ under the Riemann uniformization $\Phi : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus K$ with the property that Φ continuously extends to Q, $\Phi(Q) = a$, while the impression of any other prime end is disjoint from a.

REMARK 9.12. The above theory can be developed for a subfamily $\mathcal{P}$ of puzzle pieces satisfying the property that for any two pieces $P_1, P_2 \in \mathcal{P}$,

$$\text{either } \operatorname{int}(P_1 \cap P_2) = \emptyset \text{ or } P_1 \cap P_2 \text{ contains some puzzle piece } P \in \mathcal{P}.$$

Yoccoz puzzle will provide us with an important example of this kind.

9.1.5. *Branches and limbs.* Let us say that a cut-point $a \in K$ is *well branched with valence* $n \geq 2$ if there are n rays $\mathcal{R}_i$ landing at a such that each component of $K \smallsetminus \bigcup \overline{\mathcal{R}}_i$ is connected. Thus, we have n branches T_i of K rooted at a. Corollary 9.8 implies:

COROLLARY 9.13. *A rigid point $a \in K$ with finitely many accesses is well branched.*

Here is another useful condition for well branching:

LEMMA 9.14. *Let S be a sector rooted at $a \in K$ bounded by rays $\mathcal{R}^{\theta_\pm}$. Assume there exists a sequence of cut-points $a_n \in K \cap S$ which are landing points of rays*[25] *$\mathcal{R}_n^{\theta_\pm^n}$ such that $\theta_\pm^n \to \theta_\pm$. Then there exists only one branch of K in S (i.e., $K \cap S$ is connected).*

PROOF. Otherwise we can select a branch $T^* \subset K \cap S$ that does not contain a subsequence of the a_n's. Let us consider a ray $\mathcal{R}^\eta$ landing in T^*. Then η separates one of the $\theta_\pm$ from all of the $\theta_\pm^n$ (for that subsequence). □

EXERCISE 9.15. *Assume that a point $a \in K$ can be separated from any other point of K by a cut-line through a well branched point. Then a is perfectly rigid.*

We will often deal with pointed hulls $K \ni b$ centered at some base point b (which is usually the origin). Under these circumstances, if $a \neq b$ is a well branched cut-point, then all the branches T_i at a that do not contain b will be called *limbs* of K at a, while the corresponding sectors $S_i \supset T_i^*$ will be called *wakes*. The branch containing b will sometimes be called the *body* $\mathcal{B}$ of K at a, but sometimes this term will be used in a different way (which should not lead to confusion as the definitions will be explicitly given). Compare §25.6.5 and §37.3.

9.2. Interior components of hulls.

9.2.1. *Cut-curves and puzzle pieces for general continua.* Let us start with adjusting the terminology developed in §9.1.1 to general continua $J \subset \mathbb{C}$ (not necessarily hulls).

First, *cut-curves, cut-lines and cut-points* are defined in the same way as for hulls. (Of course, in the general case cut-curves can pass through bounded components of $\mathbb{C} \smallsetminus J$.) In particular, a cut-line $L : \mathbb{R} \to \mathbb{C}$ crossing J at points $a_i = L(t_i)$, $t_1 < \cdots < t_n$, is a concatenation of two topological rays

$$L_0 : (-\infty, t_1] \to D \cup \{a_1\}, \quad L_n : [t_n, +\infty) \to \{a_n\} \cup D,$$

[25] More rays landing at a are allowed.

and $n-1$ arcs $L_i : [t_i, t_{i+1}] \to \{a_i\} \cup D_i \cup \{a_{i+1}\}$ whose interiors lie in $\mathbb{C} \smallsetminus J$. Here D is the unbounded component of $\mathbb{C} \smallsetminus J$, while $(D_i)_{i=1}^{n-1}$ is some sequence of components of $\mathbb{C} \smallsetminus J$ (which could be repeated). In particular, if all the D_i are equal to D then L is a cut-line for the hull K of J.

Now, the Lindelöf Theorem implies that

• L_0 and L_n are respectively homotopic in D rel a_1 and a_n to the external rays landing at a_1 and a_n;

• Each L_i, $i = 1, \dots, n-1$, is homotopic in D_i rel $\{a_i, a_{i+1}\}$ to a concatenation of two arcs of internal Green rays in D_i and an arc of equipotential.

So, the whole cut-line L is homotopic in $\mathbb{C} \smallsetminus J$ rel $(L \cap J) \cup \{\infty\}$ to a line concatenated of arcs of Green rays and equipotentials in $\mathbb{C} \smallsetminus J$. We call such cut-lines *Green*.

We say that a cut-line L is *simple* for J if it is such for the hull K. We say that L is *dipole* if it crosses J at two points, a_1 and a_2, so that the arc L_1 connecting these points lies in a bounded component D_1 of $\mathbb{C} \smallsetminus J$. We call such a pair of points $a_1, a_2 \in \partial D_1$ a *dipole*.

One can proceed to define *dipole sectors* and *shadows*, as well as (*perfect*) *dipole sectors*, (*perfect*) *puzzle pieces* and their *vertices* and *corners* similarly to their ordinary counterparts introduced in §9.1.1. *Green* puzzle pieces are also defined naturally.

Exercise 9.16. *Generalize results of* §9.1.2 *to an arbitrary continuum* $J \Subset \mathbb{C}$. *Namely, let* L_i *be a finite family of disjoint simple cut-lines with* $L_i \cap J = A_i$, *where each* A_i *is either a singleton or a dipole. Then any component* S *of* $\mathbb{C} \smallsetminus \bigcup L_i$ *intersects* J, *and*

$$\overline{S} \cap J = \mathrm{cl}(S \cap J) = (S \cap J) \bigcup_{A_i \subset \partial S} A_i$$

is a subcontinuum of J.

Exercise 9.17. *If a puzzle piece* P *is perfect then* $K \cap \operatorname{int} P$ *is connected.*

We can proceed with defining (*perfectly*) *rigid* points of J.

Corollary 9.18. *If* $a \in J$ *is perfectly rigid then* J *is locally connected at* a.

9.2.2. *Limbs and local connectivity at peripheral points.* Let K be a hull, and let D be a component of $\operatorname{int} K$.

Lemma 9.19. *Assume* D *is a Jordan disk and there is a countable set* $\mathcal{A}$ *of cut-points* $a_i \in \partial D$ *such that*

$$K = \overline{D} \cup \bigcup_i \mathcal{L}_i^*, \tag{9.1}$$

where $\mathcal{L}_i^* = S_i \cap K$, S_i *being a sector bounded by two rays landing at* a_i. *Let* $\mathcal{L}_i := \overline{\mathcal{L}}_i \equiv \mathcal{L}_i^* \cup \{a_i\}$. *If* $\operatorname{diam} \mathcal{L}_i \to 0$ *then:*

(i) *Any point* $a \in \partial D \smallsetminus \mathcal{A}$ *is accessible.*

(ii) K *is perfectly rigid (and hence locally connected) at any point* $a \in \partial D \smallsetminus \mathcal{A}$, *with exactly one access (and hence with exactly one external ray landing at* a*).*

(iii) *If a limb* $\mathcal{L}_i$ *is* (*perfectly*) *rigid at its root* $a_i \in \mathcal{A}$ *then* $J \equiv \partial K$ *is* (*perfectly*) *rigid at* a_i.

PROOF. (i) It is sufficient to construct a sequence of cross-cuts σ_n whose impression is equal to $\{a\}$. Due to the Jordan-Schönflies Theorem, we can assume that $a = 0$ and D (as a domain in the Riemann sphere) is the lower half-plane $\mathbb{H}_- = \{\operatorname{Im} z < 0\}$. The case when $a \notin \overline{\mathcal{A}}$ is trivial, so assume that $a \in \overline{\mathcal{A}}$, and assume for definiteness that $\mathcal{A}$ accumulates on a on the right (while may or may not accumulate on the left).

Let us label points $a_i \in \mathcal{A}$ near a and the corresponding sets $\mathcal{L}_i$ so that $a_i > 0$ iff $i > 0$. Let ε_i be the smallest closed disk $\overline{\mathbb{D}}(a, \varepsilon_i)$ containing $\mathcal{L}_i$. Since $\varepsilon_i \to 0$, there exists a positive subsequence $i(n) \to \infty$ such that $\varepsilon_{i(n)} > \varepsilon_j$ for all $a_j \in (0, a_{n(i)})$. Then the disk $\mathbb{D}_{\varepsilon_{i(n)}}$ contains all the sets $\mathcal{L}_j$ rooted at the $a_j \in (0, a_{n(i)})$. Hence there is an arc σ_n of $\mathbb{T}_{\varepsilon_{i(n)}}$ in $\mathbb{H} \smallsetminus K$ whose right-hand endpoint belongs to $\mathcal{L}_{i(n)}$, while the left-hand endpoint belongs either to some $\mathcal{L}_{k(n)}$ with $k(n) < 0$ or to ∂D. For definiteness, assume the former (as the latter case is even either).

The cross-cuts σ_n represent some prime-end E. Let us show that its impression is equal to $\{a\}$.

Let Γ_n be the union of σ_n, the set $\mathcal{L}_{i(n)}$ and $\mathcal{L}_{k(n)}$, and the interval $[a_{k(n)}, a_{i(n)}]$. It is a continuum. Since the limbs shrink, $\operatorname{diam} \Gamma_n \to 0$. Hence $\operatorname{diam} \hat{\Gamma}_n \to 0$ as well, where $\hat{\Gamma}_n$ is the filled Γ_n.

Since the prime-end fiord D_n^+ corresponding to the cross-cut σ_n is contained in $\hat{\Gamma}_n$, we conclude that $\operatorname{diam} D_n^+ \to 0$ as well. Hence $I(E)$ is a singleton, and it cannot be anything but a.

(ii) Let $a \in \partial D \smallsetminus \mathcal{A}$. As above, we assume that $a = 0$ and $\mathbb{D} = \mathbb{H}_-$. Take nearby points $a_- < 0 < a_+$ in $\mathbb{R} \smallsetminus \mathcal{A}$, and consider rays $\mathcal{R}_\pm \subset \mathbb{H} \smallsetminus K$ landing at these points. Since $\operatorname{diam} \mathcal{L}_i \to 0$, we can truncate these rays by a horizontal interval δ on a small height that does not intersect K. We can also connect a_- to a_+ by an arc ω in $\mathbb{H}_-$. Concatenation of these four arcs, $\mathcal{R}_-$, δ, $\mathcal{R}_+$, and ω, is a Jordan cut-curve for $J = \partial K$ that bounds a small puzzle piece P around a.

Moreover, removing of $a_\pm$ from P does not disconnect the latter. Indeed, $(P \cap \overline{D}) \smallsetminus \{a_\pm\}$ is connected, and attaching to it connected sets $\mathcal{L}_i$ preserves connectivity. Thus, P is perfect, implying that a is perfectly rigid, .

Similarly, removing a from P does not disconnect the latter, implying that there is only one access to a.

(iii) Let us drop the label i, so $a \in \mathcal{A}$, S is the corresponding sector rooted at a, and $\mathcal{L} \subset S$ is the corresponding subhull attached to a. Assume $\mathcal{L}$ is rigid at a. Let γ be the boundary of small puzzle piece for $\mathcal{L}$ around a. Then it crosses both rays of the boundary ∂S, and hence it crosses nearby external rays $\mathcal{R}_\pm$ landing on ∂D. Let us consider a Jordan cut-curve for J by taking a concatenation of an arc of γ, arcs of $\mathcal{R}_\pm$ and an arc in D connecting the landing points of $\mathcal{R}_+$ and $\mathcal{R}_-$. This provides us with a small puzzle piece for J, proving rigidity of a in $J = \partial K$. Moreover, if we start with a perfect puzzle piece for $\mathcal{L}$, this construction gives us a perfect puzzle piece for J. $\square$

LEMMA 9.20. *Assume D is a Jordan disk and there is a countable dense set $\mathcal{A}$ of cut-points $a_i \in \partial D$ satisfying* (9.1). *Then* $\operatorname{diam} \mathcal{L}_i \to 0$.

PROOF. If $\operatorname{diam} \mathcal{L}_i \not\to 0$ then we can take a Hausdorff limit $\mathcal{L}_\infty = \lim \mathcal{L}_{i(k)}$ as $i(k) \to \infty$ which is not a singleton. Hence it is a continuum attached to ∂D. Since the sets $\mathcal{L}_{i(k)}$ are eventually disjoint from the sectors S_j, and the latter are open, we have

$$\mathcal{L}_\infty \subset K \setminus \bigcup S_j = K \setminus \bigcup \mathcal{L}_j^*.$$

By (9.1), $\mathcal{L}_\infty \subset \partial D$. Being a continuum, $\mathcal{L}_\infty$ is an arc of ∂D. Since $\mathcal{A}$ is dense in ∂D, the set $\mathcal{L}_\infty$ contains some point a_j in its relative interior. But then the approximating sets $\mathcal{L}_{i(k)}$ eventually cross $\mathcal{L}_j$ – contradiction. □

9.2.3. *LC hulls and Jordan disks.* We will now use the Carathéodory Theorem for further study of the topology of lc hulls.

Let K be a hull, and let (D, b) be a pointed component of $\operatorname{int} K$. (We will refer to b as the *center* of D.) Since it is simply connected, it can be uniformized by the unit disk, $\phi : (\mathbb{D}, 0) \to (D, b)$. *Internal rays* $\mathcal{R}_\theta$ of (D, b) are defined to be the images of the straight rays $\{re(\theta) : 0 \leq \ r < 1\}$ under ϕ.

Proposition 9.21. *Let $K \subset \mathbb{C}$ be a lc hull. Then any component D of* $\operatorname{int} K$ *is a Jordan disk.*

Proof. Let us consider the projection $\pi_D : K \to \overline{D}$ (1.2). Since it is continuous and K is lc, $\overline{D}$ is lc as well (Exercise 1.13,b)). By the Carathéodory-Torhorst Theorem, the boundary ∂D is lc as well and the uniformization $\phi : \mathbb{D} \to D$ extends continuously to the boundary.

This shows that ∂D is a curve. We just need to show that it is simple. If not, then there are two internal rays $\mathcal{R}_1$ and $\mathcal{R}_2$ in D that land at the same point $a \in \partial D$. Then by Lemma 9.1 (applied to the hull $\hat{\mathbb{C}} \setminus D$), the Jordan curve $\gamma := \mathcal{R}_1 \cup \mathcal{R}_2 \cup \{a\}$ surrounds a point $b \in \partial D \subset \partial K$. On the other hand, since K is full, the open Jordan disk bounded by γ is contained in $\operatorname{int} K$; in particular, $b \in \operatorname{int} K$ – contradiction. □

Exercise 9.22. *For any two components D_1 and D_2 of a lc hull K, the closures $\overline{D}_1$ and $\overline{D}_2$ are either disjoint or touch at a single point.*

9.3. Legal issues. Let us say that a lc hull K is *pointed* if every component D of $\operatorname{int} K$ is pointed. Then a point $x \in K$ is called legal if $x \in \partial K$ or x is the center of some component of $\operatorname{int} K$. An arc γ in K is called *legal* if any non-empty slice of γ by a component D of $\operatorname{int} K$ consists of one or two internal rays of D. Obviously, endpoints of a legal arc are legal. Vice versa:

Exercise 9.23. *Let K be a pointed lc hull, and let x and y be two legal points in K. Then x and y can be connected by a unique legal arc $[x, y]$.*

Exercise 9.24. *Let $\gamma : (0, 1) \to K$ be a legal arc in a lc hull K. Then any point $a \in J \cap \gamma$ is a cut-point that can be accessed from above and from below the arc.*

Exercise 9.25. *Assume that K is* 0*-symmetric. Then for any two symmetric legal points, x and $-x$, the legal arc $[-x, x]$ passes through* 0.

Let us say that a set $H \subset K$ is *legally convex* if

(i) for any two legal points $x, y \in H$, the legal arc $[x, y]$ is contained in H;

(ii) The slice of H by any component D of $\operatorname{int} K$ is a union of some internal rays.

Let now $X \subset K$ be a finite set of legal points x_i containing at least two points. The *legal hull* $H = H(X)$ of X in K is the union of the legal arcs $[x_i, x_j]$ connecting all pairs of these points. This is the smallest legally convex set containing X.

Exercise 9.26. *The legal hull $H(X)$ is a topological tree.*

9.4. Pinched disk model for a hull.

9.4.1. *Locally connected case.* The Carathéodory-Torhorst Theorem allows one to represent any hull $K \subset \mathbb{C}$ as a quotient of the unit disc $\mathbb{D}$ by a special equivalence relation $\underset{K}{\sim}$. Namely, this theorem provides us with the continuous extension ϕ : $\mathbb{C} \smallsetminus \mathbb{D} \to (\mathbb{C} \smallsetminus K) \cup \partial K$ of the Riemann uniformization. Now, the equivalence classes of $\underset{K}{\sim}$ on the unit circle $\mathbb{T}$ are defined as the fibers $\phi^{-1}(\cdot)$ of $\phi|\,\mathbb{T}$. Obviously, ∂K is homeomorphic to the quotient $\mathbb{T}/\underset{K}{\sim}$.

We will now extend it to $\mathbb{D}$. Given a non-singleton class X of $\underset{K}{\sim}$, let $\hat{X}$ stand for the hyperbolic convex hull of X, see §2.4.18. (For any singleton class $X = \{x\}$, we let $\hat{X} = X$.

Lemma 9.27. *Given a lc hull K, the convex hulls $\hat{X}$ are pairwise disjoint.*

Proof. Let us compactify the complex plane $\mathbb{C}$ with the circle $\mathbb{T}_\infty$ at infinity. Convergence of points $z_n \in \mathbb{C}$ to $\theta \in \mathbb{T}_\infty$ means that $z_n \to \infty$ and $\arg z_n \to \theta$. It is easy to check that this compactification, $\overline{\mathbb{C}}$, is homeomorphic to $\overline{\mathbb{D}}$.

The Riemann uniformization $\phi : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus K$ extends to a homeomorphism $\overline{\mathbb{C}} \smallsetminus \overline{\mathbb{D}} \to \overline{\mathbb{C}} \smallsetminus K$ in an obvious way. Since K is locally connected, it further extends to a continuous map $\overline{\mathbb{C}} \smallsetminus \mathbb{D} \to \overline{\mathbb{C}} \smallsetminus \operatorname{int} K$ by the Carathéodory-Torhorst Theorem. (We will keep notation ϕ for all these extensions.)

Given an $\underset{K}{\sim}$ equivalence class $X = \phi^{-1}(x) \subset \mathbb{T}_\infty$, $x \in \partial K$, let

$$\tilde{X} = \{re(\theta) : \ r \in [0, \infty], \ \theta \in X\} \subset \overline{\mathbb{C}} \smallsetminus \mathbb{D} :,$$

and let

$$X' = \phi(\tilde{X}) = X \cup \bigcup_{\theta \in X} \mathcal{R}_\theta \cup \{x\} \subset \mathbb{C} \smallsetminus \operatorname{int} K.$$

This is a compact set intersecting $\mathbb{T}_\infty$ by X and intersecting K by $\{x\}$.

Consider now another equivalence class, $Y = \{\phi^{-1}(y)\}$, $y \in \partial K$, $y \neq x$. Then $X \cap Y = \emptyset$, and hence $\tilde{X} \cap \tilde{Y} = \emptyset$. Since $\phi : \overline{\mathbb{C}} \smallsetminus \overline{\mathbb{D}} \to \overline{\mathbb{C}} \smallsetminus K$ is a homeomorphism, the sets $X' \smallsetminus K$ and $Y' \smallsetminus K$ are disjoint. But the intersections $X' \cap K = \{x\}$ and $Y' \cap K = \{y\}$ are also disjoint. Thus, $X' \cap Y' = \emptyset$.

By Proposition 2.65, the sets X and Y are unlinked on $\mathbb{T}_\infty \approx \mathbb{T}$, so their convex hulls $\hat{X}$ and $\hat{Y}$ are disjoint in $\overline{\mathbb{D}}$. □

Each set $\hat{X}$ is declared to be an equivalence class of $\underset{K}{\sim}$. All other equivalence classes are singletons. (This equivalence relation can be considered not only on $\overline{\mathbb{D}}$ but on the whole plane $\mathbb{C}$.)

Theorem 9.28. *A locally connected hull $K \subset \mathbb{C}$ is homeomorphic to the quotient $\overline{\mathbb{D}}/\underset{K}{\sim}$. Moreover, the inverse Riemann map $\phi : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus K$ admits am extension to a homeomorphism $(\mathbb{C}/\underset{K}{\sim}, \overline{\mathbb{D}}/\underset{K}{\sim}) \to (\mathbb{C}, K)$.*

Proof. Let $\hat{\mathbb{T}} := \bigcup \hat{X} \subset \overline{\mathbb{D}}$, where the union is taken over all equivalence classes on $X \subset \mathbb{T}$.

Step 1: The set $\hat{\mathbb{T}}$ is closed. Let $z_n \to z \in \mathbb{D}$ and $z_n \in \hat{X}_n = \phi^{-1}(\zeta_n)$ with $\zeta_n \in \partial K$. Passing to a subsequence, we can assume that $\zeta_n \to \zeta \in \partial K$. By continuity of $\Phi|\,\mathbb{T}$, we have (see Exercise 1.121)

$$\limsup_{n\to\infty} X_n \subset X := \phi^{-1}(\zeta)$$

It easily implies (see Exercise 2.28) that

$$\limsup_{n\to\infty} \hat{X}_n \subset \hat{X},$$

so $z \in \hat{X}$.

Step 2: The map $\phi : \mathbb{T} \to \partial K$ extends to a continuous map $\hat{\phi} : \hat{\mathbb{T}} \to \partial K$ by declaring $\hat{\phi}(\hat{X}) = \phi(X)$.

Let $z_n \to z \in \mathbb{D}$, $z_n \in \hat{X}_n$. Without loss of generality, we can assume that the X_n are pairwise disjoint. Then there exist points $z'_n \in \partial X_n$ converging to z as well, so we can assume that $z_n \in \partial X_n$ in the first place. But $\bigcup \partial X_n$ is the support of a geodesic lamination $\mathcal{L}$, so z_n belongs to some geodesic leaf $\gamma_n = [x_n, y_n] \in \mathcal{L}$, where $x_n, y_n \in \mathbb{T}$. But then the γ_n uniformly on $\overline{\mathbb{D}}$ converge to the geodesic leaf $\gamma = [x, y] \in \mathcal{L}$ through z (see Exercise 2.67). Since $\phi | \mathbb{T}$ is continuous,

$$\hat{\phi}(z_n) = \phi(x_n) \to \phi(x) = \hat{\phi}(z).$$

Step 3: For any gap Q in $\hat{\mathbb{T}}$ the map $\hat{\phi}$ continuously extends to a homeomorphism $\overline{Q} \to \overline{D}$, where D is a component of $\operatorname{int} K$.

The closure $\overline{Q} \subset \overline{\mathbb{D}}$ is the convex hull of its ideal boundary $\partial^I Q \subset \mathbb{T}$, which is a Jordan disk bounded by $\partial^I Q$ and a family of hyperbolic geodesics Γ_j (see Lemma 2.25 and §2.5). The quotient $\overline{Q}/ \underset{K}{\sim}$ is obtained by collapsing the Γ_j to singletons, which is also a closed Jordan disk (by the Devil Staircase in the disk, see Exercise 1.8).

Any homeomorphism between the boundaries of two Jordan discs extends continuously to the whole discs (e.g., radially). In particular, the embedding $\hat{\phi} : (\partial Q/ \underset{K}{\sim}) \to \partial K$ extends to a homeomorphism $(Q/ \underset{K}{\sim}) \to \overline{D}$, where D is the (open) Jordan disc bounded by $\hat{\phi}(\partial Q)$. This Jordan disc is contained in $\operatorname{int} D$ since K is full. Since $\partial D \subset \partial K$, D is a component of $\operatorname{int} K$.

Step 4: The map $\hat{\phi} : \overline{\mathbb{D}} \to K$ is continuous.

Given $z_n \to z \in \overline{\mathbb{D}}$, we want to show that $\hat{\phi}(z_n) \to \hat{\phi}(z)$. By the above discussion (Steps 2–3), we only need this check it in case $z_n \in Q_n$ where the Q_n are distinct gaps. Since $\operatorname{area} Q_n \to 0$, there exist points $z'_n \in \partial Q_n \subset \hat{\mathbb{T}}$ such that $\operatorname{dist}(z_n, z'_n) \to 0$, so $z'_n \to z$ as well. By Step 2, $\hat{\phi}(z'_n) \to \hat{\phi}(z)$. But $\operatorname{dist}(\hat{\phi}(z_n), \hat{\phi}(z'_n) \le \operatorname{diam} \hat{\phi}(Q_n) \to 0$ by Proposition 9.27. The conclusion follows.

Step 5: The map $\hat{\phi} : \overline{\mathbb{D}} \to K$ is onto. Here we will make use of the exterior of $\overline{\mathbb{D}}$. Let us consider some circle $\mathbb{T}_R$ with $R > 1$ and the corresponding equipotential $\mathcal{E}_R = \phi(\mathbb{T}_R)$. It goes once around K, so by the Topological Argument Principle (Proposition 3.8) all values in K must be assumed by $\hat{\phi}$ □

9.4.2. *General case.* For a general hull K, we can modify the above construction to produce a lc model K_{lc} for K. Namely, to each cut-point $a \in K$ we can associate the set $X(a) \subset \mathbb{T}$ of external angles of the rays landing at a. Take the hyperbolic convex hull $\hat{X}(a) \subset \overline{\mathbb{D}}$ of this set. Its boundary in $\mathbb{D}$ is the union of hyperbolic geodesics. Since the sets $X(a)$ are unlinked, all these geodesics are pairwise disjoint. Taking the closure and cleaning it up, we obtain a geodesic lamination $\mathcal{L}_K$ in $\mathbb{D}$.

Some gaps of this lamination are classes $\hat{X}(a)$, we call them *gaps of first kind* (or black gaps), others are called *gaps of second kind* (or white gaps). Note that no two gaps of first kind are adjacent, so we obtain a colored lamination. Taking the quotient of $\mathbb{C}$ mod this colored lamination (by collapsing the classes $\hat{X}(a)$ to single points), we obtain K_{lc}, the *lc pinched disk model* for K.

PROPOSITION 9.29. *Assume $\bigcup X(a)$ is dense in $\mathbb{T}$. Then there exists a natural continuous projection*

$$\pi : (\mathbb{C}, K) \to (\mathbb{R}^2, K_{\mathrm{lc}}).$$

This projection is a homeomorphism if and only if K is locally connected.

EXERCISE 9.30. *Let X be a Cantor set on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and let*

$$K = \bigcup_{\theta \in \mathbb{X}} [0, e(\theta)]$$

be the corresponding hedgehog. What is the lc model for K?

10. Appendix 1: Potential theory

Harmonic and subharmonic functions is a very important subject on its own right that penetrates deeply into analysis, geometry, and probability theory. From our perspective, their outstanding role comes from the fact that they lay down a foundation for a proof of the Uniformization Theorem. For readers' convenience, here we will briefly review needed basics of the theory.

10.1. Harmonic functions and differentials. Recall that a function $u : U \to \mathbb{R}$ on a domain $U \subset \mathbb{C}$ is called *harmonic* if $u \in C^2(U)$ and $\Delta u = 0$ where $\Delta = \partial_x^2 + \partial_y^2$ is the usual Euclidean Laplacian. The real and imaginary parts of any holomorphic function $f = u + iv$ on U are harmonic, which is readily seen from the *Cauchy-Riemann equations*

$$\partial_x u = \partial_y v, \quad \partial_y u = -\partial_x v.$$

They are called *conjugate harmonic functions*.

Vice versa, any harmonic function u can locally be represented as the real part of a holomorphic function. Indeed, $\Delta u = 0$ gives the integrability condition for the Cauchy–Riemann equations that allow one to recover locally the conjugate function v.

This can be nicely expressed in terms of the Hodge $*$ operator. Let $V \approx \mathbb{R}^2$ be the oriented 2D Euclidean space. By self-duality, we identify vector fields $\tau = a\partial_x + b\partial_y$ with 1-forms $\omega = adx + bdy$. The *Hodge $*$-operator* is defined as $\pi/2$-rotation of ω or τ, i.e. $*\omega = -bdx + ady$.

Then the Cauchy–Riemann equations can be written as

$$dv = d^c u, \quad \text{where} \quad d_c := *d, \quad \text{while} \quad dd^c u = \Delta u \, dz \wedge dy. \tag{10.1}$$

So, u is harmonic if and only if the form $d^c u$ is closed, and then (10.1) can be locally integrated:

$$v(z) = \int_{z_0}^{z} d^c u = \int_\gamma \frac{\partial u}{\partial n} ds, \tag{10.2}$$

where γ is a smooth (oriented) path connecting z_0 to z (within a small disk), ds is the length element on γ and n is the unit normal vector to γ rotated clockwise from the corresponding tangent vector to γ.

Globally, the integral (10.2) depends on the homotopy class of the path γ (rel the endpoints), so it defines a multi-valued harmonic function v and the corresponding multivalued holomorphic function $f = u+iv$. The monodromy for this function along a cycle γ depends only on the homology class of γ and is given by the *periods* of $d^c u$:

$$f_\gamma(z) - f(z) = i\int_\gamma d^c u = i\int_\gamma \frac{\partial u}{\partial n} ds,$$

where f_γ is the result of analytic continuation of f along along γ. In particular, if a is an isolated singularity for u, then the monodromy if f as we go around a little circle $\gamma = S_r := \{|z-a| = r\}$ is equal to

$$f_\gamma(z) - f(z) = i\int_{S_r} \frac{\partial u}{\partial r}(\zeta)\, d\theta.$$

Relation between harmonic and holomorphic functions makes the notion of harmonicity manifestly invariant under holomorphic changes of variable: if u is harmonic then so is $u\circ\phi$ for any holomorphic map ϕ. Thus, *harmonicity is well-defined on an arbitrary Riemann surface S.* This can also be seen from the original definition by expressing the Laplacian in terms of the differential operators ∂ and $\bar\partial$ (see §2.11). Indeed, we have:

$$\partial = \frac{1}{2}(d + id^c), \quad \bar\partial = \frac{1}{2}(d - id^c). \tag{10.3}$$

so,

$$\Delta u\, dx\wedge dy = dd^c\, u = 2i\,\partial\bar\partial\, u.$$

Remark 10.1. Expressions (10.3) show that d and d^c are (twice) the real and imaginary parts of the operators ∂ and $\bar\partial$.

A C^1 differential 1-form $\omega = adx + bdy$ is called *harmonic* if it is locally the differential of a harmonic function. It is called *co-closed* if $d(*\omega) = 0$. It is straightforward to check that a form ω is harmonic if and only if it is closed and co-closed.

Another characterization is that harmonic 1-forms are real part of Abelian differentials. Namely, the differential $\alpha = \omega + i\eta$ is holomorphic if and only if ω is is harmonic and $\eta = *\omega$. (Note that unlike the case of functions, this relation is *global*.)

10.2. Basic properties. Given a domain U on a Riemann surface S, let $\mathcal{H}(U)$ stand for the space of harmonic functions in U , and let $\mathcal{H}(\overline{U})$ stand for the subspace of $\mathcal{H}(U)$ consisting of functions that admit continuous extension to $\overline{U}$.

Mean Value Property. A C^2 function u on a domain $U\subset\mathbb{C}$ is harmonic is and only if for any disk $\mathbb{D}(a,r)\subset U$, we have

$$u(a) = M_u(a,r) := \frac{1}{2\pi}\int_0^{2\pi} h(a + re^{i\theta})\, d\theta;$$

Proof. The mean value property for harmonic functions immediately follows from the corresponding property for holomorphic ones. The inverse follows from the second order Taylor expansion at z averaged over a little circle:

$$M_u(z,r) - u(z) = \frac{1}{4}\Delta u(z) r^2 + o(r^2). \tag{10.4}$$

□

The Mean Value Property implies in a standard way (as for holomorphic functions):

Maximum/Minimum Principle. If a harmonic function u on a Riemann surface U has a local maximum or minimum in U then it is constant.

Corollary 10.2. *Let $U \Subset S$ be a compactly embedded domain in a Riemann surface S, and let $u \in \mathcal{H}(\overline{U})$. Then u attains its maximum and minimum on ∂U.*

Corollary 10.3. *Under the above circumstances, u is uniquely determined by its boundary values, $u|\,\partial U$.*

10.3. Poisson Formula. The Poisson Formula allows us to recover a harmonic function $h \in H(\overline{\mathbb{D}})$ from its boundary values:

Proposition 10.4. *For any harmonic function $h \in H(\overline{\mathbb{D}})$ in the unit disk, we have: formula: the following Poisson representation:*

$$h(z) = \frac{1}{2\pi}\int_0^{2\pi} h(\zeta)P(z,\zeta)\,d\theta, \quad z \in \mathbb{D},\ \zeta = e^{i\theta} \in \mathbb{T},$$

with the the Poisson kernel

$$P(z,\zeta) = \frac{1-|z|^2}{|z-\zeta|^2}. \tag{10.5}$$

Proof. For $z = 0$, this formula amounts to the Mean Value Property:

$$h(0) = \frac{1}{2\pi}\int_0^{2\pi} h(e^{i\theta})\,d\theta.$$

It implies the formula at any point $z \in \mathbb{D}$ by making a Möbius change of variable

$$\phi_z : \overline{\mathbb{D}} \to \overline{\mathbb{D}}, \quad \zeta \mapsto \frac{\zeta - z}{1 - \bar{z}\zeta}$$

that moves z to 0. Since $h \circ \phi_z^{-1} \in H(\overline{\mathbb{D}})$, we obtain:

$$h(z) = (h \circ \phi_z^{-1})(0) = \frac{1}{2\pi}\int_0^{2\pi} h \circ \phi_z^{-1}\,d\theta = \frac{1}{2\pi}\int_0^{2\pi} h\,d\theta_z,$$

where

$$d\theta_z = (\phi_z)^*(d\theta) = |(\phi_z)'(\theta)|\,d\theta,$$

and the latter derivative is equal to the Poisson kernel $P(z,\zeta)$ (check it!).

Uniqueness of the extension follows from the Maximum Principle. □

The *Dirichlet problem* (in some domain $D \subset \hat{\mathbb{C}}$) is the problem of recovery of a harmonic function $h \in \mathcal{H}(\overline{D})$ from its boundary values on ∂D. The Poisson formula provides us with an explicit solution of this problem in the unit disk:

Proposition 10.5. *Any continuous function $g \in C(\mathbb{T})$ on the unit circle admits a unique harmonic extension $h \in \mathcal{H}(\overline{\mathbb{D}})$ to the unit disk (so that $g = h|\,\mathbb{T}$). This extension is given by the Poisson formula:*

$$h(z) = \frac{1}{2\pi}\int_0^{2\pi} g(\zeta)P(z,\zeta)\,d\theta, \quad z \in \mathbb{D},\ \zeta = e^{i\theta} \in \mathbb{T}.$$

Proof. The Poisson kernel $P(z,\zeta)$ as a function of $\zeta \in \mathbb{T}$ and $z \in \mathbb{D}$ has the following properties:

(i) $P(z,\zeta) > 0$, and for any $z \in \mathbb{D}$, we have $\dfrac{1}{2\pi}\displaystyle\int_{\mathbb{T}} P(z,\zeta)\, d\theta = 1$;

(ii) For any $\zeta \in \mathbb{T}$, the kernel $P(z,\zeta)$ is harmonic in $z \in \mathbb{D}$;

(iii) For any $\zeta_0 \in \mathbb{T}$ and any $\varepsilon > 0$, we have:

$$\mathbb{P}_z(\zeta) \to 0 \text{ as } z \to \zeta_0 \text{ uniformly in } \zeta \in \mathbb{T} \smallsetminus \mathbb{D}(\zeta_0, \varepsilon).$$

Property (i) follows from the Poisson representation of the function $h(z) \equiv 1$ in $\overline{\mathbb{D}}$.

To check (ii), notice that $P(\cdot,\zeta)$ is the pullback of the function $\operatorname{Im} u$ on the upper half-plane to the unit disk under the Möbius transformation

$$\phi_\zeta : \mathbb{D} \to \mathbb{H}, \quad \phi_\zeta : z \mapsto i\,\frac{\zeta + z}{\zeta - z}.$$

Exercise 10.6. *Check this using that ϕ_z is a hyperbolic isometry.*

The last property is obvious (it corresponds to the fact the the function $\operatorname{Im} u$ vanishes on $\mathbb{R}$).

Properties (i) and (iii) imply that $P(z, e^{i\theta})\, d\theta$, viewed as measures on $\mathbb{T}$ weakly converge to δ_{ζ_0}. This implies that g gives the boundary values of h. Property (ii) implies harmonicity of h in $\mathbb{D}$. $\square$

10.4. Harnak Inequality and normality. This inequality allows one to control a positive harmonic function by its value at one point. Let us begin with the case of disk:

Lemma 10.7. *For any $r \in (0,1)$, there exists a constant $C(r) > 1$ such that for any positive harmonic function $u \in \mathcal{H}(\overline{\mathbb{D}})$, we have:*

$$C(r)^{-1}\, u(0) \le h(z) \le C(r)\, u(0), \quad |z| \le r.$$

Proof. It immediately follows from the Poisson representation since

$$C(r)^{-1} \le P(z,\zeta) \le C(r) \quad (|\zeta| = 1,\ |z| \le r) \quad \text{with } C(r) = \frac{1+r}{1-r}$$

and the Mean Value Property. $\square$

Let us now consider the general case. By a *coordinate disk* $D(a,\varepsilon)$ we mean a domain lying within some local chart and equal to the disk $\mathbb{D}(z(a),\varepsilon)$ in this coordinate.

Theorem 10.8. *Let S be a (connected) Riemann surface, and let $z_0 \in U$, $K \Subset U$. Then there exists a constant $C_K > 1$ such that for any positive harmonic function $u \in \mathcal{H}(U)$, we have:*

$$C_K^{-1}\, u(z_0) \le u(z) \le C_K\, u(z_0), \quad \textit{for any } z \in K.$$

Proof. We can find finitely many coordinate disks $D(z_i, \varepsilon_i)$ whose union $\cup D(z_i, \varepsilon_i/2)$ is connected and covers $K \cup \{z_0\}$. Applying the Lemma 10.7 consecutively to these disks, we obtain the desired inequalities. $\square$

Similarly to holomorphic functions, bounded families of harmonic functions are *normal* (i.e., precompact in the topology of uniform convergence on compact subsets):

Proposition 10.9. *A bounded family of harmonic functions on U is normal.*

Proof. The Poisson formula gives a bound on the partial derivatives of $u \in \mathcal{H}(u)$ on a compact subset $K \Subset U$ in terms of the bound on u (and the set K). By the Ascoli-Arcela, our family is precompact in the space of continuous functions on U (in topology of uniform convergence on compact subsets). But the Mean Value Property survives under taking locally uniform limits. Hence harmonicity survives as well. □

Corollary 10.10. *Let $u_n \in \mathcal{H}(U)$ be an increasing sequence of harmonic functions, and let $u_n(z_0) \leq C$ at some point $z_0 \in U$. Then the u_n converge, uniformly on compact subsets of U, to a harmonic function $u \in \mathcal{H}(\mathcal{U})$.*

Proof. Subtracting u_0 from the u_n, we see that our functions can be assumed positive. By the Harnak Inequality, the u_n are uniformly bounded on compact subsets. So, their pointwise limit $u(z)$ is finite. Moreover, by Proposition 10.9, they form a normal sequence, and hence u is harmonic. □

10.5. Subharmonic functions. Harmonic functions are analytic and hence rigid: they cannot be locally modified. Subharmonic functions are much more flexible, but at the same time, they still possess good compactness properties (an *a priori* upper bound is sufficient). This combination makes them very useful.

The basic example of a subharmonic function is $u = \log|f(z)|$ where f is a holomorphic function. In fact, this function is harmonic everywhere except for zeros of f where it assumes value $-\infty$ ("poles"of u). This suggests that in general subharmonic functions should also be allowed to have poles. Of course, $[-\infty, \infty)$ is naturally endowed with topology of a half-open interval.

Definition 10.11. A function $u : D \to [-\infty, \infty)$ on a domain $D \subset \mathbb{C}$ is called *subharmonic* if it is not identically equal to $-\infty$[26] and satisfies the following two conditions:

- *Mean Value Property (subharmonic)*: For any disk $\mathbb{D}(z, r) \Subset D$,

$$u(z) \leq M_u(z, r) \tag{10.6}$$

- u is upper-semicontinuous.

Remark 10.12. Notice that the two conditions in the above definition make the *value* of a subharmonic function well determined at a point by its values nearby. In fact, below we will be dealing only with *continuous* subharmonic functions, and mostly, assuming only *finite* values. However, the following basic subharmonic function does have a pole:

Example 10.13. *Let $u(z) = \log|z|$. This function is harmonic in $\mathbb{C}^*$, so the MVP is satisfied on an any disk $\mathbb{D}(a, r) \Subset \mathbb{C}^*$. It is also obviously satisfied on $\mathbb{D}_r$ as $-\infty < M_u(0, r)$.*

Let us check it for the disk $\mathbb{D}(a, r) \ni 0$. Making an affine change of variable, we can consider instead the Mean Value Property on $\mathbb{D}$ for a function $v(z) = \log|z - c|$, $c \in \mathbb{D}^*$. Then we have:

$$\frac{1}{2\pi}\int_{\mathbb{T}} u(z)\, d\theta = \frac{1}{2\pi}\int_{\mathbb{T}} \left(u(z) + \log\left|\frac{1 - \bar{c}z}{z - c}\right| \right) d\theta$$

[26] This convention is not completely standardized.

$$= \frac{1}{2\pi}\int_{\mathbb{T}} \log|1-\bar{z}c|\,d\theta = 0 > \log|c| = v(0).$$

For the disk $\mathbb{D}(a,|a|)$ *whose boundary passes through* 0*, MVP follows by continuity.*

REMARK 10.14. The above estimate is a particular case of the Jensen formula:

$$\frac{1}{2\pi}\int_{\mathbb{T}} \log|f(\zeta)|d\,\theta = \log|f(0)| + \sum \log\frac{1}{|a_i|},$$

where f is a holomorphic function in $\mathbb{D}$, continuous up to the boundary, that does not vanish on $\mathbb{T}$ and at 0.

We let $\mathcal{SH}(U)$ stand for the space of continuous subharmonic functions in U.

Obviously, the set of subharmonic functions is invariant under addition and multiplication by positive numbers, so it is a *cone*. Also, *Maximum of finitely many subharmonic functions is subharmonic.* For instance, the function $\log^+|z| = \max\{\log|z|, 0\}$ is subharmonic.

As for harmonic functions, the Subharmonic Mean Value Property implies:

MAXIMUM PRINCIPLE. *If a subharmonic function* u *on a Riemann surface* U *has a local maximum in* U *then it is constant.*

However, the *Minimum Principle* is not any more valid for subharmonic functions.

More generally, we can majorant a subharmonic function by a harmonic one:

LEMMA 10.15. *Let* D *be a bounded domain in* $\mathbb{C}$*, and let* u *and* h *be respectively harmonic and a continuous subharmonic functions on* D*, both admitting a continuous extensions to* $\overline{D}$*. If* $u \le h$ *on* ∂D *then* $u \le h$ *in* D*.*

Vice versa, if a function u *is continuous in a domain* $U \subset \mathbb{C}$ *and the above property is satisfied for any domain* $D \Subset U$ *and any harmonic* $h \in \mathcal{H}(U)$*, then* u *is subharmonic.*

PROOF. To check the former assertion, apply the Maximum Principle to $u-h$.

To check the latter, let us consider a coordinate disk D and let h solves the Dirichlet Problem in D with the boundary values $h|\,\partial D = u|\,\partial D$. Then $u|\,D \le h|\,D$. Evaluating it at the center of D, we obtain the Mean Value Property for subharmonic functions. □

This lemma shows that the notion of subharmonicity is bi-holomorphically invariant (at least for continuous functions[27], and hence is well defined on an arbitrary Riemann surface.

Also, let us consider a function

$$\tilde{u}_D(z) = u(z) \quad \text{for } z \in U \smallsetminus D, \quad \text{and} \quad u(z) = h(z) \quad \text{for } z \in D,$$

where h is a harmonic function in D defined in the second part of Lemma 10.15. We call $\tilde{u}_d$ the *harmonic majorant* of u rel ∂D. The first part of Lemma 10.15 implies that *the harmonic majorant of* u *is subharmonic.*

A function u is called *superharmonic* if $-u$ is subharmonic. Properties of such functions follow immediately form the corresponding properties of subharmonic ones.

[27]It is still true in general, but we will not need it

10.6. Perron method. A (non-empty) family $\mathcal{P}$ of continuous subharmonic functions on a Riemann surface U is called *Perron* if it satisfies the following properties:

(i) If $u, v \in \mathcal{P}$ then $\max(u, v) \in \mathcal{P}$;
(ii) For any $u \in \mathcal{P}$ and any coordinate disk $D \Subset U$, the harmonic majorant $\tilde{u}_D$ also belongs to $\mathcal{P}$.

Proposition 10.16. *If $\mathcal{P}$ is a Perron family on U then the function*

$$h(z) := \sup_{\mathcal{P}} u(z)$$

is either harmonic or identically equal to ∞.

Proof. Since harmonicity is a local property, it is enough to check it within coordinate disks $D \Subset U$. Fix such a disk D. Since $u \le \tilde{u} \in \mathcal{P}$, we have $h(z) := \sup_{\mathcal{P}} \tilde{u}(z)$. So, without loss of generality we can assume that all the functions $u \in \mathcal{P}$ are *harmonic* in D.

Take a countable dense subset $X \subset D$. By means of the diagonal procedure, we can select a sequence of functions $u_n \in \mathcal{P}$ such that $h(z) = \sup u_n(z)$ for any $z \in X$. Let v_n be the harmonic majorant (rel ∂D) of the function $\max(u_1, \dots, u_n)$, $n \in \mathbb{Z}_+$. This is a monotonically increasing sequence of functions of the family $\mathcal{P}$, harmonic on D, and such that $v_n(z) \to h(z)$ on X. By Corollary 10.10, $v_n \to \phi$ locally uniformly on D, where ϕ is either harmonic, or else $\phi \equiv \infty$. In either case, we have:

$$\phi(z) = h(z) \ge u(z) \quad \text{for any } z \in X,\ u \in \mathcal{P}.$$

Since both ϕ and u are continuous, we conclude that $\phi \ge u$ everywhere on D; hence $\phi \ge h$ everywhere on D. On the other hand, since $\phi = h$ on the dense set X and h is upper semicontinuous (as sup of a family of continuous functions), we conclude that $\phi \le h$ everywhere on D. Thus $\phi \equiv h$ on D. □

10.7. Dirichlet barriers. We will now apply the Perron method to solve the Dirichlet problem in an arbitrary domain (for which it is solvable at all).

Let $U \Subset S$ be a domain in a Riemann surface S, and let g be a continuous function on ∂U. Let us consider the following Perron family of subharmonic functions:

$$\mathcal{P} \equiv \mathcal{P}_U(g) = \{u \in \mathcal{SH}(U) :\ \limsup_{\zeta \to z} u(\zeta) \le g(z) \quad \forall z \in \partial U\}.$$

By Proposition 10.16, the function $h_g := \sup_{\mathcal{P}} u$ is harmonic in U. To study its boundary values, we will introduce the following notions:

A *barrier* b_a at a boundary point $a \in \partial U$ is a subharmonic function $b_a(z)$ defined on a relative neighborhood D of a in U, continuous up to ∂D,[28] and such that $b_a(a) = 0$ while $b_a(z) < 0$ for any $z \ne a$. A point $a \in \partial U$ is called *Dirichlet regular* if it has a barrier.

Example 10.17. *If ∂U near a is an arc of a smooth curve then a is regular. Indeed, then there is a wedge*

$$W = \{|\arg(z - a) - \alpha| < \varepsilon,\ 0 < |z| < 2\pi\,\varepsilon\}$$

[28]This condition can be relaxed, but it is sufficient for our purposes. In fact, *harmonic* barriers would also be good enough for us.

which is disjoint from $\overline{U}$. The complementary wedge can be mapped conformally onto the lower half-plane (by a branch of the power function $\phi(z) = e^{i\theta}(z-a)^\gamma$ with appropriate $\gamma \in (0,1)$ and θ. The function $b = \operatorname{Im}\phi(z)$ restricts to a barrier at a on U.

EXERCISE 10.18. *Show that the same is true is ∂U near a is a Jordan arc.*

THEOREM 10.19. *Let $U \Subset S$ be a domain in a Riemann surface S, and let g be a continuous function on ∂U. Let us consider the harmonic function $h = h_g$ constructed above by means of the Perron method. Then for any Dirichlet regular point $a \in \partial U$, we have: $h(z) \to g(a)$ as $z \to a$.*

PROOF. Without loss of generality, we can assume that $g(a) = 0$.

Let us first show that

$$\liminf_{z\to a} h(z) \geq 0. \tag{10.7}$$

Take a small $r > 0$ such that the barrier $b(z) = b_a(z)$ is well defined in $D_r := \mathbb{D}(a, 2r) \cap U$. Let ξ be the supremum of b on $S_r := \{|z-a| = r\} \cap U$. By definition of the barrier, $\xi < 0$.

The function $\hat{b}(z) := \max(b(z), \xi)$ is a continuous subharmonic function in $\mathbb{D}(a, r) \cap U$ equal to ξ on S_r. Hence it extends to a continuous subharmonic function in in U by letting $\hat{b} \equiv \xi$ in $U \smallsetminus D_r$.

Let now

$$\eta = \inf\{g(z) :\ z \in \partial U \smallsetminus D_r\}, \quad -\varepsilon = \inf\{g(z) :\ z \in \partial U \cap D_r\} < 0,$$

and consider

$$\beta(z) = \frac{\eta}{\xi}\,\hat{b}(z) - \varepsilon.$$

This is a subharmonic function in U with

$$\lim_{z\to p} \beta(z) = \eta \quad \text{for } p \in \partial U \smallsetminus D_r; \quad \liminf_{z\to p} \beta(z) \leq -\varepsilon \quad \text{for } p \in \partial U \cap D_r,$$

so β belongs to the Perron family $\mathcal{P}$.

It follows that $h \geq \beta$ and hence

$$\liminf_{z\to a} h(z) \geq -\varepsilon.$$

Since $\varepsilon \to 0$ as $r \to 0$, we obtain (10.7).

To obtain the opposite estimate, let us consider the negative barrier $-b(z)$. It allows us to construct, for any $\varepsilon > 0$, a *superharmonic* function α in U such that

$$\liminf_{z\to p} \alpha(z) \geq g(p)\ \forall p \in \diamond U \quad \text{and} \quad \limsup_{z\to a} \alpha(z) \leq \varepsilon.$$

By the Maximum Principle, $u \leq \alpha$ for any $u \in \mathcal{P}$, and hence $h \leq \alpha$ as well. It follows that

$$\limsup_{z\to a} h(z) \leq \varepsilon,$$

and we are done. □

We say that a domain $U \Subset S$ has a *Dirichlet regular boundary* if ∂U is non-empty and all points of ∂U are regular.

COROLLARY 10.20. *Let $U \Subset S$ be a domain with Dirichlet regular boundary. Then the Dirichlet problem is solvable in U for any continuous boundary values.*

10.8. Harmonic measure. Let $U \Subset S$ be a domain with Dirichlet regular boundary. Then any continuous function $g \in C(\partial U)$ admits a harmonic extension $\hat{g} \in \mathcal{H}(\overline{U})$ to U. Endow $\mathcal{H}(\overline{U})$ with uniform topology on the whole $\overline{U}$. It is a Banach space isomorphic to $C(\partial U)$ by means of the natural restriction and the above extension operators.

For a given $z \in U$, evaluation $\hat{g}(z)$ is a bounded linear functional on $C(\partial U)$ and hence it is represented by a Borel measure μ_z on ∂U:

$$\hat{g}(z) = \int_{\partial U} g \, d\mu_z.$$

This measure is called the *harmonic measure* for U at z. For instance, in the unit disk, we have $d\mu_z = \mathcal{P}(z, \zeta)\, d\theta$ where $\mathcal{P}$ is the Poisson kernel.

If ∂U is disconnected and $K \subset \partial U$ is a clopen subset then $\mu_z(K)$ is a harmonic function on U with boundary values 1 on K and 0 on $\partial U \smallsetminus K$. This function itself is sometimes referred to as the "harmonic measure of K" (which may sound confusing).

10.9. Green function. We will restrict our discussion to domains $U \Subset S$ with Dirichlet regular boundary. The *Green function* $G = G_p$ *on* U *with pole at* $p \in U$ is a harmonic function such that

(Gr1) $G(z) \to 0$ as $z \to \partial U$;
(Gr2) In a local coordinate z near p such that $z(p) = 0$, we have:

$$G(z) = \log \frac{1}{|z|} + O(1) \quad \text{near } p.$$

For instance, the Green function in $\mathbb{D}$ with pole at 0 is $-\log |z|$.

Remark 10.21. Obviously, existence of such a function G implies the Dirichlet regularity of U as $-G$ provides a barrier at any boundary point. In the non-regular case, condition (Gr1) can be relaxed so that the Green function still exists as long as ∂U has positive capacity.

Remark 10.22. The Green function has a clear electrostatical meaning as the potential of the unit charge placed at p in a domain bounded by a conducting material with the ground potential 0.

The level sets of the Green function G_p are called *equipotentials*, its gradient lines are called *rays* (emanated from p). They form two orthogonal foliations on $U \smallsetminus \{p\}$ with singularities at the critical points of G_p

Theorem 10.23. *Let $U \Subset S$ be a domain in a Riemann surface S with Dirichlet regular boundary. Then for any $p \in U$, there exists a unique Green function G_p with pole at p.*

Proof. Let us consider the following family $\mathcal{P} = \mathcal{P}_U[p]$ of functions on $U \smallsetminus \{p\}$:

(i) $\limsup_{z \to \partial U} u(z) \leq 0$;
(ii) In a local coordinate z near p such that $z(p) = 0$, we have:

$$u(z) = \log \frac{1}{|z|} + O(1).$$

Obviously, it is a Perron family, so the function $G = \sup_{\mathcal{P}} u$ is harmonic in $U \smallsetminus \{p\}$ unless it is identically equal to ∞. We will show that this function is actually finite, and it is the desired Green function.

First, $\mathcal{P}$ is non-empty. Indeed, for a small $r > 0$, the function $u_0 := \log^+(r/|z|)$ (equal to $\log(|z|/r)$ on the coordinate disk $D(p,r)$ and extended by 0 the whole U) is in $\mathcal{P}$. Thus,

$$G(z) \geq \log^+ \frac{|z|}{r} \geq 0. \tag{10.8}$$

Let us show that G is finite. Let S_r be the coordinate circle centered at p of radius r, and let $\|u\|_r$ be the sup-norm of a function u on S_r. Let us fix two small radii $0 < r < R$ and compare $\|u\|_r$ and $\|u\|_R$ for $u \in \mathcal{P}$.

First, let us look at u from "inside". Take a small $\varepsilon > 0$ and let

$$u_\varepsilon(z) = u(z) + (1+\varepsilon)\log|z|.$$

This function is subharmonic in $D(p,R) \smallsetminus \{p\}$ and equal to $-\infty$ at p (by property (ii) of the family $\mathcal{P}$). Hence it is subharmonic on the whole disk $D(p,R)$. By the Maximum Principle, $\|u_\varepsilon\|_r \leq \|u_\varepsilon\|_R$, so

$$\|u\|_r \leq \|u\|_R + (1+\varepsilon)\log\frac{R}{r}.$$

Letting $\varepsilon \to 0$, we obtain

$$\|u\|_r \leq \|u\|_R + \log\frac{R}{r}. \tag{10.9}$$

On the other hand, we can look at u from "outside". The Maximum Principle in $S \smallsetminus D(z,r)$ implies that for any $u \in \mathcal{P}$

$$\|u\|_R < \|u\|_r, \tag{10.10}$$

but we want to have a definite drop:

$$\|u\|_R \leq \lambda\,\|u\|_r \tag{10.11}$$

with some $\lambda < 1$ independent of u. Together with (10.9), this would imply

$$\|u\|_r \leq \frac{1}{1-\lambda}\log\frac{R}{r}$$

that would prove finiteness of G on S_r and hence everywhere on U.

To prove (10.11), let us consider the solution v of the Dirichlet problem in $U \smallsetminus D(z,r)$ with boundary values 1 on S_r and 0 on ∂U (the "harmonic measure" of S_r). Since the boundary of ∂U is regular by assumption and S_r is regular as a smooth curve, such a v exists (Corollary 10.20). By the Maximum Principe, $\lambda := \|v\|_R < 1$.

Furthermore, the function $u(z)$ is asymptotically majorated by $\|u\|_r\, v(z)$ near the boundary of $S \smallsetminus D(p,r)$. By the Maximum Principle,

$$u(z) \leq \|u\|_r\, v(z), \quad z \in \smallsetminus D(p,r). \tag{10.12}$$

Taking its sup on S_R, we obtain (10.11).

The required properties of the Green function also follow from the above estimates. Indeed, (10.8) and (10.12) imply (Gr1), while (10.8) and (10.9) imply (Gr2).

$\square$

Notice in conclusion that the Green function extends subharmonically to the whole Riemann surface S by letting $G \equiv 0$ on $S \smallsetminus U$.

EXERCISE 10.24. *The Green function has a critical point in U if and only if U is not simply connected.*

Notes

Conformal Schönflies Theorem was proven in [**Ca1**]. The theory of prime ends and Carathéodory Boundary Theorem appeared in [**Ca2**]. "Carathéodory-Torhorst Theorem" is usually attributed to Carathéodory. In fact, in the above two papers that Carathéodory wrote on the subject, there is no mentioning of local connectivity or the problem of continuity up to the boundary of the inverse Riemann mapping. The theorem was proven in Torhorst's thesis in 1918 (see [**To**]) written under advice of Hahn who introduced the notion of local connectivity in 1913. See Lasse Rempe [**Re1**] for an account of this story.

The notion of Schwarzian derivative goes back at least to Riemann.

CHAPTER 2

Quasiconformal geometry

11. Analytic definition and regularity properties

11.1. Linear discussion.

11.1.1. *Teichmüller metric on the space of conformal structures.* Let $V \approx \mathbb{R}^2$ be a real two-dimensional vector space. A *conformal structure* μ on V is a Euclidean structure $(v, w)_\mu$ up to scaling. Equivalently, it is an ellipse $E_\mu = \{\|w\|_\mu = 1\}$ centered at the origin, up to scaling (here $\|w\|_\mu$ is the associated Euclidean norm). Let $\mathrm{Conf}(V)$ stand for the space of conformal structures on V.

Let us consider two Euclidean structures, $(v, w)_\mu$ and $(v, w)_\nu$ representing conformal structures μ and ν. We define the *Teichmüller distance* between μ and ν as the distortion of one Euclidean norm with respect to the other:

$$\mathrm{dist}_{\mathrm{T}}(\mu, \nu) = \log\left(\max_{w\in V^*} \frac{\|w\|_\mu}{\|w\|_\nu} : \min_{w\in V^*} \frac{\|w\|_\mu}{\|w\|_\nu}\right) \text{ where } V^* = V \smallsetminus \{0\}.$$

Note that it is independent of the the choice of Euclidean structures representing μ and ν.

Exercise 11.1. *Check that* $\mathrm{dist}_{\mathrm{T}}$ *is a metric on* $\mathrm{Conf}(V)$.

If we simultaneously diagonalize the Euclidean structures so that

$$\|w\|_\nu^2 = x^2 + y^2,\ \|w\|_\mu^2 = x^2/a^2 + y^2/b^2,\ \text{where } w = (x, y),\ a \geq b > 0,$$

then

$$\mathrm{dist}_{\mathrm{T}}(\mu, \nu) = \log(a/b) \equiv \log K.$$

The ratio $K = a/b$ of the axes of the ellipse E_μ is called the *dilatation* of μ relative to ν. We denote it $\mathrm{Dil}(\mu : \mathbf{n})$, skipping ν if it is the standard conformal structure. Informally we can say that the Teichmüller distance measures the *relative shape* of the ellipses representing our conformal structures.

An invertible linear operator $A : V' \to V$ induces a natural *pullback* operator $A^* : \mathrm{Conf}(V) \to \mathrm{Conf}(V')$: If $(v, w)_\mu$ is the Euclidean structure representing $\mu \in \mathrm{Conf}(V)$ then the pullback $A^*\mu$ is represented by $(Av, Aw)_\mu$. It follows immediately from the definitions that the *Teichmüller metric is preserved by the pullback transformations*.

In particular, the group $\mathrm{GL}(V)$ of linear automorphisms of V isometrically acts on $\mathrm{Conf}(V)$ on the right: $\mu A := A^*\mu$. Let us restrict this action to the group $\mathrm{GL}_+(V)$ of orientation preserving automorphisms. Since this action is transitive, it turns $\mathrm{Conf}(V)$ into a $\mathrm{GL}_+(V)$-homogeneous space.

To understand this space, let us fix some reference conformal structure σ and select coordinates (x, y) on V that bring it to the standard form $x^2 + y^2$. Then

$\mathrm{GL}_+(V)$ gets identified with $\mathrm{GL}_+(2,\mathbb{R})$, and the isotropy group of σ gets identified with the group $\mathrm{Sim}(2)$ of similarities. Hence

$$\mathrm{Conf}(V) \approx \mathrm{Sim}(2) \diagdown \mathrm{GL}_+(2,\mathbb{R}) = \mathrm{SO}(2) \diagdown \mathrm{SL}(2,\mathbb{R}). \tag{11.1}$$

Recall that in §2.4.3 we endowed the symmetric space $\mathrm{SO}(2) \diagdown \mathrm{SL}(2,\mathbb{R})$ with an invariant metric.

EXERCISE 11.2. *This invariant metric coincides with the Teichmüller metric on* $\mathrm{Conf}(V)$*.*

But according to Exercise 2.19, the hyperbolic plane $\mathbb{H}$ is naturally isometric to the symmetric space

$$\mathrm{PSL}(2,\mathbb{R})/\mathrm{PSO}(2) \approx \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2).$$

Since the left and right symmetric spaces are equivariantly isometric by the inversion $A \mapsto A^{-1}$, we conclude:

PROPOSITION 11.3. *The space* $\mathrm{Conf}(V)$ *endowed with the Teichmüller metric is equivariantly isometric to the hyperbolic plane* $\mathbb{H}$*.*

REMARK 11.4. As we have already mentioned in §2.4.3, the Lie Theory provides a general underlying principle for the hyperbolicty of the symmetric space (11.1) without *a priori* familiarity with the hyperbolic plane.

In conclusion, let us give one more interpretation of the isomorphism (11.1). It is obtained by associating to an operator $A \in \mathrm{GL}_+(2,\mathbb{R})$ the conformal structure μ represented by the Euclidean structure $(v,w)_\mu = (Av, Aw)$ (where (v,w) is the standard Euclidean structure on $\mathbb{R}^2$). The corresponding ellipse E_μ is the pullback of the standard round circle: $E_\mu = A^{-1}(\mathbb{T})$.

Making use of the polar decompositions of linear operators, we can uniquely represent A as a product of a positive self-adjoint operator P and a rotation O, $A = O \cdot P$. Let $\lambda_{\max} \geq \lambda_{\min} > 0$ stands for the eigenvalues of P. The operator A is a similarity if and only if P is scalar, i.e., $\lambda_{\max} = \lambda_{\min}$. Otherwise we have two orthogonal (uniquely defined) eigenlines $l_{\max}$ and $l_{\min}$ corresponding to $\lambda_{\max}$ and $\lambda_{\min}$ respectively. These lines give the directions of maximal and minimal expansion for the operator A. Moreover, the ellipse $E_\mu = A^{-1}(\mathbb{T}) = P^{-1}(\mathbb{T})$ has the big axis of length $1/\lambda_{\min}$ in $l_{\min}$ and the small axis of length $1/\lambda_{\max}$ in $l_{\max}$.The dilatation of this ellipse (equal to $\lambda_{\max}/\lambda_{\min}$) will be also called the *dilatation* of A, $\mathrm{Dil}\, A$.

EXERCISE 11.5. *Show that* $\mathrm{Dil}\, A^{-1} = \mathrm{Dil}\, A$ *and* $\mathrm{Dil}(AB) \leq \mathrm{Dil}\, A \,\mathrm{Dil}\, B$ *with equality attained iff the eigenlines of A and B^{-1} coincide.*

11.1.2. *Beltrami coefficients.* Let now $V = \mathbb{C}_\mathbb{R}$ be the decomplixified $\mathbb{C}$. It is endowed with the standard conformal structure σ (represented by the Euclidean metric $|z|^2$) and with the standard orientation (such that $\{1, i\}$ is positively oriented). Let $A : \mathbb{C}_\mathbb{R} \to \mathbb{C}_\mathbb{R}$ be an invertible $\mathbb{R}$-linear operator (which can be also viewed as a $\mathbb{C}$-valued $\mathbb{R}$-linear form on V).

Let us describe the conformal structure $A^*\sigma$ in coordinates $z, \bar{z}$ of $\mathbb{C}_\mathbb{R}$. The operator A can be represented as

$$z \mapsto az + b\bar{z} = az(1 + \mu\frac{\bar{z}}{z}), \tag{11.2}$$

where $\mu = b/a$ is called the *Beltrami coefficient* of A. Let $\mu = |\mu|e^{2i\theta}$, where $\theta \in \mathbb{R}/\pi\mathbb{Z}$.

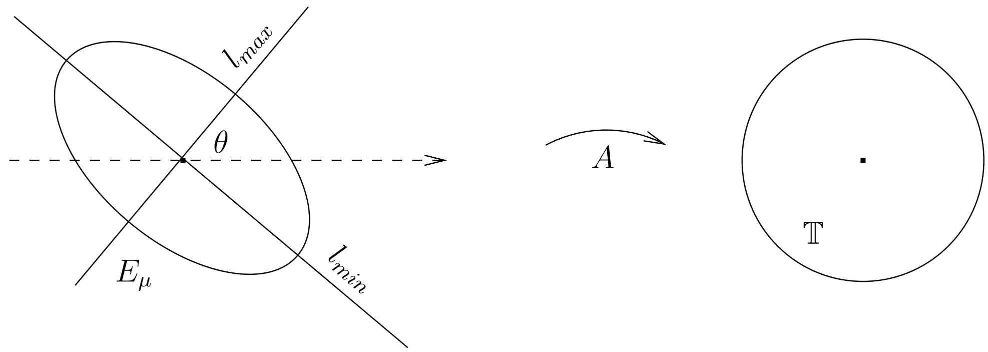

FIGURE 11.1. Ellipse E_μ.

EXERCISE 11.6. *The operator A is conformal iff $\mu = 0$; A is invertible iff $|\mu| \neq 1$; A is orientation preserving iff $|\mu| < 1$.*

REMARK 11.7. Note that the compex conjugation $\overline{Az}$ does not affect the metric $|Az|^2 = |az + b\bar{z}|^2$, but it replaces the Beltrami coefficient $\mu = b/a$ with the $\mathbb{T}$-symmetric one, $1/\bar{\mu}$. In shows that any Euclidean metric on $\mathbb{C}_R$ is conformally equivalent to a metric $|z + \mu\bar{z}|^2$ with $|\mu| < 1$ (corresponding to the orientation preserving operators A).

In what follows we assume that A is an invertible orientation preserving operator, i.e., $|\mu| < 1$. If we have another such a form $A' = a'z + b'\bar{z}$ on V then $A'/A \equiv \text{const}$ iff $\mu = \mu'$. Thus, the *conformal structures* $A^*\sigma$ are in one-to-one correspondence with the Beltrami coefficient $\mu \in \mathbb{D}$, so $\mathrm{Conf}(V) \approx \mathbb{D}$.

Let us now describe the shape of the ellipse $A^{-1}(\mathbb{T})$ in terms of μ. The maximum of $|Az|$ on the unit circle $\mathbb{T} = \{z = e^{i\phi}\}$ is attained at the direction $\phi = \theta \mod \pi\mathbb{Z}$, while the minimum is attained at the orthogonal direction $\theta + \pi/2 \mod \pi\mathbb{Z}$. These are the eigenlines $l_{\max}$ and $l_{\min}$ of the positive part P of A. The corresponding eigenvalues are equal to

$$\lambda_{\max} = |a|(1 + |\mu|) = |a| + |b|, \quad \lambda_{\min} = |a|(1 - |\mu|) = |a| - |b|.$$

Thus

$$\mathrm{Dil}\, A = \frac{1 + |\mu|}{1 - |\mu|}, \quad \det A = |a|^2 - |b|^2 = \lambda_{\min}^2 \,\mathrm{Dil}\, A. \tag{11.3}$$

This gives us a description of the dilatation and direction of the ellipse $E = A^{-1}(\mathbb{T})$ in terms of $|\mu|$ and $\arg \mu$ respectively.

Under conformal changes of variable, $z = T\zeta = \alpha\zeta$ ($\alpha \in \mathbb{C}^*$) the Beltrami coefficients is rotated: $\nu := T^*\mu = (\bar{\alpha}/\alpha)\mu$, while the $(-1, 1)$-form

$$\mu\frac{\bar{z}}{z} = \nu\frac{\bar{\zeta}}{\zeta}$$

does not change. It shows that the Beltrami coefficients in various conformal coordinates represent a single $(-1, 1)$-*Beltrami form*.

EXERCISE 11.8. *Under a general linear change of variable $z = T\zeta = \alpha\zeta + \beta\bar{\zeta}$, Beltrami coefficeints are transformed as follows:*

$$T^*\mu = \frac{\bar{\alpha}\mu + \beta}{\bar{\beta}\mu + \alpha}.$$

It follows that the map $M : \mathrm{Conf}(\mathbb{C}_{\mathbb{R}}) \to \mathbb{D}$ that associates to a conformal structure its Beltrami coefficient is $\mathrm{PSL}(2,\mathbb{R})$-*equivariant.*

EXERCISE 11.9. *The map M is an equivariant isometry between* $\mathrm{Conf}(\mathbb{C}_{\mathbb{R}})$ (*endowed with the Teichmüller metric*) *and the disk $\mathbb{D}$* (*endowed with the hyperbolic metric*)*.*

In what follows we will feel free to identify conformal structures with the corresponding Beltrami forms (and in a particular coordinate, with the corresponding Beltrami coefficients). We will often use the same notation for these objects.

11.1.3. *Infinitesimal notation.* Let us now interpret the above discussion in infinitesimal terms. Consider a map $h : U \to \mathbb{C}$ on a domain $U \subset \mathbb{C}$ differentiable at a point $z \in U$, and apply the above considerations to its differential $Dh(z) : \mathrm{T}_z U \to \mathrm{T}_{hz}\mathbb{C}$. In the $(dz, d\bar{z})$-coordinates of the tangent spaces, it assumes the form

$$\partial h + \bar{\partial} h = \partial_z h\, dz + \partial_{\bar{z}} h\, d\bar{z},$$

where the partial derivatives ∂_z and $\bar{\partial}_z$ and the operators ∂ and $\bar{\partial}$ are defined in §2.11. Moreover,

$$Dh(z) = \partial_z h(z)\, dz \left(1 + \mu_h(z) \frac{d\bar{z}}{dz}\right),$$

where $\mu_h = \partial_{\bar{z}} h / \partial_z h$ is the *Beltrami coefficient* of h at z. In fact, as was explained above, this coefficient represents a $(-1,1)$-form

$$\bar{\partial} h / \partial h = \mu_h \frac{d\bar{z}}{dz}$$

called the *Beltrami differential* of h at z. However, in what follows we will not make a notational difference between the Beltrami differential and the coefficient (and will usually use notation ∂, $\bar{\partial}$ for the partial derivatives ∂_z, $\partial_{\bar{z}}$).

Assume that $Dh(z)$ is non-singular and orientation preserving, i.e., $|\mu_h| < 1$. The map h is conformal at z if and only if $\mu_h(z) = 0$, which is equivalent to the *Cauchy-Riemann equation* $\bar{\partial} h(z) = 0$.

Let us consider an infinitesimal ellipse

$$E_h(z) \equiv Dh(z)^{-1}(\mathbb{T}_{hz}) \subset \mathrm{T}_z U, \tag{11.4}$$

where $\mathbb{T}_{hz}$ is a round circle in the tangent space $\mathrm{T}_{hz} U$. If h is not conformal at z, then $E_h(z)$ is a genuine (not round) ellipse with the small axis in the direction $\arg(\mu_h(z))/2 \mod \pi$ and the shape

$$\mathrm{Dil}(h, z) = \frac{1 + |\mu_h(z)|}{1 - |\mu_h(z)|}. \tag{11.5}$$

Moreover, by the second formula of (11.3), we have:

$$\mathrm{Jac}(h, z) = |\partial h(z)|^2 - |\bar{\partial} h(z)|^2 = \lambda_{\min}(z)^2 \ \mathrm{Dil}(h, z), \tag{11.6}$$

where $\mathrm{Jac}(h, z) \equiv \det Dh(z)$ and $\lambda_{\min}(z) = \inf\limits_{|v|=1} |Dh(z)\, v|$.

If we have a differentiable change of variable $z = \phi(\zeta)$ then infintesimal ellipses E and the corresponding Beltrami differentials μ at T_z can be pulled back to T_ζ. According to Exercise 11.8, the corresponding tranformation rule is:

$$(\phi^*\mu)(\zeta) = \frac{\overline{\partial_z\phi(\zeta)}\cdot\mu(z) + \partial_{\bar z}\phi(\zeta)}{\overline{\partial_{\bar z}\phi(\zeta)}\cdot\mu(z) + \partial_z\phi(\zeta)} \text{ or more concisely : } \phi^*\mu = \frac{\overline{\partial_z\phi}\cdot(\mu\circ\phi) + \partial_{\bar z}\phi}{\overline{\partial_{\bar z}\phi}\cdot(\mu\circ\phi) + \partial_z\phi}.$$

In the orientation preserving case, ϕ^* preserves the hyperbolic distance between Beltrami differentials with $|\mu(z)| < 1$. In the conformal case, we have:

$$(\phi^*\mu)(\zeta) = \frac{\overline{\phi'(\zeta)}}{\phi'(\zeta)}\cdot\mu(z) \quad \text{or concisely :} \quad \phi^*\mu = \frac{\overline{\phi'}}{\phi'}\cdot(\mu\circ\phi).$$

11.2. Measurable conformal structures. A (measurable) *conformal structure* on a domain $U \subset \mathbb{C}$ is a measurable family of conformal structures in the tangent planes T_zU, $z \in U$. In other words, it is a measurable family $\mathcal{E}$ of infinitesimal ellipses $E(z) \subset T_zU$ defined up to scaling by a measurable function $\rho(z) > 0$, $z \in U$. (As always in the measurable category, all the above objects are defined almost everywhere.) According to the linear discussion, any conformal structure is determined by its *Beltrami coefficient* $\mu(z)$, $z \in U$, a measurable function in z assuming its values in $\mathbb{D}$, and vice versa. Thus, conformal structures on U are described analytically as elements μ from the unit ball of $L^\infty(U)$. We say that a conformal structure has a *bounded dilatation* if the dilatation of the ellipses $E(z)$ are bounded almost everywhere. In terms of Beltrami coefficients, it means that $\|\mu\|_\infty < 1$ since

$$\operatorname{Dil}\mu := \|\operatorname{Dil}E(z)\|_\infty = \frac{1+\|\mu\|_\infty}{1-\|\mu\|_\infty}.$$

The *standard* conformal structure $\sigma \equiv \sigma_U$ is given by the family of infinitesimal circles. The corresponding Beltrami coefficient vanishes almost everywhere: $\mu = 0$ in $L^\infty(U)$.

Remark 11.10. Sometimes $\|\mu\|_\infty$ is also referred to as the *dilatation* of μ. We will reserve notation $\operatorname{dil}\mu$ for this occasion. Then "bounded dilatation" would mean that $\operatorname{dil}\mu < 1$.

The space of conformal structures on U with bounded dilatation is endowed with the *Teichmüller metric*:

$$\operatorname{dist}_{\mathrm{T}}(\mu,\nu) = \|\operatorname{dist}_{\mathrm{T}}(\mu(z),\nu(z))\|_\infty.$$

Remark 11.11. Since the right-hand side in the above formula depends only on the real structure on the tangent spaces, we do not need the reference complex structure on U to define the Teichmüller metric.

Denote by $\mathrm{DHomeo}^+(U,V)$ (standing for "differentiable homeomorphisms") the space of orientation preserving homeomorphisms $h : U \to V$ that are differentiable almost everywhere with a non-singular differential $Df(z)$ measurably depending on z.[1] Consider some homeomorphism $h \in \mathrm{DHomeo}^+(U,V)$ between two domains in $\mathbb{C}$. Then by the above linear discussion we obtain a measurable family $\mathcal{E}$ of infinitesimal ellipses $E_h(z) = Dh(z)^{-1}(\mathbb{T}_{hz}) \subset \mathrm{T}_zU$ that determines a (measurable) conformal structure $\mu_h = h^*\sigma$ on U. Analytically this structure can be described picture

[1]If we do not need to specify the domain and the range of h we write simply $h \in \mathrm{DHomeo}^+$; if we do not assume that h is orientation preserving, we skip "+".

as the Beltrami coefficient $\mu_h(z) = \bar\partial h(z)/\partial h(z)$ of h. We say that h *has a bounded dilatation* if the corresponding conformal structure $h^*\sigma_V$ does. In this case we let

$$\operatorname{Dil} h := \|\operatorname{Dil}(h,z)\|_\infty = \frac{1+\|\mu_h\|_\infty}{1-\|\mu_h\|_\infty} = \operatorname{dist}_{\mathrm{T}}(h^*\sigma_V, \sigma_U).$$

Obviously, the pullback structure $h^*\sigma$ does not change if we *postcompose* h with a conformal map ϕ. If we *precompose* h with a conformal map ϕ then the Beltrami coefficient will be transformed as follows:

$$\mu_{h\circ\phi} = \frac{\overline{\phi'}}{\phi'}\,\mu_h\circ\phi,$$

so that the Beltrami coefficients in various local charts represent a single $(-1,1)$-form $\mu\, d\bar z/dz$ called the Beltrami differential of h (compare §11.1.3).

This allows us to generalize the above discussion to arbitrary Riemann surfaces. A (measurable) conformal structure on a Riemann surface S is a measurable family $\mathcal{E}$ of infinitesimal ellipses $E(z)$ defined up to scaling. Analytically it is described as a measurable *Beltrami differential* μ with $|\mu(z)|<1$ a.e. To any homeomorphism $h\in \operatorname{DHomeo}^+(S,S')$ between two Riemann surfaces corresponds the pullback structure $h^*\sigma_{S'}$ represented by the field of ellipses $E_h(z) = Dh(z)^{-1}(\mathbb{T}_r)$.[2] The corresponding *Beltrami differential* is $\mu_h = \bar\partial h/\partial h$ (where $\bar\partial h$ and ∂h are now viewed as 1-forms).

Remark 11.12. (i) Once again, measurable conformal structures can be considered on arbitrary *smooth* surfaces as well, with the dilatation measured with respect to a reference Riemannian metric (and on a compact surface, the virtue of being bounded does not depend on the choice of the reference metric.) The space of bounded structures is endowed with the Teichmüller metric. Moreover, this discussion can be further promoted to *quasiconformal* surfaces.

(ii) A key problem is whether any conformal structure μ is associated to a certain map h. This problem has a remarkable positive solution in the category of quasiconformal maps (see §14 below).

More generally, let us consider a (non-invertible) map $f: U\to V$ which is differentiable for a.e. $z\in U$ in the classical sense with non-singular $Df(z)$. For such maps the push-forward operation is not well-defined, but the pullback $\nu = f^*\mu$ and $\operatorname{Dil} f$ are still well-defined. The property that $\operatorname{Dil}(f^*\mu)\le \operatorname{Dil}(f)\cdot\operatorname{Dil}(\mu)$ is obviously valid in this generality. This observation will be used in the context of quasiregular maps: see §29.1.1.

11.3. Analytic definition. We are now ready to give a definition of quasiconformality. An orientation preserving homeomorphism $h: S\to S'$ between two Riemann surfaces is called *quasiconformal* if

Q1. It has locally integrable distributional partial derivatives;

Q2. It has bounded dilatation, i.e., $\bar\partial h \le k\,\partial h$ a.e. for some $k\in[0,1)$.

Note that the second property makes sense because the first property implies that h is differentiable a.e. in the classical sense (by Proposition 11.18).

We will often abbreviate "quasiconformal" as "qc". A qc map h is called *K-qc* if $\operatorname{Dil} h\le K$, where $K=(k+1)/(k-1)\in[1,\infty)$ with $k\in[0,1)$ as above.

[2]Note that the ellipses $E_h(z)$ are defined only up to scaling since the round circles $\mathbb{T}_r$ on S' are (as there is no preferred metric on S').

REMARK 11.13. 1) Under the above circumstances, the quantification k-qc is sometimes more convenient (so that $\operatorname{dil} h \le k < 1$), but to avoid confusion, we will refrain from using it.

2) One of the problems with the above analytic definition is that property Q1 is not symmetric under taking the inverse h^{-1}. Neither it is invariant under taking compositions. But we will see eventually that the definition is equivalent to a geometric one, quasi-invariance of moduli (see QC2 in §12.5), that manifestly has both virtues.

11.4. Absolute continuity and Sobolev class $\mathcal{W}$. We will now prove several important regularity properties of quasiconformal maps. Recall the definition of the Sobolev class $\mathcal{W}_{\text{loc}}(U) \equiv \mathcal{W}^2_{\text{loc}}(U)$ from the Appendix to this section.

PROPOSITION 11.14. *Let $h : U \to V$ be a qc map. Then h^{-1} is absolutely continuous with respect to the Lebesgue measure,*[3] *and thus for any Borel set $X \subset U$,*

$$m(h(X)) = \int_X \operatorname{Jac}(h, z)\, dm.$$

The partial derivatives ∂h and $\bar{\partial} h$ belong to $L^2_{\text{loc}}(U)$, so $h \in \mathcal{W}^2_{\text{loc}}(U)$.

PROOF. Since both statements are local, we can restrict ourselves to homeomorphisms $h : U \to V$ between bounded domains in the complex plane. Consider the pullback of the Lebesgue measure on V, $\mu = h^* m$. It is a Borel measure defined as follows: $\mu(X) = m(h(X))$ for any Borel set $X \subset U$. Let us decompose it into absolutely continuous and singular parts: $\mu = \rho \cdot m + \nu$. By the Lebesgue Density Points Theorem, for almost all $z \in U$, we have:

$$\frac{1}{\pi\varepsilon^2} \int_{\mathbb{D}(z,\varepsilon)} \rho\, dm \to \rho(z); \quad \frac{1}{\pi\varepsilon^2} \nu(\mathbb{D}(z, \varepsilon)) \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Summing up we obtain:

$$\frac{m(h(\mathbb{D}(z, \varepsilon))}{m(\mathbb{D}(z, \varepsilon))} = \frac{\mu(\mathbb{D}(z, \varepsilon)}{m(\mathbb{D}(z, \varepsilon)} \to \rho(z) \text{ a.e.} \quad \text{as} \quad \varepsilon \to 0.$$

But if h is differentiable at z then the left hand-side of the last equation goes to $\operatorname{Jac}(h, z)$. Hence $\operatorname{Jac}(h, z) = \rho(z)$ a.e. It follows that for any Borel set X,

$$\int_X \operatorname{Jac}(h, z)\, dm = \int_X \rho\, dm \le \mu(X) = m(hX). \tag{11.7}$$

But $\operatorname{Jac}(h, z) = |\bar{\partial} h(z)|^2 - |\partial h(z)|^2 \ge (1 - k^2)\, |\partial h(z)|^2$, where $k = \|\mu_h\|_\infty$. Thus

$$\int_X |\partial h|^2\, dm \le \frac{1}{1 - k^2}\, m(hX); \quad \int_X |\bar{\partial} h|^2\, dm \le \frac{k^2}{1 - k^2}\, m(hX), \tag{11.8}$$

and we see that the partial derivatives of h are locally square integrable.

What is left is to prove the opposite to (11.7). As we have just shown, h locally belongs to the Sobolev class $\mathcal{W}$. Without loss of generality we can assume that it is so on the whole domain U, i.e., $h \in \mathcal{W}(U)$, and that h can be approximated in $\mathcal{W}(U)$ by a sequence of C^∞ functions h_n. Take a domain $D \Subset U$ with piecewise smooth boundary (e.g., a rectangle).

[3]We will show later (see Prop. 12.15) that the inverse to qc maps are also qc, making h itself absolutely continuous as well.

Let $V_n \subset h_n(D)$ be the set of regular values of h_n. By Sard's Theorem, it has full measure in $h_n(D)$. Let $R_n = h_n^{-1}(V_n) \cap D$. Note that $\int_{R_n} \operatorname{Jac} h_n \, dm$ is equal to the area of the image of $h_n | \, R_n$ counted with multiplicities:

$$\int_D \operatorname{Jac}(h_n, z) \geq \int_{R_n} \operatorname{Jac}(h_n, z)\, dm = \int_{V_n} \operatorname{card}(h_n^{-1}\zeta)\, dm \geq m(V_n) = m(h_n(D)).$$

Since $h_n \to h$ uniformly on D, $m(h_n D) \to m(h(D))$. Since $\operatorname{Jac}(h_n) \to \operatorname{Jac}(h)$ in $L^1(U)$ (as the partial derivatives converge in L^2),

$$\int_D \operatorname{Jac}(h_n, z)\, dm \to \int_D \operatorname{Jac}(h, z)\, dm.$$

Putting the last estimates together, we obtain the desired.

For an arbitrary Borel set $X \subset U$, the result follows by a simple approximation argument using a covering of X by a union of rectangles D_i with disjoint interiors such that $m(\cup D_i \smallsetminus X) < \varepsilon$. $\square$

REMARK 11.15. This proof shows that for a qc map $h : U \to \mathbb{D}_r$, where $U \subset \mathbb{C}$, the distributional derivatives belong to $L^2(U)$.

11.5. Appendix: Distributional derivatives and absolute continuity* on lines. Let U be a domain in $\mathbb{C} \equiv \mathbb{C}_{\mathbb{R}}$. All functions below are assumed to be complex valued. A *test function* ϕ on U is an infinitely differentiable function with compact support. One says that a locally integrable function $h : U \to \mathbb{C}$ has *distributional partial derivatives*[4] *of class* L^1_{loc} if there exist functions ρ and g of class L^1_{loc} on U such that for any test function ϕ,

$$\int_U h \cdot \partial\phi\, dm = -\int_U \rho\, \phi\, dm; \quad \int_U h \cdot \bar{\partial}\phi\, dm = -\int_U g\, \phi\, dm,$$

where m is the Lebesgue measure. In this case ρ and g are called $\partial-$ and $\bar{\partial}-$ derivatives of h in the sense of distributions.

This notion is obviously equivalent to the existence of distributional partial derivatives ∂_x and ∂_y in the real variables (defined analogously). Clearly, the latter property is invariant under smooth changes of variable, so that it makes sense on any smooth manifold (and in all dimensions). Below this notion is related to the absolute continuity* on lines. (See §50.4 for the meaning of the "star".)

EXERCISE 11.16. *Prove that a function h on the interval $(0, 1)$ has a distributional derivative of class L^1_{loc} if and only if it is absolutely continuous*. Moreover, its classical derivative $h'(x)$ coincides with the distributional derivative a.e..*

There is a similar criterion in the two-dimensional setting. A continuous function $h : U \to \mathbb{C}$ is called *absolutely continuous* on lines* if for any family of parallel lines in any disk $D \Subset U$, h is absolutely continuous* on almost all of them. Hence, taking a typical line γ of the above family, the curve $h : \gamma \to \mathbb{C}$ is rectifiable. Clearly such functions have classical partial derivatives almost everywhere.

PROPOSITION 11.17. *Consider a homeomorphism $h : U \to V$ between two domains in the complex plane. It has distributional partial derivatives of class L^1_{loc} if and only if it is absolutely continuous* on lines.*

[4]In the Russian literature, they are called *generalized derivatives.*

In fact, in the proof of existence of distributional partial derivatives (the easy direction of the above Proposition), just two transverse families of parallel lines are used. Thus one can relax the definition of absolutely continuity* on lines by taking any two directions ("horizontal" and "vertical").

On the other hand, since existence of distributional partial derivatives can be recognized in any local charts, a map with this property is *absolutely continuous* on almost all curves of any smooth foliation.*

PROPOSITION 11.18. *Consider a homeomorphism $h : U \to V$ that has partial derivatives a.e. Then for almost any $z \in U$, h is differentiable at z in the classical sense, i.e., $h \in$ DHomeo.*

This result can be viewed as a measurable generalization of the elementary fact that existence of continuous partial derivatives implies differentiability.

PROJECT 11.19. *Fill in details of the above discussion* (*using literature as needed*)*,* see e.g., [**A2**, Ch II B].

In conclusion, let us define the *Sobolev class* $\mathcal{W}^p \equiv \mathcal{W}^p(U)$ (on a bounded domain $U \Subset \mathbb{C}$) as the space of bounded continuous functions $h : U \to \mathbb{C}$ whose distributional partial derivatives on U belong to $L^p(U)$.[5] The norm on $\mathcal{W}$ is the maximum of the uniform norm of h and L^p-norm of its partial derivatives. Any finction $h \in \mathcal{W}^p(U)$ can be approximated by infinitely smooth functions in $\mathcal{W}^p(V)$ for any domain $V \Subset U$. This can be shown by the standard regularization procedure: convolute h with a sequence of bump-functions $\phi_n(x) = n^2\phi(nx)$, where ϕ is a non-negative test function on U with $\int \phi\, dm = 1$ (see [**Ste**, Ch V, §2.1] or [**LV**, Ch. III, Lemma 6.2]). However, if h is a homeomorphism, these approximating functions do not necessarily inherit this propery.

REMARK 11.20. The usual Sobolev class $W^{1,p}$ is defined as the space of L^p-functions with distributional derivatives of class L^p. So, our space $\mathcal{W}^p$ is the intersection of $W^{1,p}$ with the space of bounded continuous functions. Note that by the Sobolev Embedding Theorem, for $p > 2$ functions of $W^{1,p}$ are automatically continuous (see [**Ste**, Ch V, §2.2]), but the borderline $W^{1,2}$-regularity of qc maps is not sufficient for this conclusion (though in the end of the day it is known that qc maps do belong to $W^{1,p}$ with $p > 2$ [**Ge**]).

We let $\mathcal{W} \equiv \mathcal{W}^2$.

12. Geometric definitions

Besides the analytic definition given above, we will give two geometric definitions of quasiconformality, in terms of quasi-invariance of moduli, and in terms of bounded circular dilatation (or, "quasi-symmetricity").

12.1. Quasi-invariance of moduli. In this section we will show, by the length-area method, that the moduli of annuli are quasi-invariant under qc maps. This will follow from a more general result on quasi-invariance of extremal length:

LEMMA 12.1. *Let $h : U \to \tilde{U}$ be a K-qc homeomorphism. Let Γ be a smooth foliation of some domain in U and let $\tilde{\Gamma} = h(\Gamma)$. Then $\mathcal{L}(\tilde{\Gamma}) \le K\mathcal{L}(\Gamma)$.*

---

[5]Only exponents $p = 1, 2$ will be relevant for us.

PROOF. To any measurable metric $\tilde{\rho}$ on $\tilde{U}$, we are going to associate a metric ρ on U such that

$$h^*(\tilde{\rho}) \leq \rho \tag{12.1}$$

while

$$h^*(m_{\tilde{\rho}}) \geq K^{-1}\, m_\rho \tag{12.2}$$

(so, the map h is contracting with respect to these metrics, with the area contraction bounded by K). Then $\tilde{\rho}(\tilde{\gamma}) \leq \rho(\gamma)$ for almost any $\gamma \in \Gamma$ and $\tilde{\gamma} = h(\gamma) \in \tilde{\Gamma}$ (since h is absolutely continuous* on a.e. γ), while $m_{\tilde{\rho}}(\tilde{U}) \geq K^{-1} m_\rho(U)$. Hence $\mathcal{L}_{\tilde{\rho}}(\tilde{\Gamma}) \leq K\mathcal{L}_\rho(\Gamma)$. Taking the supremum over all metrics $\tilde{\rho}$, we obtain the desired estimate.

Let X be the set of full measure where h is classically differentiable. Then we let $\rho(z) = \tilde{\rho}(h(z))\lambda_{\max}(z)$ on X (recall from §11.1.3 that $\lambda_{\max}(z)$ is the maximal expansion factor of $Dh(z)$), and we let $\rho(z) = \infty$ outside X. Since Γ is a smooth foliation, h is absolutely continuous* on almost all curves of Γ. Let Γ^g be the family of such "good" curves, and let $\tilde{\Gamma}^g := h(\Gamma^g)$. By Exercise 6.2, $\mathcal{L}(\Gamma) = \mathcal{L}(\Gamma^g)$ while $\mathcal{L}(\tilde{\Gamma}) \leq \mathcal{L}(\tilde{\Gamma}^g)$, so it is enough to check the desired property for the good families.

So let $\gamma \in \Gamma^g$. Then for any $z \in X \cap \gamma$ and any unit tangent $v \in \mathrm{T}_z U$, we have:

$$|(h^*\tilde{\rho})\, v| = \tilde{\rho}(h(z)) \cdot |Dh(z)v| \leq \tilde{\rho}(h(z)) \cdot \lambda_{\max}(z) = \rho\,(v).$$

So (12.1) is satisfied for $z \in \gamma \cap X$ (while for $z \in \gamma \smallsetminus X$ it is obviously satisfied). Let dl and $d\tilde{l}$ be the length measures on γ and $\tilde{\gamma} = h(\gamma)$, respectively. Since $h : \gamma \to \tilde{\gamma}$ is absolutely continuous* with respect to these measures, $h^*(d\tilde{l}) \leq dl$. Integrating (12.1) over these measures yields: $l_{\tilde{\rho}}(\tilde{\gamma}) \leq l_\rho(\gamma)$ for any $\gamma \in \Gamma^g$. Taking the infimum of the good curves, we obtain: $l_{\tilde{\rho}}(\tilde{\Gamma}^g) \leq l_\rho(\Gamma^g)$.

On the other hand,

$$h^*(dm_{\tilde{\rho}}) = \tilde{\rho}(hz)^2 \operatorname{Jac} h(z)\, dxdy = K(z)^{-1}\rho(z)^2 dxdy \geq K^{-1} dm_\rho,$$

where the second equality comes from (11.6). This provides us with (12.2), and the conclusion follows. □

REMARK 12.2. In the above argument we had to be careful with the direction of quasiconformality (h or h^{-1}), as at this stage we do not yet know that the notion is symmetric. The next statement is exactly the moment when it gets symmetrized.

PROPOSITION 12.3. *Consider a K-qc map $h : A \to \tilde{A}$ between two topological annuli. Then*

$$K^{-1} \operatorname{mod}(\tilde{A}) \leq \operatorname{mod}(A) \leq K \operatorname{mod}(\tilde{A}),$$

PROOF. Let Γ be the genuinely vertical foliation on A, and let $\tilde{\Gamma} := h(\Gamma)$. By Proposition 6.6, $\operatorname{mod} A = \mathcal{L}(\Gamma)$, while $\operatorname{mod} \tilde{A} \leq \mathcal{L}(\tilde{\Gamma})$. By Lemma 12.1, $\mathcal{L}(\tilde{\Gamma}) \leq K\mathcal{L}(\Gamma)$, which yields the desired right hand-side estimate. The left-hand side estimate is obtained by replacing the vertical foliation with the horizontal one. □

EXERCISE 12.4. *Show that the moduli of rectangles are quasi-invariant in the same sense as for the annuli.*

EXERCISE 12.5. *Prove that $\mathbb{C}$ and $\mathbb{D}$ are not qc equivalent.*

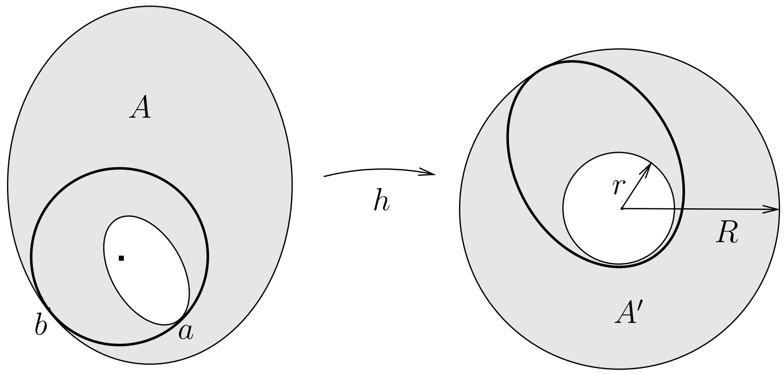

FIGURE 12.1. Bound on the circular dilatation.

12.2. Macroscopic circular dilatation. According to the original analytic definition of qc maps, they have bounded infinitesimal dilatation a.e. It turns out that this property can be substantially strengthened: in fact, qc maps have bounded macroscopic dilatation in sufficiently small scales *everywhere*.

Let $h : U \to V$ be a homeomorphism between two domains, and let $D := \mathbb{D}(z, \rho) \subset U$. Then we can define the *macroscopic circular dilatation* $\operatorname{Dil}(h, z, \rho)$ as the shape of $h(D)$ around $h(z)$ (as for conformal maps in §4.4). (Recall also from §4.4 the definitions of the inner and outer radii of a pointed domain.)

LEMMA 12.6. *Let $h : U \to V$ be a K-qc homeomorphism. Let $D = \mathbb{D}(z, \rho) \subset U$ and $\mathbb{D}(h(z), R) \subset V$, where R is the outer radius of $h(D)$. Then*

$$\operatorname{Dil}(h, z, \rho) \leq \exp CK,$$

where C an absolute constant.

PROOF. For notational convenience, let us normalize h so that $z = h(z) = 0$, and let r be the inner radius of $h(D)$. Let a and b be two points on the circle $\mathbb{T}_\rho$ for which $|h(a)| = r$ and $|h(b)| = R$. Let us consider the annulus $A' := \mathbb{A}(r, R) \subset V$ and let $A = h^{-1}(A')$. The inner component of $\mathbb{C} \smallsetminus A$ contains points 0 and $a \in \mathbb{T}_\rho$, while its outer component of $\mathbb{C} \smallsetminus A$ contains $b \in \mathbb{T}_\rho$. (See Figure 12.1.) By Lemma 6.10, $\operatorname{mod} A$ is bounded by an absolute constant C. By Lemma 12.3,

$$\frac{1}{2\pi} \log \frac{R}{r} = \operatorname{mod} A' \leq K \operatorname{mod} A \leq KC,$$

and we are done. □

The *upper circular dilatation of h at z* is defined as

$$\overline{\operatorname{Dil}}(h, z) = \limsup_{\rho \to 0} \operatorname{Dil}(h, z, \rho).$$

(Of course, if h is differentiable at z then $\overline{\operatorname{Dil}}(h, z) = \operatorname{Dil}(h, z)$.) We define the *upper circular dilatation of h* as

$$\overline{\operatorname{Dil}}\, h := \sup_{z \in U} \overline{\operatorname{Dil}}(h, z).$$

Lemma 12.6 immediately implies:

Proposition 12.7. *Any K-qc map $U \to V$ has a bounded upper dilatation:*

$$\overline{\mathrm{Dil}}\, h \le \exp CK,$$

where C is an absolute constant.

12.3. Quasisymmetry.

12.3.1. *Generalities.* We will now give a characterization of qc maps that can be applied in a very general setting. For a triple of points (x, y, z) in a metric space X, let the brackets

$$[y, z]_x := \frac{\mathrm{dist}(z, x)}{\mathrm{dist}(y, x)}$$

denote the *distance ratio centered at x.*

Let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be a function such that $\eta(t) \to 0$ as $t \to 0$. An embedding $h : X \to Y$ between two metric spaces is called *η-quasisymmetric* ("*η-qs*") if for any triple of points (x, y, z) in X we have:

$$[y, z]_x \le t \implies [h(y), h(z)]_{h(x)} \le \eta(t). \tag{12.3}$$

A map h is called *quasisymmetric* if it is η-qs for some η. Such an h distorts the ratios in a controlled way.

The function $\eta(t)$ is called the *qs dilatation* of h.

Exercise 12.8. *Show that the dilatation function η can be selected as a homeomorphism $\mathbb{R}_+ \to \mathbb{R}_+$ (which will be our standing convention in what follows).*

For instance, L-bi-Lipschitz homeomorphisms are η-qs with linear dilatation $\eta(t) = L^2 t$. However, the class of qs maps is much bigger:

Exercise 12.9. *The power homeomorphisms of $\mathbb{R}$, $x \mapsto \mathrm{sign}(x)|x|^\delta$, are quasisymmetric. What are their qs dilatations?*

QS maps can serve as *morphisms* of the category of metric spaces:

Exercise 12.10. *The inverse of a qs map is qs, with $\eta_{h^{-1}} = \sigma \circ \eta^{-1} \circ \sigma$, where $\sigma(t) = 1/t$. Compositions of qs maps are qs, with $\eta_{g\circ h} = \eta_g \circ \eta_h$.*

We conclude that quasisymemtries of any metric space form a group.

12.3.2. *QC vs QS.* The most important value of the dilatation function $\eta(t)$ is $\eta(1)$ that controls macroscopic dilatation of h on the balls and (as we will see momentarily) often controls the full $\eta(t)$.

Lemma 12.11. *An embedding $h : \mathbb{R}^m \to \mathbb{R}^n$ is η-qs if and only if (quantitatively) it has L-bounded macroscopic dilatation:* $\mathrm{Dil}(h, z, \rho) \le L$ *for all discs $\mathbb{D}(z, \rho)$.*

Proof. Obviously, quasisymmetry implies that macroscopic dilatation is bounded by $L = \eta(1)$. Vice versa, bounded macroscopic dilatation implies (12.3) with a function $\eta(t) = O(t^\alpha)$ as $t \to \infty$, where the exponent $\alpha \ge 1$ depends only on the dilatation.

Exercise 12.12. *Prove this assertion and calculate $\eta(t)$ in terms of $L = \eta(1)$.*

What is more subtle is to show that $\eta(t) \to 0$ as $t \to 0$.

Let us take a triple of points x, y, z, and let x', y', z' stand for their images under h (in what follows, the images of other points under h will be marked with the "prime" as well). Property (12.3) implies:

$$[z, y]_x \ge 1 \implies [z', y']_{x'} \ge \varepsilon = 1/L > 0. \tag{12.4}$$

By making affine changes of variable in the domain and the target, we can normalize the situation so that $x = x' = 0$, $|y| = |y'| = 1$, $z = R \in \mathbb{R}$, $z' = r \in \mathbb{R}$. Of course, we can assume that $R > 1$. We want to show that $r \to \infty$ as $R \to \infty$. Let us partition the interval $[0, z]$ by points $z_n = z/2^n$, $n = 0, 1, \dots, N$, where N is selected so that $z_N \in [1, 2)$. So, $N \geq \log_2 R - 1 \to \infty$ as $R \to \infty$.

Applying (12.4) to the triple of points $(0, 1, z_N)$, we obtain: $|z'_N| \geq \varepsilon$. Then applying it inductively (backwards) to the triples $(z_n, 0, z_{n-1})$ (centered at z_{n-1}), we conclude that

$$|z'_n - z'_{n-1}| \geq \varepsilon |z'_{n-1}| \geq \varepsilon^2,$$

so the net of points z'_n is ε^2*-separated*. On the other hand, applying (12.4) to the triple $(0, z_n, z)$, we conclude that $|z'| \geq \varepsilon |z'_n|$, so that all the points z'_n belong to the disc $\mathbb{D}_{r/e}$. Hence the discs of radius $\varepsilon^2/2$ centered at the z_n are pairwise disjoint and are contained in the disc $\mathbb{D}_{2r/e}$. It follows that

$$N \leq \frac{\operatorname{area} \mathbb{D}_{2r/\varepsilon}}{\operatorname{area} \mathbb{D}_{\varepsilon^2/2}} = \frac{16}{\varepsilon^6}\, r^2,$$

and hence $r \geq c\sqrt{\log R}$ with $c > 0$ depending only on L. □

Also, in the light of the above result, embeddings $h : \mathbb{R}^m \to \mathbb{R}^n$ with L-bounded macroscopic dilatation will also be referred to as *L-qs*. (We hope that a slight terminological inconsistency with "η-qs" will not cause confusion).

Putting together Lemmas 12.6 and 12.11, we obtain:

Proposition 12.13. *Any quasiconformal map $h : U \to V$ is quasisymmetric on compac sets $Q \Subset U$ (with the qs dilatation controlled by* $\operatorname{Dil} h$ *and a lower bound on* $\min\{\operatorname{dist}(Q, U), \operatorname{dist}(h(Q), V)$*. Moreover, there is an L depending only on K such that:*

(i) *Any K-qc homeomorphism $h : \mathbb{C} \to \mathbb{C}$ is L-qs* (*in the Euclidean metric*)*;*

(ii) *Any K-qc homeomorphism $h : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ fixing* $0, 1$ *and* ∞ *is L-qs* (*in the spherical metric*)*.*

Note that without a normalization, the last (quantitative) assertion fails as the Möbius group is not uniformly qs on the sphere.

12.4. Back to the analytic definition.

Proposition 12.14. *If a homeomorphism $h : U \to V$ between domains U and V has an L-bounded upper circular dilatation then it is L-qc.*

Proof. Since the L-bounded circular dilatation implies the L-bounded infinitesimal dilatation at any point of differentiability, all we need to show is that h has the required regularity, i.e., it is absolutely continuous* on almost all parallel lines. Since this is a local property, we can assume that U us the unit square, and that the parallel lines in question are horizontal.

Let $U_b = \{z \in U : \operatorname{Im} z \leq b\}$. Since the area function

$$\mu : b \mapsto \operatorname{area}(h(U_b))$$

is monotonic, it is differentiable for a.e. b. Let us take such a point b where μ is differentiable, and prove absolute continuity of h on the corresponding line $\gamma_b = \{z : \operatorname{Im} z = b\}$.

For $K \in \mathbb{N}$, let $X_K = \{z \in \gamma_b : \mathrm{Dil}(h, x, \varepsilon) \le K/2 \text{ for } \varepsilon \le 1/K\}$. Since the dilatation of h is bounded we have: $\bigcup X_K = \gamma_b$. Hence it is enough to prove that $h|\, X_K$ is absolutely continuous.

Let $Q \subset X_K$ be a set of zero length. We want to show that $h(Q)$ has zero length as well. By approximation, it is sufficient to show this for closed sets. Then Q can be covered with finitely many disks $D_i = \mathbb{D}(z_i, \varepsilon)$ ($z_i \in \gamma_b$, $i = 1, \dots, n$) with intersection multiplicity at most 2 and an arbitrary small total length. Hence for any $\delta > 0$, we have $n\varepsilon \le \delta$ once ε is sufficiently small.

Let us consider the outer and inner radia of the $h(D_i)$, respectively: $R_i = R_h(z_i, \varepsilon)$ and $r_i = r_h(z_i, \varepsilon)$. Then $R_i \le K r_i$, $l(h(X)) \le \sum R_i$, and by the Cauchy-Bunyakovsky inequality,

$$l(h(X))^2 \le n \sum R_i^2 \le nK^2 \sum r_i^2 \le \frac{K^2\delta}{\pi} \cdot \frac{2 \cdot \mathrm{area}(h(\bigcup D_i))}{\varepsilon}$$

(where "2" comes from the intersection multiplicity). But since $\bigcup D_i \subset U_{b+\varepsilon} \smallsetminus U_{b-\varepsilon}$, the last ratio (without "2") is bounded by

$$\frac{\mu(b+\varepsilon) - \mu(b-\varepsilon)}{\varepsilon} \to 2\,\mu'(b) \quad \text{as } \varepsilon \to 0,$$

and the desired conclusion follows. □

12.5. Summary. Thus, quasiconformality can be defined in several equivalent (non-trivially related) ways: An orientation preserving homeomorphism $h : U \to V$ between two domains in $\mathbb{C}$ is K-*quasiconformal* if one of the following equivalent properties QC1–QC3 holds:

QC1. *Analytic definition.*

(i) *Regularity*: h has distributional derivatives ∂h and $\bar\partial h$ of class L^1_{loc}. Equivalently, h is absolutely continuous* on almost all lines in any given direction (and it is sufficient to be so in two transverse directions).

(ii) *Bounded dilatation*: $\mathrm{Dil}\, h \le K$ a.e., or equivalently, $|\bar\partial h| \le k \cdot |\partial h|$ a.e., where $k = (K-1)/(K+1) < 1$.

QC2. *Quasi-invariance of moduli.* Moduli of quadrilaterals and annuli are K-quasi-invariant.

QC3. *Bounded upper circular dilatation*: $\overline{\mathrm{Dil}}\, h \le Q$ everywhere (where $Q \ge K$ can be bounded in terms of K).

A closely related notion is *quasisymmetry*:

QS. Any η-qs map h is K-qc and any K-qc map is locally η-qs, quantitatively (here we use the Euclidean metric on $\mathbb{C}$). Moreover, in case when $U = V = \mathbb{C}$, η-qs and K-qc are equivalent, quantitatively. (In Corollary 13.13, we will give a $\mathbb{D}$-version of this assertion.)

Finally, we can conclude:

PROPOSITION 12.15. *The inverse of a K-qc map is K-qc. The composition of K_1-qc map and K_2-qc map is (K_1K_2)-qc.*

COROLLARY 12.16. *The family of qc self-maps $U \to U$ of a given domain is a group.*

13. Further important properties of qc maps

13.1. Weyl's Lemma. This lemma asserts that a 1-qc map is conformal. In other words, if a qc map is infinitesimally conformal on the set of full measure (i.e., $\bar{\partial}h(z) = 0$ a.e.) then it is conformal in the classical set. Since $\bar{\partial}h(z) = 0$ is just the Cauchy-Riemann equation, this statement is classical for smooth maps.

Let us formulate a more general version of Weyl's Lemma:

Weyl's Lemma. *Assume that a continuous function $h : U \to \mathbb{C}$ belongs to the Sobolev class $\mathcal{W}^1_{\text{loc}}$. If $\bar{\partial}h(z) = 0$ a.e. then h is holomorphic.*

Proof. By approximation, Weyl's Lemma can be reduced to the classical statement. Since the statement is local, we can assume without loss of generality that the partial derivatives of h belong to $L^1(U)$. Convoluting h with smooth bump-functions we obtain a sequence of smooth functions $h_n = h * \theta_n$ converging to h uniformly on U with derivatives converging in $L^1(U)$. Let us show that $\bar{\partial}h_n = 0$. For a test function η on U, we have:

$$\int \bar{\partial}h_n(z)\,\eta(z)\,dm(z) = -\int h_n(z)\,\bar{\partial}\eta(z)\,dm(z)$$

$$= -\int h(\zeta)\,dm(\zeta)\int \theta_n(z-\zeta)\bar{\partial}\eta(z)\,dm(z)$$

$$= \int h(\zeta)\,dm(\zeta)\int \bar{\partial}\theta_n(z-\zeta)\eta(z)\,dm(z)$$

$$= \int \eta(z)\,dm(z)\int h(\zeta)\,\bar{\partial}\theta_n(z-\zeta)\,dm(\zeta)$$

$$= \int \eta(z)\,dm(z)\int \bar{\partial}h(\zeta)\,\theta_n(z-\zeta)\,dm(\zeta) = 0.$$

Here the first and the third equalities are the classical integration by parts, the next to the last one comes from the definition of the distributional derivative, and the intermediate ones come from the Fubini Theorem.

It follows that the smooth functions h_n satisfy the Cauchy-Riemann equations and hence holomorphic. Since uniform limits of holomorphic functions are holomorphic, h is holomorphic as well. $\square$

13.2. Devil's Staircase vs Weyl's Lemma. The following example shows that Weyl's Lemma is not valid for homeomorphisms of class DHomeo (i.e., differentiable a.e.). The technical assumption that the classical derivative can be understood in the sense of distributions (which allows us to integrate by parts) is thus crucial for the statement.

Take the standard Cantor set $Q \subset [0,1]$ and consider a devil's staircase $h : [0,1] \to [0,1]$, i.e., a continuous monotone function which is constant on the gaps in Q (See §1.1.2.)

Consider a strip $S = [0,1] \times \mathbb{R}$ and let $f : (x,y) \mapsto (x, y + h(x))$. This is a homeomorphism on S which is a rigid translation on every strip $G \times \mathbb{R}$ over a gap $G \subset [0,1] \smallsetminus Q$. Since $m(K \times \mathbb{R}) = 0$, this map is conformal a.e. However it is obviously not conformal on the whole strip P.

Clearly f in not absolutely continuous on the horizontal lines: it translates them to devil's staircases.

13.3. Quasiconformal removability and gluing. A closed set $Q \subset \mathbb{C}$ is called *qc removable* if any homeomorphism $h : U \to \mathbb{C}$ defined on a neighborhood U of Q which is quasiconformal on $U \smallsetminus Q$ is quasiconformal on U.

Remark. We will see later on (Proposition 16.4) that qc removable sets have zero measure and hence $\operatorname{Dil}(f|U) = \operatorname{Dil}(f|U \smallsetminus Q)$.

Exercise 13.1. *Show that isolated points are removable.*

Little Gluing Lemma (smooth version). *Piecewise smooth Jordan curves (or arcs) are removable.*

Proof. Let us consider a smooth Jordan arc $\Gamma \subset U$ and a homeomorphism $f : U \to \mathbb{C}$ which is quasiconformal on $U \smallsetminus \Gamma$. We should check that f is absolutely continuous on lines near any point $z \in \Gamma$. Take a small box B centered at z whose sides are not parallel to $T_z\Gamma$. Then any interval l in B parallel to one of its sides intersects Γ at a single point ζ. Since for a typical l, f is absolutely continuous on both sides of $l \smallsetminus \{\zeta\}$, it is absolutely continuous on the whole interval l as well.

Moreover, $\operatorname{Dil}(f)$ is bounded since it is so on $U \smallsetminus \Gamma$ and Γ has zero measure.

It proves the assertion for smooth arcs. In the piecewise smooth case, remove first smooth pieces and then remove remaining isolated points by the previous Exercise. □

The above statement is simple but important for holomorphic dynamics. It will allow us to construct global qc homeomorphisms by gluing together different pieces without spoiling dilatation. Note that it fails for 1D qs maps (see Exercise 15.1 below).

Let us now state a more delicate gluing property:

Bers' Gluing Lemma. *Consider a closed set $Q \subset \hat{\mathbb{C}}$ and two its neighborhoods U and V. Assume that we have two quasiconformal maps $f : U \smallsetminus Q \to \hat{\mathbb{C}}$ and $g : V \to \hat{\mathbb{C}}$ that match on ∂K, i.e., the map*

$$h(z) = \begin{cases} f(z), & z \in U \smallsetminus Q \\ g(z), & z \in Q \end{cases}$$

is continuous. Then h is quasiconformal and $\mu_h(z) = \mu_g(z)$ for a.e. $z \in Q$.

Proof. Consider a map $\phi = g^{-1} \circ h$. It is well-defined in a neighborhood Ω of Q, is identity on Q, and is quasiconformal on $\Omega \smallsetminus Q$. Let us show that it is quasiconformal on Ω. Again, the main difficulty is to show that h is absolutely continuous on lines near any point $z \in Q$.

Take a little box B near some point $z \in Q$ with sides parallel to the coordinate axes. Without loss of generality we can assume that $z \neq \infty$ and $\phi(B)$ is a bounded subset of $\mathbb{C}$. Let ψ denote the extension of $\partial\phi/\partial x$ from $B \smallsetminus Q$ onto the whole box B by 0. By (11.8), ψ is square integrable on B and hence it is square integrable on almost all horizontal sections of B. All the more, it is integrable on those sections. Take such a section I, and let us show that ϕ is absolutely continuous on it.

Let $I_j \subset I$ be a finite set of disjoint intervals; $\Delta\phi_j$ denote the increment of ϕ on I_j. We should show that

$$\sum |\Delta\phi_j| \to 0 \quad \text{as} \quad \sum |I_j| \to 0. \tag{13.1}$$

Take one interval I_j and decompose it as $L \cup J \cup R$ where $\partial J \subset Q$ and $\operatorname{int} L$ and $\operatorname{int} R$ belong to $B \setminus Q$. Then

$$|\Delta\phi_j| \le |J| + \int_{L\cup R} \psi\, dx \le |I_j| + \int_{I_j} \psi\, dx.$$

Summing up the last estimates over j and using integrability of ψ on I_j, we obtain (13.1).

Absolute continuity on the vertical lines is treated in exactly thesame way.

The last assertion of the lemma follows from the following remarks:

• If $z \in Q$ be a point of differentiability for h, then the differntial $Dh(z)$, and hence the dilatation $\mu_h(z)$, can be read off from two directional derivatives of h at z.

• The directional derivative of h along a line L through $z \in Q$ is determined by the restriction of $h \,|\, L \cap Q$, provided z is not isolated on $L \cap Q$.

• If z is a density point for Q, the latter property holds for a.e. line L through z.

• By the Lebesgue Theorem, almost all points of Q are density points.

(And similarly for g.) □

13.4. Compactness on $\hat{\mathbb{C}}$. We will proceed with the following fundamental property of qc maps:

Theorem 13.2. *The space of K-qc maps $h : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ fixing $0, 1$ and ∞ is compact in the topology of uniform convergence on $\hat{\mathbb{C}}$*

Proof. It will be more convenient to consider the space $\mathcal{X}$ of K-qc maps h such that $h\{0, 1, \infty\} = \{0, 1, \infty\}$. First, we will show that the family of maps $h \in \mathcal{X}$ is equicontinuous. Otherwise we would have an $\varepsilon > 0$, a sequence of maps $h_n \in \mathcal{X}$, and a sequence of points $z_n, \zeta_n \in \hat{\mathbb{C}}$ such that $d(z_n, \zeta_n) \to 0$ while $d(h_n(z_n), h_n(\zeta_n)) \ge \varepsilon$, where d stands for the spherical metric. By compactness of $\hat{\mathbb{C}}$, we can assume that the $z_n, \zeta_n \in \hat{\mathbb{C}}$ converge to some point a and the $h_n(a)$ converge to some b. Postcomposing or/and precomposing if necessary the maps h_n's with $z \mapsto 1/z$, we can make $|a| \le 1$, $|b| \le 1$.

Consider a sequence of annuli $A_n = \{z : r_n < |z - a| < 1/2\}$ where $r_n = \max(|z_n - a|, |\zeta_n - a|) \to 0$. Since the disk $\mathbb{D}(a, 1/2)$ does not contain one of the points 0 or 1, its images $h_n(\mathbb{D}(a, 1/2))$ have the same property. Hence the Euclidean distance from the point $h_n(a)$ (belonging to the inner complement of $h_n(A_n)$) to the outer complement of that annulus is eventually bounded by 3. On the other hand, the diameter of the inner complement of $h_n(A_n)$ is bounded from below by $\varepsilon > 0$. By Lemma 6.10, $\operatorname{mod}(h_n(A_n))$ is bounded from above. But

$$\operatorname{mod} A_n = \frac{1}{2\pi} \log \frac{1}{2r_n} \to \infty$$

contradicting quasi-invariance of the modulus (Proposition 12.3).

Hence $\mathcal{X}$ is precompact in the space of continuous maps $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$. Since $\mathcal{X}$ is invariant under taking the inverse $h \mapsto h^{-1}$, and the composition is a continuous operation in the uniform topology, $\mathcal{X}$ is precompact in $\operatorname{Homeo}(\hat{\mathbb{C}})$. Since $\operatorname{Homeo}^+(\hat{\mathbb{C}})$ is closed in $\operatorname{Homeo}(\hat{\mathbb{C}})$, $\mathcal{X}$ is precompact in the former space as well.

To complete the proof, we should show that the limit functions are also K-qc homeomorphisms. Let a sequence $h_n \in \mathcal{X}$ uniformly converges to some h. Given a point $a \in \hat{\mathbb{C}}$, we will show that in some neighborhood of a, f has distributional derivatives of class L^2. Without loss of generality we can assume that $a \in \mathbb{C}$. Take a neighborhood $B \ni a$ such that $h(B)$ is a bounded subset of $\mathbb{C}$. Then the neighborhoods $h_n(B)$ are eventually uniformly bounded. By (11.8), the partial derivatives ∂h_n and $\bar\partial h_n$ eventually belong to a fixed ball of $L^2(D)$. Hence they form weakly precompact sequences, and we can select limits along subsequences (without changing notations):

$$\partial h_n \underset{w}{\to} \phi \in L^2(D); \qquad \bar\partial h_n \underset{w}{\to} \psi \in L^2(D).$$

It is straightforward to show that ϕ and ψ are the distributional partial derivatives of h. Indeed, for any test functions η we have:

$$\int h\, \partial\eta\, dm = \lim \int h_n\, \partial\eta\, dm = -\lim \int \partial h_n\, \eta\, dm = -\int \phi\, \eta\, dm, \tag{13.2}$$

and the similarly for the $\bar\partial$-derivative.

What is left is to show that $|\phi(z)| \leq k|\psi(z)|$ for a.e. z, where $k = (K-1)/(K+1)$. To see this, select a further subsequence in such a way that $|\partial h_n| \underset{w}{\to} |\phi|$, $|\bar\partial h_n| \underset{w}{\to} |\psi|$ and use the fact that the weak topology respects the order (see Exercise 13.16 from the Appendix). $\square$

EXERCISE 13.3. *Fix any three points a_1, a_2, a_3 on the sphere $\mathbb{C}$. A family $\mathcal{X}$ of K-qc maps $h : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is precompact in the space of all K-qc homeomorphisms of the sphere (in the uniform topology) if and only if the reference points are not moved close to each other (or, in formal words: there exists a $\delta > 0$ such that $d(ha_i, ha_j) \geq \delta$ for any $h \in \mathcal{X}$ and $i \neq j$, where d is the spherical metric). Consider first the case $K = 0$.*

13.5. Quasi-isometries and the boundary extension.

13.5.1. *Quasi-isometries and quasi-geodesics.* A map $h : X \to X$ of a metric space (X, d) is called a (L, C)-*quasi-isometry*, where $L \geq 1$ $C \geq 0$, if

$$L^{-1} d(x, y) - C \leq d(h(x), h(y)) \leq L d(x, y) + C \quad \forall\, x, y \in X.$$

In other words, it is a bi-Lipschitz map in "big scales". (In small scales, nothing can be said, because of the additive constant C.) Note that it is not even required that h is continuous or invertible, but for the sake of our discussion, it is convenient to *assume that all quasi-isometries under consideration are homeomorphisms.*

LEMMA 13.4. *A K-quasiconformal homeomorphism $h : \mathbb{D} \to \mathbb{D}$ is a hyperbolic (C, L)-quasi-isometry (quantitatively).*

PROOF. It is sufficient to show that a geodesic arc γ of length ≤ 1 is mapped to a geodesic with a bounded distance between its endpoints. Indeed, then one can chop a geodesic arc γ of length ≥ 1 into $n \leq l_{\text{hyp}}(\gamma) + 1$ pieces of length ≤ 1 and apply the Triangle Inequality to $h(\gamma)$.

So, let $l_{\text{hyp}}(\gamma) \leq 1$. Since the group of hyperbolic motions acts transitively on $\mathbb{D}$, we can assume that γ is a straight Euclidean interval connecting 0 to some point $a \in \mathbb{D}$, where $|a| \leq \bar a < 1$ (with an absolute $\bar a$) and that h fixes 0. Then the annulus $A := \mathbb{D} \smallsetminus \gamma$ has a definite modulus $\geq \underline{\mu} > 0$. Since moduli of annuli are quasi-invariant under qc maps, $\operatorname{mod}(h(A)) \geq K^{-1}\underline{\mu} > 0$. This bounds $\operatorname{dist}(h(a)), \mathbb{T})$ from below (due to Proposition 6.14), and hence bounds the hyperbolic distance from $h(a)$ to $0 = h(0)$ from above. $\square$

A path γ in the hyperbolic plane is called am (L, C)-quasi-geodesic if for any two points $a, b \in \gamma$

$$\operatorname{diam}_{\text{hyp}}[a, b]_\gamma \leq L \operatorname{dist}_{\text{hyp}}(a, b) + C.$$

where $[a, b]_\gamma$ is the arc of γ connecting a to b.

REMARK 13.5. Once again, because of the additive constant C, this property tells us nothing about the small scale geometry of γ. The corresponding small scale notion is a *quasiarc* (see §15.3.1).

EXERCISE 13.6. *A path $\gamma \subset \mathbb{H}$ is a quasi-geodesic iff it is a quasi-isometric image of an interval $[0, t]$ (quantitatively).*

PROPOSITION 13.7. *Any (L, C)-quasigeodesic (finite or infinite) in $\mathbb{H}$ is R-shadowed by a geodesic, where R depends on (L, C) only.*

PROOF. It is sufficient to deal with finite quasi-geodesics since then one can pass to an infinite limit. Let U_R be the R neighborhood of the geodesics δ connecting the endpoints of γ. Let us show that for R big enough, any component σ of $\gamma \setminus U_R$ has a bounded diameter (where all bounds depend only on (L, C)). Then the conclusion will follows since the whole γ will be trapped in a bounded neighborhood of U_R.

So, let σ be a component of $\gamma \setminus U_R$ connecting some points $c, d \in \partial U_R$. Let c', d' be their projections to the geodesic δ (extended to a bi-infinite one). Let us consider a path ω which is a concatenation of three geodesic arcs, $[c, c']$, $[c', d']$, and $[d', d]$. By Exercise 2.32, the length $[c', d']$ is at most $e^{-R} \operatorname{diam}[c, d]_\gamma$, so the length of ω is atmost $e^{-R} \operatorname{diam}[c, d]_\gamma + 2R$. If R is big then this length is much smaller than $\operatorname{diam}[c, d]_\gamma$, (provided the latter is also big in terms of R), contradicting the quasi-geodesic quality of γ. □

13.5.2. *Boundary extension.*

LEMMA 13.8. *Any hyperbolic quasi-isometry $h : \mathbb{D} \to \mathbb{D}$ extends radially to the boundary $\mathbb{T}$.*

PROOF. Take a point $z \in \mathbb{T}$ and consider a hyperbolic geodesic γ in $\mathbb{D}$ landing at z. Then $h(\gamma)$ is a quasi-geodesic. By Proposition 13.7, it is uniformly shadowed by some geodesic δ. Then γ lands at the same $\zeta \in \mathbb{T}$ as δ does. □

THEOREM 13.9. *Any qc homeomorphism $h : \mathbb{D} \to \mathbb{D}$ extends to a homeomorphism $\overline{\mathbb{D}} \to \overline{\mathbb{D}}$.*

PROOF. The proof follows the lines of the proof of Lemma 8.1.2 from the Carathéodpry Prime Ends Theory. As in that proof, consider the family of quadrilaterals (half-annuli) $\Pi_r \subset \mathbb{D}$ with "vertical sides" on $\mathbb{T}$ and with equal moduli, shrinking to some point $b \in \mathbb{T}$. By Lemma 13.8, the images $h(\Pi_r)$ are also quadrilaterals with vertical sides on $\mathbb{T}$. By the quasi-invariance of moduli, these images have moduli of order 1. Moreover, $\operatorname{area} h(\Pi_r) \to 0$ as $r \to 0$. By the length-area estimate, the $h(\Pi_r)$ contain horizontal curves shrinking to $h(b)$. □

REMARK 13.10. More generally, quasi-isometries of $\mathbb{D}$ continuously extend to the boundary as well, and this property is valid in all dimension (see **[Th2]**).

Combining this with the Conformal Schönflies Theorem, we obtain:

COROLLARY 13.11. *Any qc homeomorphism $h : D_1 \to D_2$ between Jordan domains extends continuously to the boundary.*

13.6. Compactness on $\mathbb{D}$. Let us now state a disk version of the above Compactness Theorem:

Corollary 13.12. *The space of K-qc homeomorphisms $f : \mathbb{D} \to \mathbb{D}$ fixing 0 is compact in the topology of uniform convergence on $\mathbb{D}$.*

Proof. Let $\mathcal{Y}$ be the space of K-qc homeomorphisms $h : \mathbb{D} \to \mathbb{D}$ fixing 0, and $\mathcal{X}$ be the space of $\mathbb{T}$-symmetric K-qc homeomorphisms $H : \mathbb{C} \to \mathbb{C}$ fixing 0 and ∞. (To be $\mathbb{T}$-symmetric means to commute with the involution $\tau : \mathbb{C} \to \mathbb{C}$ with respect to the circle.) Clearly maps $H \in \mathcal{X}$ preserve the unit circle (the set of fixed points of τ); in particular, they do not move 1 close to 0 and ∞. By Theorem 13.2 (and the Exercise following it), $\mathcal{X}$ is compact.

Let us show that $\mathcal{X}$ and $\mathcal{Y}$ are homeomorphic. The restriction of a map $H \in \mathcal{X}$ to the unit disk gives a continuous map $i : \mathcal{X} \to \mathcal{Y}$. The inverse map $i^{-1} : \mathcal{Y} \to \mathcal{X}$ is given by the following extension procedure. First, extend $h \in \mathcal{Y}$ continuously to the closed disk $\overline{\mathbb{D}}$ (by Theorem 13.9), and then reflect it symmetrically to the exterior of the disk, i.e., let $H(z) = \tau \circ h \circ \tau(z)$ for $z \in \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$. Since τ is an (orientation reversing) conformal map, H is K-qc on $\hat{\mathbb{C}} \smallsetminus \mathbb{T}$. By the Little Gluing Lemma (smooth version), it is K-qc everywhere, and hence belongs to $\mathcal{X}$.

Hence $\mathcal{Y}$ is compact as well. $\square$

The extension from $\mathbb{D}$ to $\mathbb{C}$ provided in this proof also implies (via property QS from §12.5) :

Corollary 13.13. *A homeomorphism $h : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ is K-qc iff it is η-qs (quantitatively).*

Exercise 13.14. *Let $K \geq 1$, $C > 0$.*

(i) *Let $(D, a, , b)$ be a double-pointed conformal disk in $\mathbb{C}$. The space of K-qc homeomorphisms $f : D \to \mathbb{C}$ such that*

$$|f(a)|, |f(b)| \leq C \quad \text{and} \quad |f(a) - f(b)| \geq C^{-1}, \tag{13.3}$$

is compact in the topology of uniform convergence on compact subsets of D.

(ii) *More generally, let (D_n, a_n, b_n) be a sequence of disks in $\mathbb{C}$ Carathéodory converging to a hyperbolic disk (D, a, b), and let $f_n : D_n \to \mathbb{C}$ be a sequence of K-qc maps satisfying* (13.3) *at the corresponding points. Then f_n admits a subsequence converging, uniformly on compact subsets of D, to a K-qc map $f : D \to \mathbb{C}$.*

13.7. Appendix: Banach spaces preliminaries. This background can be found in any text book in Functional Analysis, see e.g., [**Lyu, Ru**].

13.7.1. *Generalized sequences.* Since we need L^∞, *we do not assume here that our spaces are separable.* This means that sequential formulations may not be sufficient. An easy way of dealing with this nuisance is to use instead *generalized sequences* $(\mu_n)_{n \in \mathcal{N}}$ labeled by directed sets. Recall that a partially ordered set $(\mathcal{N}, \succeq)$ is called *directed* if any two elements have a majorant:

$$\forall\, m, l \in \mathcal{N}\ \exists n \in \mathcal{N} \text{ s.t. } n \succeq l, n \succeq m.$$

The theory of limits for generalized sequences is identical with the standard theory. The advantage is that all the topological concepts can formulated in the generalized sequential terms, e.g., the closure of a set coincides with the set of the limits of generalized sequences. (Here relevant directed sets are sets of neighborhoods of a point ordered by inclusion: $U \succeq V$ if $U \subset V$.)

13.7.2. *Weak topologies.* The dual space to $\mathcal{B}$ (of bounded linear functionals $\Phi : \mathcal{B} \to \mathbb{C}$) is denoted by $\mathcal{B}^*$. For instance, $(L^2)^* \approx L^2$, $(L^1)^* \approx L^\infty$, $C(X)^* \approx \mathfrak{M}(X)$, where X is a compact space and $\mathfrak{M}(X)$ is the space of finite Borel measures on X (real or complex-valued depending on the main field).

The *weak topology* (*w-topology*) on $\mathcal{B}$ is defined as topology of convergence on all test functionals: $\mu_n \to \mu$ if $\Phi(\mu_n) \to \Phi(\mu)$ for all $\Phi \in \mathcal{B}^*$ (where (μ_n) is a generalized sequence). In case of a bounded sequence μ_n, it is sufficient to test convergence on a dense (in the Banach norm) family of test functionals. For instance, weak convergence of functions $\mu_n \in L^\infty(D)$ can be tested on functions $\phi \in C^\infty_{\text{comp}}(D)$:

$$\mu_n \underset{w^*}{\to} \mu \text{ iff } \int \phi\mu_n \, dm \to \int \phi\mu \, dm \ \forall \phi \in C^\infty_{\text{comp}}(D).$$

The *weak* topology* (*w^*-topology*) on $\mathcal{B}^*$ is the topology of pointwise convergence on all elements $\mu \in \mathcal{B}$. When it does not lead to confusion, we will refer to this topology as just "weak" skipping the * (for instance, in the case of the space of measures). The main virtue of this topology comes from the fact that *the unit ball $\mathcal{B}^*_1$ is w^*-compact.* Note also that vice versa, any weakly convergent sequence is bounded (Banach-Schteinhaus).

However, one should handle the weak topology with caution: for instance, product is not a weakly continuous operation:

Exercise 13.15. *Show that* $\sin nx \underset{w}{\to} 0$ *in* $L^2[0, 2\pi]$, *while* $\sin^2 nx \underset{w}{\to} 1/2$.

At least, the weak topology respects the order:

Exercise 13.16. *Let* $h_n \underset{w}{\to} h$ *in* L^2.

- *If* $h_n \geq 0$ *then* $h \geq 0$;
- *If* $h_n = 0$ *a.e. on some subset* $Y \subset X$, *then* $h = 0$ *a.e. on* Y;
- *After selecting a further subsequence, we have:*

$$h_n^+ \underset{w}{\to} h^+ \text{ and } h_n^- \underset{w}{\to} h^-, \quad \text{so that } |h_n| \underset{w}{\to} |h|.$$

Here $h^+(z) = \max(h(z), 0)$, $h^-(z) = \min(h(z), 0)$.

There is a natural embedding $\mathcal{B} \to \mathcal{B}^{**}$. It is isometric with respect to the Banach norms, but its image is dense in the w^*-topology of $\mathcal{B}^{**}$.

14. Measurable Riemann Mapping Theorem

We are now ready to prove one of the most remarkable facts of analysis: any measurable conformal structure with bounded dilatation is generated by a quasiconformal map:

Measurable Riemann Mapping Theorem. *Let μ be a measurable Beltrami differential on $\hat{\mathbb{C}}$ with $\|\mu\|_\infty < 1$. Then there is a quasiconformal map $h : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ that solves the Beltrami equation $\bar\partial h/\partial h = \mu$. This solution is unique up to post-composition with a Möbius automorphism of $\hat{\mathbb{C}}$. In particular, there is a unique solution fixing three points on $\hat{\mathbb{C}}$ (say, $0, 1$ and ∞).*

We will abbreviate this result as *MRMT*. Its local version sounds as follows:

Theorem 14.1 (Semi-local integrability). *Let μ be a measurable Beltrami differential on a domain $U \subset \mathbb{C}$ with $\|\mu\|_\infty < 1$. Then there is a quasiconformal map*

$h : U \to \mathbb{C}$ that solves the Beltrami equation $\bar{\partial}h/\partial h = \mu$. This solution is unique up to post-composition with a conformal map.

The rest of this section will be occupied with a proof of these two theorems.

14.1. Uniqueness. Uniqueness part in the above theorems is a consequence of Weyl's Lemma. Indeed, if we have two solutions h and g, then the composition $\psi = g \circ h^{-1}$ is a qc map with $\bar{\partial}\psi = 0$ a.e. on its domain. Hence it is conformal.

14.2. Local vs global. Of course, the global MRMT immediately yields the local integrability (e.g., by zero extension of μ from U to the whole sphere). Vice versa, the global result follows from the local one and the classical Uniformization Theorem for the sphere (Theorem 5.1). Indeed, by the local integrability, there is a finite covering of the sphere $S^2 \equiv \hat{\mathbb{C}}$ by domains U_i and a family of qc maps $\phi_i : U_i \to \mathbb{C}$ solving the Beltrami equation on U_i. By Weyl's Lemma, the gluing maps $\phi_i \circ \phi_j^{-1}$ are conformal. Thus, the family of maps $\{\phi_i\}$ can be interpreted as a complex analytic atlas on S^2, which endows it with a new complex analytic structure μ (compatible with the original qc structure). But by the Uniformization Theorem, all complex analytic structures on S^2 are equivalent, so there exists a biholomorphic isomorphism $h : (S^2, \mu) \to \hat{\mathbb{C}}$. It means that the maps $h \circ \phi_i^{-1}$ are conformal on $\phi_i(U_i)$. Hence h is quasiconformal on each U_i and $h_*\mu = (h \circ \phi_i^{-1})_*\sigma = \sigma$ over there. Hence h is a global quasiconformal solution of the Beltrami equation.

14.3. Strategy. The further strategy of the proof will be the following. First, we will solve the Beltrami equation locally assuming that the coefficient μ is real analytic. It is a classical (and elementary) piece of the PDE theory. By the Uniformization Theorem, it yields a global solution in the real analytic case. Approximating a measurable Beltrami coefficient by real analytic ones and using compactness of the space of normalized K-qc maps, we will complete the proof.

14.4. Real analytic case. Assume that μ is a real analytic Beltrami coefficient in a neighborhood of 0 in $\mathbb{R}^2 \equiv \mathbb{C}_{\mathbb{R}}$ with $|\mu(0)| < 1$. Then it admits a complex analytic extension to a neighborhood of 0 in the complexification $\mathbb{C}^2$. Let (x, y) be the standard coordinates in $\mathbb{C}^2$, and let $u = x + iy,\ v = x - iy$. In these coordinates the complexified Beltrami equation assumes the form:

$$\frac{\partial h}{\partial v} - \mu(u, v)\frac{\partial h}{\partial u} = 0. \tag{14.1}$$

This is a linear equation with variable coefficients, which can be solved by the standard method of characteristics. Namely, let us consider a vector field $W(u, v) = (-\mu(u, v), 1)$ near 0 in $\mathbb{C}^2$. Since the left-hand side of (14.1) is the derivative of h along X, we come to the equation $Wh = 0$. Solutions of this equation are the first integrals of the ODE $\dot{w} = W(w)$. But since W is non-singular at 0, this ODE has a non-singular local first integral $h(u, v)$. Restricting h to $\mathbb{R}^2$, we obtain a local solution $h : (\mathbb{R}^2, 0) \to \mathbb{C}$ of the original Beltrami equation. Since h is non-singular at 0, it is a local (real analytic) diffeomorphism.

By means of the Uniformization Theorem, we can now pass from local to global solutions of the Beltrami equation with a real analytic Beltrami differential $\mu(z)d\bar{z}/dz$ on the sphere (see §14.2). Note that the global solution is real analytic as well since the complex structure generated by the local solutions is compatible with the original real analytic structure of the sphere (as local solutions are real analytic).

EXERCISE 14.2. *For a real analytic Beltrami coefficient*

$$\mu(z) = \sum a_{n,m} z^n \bar{z}^m$$

on $\mathbb{C}$*, find the condition of its real analyticity at* ∞*.*

There is also a "semi-local" version of this result:

If μ *is a real analytic Beltrami differential on the disk* $\mathbb{D}$ *with* $\|\mu\|_\infty < 1$*, then there is a* (*real analytic*) *quasiconformal diffeomorphism* $h : \mathbb{D} \to \mathbb{D}$ *solving the Beltrami equation* $\bar{\partial} h / \partial h = \mu$*.*

To see it, consider the complex structure μ on the disk generated by the local solutions of the Beltrami equation. We obtain a simply connected Riemann surface $S = (\mathbb{D}, \mu)$. By the Uniformization Theorem, it is conformally equivalent to either the standard disk $(\mathbb{D}, \sigma)$ or to the complex place $\mathbb{C}$. But S is quasiconformally equivalent to the standard disk via the identical map $\mathrm{id} : (\mathbb{D}, \mu) \to (\mathbb{D}, \sigma)$. By Exercise 12.5, it is then conformally equivalent to the standard disk, and this equivalence $h : (\mathbb{D}, \mu) \to (\mathbb{D}, \sigma)$ provides a desired solution of the Beltrami equation.

By §14.1, such a solution is unique up to a post-composition with a Möbius automorphism of the disk.

14.5. Approximation. Let us consider an arbitrary measurable Beltrami coefficient μ on a disk $\mathbb{D}$ with $\|\mu\|_\infty < \infty$. Select a sequence of real analytic Beltrami coefficients μ_n on $\mathbb{D}$ with $\|\mu_n\|_\infty \le k < 1$, converging to μ a.e.

EXERCISE 14.3. *Construct such a sequence* (*first approximate* μ *with continuous Beltrami coefficients*)*.*

Applying the results of the previous section, we find a sequence of quasiconformal maps $h_n : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ solving the Beltrami equations $\bar{\partial} h_n / \partial h_n = \mu_n$. The dilatation of these maps is bounded by $K = (1+k)/(1-k)$. By Corollary 13.12, they form a precompact sequence in the topology of uniform convergence on the disk. Any limit map $h : \mathbb{D} \to \mathbb{D}$ of this sequence is a quasiconformal homeomorphism of $\mathbb{D}$. Let us show that its Beltrami differential is equal to μ.

By (11.8), the partial derivatives of the h_n belong to some ball of the Hilbert space $L^2(\mathbb{D})$. Hence we can select weakly convergent subsequences $\partial h_n \to \phi$, $\bar{\partial} h_n \to \psi$. We have checked in (13.2) that $\phi = \partial h$ and $\psi = \bar{\partial} h$. What is left is to check that $\psi = \mu\phi$. To this end, it is enough to show that $\mu_n \, \partial h_n \to \mu\phi$ weakly (to appreciate it, recall that the product is not weakly continuous, see Exercise 13.15). For any test function $\eta \in L^2(\mathbb{D})$, we have:

$$\left| \int (\eta\mu\phi - \eta\mu_n \, \partial h_n) \, dm \right| \le$$

$$\le \left| \int \eta\mu(\phi - \partial h_n) \, dm \right| + \int |\eta(\mu - \mu_n) \, \partial h_n| \, dm.$$

The first term in the last line goes to 0 since the ∂h_n weakly converge to ϕ. The second term is estimated by the Cauchy-Schwarz inequality by $\|\eta(\mu-\mu_n)\|_2 \|\partial h_n\|_2$, which goes to 0 since $\mu_n \to \mu$ a.e. (and are uniformly bounded) while the ∂h_n belong to some Hilbert ball. This yields the desired.

It proves the Measurable Riemann Mapping Theorem on the disk $\mathbb{D}$, which certainly implies the local integrability. Now the global theorem on the sphere follows from the local integrability by §14.2. This completes the proof.

14.6. Conformal and complex structures. Let us discuss the general relation between the notions of complex and conformal structures. Consider an oriented surface S endowed with a *qs structure*, i.e., supplied with an atlas of local charts $\psi_i : V_i \to \mathbb{C}$ with uniformly qc transit maps $\psi_i \circ \psi_j^{-1}$ ("uniformly qc" means "with uniformly bounded dilatation"). Note that a notion of a measurable conformal structure with bounded dilatation makes perfect sense on such a surface (in what follows we call it just a "conformal structure").

Endow S with a complex structure compatible with its qs structure. By definition, it is determined by an atlas $\phi_i : U_i \to \mathbb{C}$ on S of uniformly qc maps such that the transit maps are complex analytic. Then the conformal structures $\mu_i = \phi_i^*(\sigma)$ on U_i coincide on the intersections of the local charts and have uniformly bounded dilatations. Hence they glue into a global conformal structure on S.

Vice versa, any conformal structure μ determines by the Local Integrability Theorem (Theorem 14.1) a complex structure on the surface S compatible with its qc structure (see §14.2).

Thus, the notions of conformal and complex structures on a qc surface are equivalent. In what follows we will not distinguish them either conceptually or notationally.

Fixing a reference complex structure on S (so that S becomes a Riemann surface), complex/conformal structures on S get parametrized by measurable Beltrami differentials μ on S with $\|\mu\|_\infty < 1$.

14.7. Explicit formula. Let us now give an explicit formula for the solution of the Beltrami equation with compact support in $\mathbb{C}$:

THEOREM 14.4. *Let μ be a Beltrami differential in $\mathbb{C}$ with compact support and $\|\mu\|_\infty < 1$. Let $h : \mathbb{C} \to \mathbb{C}$ be the solution of the Beltrami equation $\bar\partial h = \mu\,\partial h$ normalized so that $h(z) - z \to 0$ as $z \to \infty$. Then*

$$h = \mathrm{id} + T(I - \mu S)^{-1}(\mu),$$

where $S : L^2 \to L^2$ is the Hilbert transform, $T : L^2_{\mathrm{comp}} \to \mathcal{W}$ is the Cauchy transform that solves the Beltrami equation $\bar\partial(T\nu) = \nu$ (see Appendix 14.10.1).

PROOF. Let $\phi := h - \mathrm{id}$. It satisfies the equation

$$\bar\partial\phi = \mu(1 + \partial\phi).$$

Let $\nu := \bar\partial\phi$. Since $\nu \in L^2$ (Theorem 11.14), we obtain:

- as ϕ is correctly normalized ($\sim c/z$) we have $\phi = T\nu$;
- the Hilbert transform S is well defined at ν and $\partial\phi = S\nu$ (see Appendix 14.10.3).

We come up with the equation

$$(I - \mu S)\nu = \mu.$$

Since the Hilbert transform is a unitary operator in L^2, the operator μS is a contraction. Hence $I - \mu S$ is invertible, so $\nu = (I - \mu S)^{-1}(\mu)$, and the desired formula follows. □

14.8. Dependence on parameters. It is important to know how the solution of the Beltrami equation depends on the Beltrami differential. It turns out that this dependence is as best possible: holomorphic. Let us start with continuity:

PROPOSITION 14.5. *Let μ_n be a sequence of Beltrami differentials on $\mathbb{C}$ with uniformly bounded dilatation, converging a.e. to a differential μ. Consider qc solutions $h_n : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ and $h : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ of the corresponding Beltrami equations fixing $0, 1$ and ∞. Then the h_n converge to h uniformly on $\mathbb{C}$.*

PROOF. By Theorem 13.2, the sequence h_n is precompact. Take any limit map g of this sequence. By the argument of §14.5, its Beltrami differential is equal to μ. By uniqueness of the normalized solution of the Beltrami equation, $g = h$. The conclusion follows. $\square$

Let us now discuss the holomorphic dependence on parameters. Beltrami differentials are elements of the complex Banach space L^∞, while qc maps $h : \mathbb{C} \to \mathbb{C}$ are elements of the complex Sobolev space $\mathcal{W}$. So, it makes sense to talk about holomorphic maps from one space to the other (see Appendix 14.11).

Let $B : L^\infty_1 \to \mathcal{W}$ be the map that associates to a Beltrami differential $\mu \in L^\infty_1(\mathbb{D}_R)$ with $\|\mu\|_\infty < 1$ the normalized solution $h_\mu : \mathbb{C} \to \mathbb{C}$ of the Beltrami equation $\bar\partial h_\mu = \mu\, \partial h_\mu$, $h_\mu(0) = 0, h_\mu(1) = 1$.

THEOREM 14.6. *For any $R > 0$, the map $B : L^\infty_1(\mathbb{D}_R) \to \mathcal{W}$ is holomorphic.*

PROOF. Let us take a look at the explicit formula of Theorem 14.4. The Hilbert and Cauchy transforms are holomorphic as they are complex linear operators. Multiplication $(\mu, S) \mapsto A = \mu S$ is holomorphic being bilinear. Moreover, since S is unitary in L^2, we have $\|A\| = \|\mu\|_\infty < 1$. Finally, the resolvent $A \mapsto (I - A)^{-1}$ is holomorphic on the unit ball of the space of operators (see §14.11.2). As the composition of holomorphic operations is holomorphic, h_μ given by the formula depends holomorphically on μ.

It is normalized differently, though. To bring it to the normal form, notice that the points $a_\mu := h_\mu(0)$ and $b_\mu := h_\mu(1)$ depend holomorphically on μ. Hence the affine transformation

$$\phi_\mu : z \mapsto \frac{z - a_\mu}{b_\mu - a_\mu}$$

is holomorphic in two variables, z and μ. It follows that the properly normalized map $\phi_\mu \circ h_\mu$ depends holomorphically on μ as well. $\square$

The above result is usually formulated in terms of one-parameter families:

COROLLARY 14.7. *Let U be a domain in $\mathbb{C}$, $R > 0$. If the Beltrami differential $\mu_\lambda \in L^\infty_1(\mathbb{D}_R)$ holomorphically depends on a parameter $\lambda \in U$, then so do the normalized solutions $h_\lambda : \mathbb{C} \to \mathbb{C}$ of the corresponding Beltrami equations.*

Note that if h_λ depends holomorphically on λ, then any point $z \in \mathbb{C}$ moves holomorphically as λ changes (in fact, holomorphic dependence on parameters is often understood in this weak sense). More generally, we have:

COROLLARY 14.8. *Let μ_t be a family of Beltrami differentials on a disk $\mathbb{D}_R$ depending smoothly on a parameter $t \in \mathbb{R}^n$. Then the corresponding normalized solutions $h_t : \mathbb{C} \to \mathbb{C}$ of the Beltrami equation depend smootly on t as well.*

In fact, the above theorem is still valid (though not needed in this book) without assuming that the Beltrami differentials have uniformly bounded support:

EXERCISE 14.9. *Prove that the map $B : L_1^\infty \to \mathcal{W}$ (that associates to a Beltrami differential $\mu \in L_1^\infty(\mathbb{C})$ with $\|\mu\|_\infty < 1$ the normalized solution $h_\mu : \mathbb{C} \to \mathbb{C}$ of the Beltrami equation $\bar\partial h_\mu = \mu\, \partial h_\mu$, $h_\mu(0) = 0, h_\mu(1) = 1$) is holomorphic.*

14.9. Quasiregular maps. A map $h : S \to S'$ is called K-*quasiregular* if for any $z \in S$ there exist K-qc local charts $\phi : (U, z) \to (\mathbb{C}, 0)$ and $\psi : (V, f(z)) \to (\mathbb{C}, 0)$ such that $\psi \circ f \circ \phi^{-1} : z \mapsto z^d$. Sometimes we will abbreviate K-quasiregular maps as K-*qr*. A map is called *quasiregular* if it is K-qr for some K. We will also use a term *quasi-holomorphic* which sounds more suggestive.

EXERCISE 14.10. *Show that any quasiregular map $f : S \to S'$ can be decomposed as $g \circ h$, where $h : S \to T$ is a qc map to some Riemann surface T and $g : T \to S'$ is holomorphic. In particular, if $S = S' = \hat{\mathbb{C}}$ then $T = \hat{\mathbb{C}}$ as well and $g : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a rational map.*

14.10. Appendix 1: $\bar\partial$-equation and Hilbert Transform.

14.10.1. *Solution of the $\bar\partial$-equation.*

THEOREM 14.11. *Let $\nu \in L^\infty_{\mathrm{comp}}$. Then the $\bar\partial$-equation $\bar\partial v = \nu$ has a unique continuous solution v of class $\mathcal{W}$ behaving as c/z at ∞. Moreover, it can be found as the Cauchy transform of ν:*

$$v(z) = T\nu(z) := -\frac{1}{\pi} \int \frac{\nu(\zeta)}{\zeta - z}\, dm(\zeta).$$

PROOF. Let $\operatorname{supp} \nu \subset \mathbb{D}_R$. First notice that for any $z \in \mathbb{C}$, $1/(z - \zeta) \in L^1(\mathbb{D}_R)$ and its L^1-norm is locally bounded, while $\nu \in L^\infty(\mathbb{D}_R)$. Hence the Cauchy transform $v(x) = T\nu(z)$ is well defined for any $z \in \mathbb{C}$ and belongs to L^1_{loc}. Moreover, v is holomorphic on $\mathbb{C} \smallsetminus \operatorname{supp} \nu$ and decays as c/z.

Let us check that v is continuous. Consider the regular representation of the additive group of $\mathbb{C}$ is L^p:

$$L_z g(u) = g(u + z), \quad g \in L^p,\ z \in \mathbb{C}.$$

It is *strongly continuous* in the sense that for any $g \in L^p$, $L_z g$ is continuous in z. Let us take $p > 1$ and $q \in (1, 2)$ such that $1/p + 1/q = 1$. Since $1/u \in L^q_{\mathrm{loc}}$, we can apply the Hölder Inequality in the space $L_p(\mathbb{D}_R)$ (for $\varepsilon > 0$ small enough):

$$|v(z + \varepsilon) - v(z)| = \left| \int_{|u| \le 3R} \frac{\nu(z + u + \varepsilon) - \nu(z + u)}{u}\, dm(u) \right|$$

$$\le \|L_\varepsilon \nu - \nu\|_p \int_{|u| \le 3R} |1/u|^q\, dm \to 0 \quad \text{as } \varepsilon \to 0,$$

and the conclusion follows.

Let us now assume that $\nu \in C^2_{\mathrm{comp}}$. Then the Lebesgue Dominated Convergence easily justifies legitimacy of differentiation under the sign of integral, implying that v is twice differentiable and hence $v \in C^1$. Moreover,

$$\bar\partial v = -\frac{1}{\pi} \int \frac{\bar\partial \nu(z + u)}{u}\, dm(u) = \nu(z),$$

where the last equality follows from the Green Formula.

The general case is obtained by approximating ν by a a sequence $\nu_k \to \nu$ a.e., where the $\nu_k \in C^2_{\rm comp}$ have uniformly bounded support and uniformly bounded L^∞-norm. Selecting a weakly convergent (in L^2) subsequence $\nu_{k(i)}$, we ensure that the $T\nu_{k(i)}$ converge in $\mathcal{W}$, implying that $v \in \mathcal{W}$ (compare with the proof of Theorem 13.2).

Uniqueness of the solution follows from Weyl's Lemma: If $\bar\partial v = 0$ a.e. for $v \in \mathcal{W}$, then v is holomorphic. Since v vanishes at ∞, it vanishes identically. $\square$

14.10.2. *Fourier transform.* Let $< z, \zeta >= \mathrm{Re}(z\bar\zeta)$ stand for the standard Hermitian structure in $\mathbb{C}$. Recall that a Fourier transform of a function $\phi \in L^1 \equiv L^1(\mathbb{C})$ is defined as

$$\hat\phi(z) \equiv \mathcal{F}\phi(z) := \frac{i}{2}\int_{\mathbb{C}} \phi(\zeta) e(- < z, \zeta >)\, d\zeta \wedge d\bar\zeta$$

Here is the list of basic properties of the Fourier transform $\mathcal{F}$:

• It is a *contracting* operator $L^1 \to C_0$, where C_0 is the space of continuous functions $\phi : \mathbb{C} \to \mathbb{C}$ such that $\phi(z) \to 0$ as $z \to 0$.

• It is an *algebra homomorphism*, where the multiplication in L^1 is convolution $*$, while the multiplication in C_0 is pointwise.

• It preserves the L^2-norm:

$$\|\hat\phi\|_2 = \|\phi\|_2, \quad \text{for any } \phi \in L^1 \cap L^2, \tag{14.2}$$

and hence extends to a *unitary operator* $L^2 \to L^2$ (for which we will keep the same notation). Equality (14.2) is called *Parseval's Identity.*

• It conjugates the *partial derivations* $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ to the *multiplication* operators by the corresponding variables (up to $2\pi i$-factor):

$$\widehat{\partial_x \phi} = 2\pi i x\, \hat\phi, \quad \widehat{\partial_y \phi} = 2\pi i y\, \hat\phi,$$

for any function $\phi \in L^1_0$ with L^1_0 distributional partial derivatives. It follows that

$$\widehat{\partial\phi} = \pi i \bar z\, \hat\phi, \quad \widehat{\bar\partial\phi} = \pi i z\, \hat\phi$$

in this class of functions.

14.10.3. *Hilbert Transform.* The *Hilbert transform* is a unitary operator $S : L^2 \to L^2$ that carries $\bar\partial\phi$ to $\partial\phi$ for any function $\phi \in \mathcal{W}$. In the Fourier chart, it is defined as the multiplication operator:

$$S\hat\phi = \frac{\bar\zeta}{\zeta}\,\hat\phi, \quad \phi \in L^2.$$

The Hilbert trasform can be explicitly defined on functions $\phi \in C^2_{\rm comp}$ as the principal value of the following singular integral:

$$S\phi(z) = -\frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|\zeta - z|>\varepsilon} \frac{\phi(\zeta)}{(\zeta - z)^2}\, dm.$$

Then it can be isometrically extended to L^2. See [**A2**, Ch V.A],[**Ste**].

14.11. Appendix 2: Holomorphic maps between Banach spaces.

14.11.1. *Smoothness in Banach spaces.* Basic notions in this generality are the same as in the finite dimensional case. Let $\mathcal{B}$ and $\mathcal{Q}$ be Banach spaces, and let $\mathcal{U}$ be a domain in $\mathcal{B}$. We refer to maps $\mathcal{U} \to \mathcal{Q}$ as *Banach maps.*

A Banach map $f : \mathcal{U} \to \mathcal{Q}$ is called *differentiable at a point* $\mu \in \mathcal{U}$ if

$$f(\mu + v) = f(\mu) + D_\mu f(v) + o(\|v\|) \quad \text{for all } v \in \mathcal{B} \text{ small enough,}$$

where $D_\mu f : \mathcal{B} \to \mathcal{Q}$ is a (bounded) linear operator called the *differential of* f. The map f is called (C^1)-*smooth* on $\mathcal{U}$ if it is differentiable at all points $\mu \in \mathcal{U}$, and the differential $D_\mu f$ depends continuously on μ. It is called a *diffeomorphism* (onto its image $\mathcal{V}$), if $\mathcal{V}$ is open in $\mathcal{Q}$, f is invertible, and the inverse map f^{-1} is smooth as well. The Banach category is appropriate for the smooth theory since the Implicit Function Theorem is still valid in this generality:

Implicit Function Theorem. *Let $f : (\mathcal{U}, 0) \to (\mathcal{Q}, 0)$ be a smooth Banach map such that $D_0 f : \mathcal{B} \to \mathcal{Q}$ is an invertible linear operator. Then f is a local diffeomorphism.*

14.11.2. *Holomorphic maps: definitions and examples.* In what follows, *all Banach spaces are assumed to be complex.* A continuous Banach map $f : \mathcal{U} \to \mathcal{Q}$ is called *holomorphic* if for any complex line $\mathcal{L} = \{x + \lambda v\}_{\lambda \in \mathbb{C}}$ (where $x, v \in \mathcal{B}$) and any (bounded) test linear functional $\Phi \in \mathcal{Q}^*$, the composition $\Phi \circ f | \mathcal{L} \cap \mathcal{U}$ is holomorphic in λ. (As we see, this is essentially one-dimensional notion.)

A Banach map $f : \mathcal{B} \to \mathcal{Q}$ is called a degree d *homogeneous polynomial* if it is the restriction of a (bounded) degree d polylinear map

$$\tilde{f} : \mathcal{B} \times \cdots \times \mathcal{B} \to \mathcal{Q}, \quad \tilde{f}(v_1 \dots v_d) \le C \|v_1\| \cdots \|v_d\|,$$

to the diagonal $\Delta = \{(v, \dots v) : \ v \in \mathcal{B}\}$. For instance, let $\operatorname{Aut}(\mathcal{B})$ be the space of (bounded) linear operators $A : \mathcal{B} \to \mathcal{B}$. Then the map $\operatorname{Aut}(\mathcal{B}) \to \operatorname{Aut}(\mathcal{B})$, $A \mapsto A^d$ is a homogeneous polynomial of degree d.

A *polynomial* is a sum of homogeneous polynomials.

Let us temporarily[6] say that a Banach map $\mathcal{U} \to \mathcal{Q}$ is *strongly holomorphic* if it admits a *Taylor expansion* near any point $\mu \in \mathcal{U}$:

$$f(\mu + v) = f(\mu) + D_\mu f(v) + D_\mu^2 f(v) + \dots,$$

where $D_\mu^d f$ is a homogeneous degree d polynomial in v. If this series converges in the whole space $\mathcal{B}$, f is called *entire.*

For instance, consider a series

$$f(A) = \sum_{d=0}^{\infty} c_d A^d, \quad \text{with } c_d \le r^d.$$

It defines a holomorphic map $\operatorname{Aut}(\mathcal{B}) \to \operatorname{Aut}(\mathcal{B})$ on the ball $\operatorname{Aut}_{1/r}(\mathcal{B})$ of radius $1/r$. (In particular, if the c_n decay super-exponentially, then it defines an entire function.) Here are two important examples: the *exponential map*

$$\exp : \operatorname{Aut} \mathcal{B} \to \operatorname{Aut} \mathcal{B}, \quad \exp A = \sum_{d=0}^{\infty} \frac{A^d}{d!},$$

[6] As we will see momentarily that this notion is equivalent to being holomorphic.

is entire, and the *resolvent*:

$$R(A) \equiv (I - A)^{-1} = \sum_{d=0}^{\infty} A^d$$

is strongly holomorphic in the unit ball of $\operatorname{Aut}(\mathcal{B})$.

EXERCISE 14.12. (i) *Complex linear and polynomial maps are holomorphic.*
(ii) *Uniform limits of holomorphic maps are holomorphic.*
(iii) *Strongly holomorphic maps* (*in particular, the exponential and the resolvent*) *are all holomorphic.*

A *holomorphic curve* in $\mathcal{B}$ is a subset Γ of $\mathcal{B}$ that locally admits a holomorphic parametrization $\gamma : \mathbb{D} \to \mathcal{B}$.

PROPOSITION 14.13. *For a continuous Banach map $f : \mathcal{U} \to \mathcal{Q}$, $\mathcal{U} \subset \mathcal{B}$, the following properties are equivalent:*
(i) *f is holomorphic;*
(ii) *f is smooth with complex linear differentials $D_\mu f : \mathcal{B} \to \mathcal{Q}$;*
(iii) *The restriction of f to any holomorphic curve in $\mathcal{B}$ is holomorphic;*
(iv) *f is strongly holomorphic.*

PROOF. Since (i) is the weakest property, while (iv) is the strongest, it is sufficient to show that (i) $\implies$ (iv). Let us first review the case of a scalar function on a finite-dimensional space, i.e, let $\mathcal{B} = \mathbb{C}^n$, $\mathcal{Q} = \mathbb{C}$. Combining the classical 1D Cauchy formula with the Fubini Formula, we obtain the n-dimensional Cauchy representation of f (for $r > 0$ sufficiently small):
(14.3)

$$f(z+v) = \frac{1}{(2\pi i)^n} \int_{\mathbb{T}_r^n} \frac{f(z+\zeta)\, d\zeta_1 \dots d\zeta_n}{(\zeta_1 - v_1) \dots (\zeta_n - v_n)}, \quad v = (v_1 \dots v_n) \in \mathbb{D}_r^n,\ \zeta = (\zeta_1 \dots \zeta_n).$$

Now the geometric series expansion (in 1D)

$$\frac{1}{\zeta - v} = \sum_{m=0}^{\infty} \frac{v^m}{\zeta^{m+1}}, \quad |v| < |\zeta|,$$

implies the Taylor expansion for f.

For a general $\mathcal{Q}$, while $\mathcal{B} = \mathbb{C}^n$, the Cauchy contour integral still makes sense (as an integral of a continuous Banach-valued function). Cauchy Formula (14.3) is still valid since it can be tested by any linear functional $\Phi \in \mathcal{Q}^*$. It implies the Taylor expansion as in the scalar case.

Let us now consider a $\mathcal{Q}$-valued function on a general $\mathcal{B}$. To define the polarized dth differential $D_z^d(v_1, \dots, v_d)$, take any finite dimensional subspace $E \subset \mathcal{B}$ containing all the vectors v_k, and use the finite dimensional result in $z + E$. The outcome is independent of the choice of E: for another subspace E' as above, we can consider $E \oplus E'$ which induces the same outcome as either E or E'. ☐

COROLLARY 14.14. *Holomorphic maps are smooth.*

COROLLARY 14.15. (i) *Composition of two holomorphic maps is holomorphic.*
(ii) *If f is holomorphic and invertible, then the inverse is holomorphic as well.*

Corollary 14.16. *A map $f : (\mathbb{D}_r, 0) \to (\mathcal{B}, 0)$ is holomorphic if and only it admits a power series representation*

$$f(\lambda) = \sum_{n=0}^{\infty} a_n \lambda^n,$$

where $a_n \in \mathcal{B}$ and $\|a_n\| \le C\rho^n$ for any $\rho > 1/r$ (with C depending on ρ).

14.11.3. *Cauchy Inequality and Normality.* As in dimension one, the *Cauchy Inequality* bounds the derivative in terms of the map:

Proposition 14.17. *If $\sup_{\mathcal{U}} \|f(\mu)\| \le M$, then*

$$\|D_\mu f\| \le \frac{M}{\operatorname{dist}(\mu, \partial\mathcal{U})}, \quad \mu \in \mathcal{U}.$$

Proof. Let $r := \operatorname{dist}(\mu, \partial\mathcal{U})$. Take a normed vector $v \in \mathcal{B}$ and a normed functional $\Phi \in \mathcal{Q}^*$. Applying the classical Cauchy Inequality to the holomorphic function $\lambda \mapsto \Phi(f(\mu + \lambda v))$, $\lambda < r$, we obtain $|\Phi(D_\mu f(v))| \le M/r$. Taking sup over Φ, we obtain $\|D_\mu f(v)\| \le M/r$ (by the Hahn-Banach Theorem). Taking sup over v, we obtain the desired. □

For a domain $\mathcal{U} \subset \mathcal{B}$, let us say that a subset $\mathcal{K} \subset \mathcal{U}$ is *strictly contained* in $\mathcal{U}$ if $\operatorname{dist}(\mathcal{K}, \partial\mathcal{U}) > 0$. We endow the space of holomorphic functions $\mathcal{U} \to \mathbb{C}$ with the topology of uniform convergence (for generalized sequences) on strictly contained subsets.

Proposition 14.18. *Given a domain $\mathcal{U}$ in a Banach space $\mathcal{B}$ and $M > 0$, the family of all holomorphic functions $\mathcal{U} \to \overline{\mathbb{D}}_M$ is compact.*

Proof. By the Cauchy Inequality, such a family is equicontinuous on any strict subset $\mathcal{K}$ of $\mathcal{U}$. The Ascoli-Arcela criterion implies precompactness of our family in the space of continuous functions $\mathcal{U} \to \overline{\mathbb{D}}_M$. Since uniform limits of holomorphic functions are holomorphic, the conclusion follows. □

This validates the Montel Theorems for families of functions on Banach domains.

Proposition 14.19. *Let $f_n : \mathcal{U} \to \mathcal{Q}$ be a bounded (generalized) sequence of holomorphic maps pointwise converging to a map f, i.e., $f_n(\mu) \to f$ for any $\mu \in \mathcal{U}$. Then f is holomorphic as well.*

Proof. For any functional $\Phi \in \mathcal{Q}^*$, the sequence of functions $g_n := \Phi \circ f_n$ pointwise converges to $\Phi \circ f$. By the last Proposition, the functions g_n form a normal family. Hence there is a generalized subsequence converging (uniformly on strict subsets $\mathcal{K} \subset \mathcal{U}$, and hence pointwise) to a holomorphic function g. Necessarily, $f = g$, and we are done. □

14.11.4. *Sufficient supply of test functionals.* Here we will see that holomorphicity can be tested by dense sets of functionals.

Lemma 14.20. *Let $f : \mathcal{U} \to \mathcal{Q}$ be a a locally bounded Banach map, $\mathcal{U} \subset \mathcal{B}$. If the function $\Phi \circ f$ is holomorphic for a w^*-dense set $\mathcal{Q}_0^*$ of functionals $\Phi \in \mathcal{Q}^*$, then f is holomorphic.*

Proof. We need to show that $\Phi \circ f$ is holomorphic for all $\Phi \in \mathcal{Q}^*$. By assumption, any $\Phi \in \mathcal{Q}^*$ is the w^*-limit of some directed family $(\Phi_n) \subset \mathcal{Q}_0^*$. Hence $\Phi_n \circ f \to \Phi \circ f$ pointwise on $\mathcal{U}$. Since the family of functions $(\Phi_n \circ f)$ is locally bounded, the conclusion follows from Proposition 14.19. □

Corollary 14.21. *Let $\mu : \mathcal{U} \to \mathcal{Q}^*$, $\lambda \mapsto \mu_\lambda$, be a continuous family in the dual space such that $\lambda \mapsto \mu_\lambda(\phi)$ is holomorphic for a w-dense set of test elements $\phi \in \mathcal{Q}$. Then μ is holomorphic.*

Proof. Since the natural embedding $\mathcal{Q} \to \mathcal{Q}^{**}$ has a dense image in the w^*-topology of $\mathcal{Q}^{**}$, we obtain a w^*-dense set of functionals on $\mathcal{Q}^*$ to test holomorphicity. □

The above lemmas allow us to test holomorphicity on C^∞-smooth functions only:

Corollary 14.22. *Holomorphicity of a map $f : \mathcal{U} \to \mathcal{Q}$ to any of the functional spaces $\mathcal{Q} = L^p(D)$ or $\mathcal{W}^p(D)$, $p \in [1, \infty]$, can be tested by pairing of f with functions $\phi \in C^\infty_{\mathrm{comp}}(D)$.*

14.11.5. *Holomorphic curves in functional spaces.* The space L^∞ is particularly important for us since Beltrami differentials belong to this class.

Lemma 14.23. *Let μ_λ be a family in $L^\infty(D)$ over a domain $\Lambda \subset \mathbb{C}$. It is holomorphic in λ if and only if it is locally bounded and the functions $\lambda \mapsto \mu_\lambda(z)$ are holomorphic in λ for a.e. z (after making an appropriate choice of representatives of the μ_λ).*

Proof. Without loss of generality we can assume that $\Lambda = \mathbb{D}$ is the unit disk.

Assume $\lambda \mapsto \mu_\lambda$ is holomorphic over $\mathbb{D}$. Then by Corollary 14.16, it admits a power series representation

$$\mu_\lambda(z) = \sum_{n=0}^{\infty} \nu_n(z)\, \lambda^n, \tag{14.4}$$

where $\nu_n \in L^\infty$ and $\|\nu_n\|_\infty \leq C\rho^n$ for any $\rho > 1$. Hence there exists a subset $X \subset D$ of full measure such that for any $\rho > 1$ we have:

$$\nu_n(z) \leq C\rho^n \quad \forall\, z \in X.$$

It follows that for any $z \in X$, the function $\lambda \mapsto \mu_\lambda(z)$ is holomorphic over $\mathbb{D}$ (where the representative of μ_λ on X is chosen by the power series (14.4)).

Vice versa, assume that for a.e. $z \in D$, the function $\lambda \mapsto \mu_\lambda(z)$ is holomorphic over $\mathbb{D}$. Then (14.4) holds for a.e. $z \in D$, with

$$\nu_n(z) = \frac{1}{2\pi i} \int_{|\lambda| = r} \frac{\mu_\lambda(z) d\lambda}{(-\lambda)^{n+1}} \quad \text{for any } r \in (0, 1).$$

But since the family (μ_λ) is locally bounded, it is bounded over the circle $\{|\lambda| = r\}$, implying that

$$|\nu_n(z)| \leq \frac{C}{r^n},$$

with $C = C(r)$ independent of z. Hence $\|\nu_n\|_\infty = O(r^{-n})$, and the map $\lambda \mapsto \mu_\lambda$ is holomorphic by Corollary 14.16. □

EXERCISE 14.24. *Let $f : S \to T$ be a holomorphic map between two Riemann surfaces, and let (μ_λ) be a holomorphic family of Beltrami differentials on T. Then $(f^*(\mu_\lambda))$ is a holomorphic family of Beltrami differentials on S.*

Functions of class L^2 appear in our context as derivatives of qc maps.

LEMMA 14.25. *Let ϕ_λ be a complex one-parameter family in $L^2(D)$ over a domain $\Lambda \subset \mathbb{C}$. It is holomorphic in λ if and only if it is locally bounded in L^2 and the functions $\lambda \mapsto \phi_\lambda(z)$ are holomorphic in λ for a.e. z (after making an appropriate choice of representatives of the μ_λ).*

PROOF. As in the previous lemma, assume that $\Lambda = \mathbb{D}$ is the unit disk and consider a power series representation

$$\phi_\lambda(z) = \sum_{n=0}^{\infty} \psi_n(z)\, \lambda^n,$$

where $\psi_n \in L^2(D)$ and $\|\psi_n\|_2 \le C\rho^n$ for any $\rho > 1$. Let $\sigma > \rho$. By the Chebyshev Inequality,

$$\text{area}\{z : \ |\psi_n(z)| \ge \sigma^n\} \le C^2 \left(\frac{\rho}{\sigma}\right)^{2n}.$$

By the Borel-Cantelli Lemma, $|\psi_n(z)| = O(\sigma^n)$ a.e. The conclusion follows as in the previous lemma.

The inverse statement we leave to the reader. □

Finally, let us consider the space $\mathcal{W}$ corresponding to qc maps themselves:

LEMMA 14.26. *Let h_λ be a complex one-parameter family in $\mathcal{W}(D)$ over a domain $\Lambda \subset \mathbb{C}$. It is holomorphic in λ if and only if it is locally bounded in $\mathcal{W}$ and for any $z \in D$, the evaluation function $\lambda \mapsto h_\lambda(z)$ is holomorphic in λ. Moreover, in this case the partial derivatives $(\partial h_\lambda)_{\lambda\in\Lambda}$ and $(\bar\partial h_\lambda)_{\lambda\in\Lambda}$ form holomorphic curves in L^2.*

PROOF. The necessity is obvious since for any $z \in D$, the evaluation $h \mapsto h(z)$ is a linear functional on $\mathcal{W}$. Vice versa, if all the evaluations are holomorphic in λ, then for any test function $\psi \in C^\infty_{\text{comp}}(D)$, the pairing

$$\int_D h_\lambda(z)\, \psi(z)\, dm(z)$$

is holomorphic in λ as well, and Corollary 14.22 implies that $\lambda \mapsto h_\lambda$ is a holomorphic curve in $\mathcal{W}$.

Plugging $\partial\psi$ is place of ψ, we see that

$$\int_D h_\lambda(z)\, \partial\psi(z)\, dm(z) = -\int_D \partial h_\lambda(z)\, \psi(z)\, dm(z),$$

depends holomorphically on λ. Applying Corollary 14.22 again, we conclude that $\lambda \mapsto \partial h_\lambda$ is a holomorphic curve in L^2. Similarly, $\lambda \mapsto \bar\partial h_\lambda$ is. □

15. One-dimensional qs maps, quasicircles and qc welding

In this section, we will develop further the idea of quasisymmetry (see §12.3) for one-dimensional maps and plane curves.

15.1. Quasisymmetric 1D maps.

15.1.1. *QS maps of the line.* Let us first consider the rel line $\mathbb{R}$ in the Euclidean metric. According to Lemma 12.11, L-qs maps $h : \mathbb{R} \to \mathbb{R}$ can be defined as in terms of bounded macroscopic dilatation. Namely, for any two adjacent intervals $I, J \subset \mathbb{R}$ of equal length, we require:

$$\frac{|f(I)|}{|f(J)|} \leq L. \tag{15.1}$$

It looks at first glance that the class of 1D qs maps is a good analogue of the class of 2D qc maps. However, this impression is superficial: two-dimensional qc maps are fundamentally better than one-dimensional qs maps. For instance, qc maps can be glued together without any loss of dilatation (the Gluing Lemma) while qs maps cannot:

EXERCISE 15.1. (i) *For any $\delta > 0$, the power map $h : [0, 1] \to [0, 1]$ is qs.*
(ii) *Consider a map $h : \mathbb{R} \to \mathbb{R}$ equal to* id *on the negative axis, and equal to $x \mapsto x^2$ on the positive one. This map is not quasi-symmetric, though its restrictions to the both positive and negative axes are.*

Another deficiency of one-dimensional qs maps is that they can well be singular (and typically are in the dynamical setting), while 2D qc maps are always absolutely continuous (Proposition 11.14).

These advantages of qc maps makes them much more efficient tool for dynamics than one-dimensional qs maps. This is one of the reasons why complexification of one-dimensional dynamical systems is so powerful.

15.1.2. *QS circle maps.* Of course, an L-qs circle homeomorphism $h : \mathbb{T} \to \mathbb{T}$ can be defined is the same way as in the case of $\mathbb{R}$, with understanding of (15.1) in terms of the circle metric. However, there is a subtle difference between these two cases. Namely, in the line case, the group of 1-qs maps coincides with the group of affine maps $x \mapsto ax + b$, which is equal to the group of Möbius automorphisms of $\mathbb{R}$. On the other hand, in the circle case, only rotations are 1-qs, and in fact,

EXERCISE 15.2. *The group of Möbius automorphisms ϕ of the circle $\mathbb{T}$ is not uniformly qs. However, if $\phi(0) \leq r < 1$ then ϕ is $L(r)$-qs.*

15.1.3. *Tilings with bounded geometry.* Let I be a closed interval or a circle. Assume we have an increasing nest $(\mathcal{T}^n)_{n=0}^{\infty}$ of tilings of I,

$$\mathcal{T}^0 \succ \mathcal{T}^1 \succ \dots,$$

by intervals T_k^n, $k = 0, 1, \dots, p_n - 1$. One says that the nest has *bounded combinatorics* if each interval T_k^n is tiled by a bounded number of intervals T_j^{n+1} of the next level. It has *bounded geometry* if all such nested intevals $T_k^n \supset T_j^{n+1}$ are comparable in size. Obviously, bounded geometry implies bounded combinatorics.

EXERCISE 15.3. *Assume we have two intervals I and $\tilde{I}$ (or two circles) supplied with two nests of tilings as above, $(\mathcal{T}^n)$ and $(\tilde{\mathcal{T}}^n)$, with bounded geometry. Let $h : I \to \tilde{I}$ be a homeomorphism respecting these tilings, i.e., $h(T_k^n) = \tilde{T}_k^n$ for any tile T_k^n. Then h is qs (quantitatively).*

This statement will be useful in the dynamical setting where nests of tilings as above appear naturally.

15.2. Ahlfors-Beurling Extension.

15.2.1. *Extension from* $\mathbb{R}$. As we know, the class of orientation preserving qs maps on the plane coincides with the class of qc maps (Propositions 12.13 and 12.14). In particular, if we consider a quasiconformal map $h : \mathbb{C} \to \mathbb{C}$ preserving the real line $\mathbb{R}$, it restricts to a quasisymmetric map on the latter. Remarkably, the inverse is also true:

THEOREM 15.4. *Any L-qs orientation preserving map $h : \mathbb{R} \to \mathbb{R}$ extends to a $K(L)$-qc map $H : \mathbb{C} \to \mathbb{C}$. Moreover, this extension can be selected to be affinely equivariant (i.e, so that it commutes with the action of the affine group $z \mapsto az + b$, $a \in \mathbb{R}_+,\ b \in \mathbb{R}$). Moreover, this map is smooth outside $\mathbb{R}$.*

PROOF. An extension to the upper half-plane $\mathbb{H}$ can be given by an explicit formula:

$$H(x + iy) = \frac{1}{2y} \int_{x-y}^{x+y} h(t)dt + \frac{i}{y} \left(\int_{x}^{x+y} h(t)dt - \int_{x-y}^{x} h(t)dt \right).$$

It is clearly smooth in $\mathbb{H}$ and is continuous up to the boundary with boundary values h. By a fairly direct calculation, one can check that it has a positive Jacobian (so it is a local orientation preserving diffeomorphism) and to bound its dilatation in terms of L. One should also check that $H(z) \to \infty$ as $z \to \infty$ in $\mathbb{H}$, so H is proper. As $h|\, \mathbb{R} \cup \{\infty\}$ is a homeomorphism, we conclude that $\deg H = 1$, and hence H is a homeomorphism as well.

EXERCISE 15.5. *Supply omitted technical details.*

Finally, the transform $h \mapsto H$ is manifestly affinely equivariant, and it extends to the lower half-plane by reflection. □

15.2.2. *Extension from* $\mathbb{T}$. As the group of Möbius automorphisms of the circle is not uniformly qs, the circle version of the Ahlfors-Beurling Theorem requires some extra care:

LEMMA 15.6. *Let $H : \mathbb{D} \to \mathbb{D}$ be a K-qc map with $H(0) \le r < 1$. Then H admits an extension to a $L(K, r)$-qs circle homeomorphism.*

Vice versa, any L-qs circle homeomorphism h admits and extension to a $K(L)$-qc map $H : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$.

PROOF. H can be continuously extended to $\mathbb{T}$, and then by symmetry to the whole Riemann sphere.

Since Möbius automorphisms $\phi : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ with $|\phi(0)| \le r$ are $L(r)$-qs on $\mathbb{T}$, H can be normalized so that $H(0) = 0$, and by symmetry, $H(\infty) = \infty$. Then $H(\mathbb{C}) = \mathbb{C}$, and Lemma 12.6 implies that H is $L(K)$-qs. □

15.2.3. *Interpolation in an annulus.*

LEMMA 15.7. *Let us consider two round annuli $A = \mathbb{A}[1, r]$ and $\tilde{A} = \mathbb{A}[1, \tilde{r}]$, with $0 < \varepsilon \le \operatorname{mod} A \le \varepsilon^{-1}$ and $\varepsilon \le \operatorname{mod} \tilde{A} \le \varepsilon^{-1}$. Then any κ-qs map $h : (\mathbb{T}, \mathbb{T}_r) \to (\tilde{\mathbb{T}}, \tilde{\mathbb{T}}_{\tilde{r}})$ admits a $K(\kappa, \varepsilon)$-qc extension to a map $H : A \to \tilde{A}$.*

PROOF. Since A and $\tilde{A}$ are ε^2-qc equivalent, we can assume without loss of generality that $A = \tilde{A}$. Let us cover A by the upper half-plane, $\theta : \mathbb{H} \to A$, $\theta(z) = z^{\frac{-\log ri}{\pi}}$, where the covering group generated by the dilation $T : z \mapsto \lambda z$, with

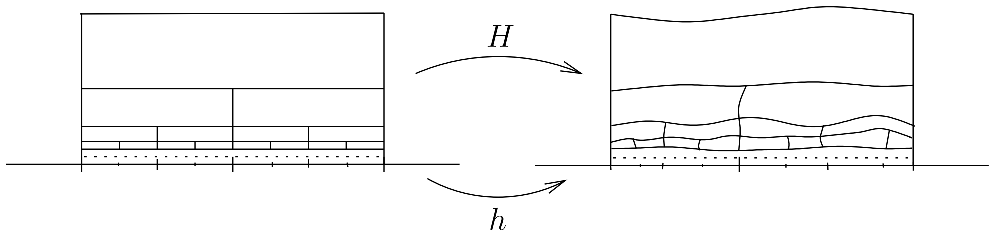

FIGURE 15.1. QC extension of a qs map by means of Calreson-Whitney tilings.

$\lambda = e^{\frac{2\pi^2}{\log r}}$. Let $\bar{h} : (\mathbb{R}, 0) \to (\mathbb{R}, 0)$ be the lift of h to $\mathbb{R}$ such that $\bar{h}(1) \in [1, \lambda) \equiv I_\lambda$ and $\bar{h}(1) \in (-\lambda, -1]$ (note that $\mathbb{R}_+$ covers $\mathbb{T}_r$, while $\mathbb{R}_-$ covers $\mathbb{T}$). Moreover, since $\deg h = 1$, it commutes with the deck transformation T.

A direct calculation shows that the dilatation of the covering map θ on the fundamental intervals I_λ and $-I_\lambda$ is comparable with $(\log r)^{-1}$. Hence $\bar{h}$ is $C(\kappa, r)$-qs on this interval. By equivariance it is $C(\kappa, r)$-qs on the rays $\mathbb{R}_+$ and $\mathbb{R}_-$.

It is also quasi-symmetric near the origin. Indeed, by the equivariance and normalization,

$$(1 + \lambda)^{-1}|J| \le |\bar{h}(J)| \le (1 + \lambda)|J|$$

for any interval J containing 0, which easily implies quasi-symmetry.

Since the Ahlfors-Beurling extension is affinely equivariant, the map $\bar{h}$ extends to a $K(\kappa, r)$-qc map $\bar{H} : \mathbb{H} \to \mathbb{H}$ commuting with T. Hence $\bar{H}$ descends to a $K(\kappa, r)$-qc map $H : A \to A$. $\square$

15.2.4. *QS equivalence between Cantor sets with bounded geometry.* Recall from §1.1.1 basics about combinatorics and geometry of real Cantor sets. The following assertion will have important dynamical ramifications (see §38.9):

EXERCISE 15.8. *Any two Cantor sets $K, \tilde{K} \subset \mathbb{R}$ with the same combinatorics and bounded geometry qs equivalent, i.e., there exists a quasisymmetric homeomorphism $h : (\mathbb{R}, K) \to (\mathbb{R}, \tilde{K})$ respecting the combinatorics. Moreover, the dilatation of h depends only on the bounds on the geometry of K and $\tilde{K}$.*

15.3. Quasicircles.

15.3.1. *Geometric definition.* Let us start with an intrinsic geometric definition of quasicircles:

DEFINITION 15.9. A Jordan curve $\gamma \subset \mathbb{C}$ is called a κ-*quasicircle* if for any two points $x, y \in \gamma$ there is an arc $\delta \subset \gamma$ bounded by these points such that

$$\operatorname{diam} \delta \le \kappa |x - y|. \tag{15.2}$$

A curve is called a quasicircle if it is a κ-quasicircle for some κ. The best possible κ in the above definition is called the *geometric dilatation* of the quasicircle. Let us emphasize that this notion is *global* in the sense that (15.2) should be satisfied in all scales. However, it can be localized as follows:

EXERCISE 15.10. *If* (15.2) *is satisfied for all pairs of points with $|x - y| \le \varepsilon$, then γ is a κ'-quasicircle with κ' depending only on κ and N, where N is the number of arcs of* $\operatorname{diam} \varepsilon$ *needed to cover γ.*

A Jordan disk (either open or closed) is called (κ-)*quasidisk* if it is bounded by a (κ-)quasicircle.

EXERCISE 15.11. *A Jordan disk D is a κ-quasidisk if and only if the Euclidean path metric on $\overline{D}$ is κ-Lipschitz equivalent to the Euclidean chordal metric.*

A *C-quasi-center* of a Jordan curve γ (or, of the corresponding Jordan disk D) is a point $a \in D$ such that D has a C-bounded shape around a:

$$\frac{R_D(a)}{r_D(a)} \le C.$$

(Here $R_D(a)$ and $r_D(a)$ are outer and inner radii of D around a, see §4.4.)

EXERCISE 15.12. *Any κ-quasidisk has a $C(\kappa)$-quasi-center.*

The shape bound $C(\kappa)$ will often be implicit in our discussion, and sometimes we will even say that D is "centered at a".

EXERCISE 15.13. *Let γ be a 0-symmetric κ-quasicircle. Then 0 is $(2\kappa+1)$-quasi-center of γ.*

For a simply connected domain $D \subset \hat{\mathbb{C}}$, let us say that a point $z_0 \in \partial D$ is a *cusp* if $\operatorname{dist}(z, z_0)/\operatorname{dist}(z, \partial D) \to 0$ as $z \to z_0$ in D. The following simple assertion gives the best intuitive characterization of quasicircles:

EXERCISE 15.14. *Quasicircles do not have cusps.*

15.3.2. *Quasi-rectangles and the cross-ratios.* Given four points a, b, c, d on a Jordan curve γ, let $\Pi_\gamma(a, b, c, d)$ stand for the corresponding quadrilateral. In case when γ is a quasicircle, this quadrilateral will be called a *quasi-rectangle*.

LEMMA 15.15. *The modulus of a quasi-rectangle,* $\operatorname{mod}(\Pi_\gamma(a, b, c, d))$*, is controlled by the cross-ratio $R := [a, b, c, d]$. More precisely,*

$$0 < \theta_1(R) \le \operatorname{mod} \Pi_\gamma(a, b, c, d) \le \theta_2(R),$$

where the functions θ_i depend only on the geometric dilatation of γ, and $\theta_1(R) \to \infty$ as $R \to \infty$.

15.3.3. *Quasitriangles and ratios.* A Jordan domain D with four marked points a, b, c, d such that $a, b, c \in \gamma = \partial D$ while $d \in \operatorname{int} D$ is called a *pointed topological triangle* $\Delta_\gamma(a, b, c; d)$. Let as define $\operatorname{mod} \Delta_\gamma(a, b, c; d)$ as the extremal length of the family of proper paths $\gamma \subset D$ connecting $[a, b]$ to $[c, a]$ and separating d from $[b, c]$. In case when γ is a quasicircle centered at d, $\Delta_\gamma(a, b, c; d)$ will be called *pointed quasitriangle*.

LEMMA 15.16. *The modulus of a quasitriangle,* $\operatorname{mod} \Delta_\gamma(a, b, c; d)$*, is controlled by the ratio $R := |b - c|/|b - a|$, in the same sense as above.*

15.3.4. *The Riemann mapping.* What makes quasicircles so important is their characterization as qc images of the circle:

THEOREM 15.17. *Let a be a quasi-center of a κ-quasidisk D, and let $\psi : (\mathbb{D}, 0) \to (D, a)$ be the normalized Riemann mapping. Then ψ admits a $K(\kappa)$-qc extension to the whole complex plane.*

Vice versa, let (D, a) be a pointed Jordan disk such that there exists a K-qc map $h : (\mathbb{C}, \mathbb{D}, 0) \to (\mathbb{C}, D, a)$. Then D is a κ-quasidisk with a quasi-center a.

PROOF. The last assertion follows immediately from the fact that h has $L(K)$-bounded macroscopic dilatation (by Lemma 12.6). □

15.3.5. *Quasi-annuli.* A *C-quasi-annulus* is a conformal annulus $A \subset \mathbb{C}$ such that there is a C-qc map $h : (\mathbb{C}, A) \to (\mathbb{C}, \mathbb{A}(1, r))$. By the second part of Theorem 15.17, a C-quasi-annulus is bounded by $\kappa(C)$-quasicircles. Vice versa, we have:

LEMMA 15.18. *Let A be a conformal annulus with $\operatorname{mod} A = \log r \geq \mu > 0$ bounded by κ-quasicircles. Then A is a $C(\mu, \kappa)$-quasi-annulus. In fact, the Riemann mapping $\phi : A \to \mathbb{A}(1, r)$ admits a $K(\mu, \kappa)$-qc extension to the whole plane.*

LEMMA 15.19. *Let A and $\tilde{A}$ be C-quasi-annuli with $\min(\operatorname{mod} A, \operatorname{mod} \tilde{A}) \geq \mu > 0$. Then any L-qs map $h : \partial A \to \partial \tilde{A}$ admits a $K(C, \mu, L)$-extension to $\mathbb{C}$ (quantitatively).*

EXERCISE 15.20. *Assume an annullus A is partitioned by a $\kappa-$quasicircle γ into two homotopic sub-annuli A_i with $\operatorname{mod} A_i \geq \mu > 0$. Then*

$$\operatorname{mod} A \leq C(\mu, \kappa)\,(\operatorname{mod} A_1 + \operatorname{mod} A_2).$$

15.3.6. *Little Gluing Lemma.*

LITTLE GLUING LEMMA. *Let Γ be a piecewise quasicircle (or quasiarc) contained in a domain $U \subset \mathbb{C}$, and let $h : U \to V$ be a homeomorphism. If h is K-qc on $U \smallsetminus \Gamma$ then h is K-qc on the whole domain U.*

15.3.7. *Compactness in the space of quasicircles.* Let $\mathcal{QD}_{\kappa,r}$, $r > 0$, denote the space of pointed κ-quasidisks $(D, 0)$ with $r \leq r_{D,0} \leq R_{D,0} \leq 1/r$, endowed with the Carathéodory topology.

PROPOSITION 15.21. *The space $\mathcal{QD}_{\kappa,r}$ is compact.*

PROOF. Consider a quasidisk $(D, 0) \in \mathcal{QD}_{\kappa,r}$. By Theorem 15.17, the normalized Riemann mapping $h : (\mathbb{D}, 0) \to (D, 0)$ admits a K-qc extension to the whole complex plane $\mathbb{C}$, where K depends only on κ and r. Moreover, $r \leq |h(1)| \leq 1/r$. By the Compactness Theorem (see Exercise 13.3), this family of qc maps is compact in the uniform topology on $\mathbb{C}$. Since uniform limits of κ-quasidisks are obviously κ-quasidisks, the conclusion follows. □

15.4. QC welding. Recall from §1.7.2 and §2.1 the discussion of the connected sum $\overline{\mathbb{D}} \sqcup_h (\hat{\mathbb{C}} \smallsetminus \mathbb{D})$ of two disks along the circle $\mathbb{T}$ by means of an orientation preserving [7] homeomorphism $h : \mathbb{T} \to \mathbb{T}$. The outcome is a topological sphere S^2.

[7] As we know from §1.7.2, h should be "orientation reversing". There is no contradiction here because our h is indeed orientation reversing with respect to the orientations that $\mathbb{T}$ inherits from $\overline{\mathbb{D}}$ and from $\mathbb{C} \smallsetminus \mathbb{D}$.

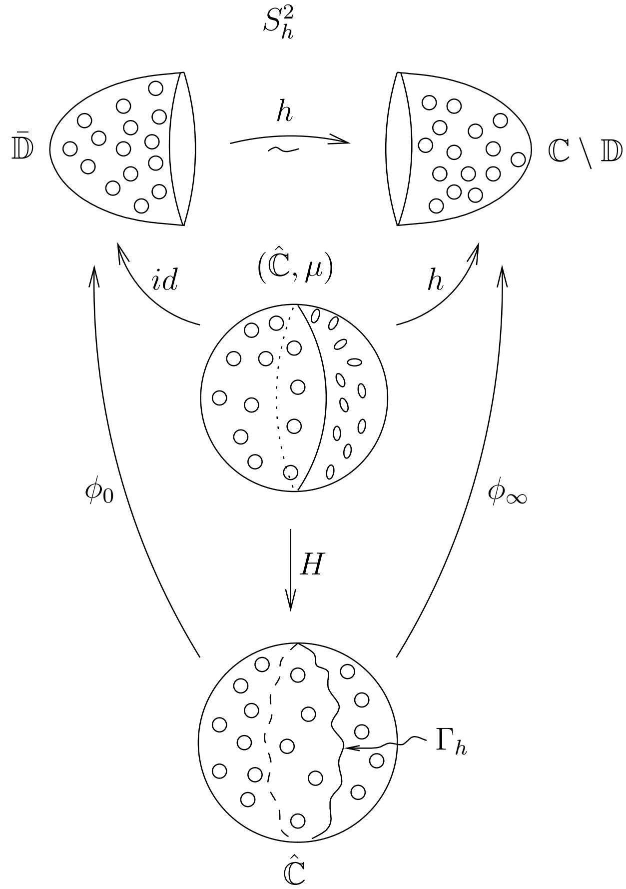

FIGURE 15.2. Quasiconformal welding.

As we know from Exercise 2.3, if h is real analytic then this sphere has a natural complex structure. By the Uniformization Theorem, there is a conformal isomorphism $H : S^2 \to \hat{\mathbb{C}}$. Then $\Gamma \equiv \Gamma_h := H(\mathbb{T})$ is an analytic Jordan curve in $\hat{\mathbb{C}}$.

Vice versa, given an (oriented) analytic Jordan curve $\Gamma \subset \hat{\mathbb{C}}$, let U_0 and U_∞ be the components of $\hat{\mathbb{C}} \smallsetminus \Gamma$ (where Γ is positively oriented with respect to U_0), and let $\phi_0 : \overline{U}_0 \to \overline{\mathbb{D}}$, $\phi_\infty : \overline{U}_\infty \to \mathbb{C} \smallsetminus \mathbb{D}$ be the corresponding Riemann mappings. Then $\phi_0 \circ \phi_\infty^{-1} | \mathbb{T}$ is an orientation preserving analytic homeomorphism of $\mathbb{T}$.

EXERCISE 15.22. *Show that these constructions provide us with a one-to-one correspondence between analytic orientation preserving homeomorphisms $h : \mathbb{T} \to \mathbb{T}$, up to two-sided action of* $\mathrm{Möb}(\mathbb{D})$, *and analytic Jordan curves $\Gamma \subset \hat{\mathbb{C}}$, up to the action of* $\mathrm{Möb}(\hat{\mathbb{C}})$.

We are now prepared for a far-reaching generalization of this assertion:

THEOREM 15.23. *Let $h : \mathbb{T} \to \mathbb{T}$ be a quasisymmetric orientation preserving homeomorphism. Then the connected sum $S_h^2 = \overline{\mathbb{D}} \sqcup_h (\mathbb{C} \smallsetminus \mathbb{D})$ can be endowed with a unique complex structure compatible with the complex structures of $\mathbb{D}$ and $\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$. This gives us a one-to-one correspondence between orientation preserving qs* *homeomorphisms* $h : \mathbb{T} \to \mathbb{T}$*, up to the two-sided action of* $\text{Möb}(\mathbb{D})$*, and quasicircles* $\Gamma \subset \hat{\mathbb{C}}$*, up to the action of* $\text{Möb}(\hat{\mathbb{C}})$*. The inverse map is obtained as the composition* $\phi_0 \circ \phi_\infty^{-1} | \mathbb{T}$*, where*

$$\phi_0 : \overline{U}_0 \to \overline{\mathbb{D}}, \quad \phi_\infty : \overline{U}_\infty \to \mathbb{C} \smallsetminus \mathbb{D} \tag{15.3}$$

are Riemann mappings to the components of $\hat{\mathbb{C}} \smallsetminus \Gamma$ *(where* Γ *is positively oriented with respect to* U_0*).*

Moreover, there is a quantitative one-to-one correspondence between orientation preserving L-qs homeomorphisms $h : \mathbb{T} \to \mathbb{T}$ *(up to the two-sided action of the rotation group* $\mathbb{T}$*), and* κ*-quasicircles* $\Gamma \subset \mathbb{C}^*$ *that are* κ*-quasi-centered at* 0 *(up to the action of the complex scaling group* $\mathbb{C}^*$*).*

Proof. By the Ahlfors-Beurling Theorem, h extends to a qc map

$$\hat{h} : \hat{\mathbb{C}} \smallsetminus \mathbb{D} \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}.$$

Define a measurable conformal structure μ on $\hat{\mathbb{C}}$ by letting $\mu = h^*(\sigma)$ on $\hat{\mathbb{C}} \smallsetminus \mathbb{D}$ and $\mu = \sigma$ on $\mathbb{D}$. By MRMT, it determines a new complex structure on $\mathbb{C}$, which can be uniformized by the standard Riemann sphere by means of a qc map

$$H : (\mathbb{C}, \mu) \to (\mathbb{C}, \sigma).$$

This gives us a quasicircle $\Gamma_{\hat{h}} := H(\mathbb{T})$, with the orientation induced from $\mathbb{T}$. Since H is defined up to the post-composition with a Möbius map, the quasicircle $\Gamma_{\hat{h}}$ is defined up to the action of $\text{Möb}(\hat{\mathbb{C}})$ (given the extension $\hat{h}$).

Let us show that $\Gamma_{\hat{h}}$ is actually independent of the choice of the extension $\hat{h}$. Indeed, if $\hat{h}' : \hat{\mathbb{C}} \smallsetminus \mathbb{D} \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}$ is another qc extension, then $h' = h \circ \psi$, where $\psi : \hat{\mathbb{C}} \smallsetminus \mathbb{D} \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}$ is a qc homeomorphism equal to id on $\mathbb{T}$. Let us consider a homeomorphism $\Psi : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ which is equal to ψ on $\hat{\mathbb{C}} \smallsetminus \mathbb{D}$ and equal to id on $\overline{\mathbb{D}}$. By the Gluing Lemma, it is qc on the whole sphere. Moreover, the homeomorphism $H' := H \circ \Psi$ solves the Beltrami equation for the differential μ' associated with h', and $H' | \mathbb{T} = H | \mathbb{T}$. The conclusion follows.

Thus, we can write $\Gamma_{\hat{h}} \equiv \Gamma_h$. In fact, Γ_h is invariant under the two-sided action of $\text{Möb}(\mathbb{D})$: it does not change if h is replaced with $A \circ h \circ B^{-1} | \mathbb{T}$, where $A, B \in \text{Möb}(\mathbb{D})$. Indeed, replacing h with $A \circ h$ does not change the conformal structure μ, so it does not change H. Replacing h with $h \circ B^{-1} | \mathbb{T}$ amounts to replacing H with $H \circ B^{-1}$, which does not affect Γ_h either. So we have constructed a map

$$h \,(\text{modulo}\, \text{Möb}(\mathbb{D})) \mapsto \Gamma_h \,(\text{modulo } \text{Möb}(\hat{\mathbb{C}})) \tag{15.4}$$

Let U_0 and U_∞ be the complementary components of $\mathbb{C} \smallsetminus \Gamma_h$. By construction, the maps

$$\phi_0 := H^{-1} : \overline{U}_0 \to \overline{\mathbb{D}} \quad \text{and} \quad \phi_\infty := \hat{h} \circ H^{-1} : \overline{U}_\infty \to \mathbb{C} \smallsetminus \mathbb{D} \tag{15.5}$$

are conformal, so they are equal to the Riemann mappings for U_0 and U_∞ respectively. Moreover, their composition brings us back the original map h:

$$\phi_\infty \circ \phi_0^{-1} | \mathbb{T} = h. \tag{15.6}$$

Let ψ_0 and ψ_∞ be the inverse maps. Then $\psi_0 = \psi_\infty \circ h$, so the map

$$\Psi : (S^2_h, \mathbb{T}) \to (\hat{\mathbb{C}}, \Gamma) \quad \text{given by} \quad \Psi | \overline{\mathbb{D}} = \psi_0, \ \Psi | \hat{\mathbb{C}} \smallsetminus \mathbb{D} = \psi_\infty, \tag{15.7}$$

is a well defined homeomorphism. As it is conformal on $\mathbb{D}$ and $\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$, it induces a desired complex structure $\Psi^*(\sigma)$ on S^2_h.

Let us show that such a complex structure is unique. Indeed, assume there are two such structures, and let Γ, Γ' be the corresponding quasicircles. Then there exists a homeomorphism $\Phi : (\hat{\mathbb{C}}, \Gamma) \to (\hat{\mathbb{C}}, \Gamma')$ conformal on $\hat{\mathbb{C}} \smallsetminus \Gamma$. Since quasicircles are removable, Φ is Möbius, so the structures are conformally equivalent.

Furthermore, application of a Möbius transformation $A \in \text{Möb}(\hat{\mathbb{C}})$ to Γ leads to pre-composition of the Riemann mappings (15.5) with A^{-1}, which does not affect the gluing map h in (15.6). Also, as the above Riemann mappings are well defined up to post-composition with a Möbius map $A \in \text{Möb}(\mathbb{D})$, the gluing map h is well defined up to two-sided action of $\text{Möb}(\mathbb{D})$. Hence the above construction provides us with the left inverse for the map (15.4), showing that the latter is one-to-one onto the image.

What is left, is to show that (15.4) is surjective. It amounts to the repetition of the construction of the gluing map (15.6) for a general oriented quasicircle $\Gamma \subset \hat{\mathbb{C}}$. So, let U_0 and U_∞ be the components of $\hat{\mathbb{C}} \smallsetminus \Gamma$, where Γ is positively oriented with respect to U_0. Let us consider the corresponding Riemann mappings (15.3), and let ψ_0, ψ_∞ be the inverse maps. Then $h := \phi_\infty \circ \phi_0^{-1} | \mathbb{T}$ is an orientation preserving circle homeomorphism.Moreover, as $\psi_0 = \psi_\infty \circ h$, the map defined as (15.7) is a well defined homeomorphism.

By Theorem 15.17, ψ_0 and ψ_∞ admit qc extensions $\Psi_0, \Psi_\infty : (\hat{\mathbb{C}}, \mathbb{T}) \to (\hat{\mathbb{C}}, \Gamma)$. Hence their restrictions $\psi_0, \psi_\infty : \mathbb{T} \to \Gamma$ are quasisymmetric, implying that $h : \mathbb{T} \to \mathbb{T}$ is qs as well.

By what we have already shown, the connected sum $S^2_h := \overline{\mathbb{D}} \cup_h (\hat{\mathbb{C}} \smallsetminus \mathbb{D}), \mathbb{T})$ has a unique complex structure compatible with the complex structures on $\mathbb{D}$ and $\mathbb{C} \smallsetminus \overline{\mathbb{D}}$. Since the homeomorphism Ψ (15.7) is conformal on $\mathbb{D}$ and $\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$, it is a biholomorphic isomorphism between $(S^2_h, \mathbb{T})$ and $(\mathbb{C}, \Gamma)$, so Γ is realized by a qc welding.

For the last quantitative assertion, recall from Lemma 15.6 that any L-qc homeomorphism $h : \mathbb{T} \to \mathbb{T}$ extends to a $K(L)$-qc homeomorphism $h : (\overline{\mathbb{D}}, 0) \to (\overline{\mathbb{D}}, 0)$. Then the solution H of the Beltrami equation is also $K(L)$-qc. Normalizing it so that $H(0) = 0$ and $H(\infty) = \infty$, we obtain by Proposition 12.13 the desired $\kappa(L)$-quasicircle Γ (modulo the action of $\mathbb{C}^*$).

Vice versa, let $\Gamma \subset \mathbb{C}^*$ be a κ-quasicircle centered at 0 (modulo the action of $\mathbb{C}^*$). Normalizing the corresponding Riemann mappings ϕ_0 and ϕ_∞ (15.3) so that they fix 0 and ∞ respectively, we make them well defined up to post-composition with a rotation and precomposition with a complex scaling. Hence the transit map $h = \phi_\infty \circ \phi_0^{-1} | \mathbb{T}$ is well defined up to the two-sided action of the rotation group. Moreover, by Theorem 15.17, each of them admits a $K(\kappa)$-extension to the whole sphere $\hat{\mathbb{C}}$ fixing 0 and ∞. Applying Proposition 12.13 once again, we conclude that both maps are $L(\kappa)$-qs. Hence so is h (with a different L). □

This construction is called the *qc welding* of $\overline{\mathbb{D}}$ and $\hat{\mathbb{C}} \smallsetminus \mathbb{D}$ by means of a qs homeomorphism $h : \mathbb{T} \to \mathbb{T}$. More generally, *any Jordan curve* $\Gamma \subset \hat{\mathbb{C}}$ can be viewed as a qc welding by means of some homeomorphism $h : \mathbb{T} \to \mathbb{T}$ (namely, by the transit map from the interior to the exterior Riemann mappings). However, not all homeomorphisms appear this way.

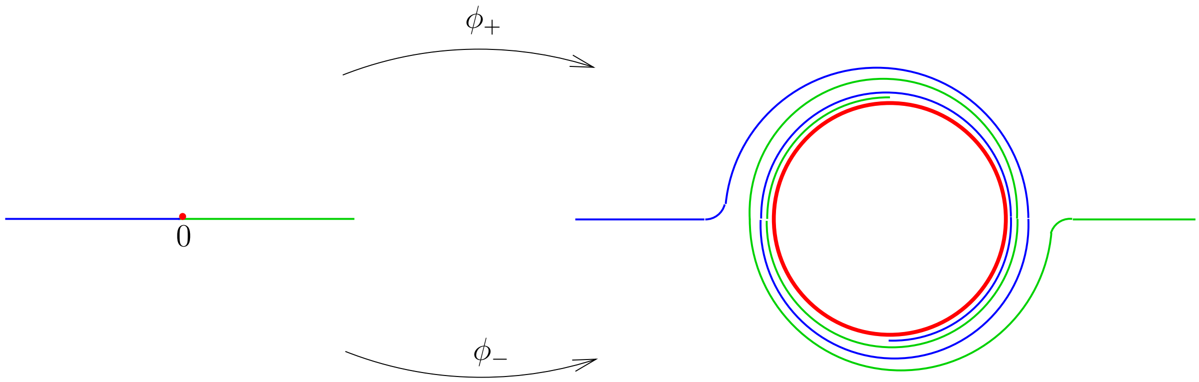

FIGURE 15.3. A homeomorphism of $\mathbb{R}$ that cannot be realized as the welding for a Jordan curve.

16. Removability

16.1. Conformal vs quasiconformal. Similarly to the notion of qc removability introduced in §13.3 we can define conformal removability:

DEFINITION 16.1. A compact subset $X \subset \mathbb{C}$ is called *conformally removable* if for any open sets $U \supset X$ in $\mathbb{C}$, any homeomorphic embedding $h : U \hookrightarrow \mathbb{C}$ which is conformal on $U \smallsetminus X$ is conformal/qc on U.

In fact, these two properties are equivalent:

PROPOSITION 16.2. *Conformal removability is equivalent to qc removability.*

Thus, we can unambiguously call a set "removable".

It is classical that isolated points and smooth Jordan curves are conformally removable. Proposition 16.2 implies that they are qc removable as well (which was also shown directly in §13.3 of Ch. 2). Since qc removability is invariant under qc changes of variable, we obtain:

LEMMA 16.3. *Quasicircles are removable.*

.

16.2. Removability and area. The Measurable Riemann Mapping Theorem yields:

PROPOSITION 16.4. *Removable sets have zero area.*

PROOF. Assume that $m(X) > 0$. Then there exists a non-trivial Beltrami differential μ supported on X. Let $h : \mathbb{C} \to \mathbb{C}$ be a solution of the corresponding Beltrami equation. Then h is conformal outside X but is not conformal on X. $\square$

The reverse is false:

EXAMPLE 16.5.

16.3. Divergence property.

Definition 16.6. Let us say that a compact set $X \subset \mathbb{C}$ satisfies the *divergence property* if for any point $z \in X$ there exists a nest of annuli $A^n(z)$ around z such that

$$\sum A^n(z) = \infty.$$

Without loss of generality we can assume (and we will always do so) that each annulus in this definition is bounded by two Jordan curves.

Lemma 16.7. *Compact sets satisfying the divergence property are Cantor.*

Proof. Consider any connected component X_0 of X, and let $z \in X_0$. Then the annuli $A^n(z)$ are nested around X_0. By Corollary 6.20 of the Grötzsch Inequality, X_0 is a single point. $\square$

Lemma 16.8. *Let $X \subset \mathbb{C}$ be a compact set satisfying the divergence property. Then for any neighborhood $U \supset X$, any qc embedding $h : U \smallsetminus X \hookrightarrow \mathbb{C}$ admits a homeomorphic extension through X.*

Proof. Let $h : U \smallsetminus X \hookrightarrow \mathbb{C}$ be a K-qc embedding. If $X \subset U' \Subset U$ then $h(U')$ is bounded in $\mathbb{C}$. So, without loss of generality we can assume that $h(U)$ is bounded in $\mathbb{C}$.

For $z \in X$, let us consider the nest of annuli $h(A^n(z))$. Since h is quasiconformal,

$$\sum \operatorname{mod} h(A^n(z)) \geq K^{-1} \sum \operatorname{mod} A^n(z) = \infty.$$

Let $\Delta^n(z)$ be the bounded component of $\mathbb{C} \smallsetminus h(A^n(z))$, and let

$$\Delta^\infty(z) = \bigcap_n D^n(z).$$

By Corollary 6.20 of the divergence property, $\Delta^\infty(z)$ is a single point $\zeta = \zeta(z)$. Let us extend h through X by letting $h(z) = \zeta$.

This extension is continuous. Indeed, let $D^n(z)$ be the bounded component of $\mathbb{C} \smallsetminus A^n(z)$. Then by Corollary 6.20, $\operatorname{diam} D^n(z) \to 0$, so that $D^n(z)$ is a base of (closed) neighborhoods of z. But

$$\operatorname{diam} h(D^n(z)) = \operatorname{diam} \Delta^n(z) \to 0,$$

which yields continuity of h at z.

Switching the roles of (U, X) and $(h(U), h(X))$, we conclude that h^{-1} admits a continuous extension through $h(X)$. Hence the extension of h is homeomorphic. $\square$

It is worthwhile to note that, in fact, general homeomorphisms extend through Cantor sets:

Exercise 16.9. (i) *Let us consider two Cantor sets X and $\tilde{X}$ in $\mathbb{C}$ and their respective neighborhoods U and $\tilde{U}$. Then any homeomorphism $h : U \smallsetminus X \to \tilde{U} \smallsetminus \tilde{X}$ admits a homeomorphic extension through X.*

(ii) *It was essential to assume that* both *sets X and $\tilde{X}$ are Cantor! For* any *compact set $X \subset \mathbb{C}$, give an example of an embedding $h : C \smallsetminus X \hookrightarrow \mathbb{C}$ which does not admit a continuous extension through X.*

Lemma 16.10. *Compact sets satisfying the divergence property have zero area.*

We will show now that sets satisfying the divergence property are removable, and even in the following stronger sense:

THEOREM 16.11. *Let $X \subset \mathbb{C}$ be a compact set satisfying the divergence property. Then for any neighborhood $U \supset X$, any conformal/qc embedding $h : U \smallsetminus X \hookrightarrow \mathbb{C}$ admits a conformal/qc extension through X.*

PROOF. Let $h : U \smallsetminus X \hookrightarrow \mathbb{C}$ be a K-qc embedding. By Lemma 16.8, h extends to an embedding $U \hookrightarrow \mathbb{C}$, which will be still denoted by h. Let us show that h belongs to the Sobolev class $H(U)$.

Since X is a Cantor set, it admits a nested base of neighborhoods U^n such that each U^n is the union of finitely many disjoint Jordan disks. Take any $\mu > 0$. By the Grẗzsch Inequality, for any $n \in \mathbb{N}$ there is $k = k(\mu, l) > 0$ such that $\mathrm{mod}(\partial U^{n+k}, \partial U^n) \geq \mu > 0$. Let χ_n be the solution of the Dirichlet problem in $U^n \smallsetminus U^{n+k}$ vanishing on ∂U^{n+k} and equal to 1 on ∂U^n. By Theorem 6.30, $D(\chi_n) \leq 1/\mu$.

Let us continuously extend χ to the whole plane in such a way that it vanishes on U^{n+k} and identically equal to 1 on $\mathbb{C} \smallsetminus U^n$. We obtain a piecewise smooth function $\chi : \mathbb{C} \to [0, 1]$, with the jump of the derivative on the boundary of the domains U^n and U^{n+k}.

Let $h_n = \chi_n h$. These are piecewise smooth functions with bounded Dirichlet integral. Indeed,

$$D(h_n) = \int (|\nabla \chi_n|^2 |h|^2 + |\chi_n|^2 |\nabla h|^2) dm \leq \mathrm{diam}(h(U))/\mu + C(K) m(h(U)),$$

where $C(K) = (1 + k^2)/(1 - k^2)$ comes from the area estimate (area estimate). By weak compactness of the unit ball in $L^2(U)$, we can select a converging subsequence $\partial h_n \to \phi$, $\bar{\partial} h_n \to \psi$. But $h_n \to h$ pointwise on $U \smallsetminus X$, so that by Lemma 16.10, $h_n \to h$ a.e. It follows that ϕ and ψ are distributional partial derivatives of h (see (13.2)).

Finally, if h is conformal on $U \smallsetminus X$ then by Weyl's Lemma it is conformal on U. □

17. Holomorphic motions

17.1. Definition. Let $(\Lambda, \lambda_\circ)$ be a pointed complex Banach manifold[8] and let $X \equiv X_\circ \subset \hat{\mathbb{C}}$ be an arbitrary subset of the Riemann sphere (can be non-measurable). A *holomorphic motion* $\mathbf{h}$ *over* $(\Lambda, \lambda_\circ)$[9] is a family of injections

$$h_\lambda : X \to \hat{\mathbb{C}}, \quad \lambda \in \Lambda,$$

depending holomorphically on λ (in a weak sense that the functions $z \mapsto h_\lambda(z)$ are holomorphic in λ for all $z \in X$) and such that $h_{\lambda_\circ} \equiv h_\circ = \mathrm{id}$. In this situation, we let $X_\lambda := h_\lambda(X_\circ)$.[10]

For $z \in X$, holomorphic functions $\phi_z : \Lambda \to \hat{\mathbb{C}}$, $\lambda \mapsto h_\lambda(z)$, are called *orbits* of the holomorphic motion. Since the functions h_λ are injective, the orbits do not

[8]We will eventually deal with infinite dimensional parameter spaces, so we need to prepare the background in this generality. However, in the first reading the reader can safely assume that the space Λ is a one-dimensional disk (which is the main case to consider anyway).

[9]We will often make a point $\lambda_\circ$ implicit in the notation and terminology.

[10]We will sometimes say briefly that "the sets X_λ move holomorphically" or "the set X moves holomorphically" without mentioning explicitly the maps h_λ.

collide, or equivalently, their graphs $\mathcal{L}_z \subset \Lambda \times \hat{\mathbb{C}}$ (also called *leaves* of the motion) are disjoint. Thus, a holomorphic motion provides us with a family of disjoint holomorphic graphs over Λ. We refer to such a family as a (*trivial*) *holomorphic lamination* $\mathfrak{F}$. Of course, the above reasoning can be reversed, so that, trivial holomorphic laminations give us an equivalent (dual) way of describing holomorphic motions.

A *regularity* of a holomorphic motion is the regularity of the maps h_λ on X. For instance, a holomorphic motion is called *continuous, qc, smooth* or *biholomorphic* if all the maps h_λ, $\lambda \in \Lambda$, have the corresponding regularity on X (to make sense of it in some cases we need extra assumptions on X, e.g., openness). The regularity of $\mathbf{h}$ can also be interpreted as the *transverse regularity* of the corresponding lamination $\mathfrak{F}$, see §17.4.2 below.

Notice that *a priori* we do not impose any regularity on the maps h_λ (not even measurability!). A remarkable property of holomorphic motions is that they automatically have nice regularity properties and that they automatically extend to motions of the whole Riemann sphere. This set of properties are usually referred to as the λ-*lemma*. It will be the theme of the rest of this section.

While dealing with a holomorphic motion of a set X, Y, etc., we let $X_\lambda := h_\lambda(X)$, $Y_\lambda := h_\lambda(Y)$, etc. We will refer to the z-variable of a holomorphic motion as the *dynamical variable* (though in general, there is no dynamics in the z-plane). The λ-variable is naturally referred to as the *parameter*.

We let

$$\mathbf{X} := \bigcup_{\lambda\in\Lambda} X_\lambda \subset \mathbb{C}^2, \quad \mathbf{Y} := \bigcup_{\lambda\in\Lambda} Y_\lambda \subset \mathbb{C}^2, \text{ etc.} \tag{17.1}$$

be the *total space* of the corresponding motion. It has two transverse structures: It is *fibered* over Λ with *fibers* X_λ (resp., Y_λ, etc.) and it is *foliated* by the leaves of the motion. For a subdomain $\Lambda' \subset \Lambda$, we let

$$\mathbf{X}|\,\Lambda' := \bigcup_{\lambda\in\Lambda'} X_\lambda$$

be the total space of the restricted motion.

In case when X_λ are Jordan disks, we will refer to $\mathbf{X}$ as a *foliated tube*.

17.2. Extension to the closure and continuity.

First λ-lemma (Extension to the closure). *A holomorphic motion* $\mathbf{h}$ *of any set* $X \subset \hat{\mathbb{C}}$ *extends to a continuous holomorphic motion of its closure* $\overline{X}$.

Proof. If X is finite, there is nothing to prove, so assume it is infinite.

Let us show that the family of orbits ϕ_z, $z \in X$, of our holomorphic motion is normal. To this end, let us remove from X three points $z_i \in X$; let $X' = X \smallsetminus \{z_i\}$ and let ψ_i be the orbits of the points z_i. Since the orbits of a holomorphic motion do not collide, the family of orbits of points $z \in X'$ satisfies the condition of the Refined Montel Theorem, 4.9, with exceptional functions ψ_i, and the normality follows.[11]

Let Φ be the closure of the family of orbits in the space $\mathcal{M}(\Lambda)$ of meromorphic functions on Λ. By the Hurwitz Theorem, the graphs of these functions are disjoint, so they form a holomorphic lamination representing a holomorphic motion of $\overline{X}$.

[11]Results of §14.11.3 allow us to apply the Montel Theorem on Banach domains.

Let us keep notation h_λ for the extended holomorphic motion, and notation ϕ_z, $z \in \overline{X}$, for its orbits.

Let us show that this motion is continuous. Let $\lambda \in \Lambda$, let $z_n \to z$ be a converging sequence of points in $\overline{X}$, and let $\phi_n \in \Phi$ and $\phi \in \Phi$ be their respective orbits. We want to show that $h_\lambda(z_n) \to h_\lambda(z)$, or equivalently $\phi_n(\lambda) \to \phi(\lambda)$. But otherwise, the sequence ϕ_n would have a limit point $\psi \in \mathcal{M}(\Lambda)$ such that $\psi(\lambda_\circ) = \phi(\lambda_\circ)$ while $\psi(\lambda) \neq \phi(\lambda)$, which would contradict to the laminar property of the family Φ. $\square$

In particular, for a holomorphic motion of any compacts set X, the maps $h_\lambda : X \to X_\lambda$ are automatically *homeomorphisms*.

17.3. Extension of smooth holomorphic motions. In this short section we will prove a simple extension lemma for smooth holomorphic motions.

LEMMA 17.1 (Local extension). *Let us consider a compact set $Q \subset \mathbb{C}$ and a smooth holomorphic motion h_λ of a neighborhood U of Q over a Banach domain $(\Lambda, \lambda_\circ)$. Then there is a smooth holomorphic motion H_λ of the whole complex plane $\mathbb{C}$ over some neighborhood $\Lambda' \subset \Lambda$ of $\lambda_\circ$ whose restriction to Q coincides with h_λ.*

PROOF. We can certainly assume that $\overline{U}$ is compact. Take a smooth function $\phi : \mathbb{C} \to \mathbb{R}$ supported in U such that $\phi|\, Q \equiv 1$, and let

$$H_\lambda = \phi\, h_\lambda + (1 - \phi)\, \mathrm{id}\,.$$

Clearly H is smooth in both variables, holomorphic in λ, and identical outside U. As $H_\circ = \mathrm{id}$, $H_\lambda : \mathbb{C} \to \mathbb{C}$ is a diffeomorphism for λ sufficiently close to $\lambda_\circ$, and we are done. $\square$

We will sometimes refer to this statement as the *Elementary λ-Lemma*.

17.4. Transverse quasiconformality.

17.4.1. *Quasiconformality of h_λ.*

SECOND λ-LEMMA (Quasiconformality). *Let $h_\lambda : X \to X_\lambda$ be a holomorphic motion of a set $X \subset \mathbb{C}$ over the disk $\mathbb{D}$. Then for $|\lambda| \le r < 1$, the maps h_λ are η-quasisymmetric with dilatation η depending only on r. Consequently, if X is open then the maps h_λ are K-qc with dilatation K depending only on r.*

PROOF. Let $t > 1$. Take three distinct points $a, b, c \in \mathbb{C}$ such that

$$t^{-1} \le \frac{|c-a|}{|b-a|} \le t.$$

We need to show that

$$\eta^{-1} \le \frac{|h_\lambda(c) - h_\lambda(a)|}{|h_\lambda(b) - h_\lambda(a)|} \le \eta,\ |\lambda| \le r, \quad \text{for some } \eta = \eta_r(t) > 1.$$

Let us normalize the holomorphic motion by affine changes of variables so that $h_\lambda(a) \equiv 0$, $h_\lambda(b) \equiv 1$. Since affine maps do not distort ratios, it is enough to prove the assertion for the normalized motion, for which it assumes the form:

$$t^{-1} \le |c| \le t,\ |\lambda| \le r,\ c \neq 1 \implies \eta^{-1} \le |h_\lambda(c)| \le \eta.$$

But the orbit $\lambda \mapsto h_\lambda(c)$ is a holomorphic map $\mathbb{D} \to \mathbb{C} \smallsetminus \{0, 1\}$, and the conclusion follows from the normality of the family of all such maps (Big Montel Theorem). $\square$

EXERCISE 17.2. *Check a slightly more general assertion, for holomorphic motions $h_\lambda : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$, $\lambda \in \mathbb{D}$, of the Riemann sphere.*

Given a holomorphic motion $\mathbf{h}$ over Λ, let

$$\operatorname{Dil} \mathbf{h} = \sup_{\lambda \in \Lambda} \operatorname{Dil} h_\lambda$$

(which can be infinite). We say that the holomorphic motion $\mathbf{h}$ is K-qc if

$$\operatorname{Dil} \mathbf{h} \le K.$$

In these terms, the Second λ-Lemma tells us that if $\mathbf{h}$ is a holomorphic motion over $\mathbb{D}$ then for $r < 1$, $\operatorname{Dil}(\mathbf{h}|\,\mathbb{D}_r) \le K(r)$. Since biholomorphic reparametrizations of the parameter domain do not affect $\operatorname{Dil} \mathbf{h}$, we can formulate the Second λ-Lemma in an invariant form:

COROLLARY 17.3. *Any holomorphic motion* $\mathbf{h}$ *over* $\mathbb{D}$*, restricted to a hyperbolic disk* $\mathbb{D}_{\text{hyp}}(a, \rho) \subset \mathbb{D}$*, has a bounded dilatation (in terms of* ρ*):*

$$\operatorname{Dil}(\mathbf{h}|\,\mathbb{D}_{\text{hyp}}(a, \rho)) \le K(\rho).$$

Note finally that the Second λ-Lemma is valid over any Banach ball as well (by restricting the holomorphic motion to one-dimensional complex disks):

COROLLARY 17.4. *Let $h_\lambda : X \to X_\lambda$ be a holomorphic motion over a complex Banach ball $\mathcal{B}_1$. Then for $|\lambda| \le r < 1$, the maps h_λ are η-quasisymmetric with dilatation η depending only on r.*

17.4.2. *Holonomy.* Take two point $p_\circ = (\lambda_\circ, z_\circ)$ and $p = (\lambda, z = h_\lambda(z_0))$ on the same leaf $\mathcal{L}(p_0)$ and consider local transversals $\Gamma_\circ \ni p_\circ$ and $\Gamma \ni p$ to $\mathcal{L}$ through these points (i.e., local holomorphic curves transverse to $\mathcal{L}$). Then for $q_\circ = (\lambda_\circ, \zeta) \in \Gamma_\circ$ near p_0, the leaf $\mathcal{L}(q_0)$ intersects Γ transversely at a single point, so there is a well defined local map $\mathfrak{h} : (\Gamma_\circ, p_\circ) \to (\Gamma, p)$ called the *holonomy* from $\Gamma_\circ$ to Γ. The lamination $\mathfrak{F}$ is called *transversely (locally qc)/smooth/biholomorphic* if all the holonomy maps are such.[12]

EXERCISE 17.5. *The holomorphic motion* $\mathbf{h}$ *is smooth/biholomorphic iff the corresponding lamination* $\mathfrak{F}$ *is transversely smooth/biholomorphic.*

LEMMA 17.6. *Let $\mathbf{h} = (h_\lambda)$ be a holomorphic motion of an open set U over the disk $(\mathbb{D}, 0)$, and $\mathfrak{F}$ be the corresponding holomorphic lamination. Then $\mathfrak{F}$ is transversely locally quasiconformal. Moreover, if Γ_0 and Γ are local transversals though points $p_0 = (0, z_0)$, and $p = (\lambda, z = h_\lambda(z_0))$, then the dilatation $\operatorname{Dil} \mathfrak{h}(z_0)$ of the holonomy $\mathfrak{h} : (\Gamma_0, p_0) \to (\Gamma, p)$ is bounded by $\operatorname{Dil} h_\lambda(z_0)$ (which in turn, depends only on an upper bound $r \in (0, 1)$ for $|\lambda|$).*

PROOF. If the transversals are vertical lines $\{0\} \times \mathbb{C}$ and $\lambda \times \mathbb{C}$ then the result follows from the Second λ-Lemma.

Furthermore, the holonomy from the vertical line $\lambda \times \mathbb{C}$ to the transversal Γ is locally conformal at point p. To see this, let us select a holomorphic coordinates (θ, z) near p in such a way that $p = 0$ and the leaf via p becomes the parameter axis. Let $z = \psi(\theta) = \varepsilon + \dots$ parameterizes a nearby leaf of the foliation, while $\theta = g(z) = bz + \dots$ parameterizes the transversal Γ.

[12] Notice that smoothness and holomorphicity are local properties, while quasiconformality is not: that is why we need to say "locally qc" but not "locally smooth" or "locally biholomorphic".

Let us do the rescaling $z = \varepsilon\zeta, \theta = \varepsilon\nu$. In these new coordinates, the above leaf is parametrized by the function $\Psi(\nu) = \varepsilon^{-1}\psi(\varepsilon\nu)$, $|\nu| < R$, where R is a fixed parameter. Then $\Psi'(\nu) = \psi'(\varepsilon\nu)$ and $\Psi''(\nu) = \varepsilon\psi''(\varepsilon\nu)$. By the Cauchy Inequality, $\Psi''(\nu) = O(\varepsilon)$. Moreover, ψ uniformly goes to 0 as $\psi(0) \to 0$. Hence $|\Psi'(0)| = |\psi'(0)| \le \delta_0(\varepsilon)$, where $\delta_0(\varepsilon) \to 0$ as $\varepsilon \to 0$. Thus $\Psi'(\nu) = \delta_0(\varepsilon) + O(\varepsilon) \le \delta(\varepsilon) \to 0$ as $\varepsilon \to 0$ uniformly for all $|\nu| < R$. It follows that $\Psi(\nu) = 1 + O(\delta(\varepsilon)) = 1 + o(1)$ as $\varepsilon \to 0$.

On the other hand, the manifold Γ is parametrized in the rescaled coordinates by a function $\nu = b\zeta(1+o(1))$. Since the transverse intersection persists, $\mathcal{S}$ intersects the leaf at the point $(\nu_0, \zeta_0) = (1, b)(1 + o(1))$ (so that R should be selected bigger than $\|b\|$). In the old coordinates the intersection point is $(\theta_0, z_0) = (\varepsilon, b\varepsilon)(1+o(1))$.

Thus the holonomy from $\lambda \times \mathbb{C}$ to Γ transforms the disc of radius $|\varepsilon|$ to an ellipse with small eccentricity, which means that this holonomy is asymptotically conformal. As the holonomy from $\{0\} \times \mathbb{C}$ to Γ_0 is also asymptotically conformal, the conclusion follows. $\square$

COROLLARY 17.7. *Under the above circumstances, if* $\operatorname{Dil}\mathbf{h} \le K$*, then* $\mathfrak{F}$ *is transversely* K*-qc.*

Again, the above discussion is valid over Banach balls, by restricting the motion to one-dimensional complex disks:

LEMMA 17.8. *For any holomorphic motion* $\mathbf{h} = (h_\lambda)$ *of an open set* U *over a Banach ball* $\mathcal{B}_1$*, the corresponding lamination* $\mathfrak{F}$ *is transversely locally quasiconformal, with the dilatation* $\operatorname{Dil}\mathfrak{h}(z_0)$ *of the holonomy* $\mathfrak{h} : (\Gamma_0, p_0) \to (\Gamma, p)$*, where* $p_0 = (0, z_0) \in \Gamma_0$*,* $p = (\lambda, z = h_\lambda(z_0)) \in \Gamma$*, bounded by* $\operatorname{Dil} h_\lambda(z_0)$ *(which in turn, depends only on the upper bound on* $|\lambda|$*). If* $\operatorname{Dil}\mathbf{h} \le K$ *then* $\mathfrak{F}$ *is transversely* K*-qc.*

More generally, we can consider a holonomy on holomorphic curves that are not necessarily transverse to the motion:

LEMMA 17.9. *Let* $\mathcal{T}$ *be a transversal to a holomorphic motion* $\mathbf{h}$ *of an open set, while* $\mathcal{S}$ *be an arbitrary holomorphic curve in the domain of the motion. Then the holonomy* $\gamma : \mathcal{S} \to \mathcal{T}$ *is locally quasiregular (with the same quantification as above).*

PROOF. It is locally K-qc at the points where $\mathcal{S}$ crosses $\mathcal{F}$ transversally, with dilatation depending only to the hyperbolic distance of the corresponding parameter to $\mathcal{S}$. By removability of isolated singularities, it is locally quasiregular at the tangency points. $\square$

Quasiconformality is apparently the best regularity of holomorphic motions which is satisfied automatically.

17.4.3. *Lifts of holomorphic motions.*

LEMMA 17.10. *Let* $h_\lambda : V_\circ \to V_\lambda$ *be a holomorphic motion of a domain* $V_\circ \subset \mathbb{C}$ *over a simply connected parameter domain* Λ*. Let* $f_\lambda : U_\lambda \to V_\lambda$ *be a holomorphic family of proper maps with critical points* c_λ^k *such that the critical values* $v_\lambda^k = f_\lambda(c_\lambda^k)$ *form orbits of* h_λ*.*[13] *Then* h_λ *uniquely lifts to a holomorphic motion* $H_\lambda : U_\circ \to U_\lambda$ *such that*

$$f_\lambda \circ H_\lambda = h_\lambda \circ f_\circ. \tag{17.2}$$

[13]In particular, any holomorphic family of univalent maps $f_\lambda : U_\lambda \to V_\lambda$ is allowed.

PROOF. Notice that (17.2) means that the lamination associated with the motion $\mathbf{H}$ is the pullback of the lamination associated with the motion $\mathbf{h}$ under the map $\mathbf{f}$ (23.18). Clearly, such a pullback unique if exists.

Let us take any regular value $\zeta_\circ = f_\circ(z_\circ) \in V_\circ$, and let $\phi(\lambda) = h_\lambda(\zeta_\circ)$ be its orbit. We would like to lift this orbit to a desired orbit of $z_\circ$, so we are looking for a holomorphic solution $z = \psi(\lambda)$ of an equation

$$f_\lambda(z) = \phi(\lambda) \tag{17.3}$$

with $\psi(z_\circ) = \zeta_\circ$. Since $\phi(\lambda)$ is a regular point of f_λ for any $\lambda \in \Lambda$, the Implicit Function Theorem implies that near any point (λ', z') satisfying (17.3), it admits a unique local analytic solution $z = \psi(\lambda)$. Since the maps f_λ are proper, this continuation along any path compactly contained in Λ cannot escape the domain U_λ. Since Λ is simply connected, $\psi(\lambda)$ extends to the whole domain Λ as a single valued holomorphic function.

Two different orbits $\lambda \mapsto \psi(\lambda)$ obtained in this way do not collide, for (17.3) would have two different solutions near the collision point. Hence they form a holomorphic motion of $V_\circ \smallsetminus \{v_\circ^k\}$ over Λ. By the First λ-lemma, this motion extends to the whole domain $V_\circ$.

Finally,

$$f_\lambda(H_\lambda(z_\circ)) = f_\lambda(\psi(\lambda)) = \phi(\lambda) = h_\lambda(\zeta_\circ) = h_\lambda(f_\circ(z_\circ))$$

holds for any point $z_\circ \in U_\circ$ except perhaps finitely many exceptions (preimages of the critical values of $f_\circ$). By continuity, it holds for all $z_\circ \in U_\circ$. $\square$

17.4.4. *Global transversal.* A *global transversal* Γ to a holomorphic motion $h_\lambda : X_\circ \to X_\lambda$ over Λ is the graph of a holomorphic function $\phi : \Lambda \to \mathbb{C}$ that intersects every leaf of the motion transversely at a single point. In fact, the transversality is automatic under a mild assumption:

LEMMA 17.11. *Assume that the moving set X_λ has dense interior. If a holomorphic graph Γ intersects every leaf of $\mathbf{h}$ at a single point then the intersection is transverse.*

LEMMA 17.12. *Let U be a Jordan disk and let $X \subset U$. Let $h_\lambda : \partial U \cup X \to \partial U_\lambda \cup X_\lambda$ be a holomorphic motion of these sets over Λ. If Γ is a global transversal to $\partial_{\rm hor}\mathbf{U}$ then it is a global transversal to $\mathbf{X}$.*

LEMMA 17.13. *Let U_λ and V_λ be two Jordan disks holomorphically moving over Λ. Let $F_\lambda : U_\lambda \to V_\lambda$ be a fibered conformal isomorphism between these disks. Let $\boldsymbol{\delta} : \Lambda \to \mathbb{C}^2$ and $\boldsymbol{\gamma} : \Lambda \to \mathbb{C}^2$ be two holomorphic curves such that $\mathbf{F}(\boldsymbol{\gamma}) = \boldsymbol{\delta}$. If $\boldsymbol{\gamma}$ is a global transversal to $\partial\mathbf{U}$ then $\boldsymbol{\delta}$ is a global transversal to $\partial\mathbf{V}$.*

17.5. Phase-Parameter Relation (without dynamics). :et $\gamma : X \to \Gamma$ (where $X \equiv X_\circ$) be the holonomy along the leaves of the motion, and let $\pi : \Gamma \to \Lambda$ be the projection of Γ onto Λ. Their composition,

$$\chi : X \to \Lambda, \quad \chi := \pi \circ \gamma, \tag{17.4}$$

is a homeomorphism onto the image called the *phase-parameter map*. Since the projection π is holomorphic, Corollary 17.7 implies:

LEMMA 17.14. *Let $\mathbf{h}$ be a holomorphic motion of a domain $U \subset \mathbb{C}$ over $\Lambda \subset \mathbb{C}$ with* $\operatorname{Dil}\mathbf{h} \le K$. *Then for any global transversal Γ, the corresponding phase-parameter map $\chi : U \to \Lambda$ is K-qs.*

17.5.1. *Holomorphic dependence of the Beltrami differential on λ.*

LEMMA 17.15. *Let $h_\lambda : U \to U_\lambda$ be a holomorphic motion of a domain $U \subset \mathbb{C}$ over a disk $\Lambda \subset \mathbb{C}$. Then the Beltrami differential $\mu_\lambda = \bar\partial h_\lambda / \partial h_\lambda$ depends holomorphically on λ (as an element of $L^\infty(U)$).*

PROOF. Let us take subdomains $U' \Subset U$ and $\Lambda' \Subset \Lambda$. Since the family of functions $\lambda \mapsto h_\lambda(z)$, $z \in U'$, is normal, it is uniformly bounded over Λ'. Hence the maps $h_\lambda : U' \to \mathbb{C}$, $\lambda \in \Lambda'$, are uniformly bounded. Moreover, by the Second λ-Lemma, they are uniformly K-qc. By (11.8), the $L^2(U')$-norms of the partial derivatives ∂h_λ, $\bar\partial h_\lambda$, are uniformly bounded as well. Thus, the family of maps h_λ, $\lambda \in \Lambda'$, is bounded in the Sobolev space $H^2(U')$.

By Lemma 14.26, $(h_\lambda)_{\lambda\in\Lambda}$ is a holomorphic curve in $H^2(U')$, and moreover, the partial derivatives $(\partial h_\lambda)_{\lambda\in\Lambda}$, $(\bar\partial h_\lambda)_{\lambda\in\Lambda}$, form holomorphic curves in $L^2(U')$. By Lemma 14.25, the functions $\lambda \mapsto \partial h_\lambda(z)$ and $\lambda \mapsto \bar\partial h_\lambda(z)$ are holomorphic over Λ for a.e. $z \in U$. Hence so is the Beltrami differential $\mu_\lambda(z) = \bar\partial h_\lambda(z)/\partial h_\lambda(z)$. Moreover, $\|\mu_\lambda\|_\infty < 1$. By Lemma 14.23, μ_λ as an element of $L^\infty(U)$ depends holomorphically on λ. □

17.6. Further λ-lemmas. Let us say that an extension of a holomorphic motion to some domain $D \subset \mathbb{C}$ is *canonical* if it behaves naturally under various conformal representations of D.

THIRD λ-LEMMA (Canonical Extension). *Let $h_\lambda : X \to X_\lambda$, $\lambda \in \mathcal{B}_1$, be a holomorphic motion of some set $X \subset \mathbb{C}$ over a Banach ball $\mathcal{B}_1$. Then it admits a canonical extension to a motion $\hat h_\lambda : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ over $\mathcal{B}_{1/3}$.*

This result is based upon quite advanced Teichmüller theory: it will be proved (and used) in vol. III.

We say that an extension of a holomorphic motion over Λ s *parameter global* if it is defined over the whole parameter domain Λ

FOURTH λ-LEMMA (Parameter Global Extension). *Let $h_\lambda : X \to X_\lambda$, $\lambda \in \mathbb{D}$, be a holomorphic motion of some set $X \subset \mathbb{C}$ over the disk $\mathbb{D}$. Then it admits a parameter global extension to a holomorphic motion $\hat h_\lambda : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ (over the same disk $\mathbb{D}$).*

This result needs some preparation in Complex Analysis in Several Variables: it will also be dealt with (and used) in vol. III.

18. Moduli and Teichmüller spaces of punctured spheres

18.1. Moduli spaces: preview. Consider some qc surface S (with or without boundary, possibly marked or partially marked).

The *moduli space* $\mathcal{M}(S)$, or the *deformation space of* S is the space of all conformal structures on S compatible with the underlying qc structure, up to the action of qc homeomorphisms preserving the marked data. In other words, $\mathcal{M}(S)$ is the space of all Riemann surfaces qc equivalent to S, up to conformal equivalence relation (respecting the marked data).

If we fix a reference Riemann surface S_0, then its deformations are represented by qc homeomorphisms $h : S_0 \to S$ to various Riemann surfaces S. Two such homeomorphisms h and $\tilde{h}$ represent the same point of the moduli space if there exists a conformal isomorphism $A : S \to \tilde{S}$ such that the composition

$$H = \tilde{h}^{-1} \circ A \circ h : S_0 \to S_0$$

respects all the marked data. In particular, $H = \mathrm{id}$ on the marked boundary. In the case when the whole fundamental group is marked, H must be homotopic to the id relative to the marked boundary.

For instance, if S has a finite conformal type, i.e., S is a Riemann surface of genus g with n punctures (without marking), then $\mathcal{M}(S)$ is the classical moduli space $M^{g,n}$. If S is fully marked then $\mathcal{M}(S)$ is the classical Teichmüller space $T^{g,n}$. This space has a natural complex structure of complex dimension $3g - 3 + n$ for $g > 1$. For $g = 1$ (the torus case), $\dim T^{1,0} = 1$ (see §2.6.3) and $\dim T^{1,n} = n$ for $n \geq 1$. For $g = 0$ (the sphere case), $\dim T^{0,n} = 0$ for $n \leq 3$ (by the Riemann-Koebe Uniformization Theorem and 3-transitivity of the Möbius group action) and $\dim T^{0,n} = n - 3$ for $n > 3$.

Exercise 18.1. *What is the complex modulus of the four punctured sphere?*

There is a natural projection (forgetting the marking) from $T^{g,n}$ onto $M^{g,n}$. The fibers of this projection are the orbits of the so called *Teichmüller modular group* acting on $T^{g,n}$ (it generalizes the classical modular group $\mathrm{PSL}(2, \mathbb{Z})$, see §2.6.3).

By the Riemann Mapping Theorem, the disk $\mathbb{D}$ does not have moduli. However, if we mark its boundary $\mathbb{T}$, then the space of moduli, $\mathcal{M}(\mathbb{D}, \mathbb{T})$, becomes infinitely dimensional! By definition, $\mathcal{M}(\mathbb{D}, \mathbb{T})$ is the space of all Beltrami differentials μ on $\mathbb{D}$ up to the action of the group of qc homeomorphisms $h : \mathbb{D} \to \mathbb{D}$ whose boundary restrictions are Möbius: $h|\mathbb{T} \in \mathrm{PSL}(2, \mathbb{R})$. It is called the *universal Teichmüller space*, since it contains all other deformation spaces. It plays an important role in holomorphic dynamics.

18.2. Definitions. Let us consider the Riemann sphere with a tuple of n marked points $\mathcal{Z} = (z_1, \ldots, z_n)$ (or, equivalently, n punctures). The punctures are considered to be "colored", or, in other words, the set $\mathcal{P}$ is ordered. Two such spheres $(\mathbb{C}, \mathcal{Z})$ and $(\mathcal{C}, \mathcal{Z}')$ are considered to be equivalent if there is a Möbius transformation $\phi : (\mathbb{C}, \mathcal{Z}) \to (\mathbb{C}, \mathcal{Z}')$ (preserving the colors of the punctures, i.e., $\phi(z_i) = z_i'$). The space of equivalence classes is called the *moduli space* $\mathcal{M}_n$.

If $n \leq 3$ then the moduli space $\mathcal{M}_n$ is a single point. If $n \geq 4$, we can place the last three points to $(0, 1, \infty)$ by means of a Möbius transformation. With this normalization $(\mathbb{C}, \mathcal{Z}) \sim (\mathbb{C}, \mathcal{Z}')$ if and only if $\mathcal{Z} = \mathcal{Z}'$, and we see that

$$\mathcal{M}_n = \{\mathbf{z} = (z_1, \ldots, z_{n-3}) : \; z_i \neq 0, 1; \; z_i \neq z_j\}.$$

This shows that $\mathcal{M}_n$ an $(n-3)$-dimensional complex manifold.

Let us fix some reference normalized tuple $\mathcal{Z}_\circ = (a_1, \ldots a_{n-3}, 0, 1, \infty)$. Then we can also define $\mathcal{M}_n$ as the space of homeomorphisms $h : (\mathbb{C}, \mathcal{Z}_\circ) \to \mathbb{C}$ normalized by $h(0) = 0$, $h(1) = 1$, up to equivalence: $h \sim h'$ if $h(\mathcal{Z}_\circ) = h'(\mathcal{Z}_\circ)$.

Let us now refine this equivalence relation by declaring that $h \simeq h'$ if h is homotopic (or, equivalently, isotopic) to h' rel $\mathcal{Z}_\circ$, and let $[h]$ stand for the corresponding equivalence classes. It inherits the quotient topology from the space of homeomorphisms (endowed with the uniform topology). This quotient space is called the *Teichmüller space* $\mathcal{T}_n$. Since the equivalence relation $\simeq$ is obviously stronger than $\sim$ we have a natural projection $\pi : \mathcal{T}_n \to \mathcal{M}_n$.

18.3. Spiders. The homotopy class $[h]$ can be visualized as the punctured sphere marked with a "spider". A *spider* $\mathcal{S}$ on the punctured sphere $(\mathbb{C}, \mathcal{Z})$ is a family of disjoint paths σ_i in $\mathbb{C} \smallsetminus \mathcal{Z}$ connecting z_i to ∞, $i = 1, \dots n-1$. We let $[\mathcal{S}]$ be the class of isotopic spiders (rel $\mathcal{Z}$).

LEMMA 18.2. *There is a natural one-to-one correspondence between points of $\mathcal{T}_n$ and classes of isotopic spiders,* $(\mathbb{C}, \mathcal{Z}, [\mathcal{S}])$.

PROOF. Let us fix a reference spider $(\mathbb{C}, \mathcal{Z}_\circ, \mathcal{S}_\circ)$. Then to each homeomorphism $h \in \mathcal{T}_n$ we can associate a spider $\mathcal{S} = h(\mathcal{S}_\circ)$. Isotopy h_t rel $\mathcal{Z}_\circ$ induces isotopy of the corresponding spiders rel $\mathcal{Z}$. Hence we obtain a map $[h] \mapsto [\mathcal{S}]$.

Vice versa, let us have a spider $(\mathbb{C}, \mathcal{Z}, \mathcal{S})$. Then there exists a homeomorphism $h : (\mathbb{C}, \mathcal{Z}_\circ, \mathcal{S}_\circ) \to (\mathbb{C}, \mathcal{Z}, \mathcal{S})$. If $(\mathbb{C}, \mathcal{Z}, \mathcal{S}')$ is an isotopic spider then the isotopy $\mathcal{S}_t$ rel $\mathcal{Z}$, $0 \le t \le 1$, lifts to an isotopy h_t rel $\mathcal{Z}_0$. Given any parameterizing homeomorphism $h' : \mathcal{S}_\circ \to \mathcal{S}'$, we can isotopy h_1 so that it will coincide with h' on $\mathcal{S}_\circ$. Since two homeomorphisms of a topological disk coinciding on the boundary are isotopic rel the boundary, we are done. □

18.4. Universal covering. The spiders can be labeled by tuples of $n-1$ elements of the fundamental group $\pi_1(\mathbb{C} \smallsetminus \mathcal{Z}) \approx \mathbb{F}_{n-1}$ (where the latter stands for the free group in $n-1$ generators). Indeed, let us consider a bouquet of circles $\bigvee_{i=1}^{n-1} C_i$ in $\mathbb{C}_i \smallsetminus \mathcal{Z}$ based at some point $a \in \mathbb{C} \smallsetminus \mathcal{Z}$ and such that the circle C_i surrounds z_i but not the other points of $\mathcal{Z}$. These circles oriented anti-clockwise represent generators of the fundamental group $\pi_1(\mathbb{C} \smallsetminus \mathcal{Z},\, a)$. Accordingly, any loop in $\bigvee C_i$ is homotopic to a concatenation of the loops C_i and their inverse. Let us select a proper arc γ_∞ connecting a to ∞ in the complement of $\bigvee C_i$, and $n-1$ arcs γ_i in the punctured disks bounded by the C_i. Since $\bigvee C_i$ is a homotopy retract for $\mathbb{C} \smallsetminus \mathcal{Z}$, any arc connecting z_i to ∞ is homotopic to the concatenation of the γ_i, a loop in $\bigvee C_i$, and γ_∞. Thus, any spider leg is labeled by an element of $\pi_1(\mathbb{C} \smallsetminus \mathcal{Z},\, a)$.

PROPOSITION 18.3. *The natural projection $\pi : \mathcal{T}_n \to \mathcal{M}_n$ is the universal covering over $\mathcal{M}_n$.*

PROOF. Let us first show that π is a covering. Take some base tuple $\mathcal{Z}_\circ = (z_1^\circ, \dots z_{n-1}^\circ) \in \mathcal{M}_n$, and consider a bouquet of circles C_i and the paths γ_i°, γ_∞° in $\mathbb{C} \smallsetminus \mathcal{Z}$ as above. Consider a neighborhood $U_1 \times \dots \times U_{n-3}$ of $\mathcal{Z}_\circ$ in $\mathcal{M}_n$, where the U_i are little round disks around z_i° fully surrounded by the circle C_i. Let us connect any point $z_i \in U_i$ to z_i° with a straight interval. Concatenating them with γ_i°, we obtain a path γ_i connecting z_i to a and continuously depending on $z_i \in U_i$.

Select now any element $\tau \in \pi(\mathbb{C} \smallsetminus \mathcal{Z}, a)$.

□

18.5. Infinitesimal theory. A tangent vector to the moduli space $\mathcal{M}_n$ at point $z = (z_1 \dots, z_{n-3}, 0, 1, \infty)$ can be represented as a tuple

$$v = (v(z_1), \dots v(z_{n-3}))$$

of tangent vectors to $\mathbb{C}$ at points z_i. Since the natural projection $\mathcal{T}_n \to \mathcal{M}_n$ is a covering, tangent vectors to $\mathcal{T}_n$ can be represented in the same way.

Any such tuple of vectors admits an extension to a smooth vector field v vanishing at points $(0, 1, \infty)$ (such vector field will be called "normalized"). So, we can view the tangent space to $\mathcal{M}_n$ (and $\mathcal{T}_n$) as the space $\mathrm{Vect} = \mathrm{Vect}(\hat{\mathbb{C}}, \mathcal{Z})$ of smooth normalized vector fields modulo equivalence relation: $v \sim w$ if $v(z_i) = w(z_i)$, $i = i, \dots, n-3$.

With this in mind, we can give a nice description of the cotangent space to $\mathcal{M}_n$ and $\mathcal{T}_n$. Let us consider the space $\mathcal{Q}^1 = \mathcal{Q}^1(\hat{\mathbb{C}}, \mathcal{Z})$ of integrable holomorphic quadratic differentials $q = q(z)\,dz^2$ on $\hat{\mathbb{C}}$ with poles in $\mathcal{Z}$. Such differentials must have at most simple poles (at ∞ it amounts to $q(z) = O(1/|z^3|)$).

Exercise 18.4. *Show that this space $\mathcal{Q}^1$ of quadratic differentials has complex dimension $n-3$. Moreover, the map $q \mapsto (\lambda_1, \dots, \lambda_{n-3})$, where $\lambda_i = \mathrm{Res}_{z_i} q$, is an isomorphism between $\mathcal{Q}^1$ and $\mathbb{C}^{n-3}$.*

It turns out that it is not an accident that $\dim \mathcal{Q}^1 = \dim \mathcal{M}_n$.

Proposition 18.5. *The space $\mathcal{Q}^1(\hat{\mathbb{C}} \smallsetminus \mathcal{Z})$ of quadratic differentials is naturally identified with the cotangent space to $\mathcal{M}_n$ (and $\mathcal{T}_n$). The pairing between a cotangent vector $q \in \mathcal{Q}^1(\hat{\mathbb{C}} \smallsetminus \mathcal{Z})$ and a tangent vector $v \in \mathrm{Vect}(\hat{\mathbb{C}}, \mathcal{Z})$ is given by the formula:*

$$< q, v >= -\frac{1}{\pi} \int\int q\, \bar{\partial} v = \sum_{i=1}^{n-3} \lambda_i v_i, \tag{18.1}$$

where $v_i = v(z_i)$, $\lambda_i = \mathrm{Res}_{z_i} q$.

Proof. Let us first note that this pairing is well defined. Indeed, as we saw in §2.11, $\bar{\partial} v$ is a Beltrami differential, and the product $q\,\bar{\partial} v$ as a conformal Riemannian metric that can be identified with its area form

$$q\,\bar{\partial} v \sim \frac{i}{2} q(z)\, \bar{\partial} v(z)\, dz \wedge d\bar{z}.$$

Moreover, this area form is integrable since q is integrable and $\bar{\partial} v$ is bounded.

Let us calculate the integral. Since q is holomorphic, we have[14]:

$$q\, \partial_{\bar{z}} v\, dz \wedge d\bar{z} = \partial_{\bar{z}}(q\, v)\, dz \wedge d\bar{z} = -\bar{\partial}(q\, v\, dz) = -d(q\, v\, dz).$$

Let $\gamma_\varepsilon(z_i)$ be the ε-circles centered at finite points of $\mathcal{Z}$, $i = 1, \dots, n-1$, and let Γ_ε be the ε^{-1}-circle centered at 0 (where all the circles are anti-clockwise oriented), and let D_ε be the domain of $\mathbb{C}$ bounded by these circles. Then by the Stokes formula

$$-\frac{1}{2\pi i} \int\int_{D_\varepsilon} d(q\, v dz) = \frac{1}{2\pi i} \sum \int_{\gamma_\varepsilon(z_i)} q\, v\, dz - \frac{1}{2\pi i} \int_{\Gamma_\varepsilon} q\, v\, dz.$$

But near any $z_i \in \mathbb{C}$ we have:

$$q\, v\ = \frac{\lambda_i v(z_i)}{z - z_i} + O(1).$$

Hence

$$\frac{1}{2\pi i} \int_{\gamma_\varepsilon(z_i)} q\, v\, dz \to \lambda_i v(z_i) \text{ as } \varepsilon \to 0.$$

[14]In this calculation ∂ and $\bar{\partial}$ are interpreted as *external* derivatives rather than their tensor counterparts

Note that these integrals asymptotically vanish at $z_{n-2} = 0$ and $z_{n-1} = 1$ since v vanishes at these points. The integral over Γ_ε asymptotically vanishes as well since $q(z) = O(|z|^{-3})$ while $v(z) = o(|z|^2)$ near ∞ (as the vector field v/dz vanishes at ∞).

Finally, we obtain:

$$\frac{1}{2\pi i}\int\int q\,\partial_{\bar z} v\,dz\wedge d\bar z = \sum_{i=1}^{n-3}\lambda_i v(z_i)$$

So, the pairing (18.1) depends only on the values of v at the points $z_1, \dots, z_{n-3}$, and hence defines a functional on tangent space $\mathrm{T}\mathcal{M}_n$. This gives an isomorphism between $\mathcal{Q}^1$ and the cotangent space $\mathrm{T}^*\mathcal{M}_n$ since $(\lambda_1, \dots, \lambda_{n-3})$ are global coordinates on the both spaces (see Exercise 18.4). $\square$

18.6. Teichmüller metric. Let us endow the space $\mathcal{Q}^1(\hat{\mathbb{C}} \smallsetminus \mathcal{Z})$ with the L^1-norm:

$$\|q\|_1 = \int |q|,$$

and the dual space $\mathrm{Vect}(\hat{\mathbb{C}}, \mathcal{Z})$ with the dual norm:

$$\|v\|_{\mathrm{T}} = \inf \|\bar\partial v\|_\infty,$$

where the infimum is taken over all smooth vector fields v with $v(z_i) = v_i$, $i = 1, \dots, n-3$, that vanish at $0, 1$ and ∞.

EXERCISE 18.6. *Check that the above two metrics are dual in the usual sense:*

$$\|q\|_1 = \sup_{\|v\|=1} | < q, v > |.$$

Recall that a *Finsler metric* on a manifold X is a continuous family of norms $\|v\|_x$ on the tangent spaces $\mathrm{T}_x X$ (where continuity means that the function $(x, v) \mapsto \|v\|_x$ is continuous on the tangent bundle). Equivalently, it is a continuous family of norms on the cotangent bundle.

Given a Finsler metric, we can measure the length of rectifiable paths:

$$l(\gamma) = \int_\gamma \|\dot\gamma(t)\|\,dt,$$

which induces the *Finsler distance* on X:

$$\mathrm{dist}(x, y) = \inf_\gamma l(\gamma),$$

where the infimum is taken over all rectifiable curves γ connecting x to y.

EXERCISE 18.7. *Show that the above norms on Q^1 and* Vect *endow $\mathcal{T}_n$ (and $\mathcal{M}_n$) with a Finsler metric (i.e., check continuity).*

Obviously, the projection $\pi : \mathcal{T}_n \to \mathcal{M}_n$ is a local isometry with respect to the corresponding Finsler metrics.

The *Teichmüller metric* on $\mathcal{T}_n$ is defined as follows. Let us consider two marked Riemann surfaces $h : (S^2, \mathcal{P}) \to (\hat{\mathbb{C}}, \mathcal{Z})$ and $h' : (S^2, \mathcal{P}) \to (\hat{\mathbb{C}}, \mathcal{Z}')$ representing points $\tau = [h]$ and $\tau' = [h']$ of $\mathcal{T}_n$. Then

$$\mathrm{dist}(\tau, \tau') = \inf_\phi \log \mathrm{Dil}\,\phi,$$

where ϕ runs over all qc maps $(\hat{\mathbb{C}}, \mathcal{Z}) \to (\hat{\mathbb{C}}, \mathcal{Z}')$ such that $h' \simeq \phi \circ h$.

Exercise 18.8. *Check that this is a metric.*

Theorem 18.9. *The above Finsler metric on $\mathcal{T}_n$ coincides with the Teichmüller metric.*

18.7. Compactness in $\mathcal{M}_n$.

Lemma 18.10. *A subset $\mathcal{K} \subset \mathcal{M}_n$ is precompact if and only if there exists an $\varepsilon > 0$ such that for any $(\hat{\mathbb{C}}, \mathcal{Z}) \in \mathcal{K}$ the marked points $z_i \in \hat{\mathbb{C}}$ are ε-separated in the spherical metric.*

Proof. The space $\hat{\mathbb{C}}^n$ is a natural compactification of $\mathcal{M}_n$: a point $\mathbf{z} = (z_1, \dots, z_n) \in \hat{\mathbb{C}}^n)$ belongs to $\mathcal{M}_n$ if and only if $z_i \neq z_j$ for any $i \neq j$. For any sequence $(\mathbf{z}^k)$ in $\mathcal{M}_n$ we can take a limit in $\hat{\mathbb{C}}^n$. This limit belongs to $\mathcal{M}_n$ if and only if the coordinates of the $\mathbf{z}^k$ are ε-separated for some $\varepsilon > 0$. □

A topological annulus $A \subset S^2 \smallsetminus \mathcal{Z}$ is called *trivial/peripheral* if so are its horizontal curves (see §1.7.10).

It is important to formulate the above compactness criterion in the conformally invariant/hyperbolic terms:

Lemma 18.11. *A subset $\mathcal{K} \subset \mathcal{M}_n$ is precompact if and only if one of the following equivalent properties hold:*

- *There exists $\mu > 0$ such that $\operatorname{mod} A \le \mu$ for any non-peripheral annulus $A \subset \hat{\mathbb{C}} \smallsetminus \mathcal{Z}$;*
- *There exists $\delta > 0$ such that $l_{\text{hyp}}(\gamma) \ge \delta$ for any closed hyperbolic geodesic in $\hat{\mathbb{C}} \smallsetminus \mathcal{Z}$.*

18.8. Appendix 1: General Teichmüller spaces.

18.8.1. *Marked Riemann surfaces.* The previous discussion admits an extension to an arbitrary qc class $\mathcal{QC}$ of Riemann surfaces that we will outline in this section. Take some *base* Riemann surface $S_0 \in \mathcal{QC}$ (without boundary), and let $\overline{S}_0$ be the ideal boundary compactification of S_0. Given another Riemann surface $S \in \mathcal{QC}$ (with compactification $\overline{S}$), a *marking* of S is a choice of a qc homeomorphism $\phi : \overline{S}_0 \to \overline{S}$ (*parametrization* by S_0) up to the following equivalence relation. Two parametrized surfaces (S, ϕ) and (S', ϕ') are equivalent if there is a conformal isomorphism $h : S \to S'$ that makes the following diagram *homotopically commutative rel the ideal boundary* (i.e., there is a qc homeomorphism $\tilde{\phi} : S_0 \to S$ homotopic to ϕ rel $\partial\overline{S}_0$ such that $h \circ \tilde{\phi} = \phi'$). A *marked Riemann surfaces* is an equivalence class $\tau = [S, \phi]$ of this relation. The space of all marked Riemann surfaces is called the *Teichmüller space* $\mathcal{T}(S_0)$.

Remark 18.12. Fixing a set Δ_0 of generators of $\pi_1(S_0)$ and parametrizations of the boundary components of $\partial\overline{S}_0$ by the standard circle $\mathbb{T}$, we naturally endow any marked Riemann surface $[S, \phi]$ with a set of generators of $\pi_1(S)$ (up to an inner automorphism of $\pi_1(S)$) and with a parametrization of the components ∂S by $\mathbb{T}$. Thus, we obtain a marked surface in the sense of §1.7.15.

18.8.2. *Representation variety.* Let us now uniformize the base Riemann surface S_0 by a Fuchsian group Γ_0. The *(Fuchsian) representation variety* $\mathrm{Rep}(\Gamma_0)$ is the space of faithful[15] Fuchsian representations $i : \Gamma_0 \to \mathrm{PSL}(2,\mathbb{R})$ up to conjugacy in $\mathrm{PSL}(2,\mathbb{R})$ endowed with the *algebraic* topology. In this topology $i_n \to i$ if after a possible replacement of the i_n with conjugate representations, we have: $i_n(\gamma) \to i(\gamma)$ for any $\gamma \in \Gamma_0$.

LEMMA 18.13. *There is a natural embedding* $e : \mathcal{T}(S_0) \to \mathrm{Rep}(S_0)$.

PROOF. Let $\phi : S_0 \to S$ be a qc parametrization of some Riemann surface $S \in \mathcal{QC}$, and let Γ be a Fuchsian group uniformizing S. Then ϕ lifts to an equivariant qc homeomorphism $\Phi : (\mathbb{H}, \Gamma_0) \to (\mathbb{H}, \Gamma)$, so there is an isomorphism $i : \Gamma_0 \to \Gamma$ such that $\Phi \circ \gamma_0 = \gamma \circ \Phi$ for any $\gamma_0 \in \Gamma_0$ and $\gamma = i(\gamma_0)$.

If we replace Φ with another lift $T \circ \Phi$, where $T \in \Gamma$, then i will be replaced with a conjugate representation $\gamma_0 \mapsto T^{-1} \circ i(\gamma_0) \circ T$.

If we replace ϕ with a homotopic parametrization $\tilde{\phi} : S_0 \to S$ then the induced representation $\Gamma_0 \to \Gamma$ will not change. Indeed, a homotopy ϕ_t connecting ϕ to $\tilde{\phi}$ lifts to an equivariant homotopy $\Phi_t : (\mathbb{H}, \Gamma_0) \to (\mathbb{H}, \Gamma)$ inducing a path of representations $i_t : \Gamma_0 \to \Gamma$. Then for any $\gamma_0 \in \Gamma_0$, the image $i_t(\gamma_0) \in \Gamma$ moves continuously with t. Since Γ is discrete, $i_t(\gamma)$ cannot move at all.

If we further replace $\tilde{\phi}$ with $h \circ \tilde{\phi}$, where $h : S \to S'$ is a conformal isomorphism then the representation $i : \Gamma_0 \to \Gamma$ will be replaced with a conjugate by $T : \mathbb{H} \to \mathbb{H}$ where $T \in \mathrm{PSL}(2,\mathbb{R})$ is a lift of h.

Thus, we obtain a well defined map $e : \mathcal{T}(S_0) \to \mathrm{Rep}(S_0)$ that associates to a marked surface $[S, \phi]$ the induced representation $i : \Gamma_0 \to \Gamma$ up to conjugacy in $\mathrm{PSL}(2,\mathbb{R})$.

Let us now show that e is injective. Let $\phi : S_0 \to S$ and $\phi' : S_0 \to S'$ be two parametrizations whose lifts Φ and Φ' to $\mathbb{H}$ induce two representations i and i' of Γ_0 that are conjugate by $T \in \mathrm{PSL}(2,\mathbb{R})$. Then Φ and $\Psi = T^{-1} \circ \Phi$ are two equivariant homeomorphisms $(\mathbb{H}, \Gamma_0) \to (\mathbb{H}, \Gamma)$ that induce the same representation $i : \Gamma_0 \to \Gamma$. We need to show that they are equivariantly homotopic.

To this end let us consider the following diagram encoding equivariance of Φ and Ψ:

Let $\delta(x)$ be the hyperbolic geodesic connecting $\Phi(x)$ to $\Psi(x)$. Since γ is a hyperbolic isometry, it isometrically maps $\delta(x)$ to $\delta(\gamma_0 x)$. Let $t \mapsto \Phi_t(x)$ be a uniform motion along $\delta(x)$ from $\Phi(x)$ to $\Psi(x)$ with such a speed that at time $t = 1$ we reach the destination (in other words, $\Phi_t(x)$ is the point on $\delta(x)$ on hyperbolic distance $t\,\mathrm{dist}_{\mathrm{hyp}}(\Phi(x), \Psi(x))$ from $\Phi(x)$). Then $\gamma(\Phi_t x) = \Phi_t(\gamma_0 x)$, and we obtain a desired equivariant homotopy. $\square$

18.8.3. *Teichmüller metric.* Let us endow the space $\mathcal{T}(S_0)$ with the following *Teichmüller metric.* Given two marked surfaces $\tau = [S, \phi]$ and $\tau' = [S', \phi']$, we let $\mathrm{dist}_{\mathrm{T}}(\tau, \tau')$ be the infimum of dilatations of qc maps $h : S \to S'$ that make the above diagram homotopically commutative.

LEMMA 18.14. $\mathrm{dist}_{\mathrm{T}}$ *is a metric.*

PROOF. Triangle inequality for $\mathrm{dist}_{\mathrm{T}}$ follows from submultiplicativity of the dilatation under composition. So, $\mathrm{dist}_{\mathrm{T}}$ is a pseudo-metric. Let us show that it

[15] i.e., injective

is a metric, Indeed, if $\mathrm{dist}_{\mathrm{T}}(\tau, \tau') = 0$ then there exists a sequence $h_n : S \to S'$ of qc maps in the right homotopy class with $\mathrm{Dil}(h_n) \to 0$. Let $H_n : \mathbb{H} \to \mathbb{H}$ be the lifts of the h_n that induce the same isomorphism between Γ and Γ'. Then the H_n is a sequence of qc maps with uniformly bounded dilatation whose extensions to $\mathbb{R} = \partial\mathbb{H}$ all coincide. Now Compactness Theorem 13.2 implies that the H_n uniformly converge to an equivariant conformal isomorphism $T : (\mathbb{H}, \Gamma_0) \to (\mathbb{H}, \Gamma)$. It descends to a conformal isomorphism $h : S \to S'$ in the same homotopy class as the h_n. $\square$

EXERCISE 18.15. *Show that the embedding $e : \mathcal{T}(S_0) \to \mathrm{Rep}(\Gamma_0)$ is continuous. (from the Teichmüller metric to the algebraic topology).*

Notes

The local version of the "MRMT" goes back to Gauss who proved that any real analytic metric can be locally brought to a conformal form $\rho(z)|dz|^2$ thus, solving the "Beltrami equation" with an analytic coefficient). Once the Uniformization Theorem becomes available, the global version follows (with the corresponding regularity of the metric.)

Apparently, the theory of quasiconformal maps originated in the work on cartography by Tissot in mid XIXth century: see a historical account in [**Pap**]. The official birth is usually associated with Gröztsch's 1928 paper [**Gr**], where the extremal problem for rectangle diffeomorphisms was considered (making one of the first applications of the length-area method). Probably, the Koebe Distortion Theorem played a motivating role in this story.[16] Gröztsch developed this circle of ideas through the early 1930s. It was further advanced in the work of Teichmüller around 1940 who connected extremal maps on Riemann surfaces to quadratic differentials. (See [**Ku**] for more comments on this early history.)

In 1935, the notion of quasiconformality was rediscovered by Lavrentiev (already for homeomorphisms) who proved solvability of the Beltrami equation with a continuous coefficient. In 1938, Morrey proved the measurable version. (Lavrentiev was motivated by the geometric problem of bringing Riemannian metrics to a conformal form, while Morrey's interest came from the PDE side). Lavrentiev called these maps "pseudo-analytic"; the name "quasiconformal" was given by Ahlfors [**A4**] (p. 185).

However, a systematic development of the theory (under proper regularity assumptions) was undertaken only after the war, by Lavrentiev, Bers and Nierenberg, Vekua and Bojarski, Ahlfors and Bers, Volkovyskii, Belinskii and Pesin, Strebel, Pfluger and Mori, followed by many others.

Let us mention, in particular, the following contributions:[17]

Definition of quasiconformality in terms of the uniform bound on the upper circular dilatation was introduced by Lavrentiev [**Lav**]. This class of maps was systematically studied by I.N. Pesin [**Pes**] who proved that the absolutely continuity property, as well as compactness of the space of K-qc maps. The quasisymmetry property

---

[16] Gröztsch was a student of Koebe.

[17] Unfortunately, the history of the theory of qc maps, though quite recent, is not readily decipherable. The author has browsed through the literature of those days in a non-systematic way, so some important contributions (and causality relations between them) can certainly be overlooked.

also appeared in [**Pes**] with attribution to Belinskii. It is also mentioned that such maps are absolutely continuous on lines, with a reference to Menshov (1937).

With this regularity in hands, the Grötzsch method implies the quasi-invariance of moduli. Mori's paper [**Mori**, Lemma 4] closes up the loop by showing that the quasi-invariance of moduli implies local quasisymmetry (and hence a uniform bound on the upper circular dilatation).

As far as we can tell, distributional derivative were introduced into the subject by Vekua [**Ve1**].

The Ahlfors-Beurling criterion appeared in [**ABeu**].

As we have already mentioned, the local version of the MRMT is due to Morrey [**Mor**]. A global approach via integral representations was developed by Vekua [**Ve2**] and Bojarski [**Bo**]. It was further explored by Ahlfors and Bers [**AB**] (see also [**A2**]), with the emphasis on the parameter dependence, who followed up with numerous deep applications to Teichmüller theory and theory of Kleinian groups. Various contemporary views appeared in [**DB, IM**]. We have taken as qualitative path as we could get, making a minimalistic use of integral representations.

The idea of the Moduli spaces of Riemann surfaces goes back to Riemann. The idea of marked surfaces leading to "Teichmüller spaces" is usually attributed to Teichmüller, though apparently it had appeared already in Fricke's work early in the of the 20th century. (In fact, sometimes the spaces are called *Fricke*, see [**Ab**]).

As we have mentioned above, quasiconformal maps found deep applications in the work of Ahlfors and Bers (in the 1960-70's) to the theory of Kleinian groups (with a feedback to the qc theory). In particular, qc welding appeared in [**Bers1**] in this context.

Quasiconformal maps attained even greater prominence in the work of Mostow (late 1960s), Thurston, and Sullivan in the 1970s, relating them in a deep way to Hyperbolic Geometry and Ergodic Theory. They were introduced to Holomorphic Dynamics by Sullivan in the early 1980s, and have become an indispensable tool in this field ever since.

The First λ-Lemma (extension to the closure) appeared in [**L7**] and [**MSS**] in the dynamics context. The Second λ-Lemma (quasiconformality) is due to Mañé-Sad-Sullivan [**MSS**]. The Third λ-Lemma (the canonical extension) is due to Bers and Royden [**BR**]. Independently, existence of some extension to the whole sphere over some ball $\mathcal{B}_r$ (of a universal radius $r \in (0, 1)$) was proved by Sullivan and Thurston [**STh**]. The Fourth λ-Lemma was proved by Slodkovski [**Sl**], based on the Forstenrič machinery [**For**]

The first text book on the basic theory of qc maps was written by Volkovysski [**V**] (who applied them to the type problem for Riemann surfaces). It followed with many more, in Russian and English, see [**A2, Bel, Kr, LV**], with the book by Ahlfors remaining the most popular source. Among more recent sources let us mention [**GaL, He**], where the former focuses on applications to the Teichmüller theory, while the latter develops a contemporary general theory of quasisymmetric maps on metric spaces.

Quasiconformal maps remain an active area of research, with many important applications. In the upcoming volumes, we will encounter them frequently.

Acknowledgement. Our take on the theory of quasiconformal maps (like language of conformal structures and base free viewpoint at the Teichmüller metric) is largely inspired by Dennis Sullivan's insights. We also thank A. Eremenko and S. Krushkal for helpful comments on the early history of qc maps and Teichmüller spaces.

Part 2

Complex and real quadratic family

19. Glossary of Dynamics

This glossary collects some basic notions, examples and results of Ergodic Theory and Dynamics. In particular, we give a nearly complete account of the theory of expanding circle maps that serves as a good prototype for many dynamical themes.

19.1. Orbits and invariant sets. Consider a continuous endomorphism $f : X \to X$ of a topological space X. The n-fold iterate of f is denoted by f^n, $n \in \mathbb{N}$. A *topological dynamical system* (with discrete positive time) is the $\mathbb{N}$-action generated by f, $n \mapsto f^n$. The *orbit* or *trajectory* of a point $x \in X$ is $\operatorname{orb} x = \{f^n x\}_{n\in\mathbb{N}}$. (We often let $x_n \equiv f^n x$, $z_n \equiv f^n z$, etc.) The subject of *Topological Dynamics* is to study qualitative behavior of orbits of a topological dynamical system.

Here is the simplest possible behavior: a point α is called *fixed* if $f\alpha = \alpha$. More generally, a point α is called *periodic* if it has a finite orbit, i.e., there exists a $p \in \mathbb{Z}_+$ such that $f^p\alpha = \alpha$. Any such moment p is called a *period* of α, but we will reserve this term for the *minimal period* , unless otherwise is explicitly assumed. The orbit of α (consisting of p permuted points) is naturally called a *periodic orbit* or a *cycle* (of period p). We will write periodic orbits in bold: $\boldsymbol{\alpha} = \operatorname{orb}\alpha$, $\boldsymbol{\beta} = \operatorname{orb}\beta$, etc. The sets of fixed and periodic points are denoted $\operatorname{Fix}(f)$ and $\operatorname{Per}(f)$ respectively.

Exercise 19.1. *Count the number of periodic points of minimal period p in terms of the numbers $|\operatorname{Fix}(f^n)|$.*

A point α is called *preperiodic* if $f^n\alpha$ is periodic for some $n > 0$ (but α itself is not periodic). The minimal such n is called the *preperiod* of α.

A subset $Z \subset X$ is called *(forward) invariant* under f if $f(Z) \subset Z$ (or equivalently, $f^{-1}(Z) \supset Z$). It is called *backward invariant* if $f^{-1}(Z) \subset Z$. If Z is simultaneously forward and backward invariant (so that $f^{-1}(Z) = Z$), it is called *completely invariant*.

A set Z is called *wandering* if $f^n Z \cap f^m Z = \emptyset$ for any $n > m \geq 0$. It is called *weakly wandering*[18] if $f^{-n}(Z) \cap Z = \emptyset$ for any $n > 0$.

Exercise 19.2. *Show that wandering sets are weakly wandering but not the other way around (in general). Show that Z is weakly wandering if and only if either of the following properties is satisfies:*

- $f^{-n}(Z) \cap f^{-m}(Z) = \emptyset$ *for any* $n > m \geq 0$*;*
- *No point $z \in Z$ returns back to Z under iterates of f.*

Topologically, the asymptotical behavior of an orbit can be studied in terms of its limit set. The ω-*limit set* $\omega(x)$ of a point x is the set of all accumulation points of $\operatorname{orb}(x)$. We say that the orbit of x *converges to a cycle* $\boldsymbol{\alpha}$ if $\omega(x) = \boldsymbol{\alpha}$.

A point x is called *recurrent* if $\omega(x) \ni x$. Existence of non-periodic recurrent points is a feature of non-trivial dynamics.

A *backward orbit* of a point x is a sequence of points $(x_{-n})_{n\in\mathbb{N}}$ such that $x_0 = x$ and $f(x_{-n-1}) = x_{-n}$ for all $n \in \mathbb{N}$.

Exercise 19.3. *If X is compact that the limit set for any forward or backward orbit is a non-empty compact invariant subset $\mathcal{O}$ of X, and the restriction $f|\mathcal{O}$ is surjective .*

---

[18] Actually, in the literature these sets are usually called “wandering” as well.

The sets

$$\mathrm{Orb}_-(x) := \bigcup_{n\ge 0} f^{-n}(x), \quad \mathrm{Orb}(x) := \bigcup_{m\ge 0} \mathrm{Orb}_-(f^m x) \equiv \mathrm{Orb}_-(\mathrm{orb}\, x)$$

are called respectively the *grand backward orbit* and the *grand orbit* of a point x. The latter is an equivalence class of the following equivalence relation:

$$x \sim y \text{ if } f^n x = f^m y \quad \text{for some } m, n \in \mathbb{N}.$$

Note that the usual forward orbits $\mathrm{orb}\, x$ are not classes of any equivalence relation. In fact, the grand orbits relation is the minimal one generated by the forward orbits.

There is a smaller equivalence relation

$$z \sim \zeta \text{ if } f^n z = f^n \zeta \quad \text{for some } n \in \mathbb{N}.$$

These equivalence classes will be called *petit orbits*[19] of f.

Given a connected set U and a point x such that $f^n x \in \mathrm{int}\, U$ for some $n \ge 0$, let V be the connected component of $f^{-n}U$ containing x. It is called i the *pullback* of U (along the n-orbit of x).

A point x is called *Lyapunov stable* if the orbits of nearby points stay close to $\mathrm{orb}\, x$, i.e.,

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that } d(x,y) < \delta \implies d(f^n x, f^n y) < \varepsilon,\ n = 0, 1, 2, \dots$$

This notion is particularly useful in the case when x is fixed or periodic.

19.2. Return map and its relatives. Let $Y \subset X$. The *first return map* to Y is a partially defined map $R \equiv R_Y : Y \to Y$ such that $Rx = f^n x$, where $n = n(x) > 0$ is the first positive moment such that $f^n x \in Y$ (if exists). Such $n(x)$ is called the *first return time*.

Similarly, the *first landing map* to Y is a partially defined map $L \equiv L_Y : X \to Y$ such that $Lx = f^n x$, where $n = n(x) \ge 0$ is the first non-negative moment such that $f^n x \in Y$. Such $n(x)$ is called the *first landing time*. Note that $L \,|\, Y = \mathrm{id}$ and $R = L \circ f \,|\, Y$.

The first return map to $X \smallsetminus Y$ is also called the *transit map* through Y,

$$T \equiv T_Y := R_{X \smallsetminus Y}, \quad fx = f^n x,$$

where $n = n(x)$ is called the *first transit time* of $x \in X \smallsetminus Y$ through Y.

Each of these maps is usually discontinuous.[20]

19.3. Topological transitivity and related notions. A map $f : X \to X$ is called *topologically transitive* if it has a dense orbit.

In case of a complete metric space X (in particular, for a compact metrizable space with an arbitrary metric) a property of a point $x \in X$ is called *generic* if it is satisfied on the countable intersection of dense open sets. By the Baire Theorem, the complementary set is nowhere dense.

EXERCISE 19.4. *Assume X is a complete metric space.*

(i) *Show that topological transitivity is equivalent to the following property: for any open sets U and V, there exists an $n \in \mathbb{N}$ such that $f^{-n}(U) \cap V \ne \emptyset$.*

(ii) *If f is topologically transitive then the orbit of a generic point $x \in X$ is dense.*

---

[19] Usually they are called *small orbits* but we reserve this term for another purpose.

[20] To make these maps better, we will sometimes slightly modify the definitions (see §31.2), but this will always be explicitly pointed out.

A map $f : X \to X$ is called *minimal* if *all* its orbits are dense.

EXERCISE 19.5. (i) *If f is minimal then the first landing times to any neighborhood U are bounded (of course, with a bound depending on U).*
(ii) *Vice versa, assume f is topologically transitive, and let $x \in X$ be a point with dense orbit. If the first landing times to any neighborhood $U \ni x$ are bounded (for points that land in U), then f is minimal.*

A map f is called *topologically mixing* if for any open sets U and V, there exists an $N \in \mathbb{N}$ such that $f^{-n}(U) \cap V \neq \emptyset$ for *all* $n \geq N$. It follows from the above Exercise that mixing is stronger than transitivity. An even stronger property is *topological exactness* asserting that for any open set U, there exists an $n \in \mathbb{N}$ such that $f^n(U) = X$. (Of course, this property makes sense only for endomorphisms.) This property is also called *leo* ("locally eventually onto").

19.4. Equivariant maps. Two dynamical systems $f : X \to X$ and $g : Y \to Y$ are called *topologically conjugate* (or topologically equivalent) if there exists a homeomorphism $h : X \to Y$ such that $h \circ f = g \circ h$, i.e., the following commutative diagram is valid:

$$\begin{array}{ccc} X & \xrightarrow[f]{} & X \\ h \downarrow & & \downarrow h \\ Y & \xrightarrow[g]{} & Y \end{array}$$

Classes of topologically equivalent dynamical systems (within an *a priori* specified family) are called *topological classes*. If X and Y are endowed with an extra structure (e.g., smooth, conformal, quasiconformal etc.) respected by h, then f and g are called *smoothly/conformally/quasiconformally conjugate* (or equivalent). The corresponding equivalence classes are called smooth/conformal/quasiconformal classes.

Topological conjugacies respect all properties which can be formulated in terms of topological dynamics: orbits go to orbits, cycles go to cycles of the same period, ω-limit sets go to the corresponding ω-limit sets, converging orbits go to converging orbits, etc.

A homeomorphism $h : X \to X$ commuting with a dynamical system $f : X \to X$ (i.e., conjugating f to itself) is called an *automorphism* of f.

A continuous map which makes the above diagram commutative is called *equivariant* (with respect to the actions of f and g). A *surjective* equivariant map is called a *semi-conjugacy* between f and g. In this case g is also called a *quotient* of f.

It will be convenient to extend the above terminology to *partially defined* maps. Let f and g be partially defined maps on the spaces X and Y respectively (i.e., f maps its domain $\mathrm{Dom}(f) \subset X$ to X, and similarly does g). Let $A \subset X$. A map $h : A \to Y$ is called *equivariant* (with respect to the actions of f and g) if for any $x \in A \cap \mathrm{Dom}(f)$ such that $fx \in A$ we have: $hx \in \mathrm{Dom}(g)$ and $h(fx) = g(hx)$. (Briefly speaking, "the equivariance is satisfied whenever it makes sense".)

19.5. Multipliers. Let $f : (M, \alpha) \to (M, \alpha)$ be a local smooth map of a manifold fixing a point α.. The eigenvalues of the differential $Df(\alpha)$ (calculated in any local chart near α) are called the *multipliers* ρ_k of α. A basic observation is that the multipliers remain invariant under smooth conjugacies, and in particular, are independent of the choice of a local chart near α:

PROPOSITION 19.6. *If two local diffeomorphisms*[21] $f : (M, \alpha) \to (M, \alpha)$ *and* $\tilde{f} : (\tilde{M}, \tilde{\alpha}) \to (\tilde{M}, \tilde{\alpha})$ *are conjugate by a local homeomorphism* $h : (M, \alpha) \to (\tilde{M}, \tilde{\alpha})$ *which is differentiable at* α, *then they have the same multipliers.*

PROOF. By the Chain Rule, the differentials $Df(\alpha)$ and $D\tilde{f}(\tilde{\alpha})$ are conjugate by the differential $Dh(\alpha)$. $\square$

More generally, the *multipliers* of a periodic point α of minimal period p (and of its cycle $\boldsymbol{\alpha}$) are defined as the multipliers of α as a fixed point for f^p. By the previous statement, they remain invariant under conjugacies differentiable at α. Moreover, by the Chain Rule, we can calculate them as the eigenvalues of the matrix product

$$Df^p(\alpha) = \prod_{k=0}^{p-1} Df(\alpha_k), \quad \alpha_k = f^k\alpha,$$

where the differentials are calculated at any local charts near the α_k. (Note that this formula shows that the multipliers $\rho(\boldsymbol{\alpha})$ are independent of the choice of a periodic point α_k within the cycle $\boldsymbol{\alpha}$.)

In the case of a local holomorphic map $f : (M, \alpha) \to (M, \alpha)$, the multipliers of α are defined as the *complex* eigenvalues, so there are only n (rather than $2n$) of them. They remain invariant under biholomorphic conjuagacies, and in particular, are independent of the choice of a holomorphic local chart.

This definition extends naturally to the case of periodic points.

19.6. Elements of Ergodic Theory.

19.6.1. *Space of probability measures.* Let X be a compact space. The space of Borel measures on X is endowed with the *weak topoology* (w^*), see §13.7.2. Recall that convergence $\mu_n \to \mu$ in this topology means that

$$\int \phi \, d\mu_n \to \int \phi \, d\mu \text{ as } n \to \infty$$

for any continuous function $\phi \in C(X)$. Equivalently

$$\mu_n(D) \to \mu(D) \quad \text{for any domain } D \text{ with } \mu(\partial D) = 0.$$

(It is sufficient to check this for a basis of topology, e.g. for a suitable family of balls.)

Let $\mathfrak{M}(X)$ be the subspace of probabilty measures on X. It is a convex compact space.

Any continuous map $f : X \to X$ induces the *push-forward operator* $f_* : \mathfrak{M} \to \mathfrak{M}$ that can be defined in two equivalent ways:

$$(f_*\mu)(Y) = \mu(f^{-1}(Y)) \quad \text{for any measurable subset } Y \subset X$$

or

$$\int \phi \, d(f_*\mu) = \int \phi \circ f \, d\mu \quad \text{for any continuous function } \phi \in C(X).$$

EXERCISE 19.7. *Assume that* $f : X \to X$ *is a homeomorphism, and let* $\rho \in L^1(\mu)$. *Then* $f_*(\rho \, d\mu) = (\rho \circ f^{-1}) \, d\mu$.

[21] In fact, it is enough to assume differentiability at the marked fixed points only.

19.6.2. *Invariant measures: averaging and equidistribution.* A measure $\mu \in \mathfrak{M}(X)$ is called *invariant* if $f_* \mu = \mu$, i.e.,

$$\int \phi \circ f \, d\mu = \int \phi \, d\mu$$

for any continuous function $\phi \in C(X)$. Equivalently, for any measurable set $Y \subset X$, we have $\mu(f^{-1}(Y)) = \mu(Y)$.

The simplest example of an invariant measure is the δ-measure δ_α supported on a fixed point α. More generally, one can consider a measure

$$\delta_{\boldsymbol{\alpha}} = \frac{1}{p} \sum_{k=0}^{p-1} \delta_{f^k \alpha}$$

equidistributed over a periodic cycle $\boldsymbol{\alpha} = \{f^k \alpha\}_{k=0}^{p-1}$.

Exercise 19.8. *Show that if δ_α is an* atom *of an invariant measure μ (i.e., $\mu(\{\alpha\}) > 0$) then α is a periodic point.*

We say that an orbit of x is *equidistributed with respect to* μ if

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x} \to \mu \text{ as } \quad n \to \infty$$

in the space $\mathfrak{M}(X)$. In other words,

$$\frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k x) \to \int \phi \, d\mu \quad \text{as } n \to \infty$$

for any continuous function $\phi \in C(X)$. Equivalently, for any open subset $D \subset X$ with $\mu(D) = 0$, the *orbit of x visits D with asymptotic frequency equal to $\mu(D)$.*

Under these circumstances, we also say that *the orbit of x is governed by μ.*

Bogolyubov-Krylov Theorem. *Any continuous map $f : X \to X$ on a compact space X has an invariant probability measure.*

Proof. It can be constructed by an *averaging procedure.* Namely, start with an arbitrary probability measure m and consider its *ergodic averages*

$$m_n = \frac{1}{n} \sum_{k=0}^{n-1} f_*^k m. \tag{19.1}$$

Since the space $\mathfrak{M}(X)$ is compact, we can select a converging subsequence $m_{n(i)} \to \mu$, which is a desired invariant measure, since

$$f_* m_n - m_n = \frac{1}{n}(f_*^n m - m) \to 0 \quad \text{as } n \to \infty.$$

$\square$

Non-atomic invariant measures immediately create an interesting dynamics:

Poincaré Recurrence Theorem. *For any finite invariant measure μ, almost all points are recurrent.*

Proof. It follows from a stronger assertion that for any measurable set Y, almost all points $x \in Y$ return to Y. Indeed, let $Z \subset Y$ be the set of points that never return to Y. This set is weakly wandering, hence the full preimages $f^{-n}(Z)$, $n = 0, 1, \dots,$ are pairwise disjoint. Since μ is invariant, all these preimages have the same measure m. Since μ is finite, $m = 0$. $\square$

A measure μ is called *ergodic* if for any any measurable decomposition $X = Y_1 \sqcup Y_2$ into two invariant sets, either $\mu(Y_1) = 0$ or $\mu(Y_2) = 0$.[22] Equivalently, any completely invariant measurable subset $Y \subset X$ has either zero or full measure.

Exercise 19.9. *Non-ergodicity of a measure μ is equivalent to the existence of an invariant function $\phi \in L^1(\mu)$ (i.e., $\phi \circ f = \phi$ μ-a.e.)*

The following fundamental result asserts that *an ergodic measure μ governs behavior of μ-almost all orbits*:

Birkhoff Ergodic Theorem. *For any ergodic invariant probability measure μ, almost all orbits are equidistributed with respect to μ.*

In other words, *μ-typical points are equidistributed with respect to an invariant measure μ.* In fact, the equidistribution property is often used as a definition of a *typical point* for μ.

Exercise 19.10. *Show that for any ergodic invariant probability measure μ and any function $\phi \in L^1(\mu)$, we have*

$$\frac{1}{n}\sum_{k=0}^{n-1} \phi(f^k x) \to \int \phi\, d\mu \quad \text{as } n \to \infty$$

for μ-a.e. x.

In particular, if we take $\phi = \chi_Y$ to be the characteristic function of some measurable set Y then *μ-almost all orbits visit Y with the asymptotic frequency equal to $\mu(Y)$.*

Proposition 19.11. *Two different ergodic invariant probability measures (for the same transformation) are mutually singular.*

Proof. Let us first show that *if $\mu \in \mathfrak{M}_f(X)$ is ergodic and ν is absolutely continuous with respect to μ then $\nu = \mu$.* Indeed, such measures share a typical point x, and both of them can be recovered from this point by taking the limit of its Birkhoff averages.

Exercise 19.12. *Justify the above assertion in the case of invertible f without using the Ergodic Theorem.*

In general, let $\nu = \nu_a + \nu_s$ be the decomposition of ν into the absolutely continuous and singular parts with respect to μ. If both of them are not vanishing then their measurable supports[23] can be selected to be invariant and disjoint, contradicting ergodicity of ν.

Exercise 19.13. *Justify this assertion. Show that both components ν_a and ν_s are invariant measures.*

[22] Note that this definition makes sense for non-invariant measures, too.

[23] A *measurable support* is any measurable set of full measure.

Hence ν is either absolutely continuous or singular with respect to μ. But in the former case, $\nu = \mu$, as we have shown above. The conclusion follows. $\square$

Let $\mathfrak{M}_f(X)$ stand for the space of f-invariant probability measures. It is a non-empty convex compact subset of $\mathfrak{M}(X)$.

EXERCISE 19.14. *Show that ergodic measures can be characterized as extreme points of $\mathfrak{M}_f(X)$. (Recall that μ is extreme if it cannot be represented as a convex combination of two other measures.)*

As extreme points exist by the *Krein-Milman Theorem*, we conclude:

COROLLARY 19.15. *Any continuous map $f : X \to X$ on a compact space has an ergodic invariant measure.*

By the *Choquet Theorem*, for any invariant measure μ there is a probability distribution dP on the space of extreme points such that

$$\mu = \int \nu \, dP(\nu).$$

This representation is called the *ergodic decomposition* of μ. In the case when dP is supported on a finite or countable set of measures, it assumes a simple form

$$\mu = \sum p_i \nu_i, \quad \sum p_i = 1, \; p_i > 0$$

where the ν_i are mutually singular ergodic measures, called the *ergodic components* of μ. The meaning of a continuous ergodic decomposition is less obvious. A suggestive example is provided by a skew map of the cylinder (endowed with the flat area):

$$f : \mathbb{T} \times I \to \mathbb{T} \times I, \quad (x, y) \mapsto (x + y, y)$$

decomposed a.e. into ergodic circle rotations (corresponding to irrational y).

See §... for more comments on this notion.

19.6.3. *Unique ergodicity.* A map f is called *uniquely ergodic* if it has a unique (Borel probability) invariant measure μ. It amounts to the following *Uniform Ergodic Theorem*:

EXERCISE 19.16. *A measure μ is a unique invariant measure for $f : X \to X$ (where X is compact) iff for any continuous function $\phi \in C(X)$ we have*

$$\frac{1}{n} \sum_{k=0}^{n-1} \phi \circ f^k \to \int \phi \, d\mu \quad \text{as } n \to \infty$$

uniformly on X.

19.6.4. *Mixing.* An invariant measure μ is called *mixing* if for any measurable sets Y and Z we have:

$$\mu(f^{-n}(Y) \cap Z) \to \mu(Y) \cdot \mu(Z) \quad \text{as } n \to \infty.$$

EXERCISE 19.17. *Show that mixing is equivalent to the following property: For any continuous functions ϕ and ψ we have*

$$\int (\phi \circ f^n) \cdot \psi \, d\mu \to \int \phi \, d\mu \int \psi \, d\mu.$$

This also holds for any integrable ϕ and ψ.

Obviously, mixing implies ergodicity.

19.6.5. *Quasi-invariant measures and acim's.* In Dynamics, one often encounters the following situation: there is a natural *geometric* or *physical* measure m (e.g., Lebesgue or Hausdorff) that provides a meaningful sense of typicality, so it is interesting to understand the behavior of m-typical orbits. Here we supply a conceptual frame for this discussion.

A measure m is called *quasi-invariant* if $f_* m \sim Lm$ (i.e., these two measures are in the same measure class). In other words, for any measurable set Y, we have

$$m(Y) = 0 \iff m(f^{-1}(Y)) = 0.$$

Ergodicity of a quasi-invariant measure is defined exactly as above (in terms of indecomposibility). (Equivalently, there are no non-constant measurable functions $\phi : X \to \mathbb{R}$ invariant under f, i.e., such that $\phi \circ f = \phi$.) However, it does not imply equidistribution of *m-typical* orbits. To address this problem, one can try to find an *(ergodic) invariant measure μ which is absolutely continuous* with respect to m (abbreviated *acim*). By the Birkhoff Ergodic Theorem, *μ would govern the behavior of m-typical orbits.* This makes the problem of existence of an acim central in the field. Here is a useful general criterion:

Proposition 19.18. *Let m be a probability quasi-invariant measure. Assume that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$m(Y) < \delta \implies m(f^{-n}(Y)) < \varepsilon, \quad n = 0, 1, 2, \dots$$

Then there exists an acim μ.

Proof. Let us construct an invariant measure μ by the Bogolyubov-Krylov averaging procedure applied to m. Then for the measures m_n (19.1), we have:

$$m(Y) < \delta \implies m_n(Y) < \varepsilon, \quad n = 0, 1, 2, \dots.$$

If Y is an open set with $m(\partial Y) = 0$, then we can pass to the limit and conclude that $\mu(Y) < \varepsilon$ as well. This implies absolute continuity. □

Assume now that M is a Riemannian manifold, $f : M \to M$ is a smooth map with Jacobian $\operatorname{Jac} f \equiv \operatorname{Jac}_m f$ evaluated with respect to the Riemannian volume m.

Exercise 19.19. *Under these circumstances, assume the critical locus of f has zero volume. Then m is quasi-invariant under f. Moreover, if $d\mu = \rho\, dm$ is an acim with density ρ, then $f_*(d\mu) = h\, dm$, where*

$$h(y) = \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{\operatorname{Jac}_m f(x)} \tag{19.2}$$

(*as long as y is not a critical value*).

Exercise 19.20. *Let $f : \mathbb{R} \to \mathbb{R}$, $y = x^2$, and let m be the Lebesgue mesure on $\mathbb{R}$. Then $f_*(m) = \dfrac{m}{2\sqrt{y}}$.*

19.7. Attractors. Let $f : M \to M$ be a smooth map of a manifold. A periodic point $\alpha \in M$ (and its cycle $\boldsymbol{\alpha}$) is called *attracting* if $|\rho_k| < 1$ for all its multipliers.

Exercise 19.21. (i) *Any attracting fixed point α has an invariant neighborhood U such that $\displaystyle\bigcap_{n=0}^{\infty} f^n(U) = \{\alpha\}$.*

(ii) *Generalize it to the case of an attracting cycle.*

Attracting cycles provide us with simplest examples of attractors. In general, we want to define an "attractor" as a compact invariant subset $A \subset M$ that "attracts" an "essential" set of points. This notion vary depending on the meaning of the words "attracts" and "essential". There are two viewpoints on the latter: Lebesgue (measure-theoretical) and Baire (topological). Let us start with the former.

Let us endow M with a Riemannian metric and the corresponding Lebesgue measure m. Let

$$\mathcal{R}(A) := \{x \in M : \ \omega(x) \subset A\}$$

be the *realm of attraction* of A.

A measurable set is *neglectable* from the measure-theoretical point of view if it has zero Lebesgue measure; otherwise, it is *essential*. Some property depending on a point $x \in M$ is called *typical* if it is satisfied on a subset of full measure.

A *measure-theoretic attractor* (in the sense of Milnor) is a compact invariant subset $A \subset M$ such that $m(\mathcal{R}(A)) > 0$, and for any proper compact invariant subset $A' \subset A$, we have $m(\mathcal{R}(A)) \smallsetminus \mathcal{R}(A')) > 0$ (i.e., we cannot shrink A without an essential loss of the realm). For instance, it is sufficient that $\omega(x) = A$ for a set of x of positive measure.

An attractor A is called *global* if $\omega(x) \subset A$ for almost all $x \in M$.

Exercise 19.22. *For any smooth endomorphism $f : M \to M$ of a compact manifold, there is a unique global measure-theoretic attractor.*

An attractor A is called *minimal* if it does not contain any smaller attractors.

Exercise 19.23. (i) *A measure-theoretic attractor A is minimal iff $\omega(x) = A$ for almost all $x \in \mathcal{R}(A)$.*

(ii) *The global attractor A is minimal iff it is a unique attractor of f. In this case, $\omega(x) = A$ for almost all $x \in M$.*

A *topological attractor* is defined by replacing the conditions of having positive Lebesgue measure in the above definitions with conditions being topologically essential.

Exercise 19.24. (i) *Develop a little theory for topological attractors along the lines of the above measure-theoretic discussion.*

(ii) *Let A be a compact invariant subset with non-empty interior. If $f|\, A$ is topologically transitive then A is a minimal topological attractor.*

We define the *basin of attraction* $\mathcal{D}(A)$ as the *interior* of the realm. In the simplest case of an attracting cycle $\boldsymbol{\alpha}$, the basin and the realm coincide. However, for the map $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x + x^2$, with fixed point 0, we have: $\mathcal{R}(0) = [-1, 0]$, while $\mathcal{D}(0) = (-1, 0)$. (It is an example of a *parabolic* point: they will be important players in what follows.)

In this book we will encounter interesting examples of attractors with an empty basin.

Let us finish with a useful related notion. Let X be a topological space that can be exhausted with compact subsets,[24] and let $f : X \to X$ be a continuous map.

[24]For time being, the reader can think of *locally compact* spaces, but eventually we will need more general ones (that would include the space of "quadratic-like germs").

A pre-compact subset $K \Subset X$ is called *absorbing* if for any compact subset $Y \Subset X$, there exists an N such that $f^n(Y) \subset K$ for all $n \geq N$.

This notion can also be applied to *partially defined* maps f. Then $f^n(Y)$ should be understood as $f^n(Y \cap \operatorname{Dom} f^n)$.

19.8. Invertible one-dimensional maps.

19.8.1. *Local topological classification.* Let us consider a local homeomorphism $f : (\mathbb{R}, 0) \to (\mathbb{R}, 0)$ (a *germ*) with an isolated fixed point at the origin. Let $I^+ = [0, \varepsilon]$ and $I^- = [-\varepsilon, 0]$ be one-sided closed neighborhoods of 0 that do not contain any other fixed points.

Assume first that f is orientation preserving (in which case the fixed point is also called *orientation preserving).* Then two different scenarios can occur in I^+:

$$0 < f(x) < x \quad \text{or} \quad f(x) > x \quad \forall\, x \in I^+.$$

In the former case, 0 (and the germ f) are called (*topologically*) *attracting on the right*, while in the latter case, they are called (*topologically*) *repelling on the right.*

EXERCISE 19.25. (i) *If* 0 *is topologically attracting on the right then*

$$f(I^+) \Subset I^+ \text{ and } f^n x \to 0 \text{ as } n \to \infty \text{ for any } x \in I^+.$$

(ii) *In the topologically repelling case, $f(I^+) \Supset I^+$ and all orbits escape: For any $x \in I^+$ there exists the escape moment $n = n(x) \in \mathbb{Z}_+$ such that $f^n x \notin I^+$.*

(iii) *All maps as above that are attracting* (*respectively: repelling*) *on the right are topologically conjugate on their I^+−half-neighborhoods.*[25]

A similar discussion can obviously be carried on the left-hand side. If 0 is (*topologically*) *attracting/repelling* on both sides then it (and the germ f) are naturally called (*topologically*) *attracting/repelling* . Otherwise, we say that 0 (and the germ f) are of *mixed type.*

COROLLARY 19.26. *Any two orientation preserving topologically attracting* (*resp., topologically repelling*) *local homeomorphisms as above are topologically conjugate near the origin* (*by an orientation preserving homeomorphism*). *And so are all the maps of mixed type* (*perhaps, by an orientation reversing conjugacy*).

If f is orientation reversing then 0 (and the germ itself) are called a *flip.* The dynamics in this case can be analysed by taking the 2nd iterate of f:

EXERCISE 19.27. *Assume* 0 *is a flip and there are no periodic points of period 2 near it. Then there is a neighborhood $I \ni 0$ such that either $f(I) \Subset I$ or $f(I) \Supset I$. In the former case,* 0 *is called* (*topologically*) *attracting: $f^n x \to 0$ for all $x \in I$. In the latter case, it is called* (*topologically*) *repelling: all orbits escape I. Any two attracting* (*resp., repelling*) *flips are topologically conjugate.*

Notice that in the attracting flip case, the components $I^\pm$ of $I \setminus \{0\}$ form a *cycle of intervals of period* 2:

$$f(I^+) \subset I^- \quad \text{and} \quad f(I^-) \subset I^+.$$

[25]With the understanding that in the repelling case the equivariance relation is satsfied on $f^{-1}(I^+)$.

If f is smooth then it can be locally written as $f(x) = \rho x + o(|x|)$, where ρ is the multiplier of 0. If $\rho \neq 0$, then f is a local diffeomorphism; moreover, it is a flip iff $\rho < 0$.

According to the previous discussion, the following germs represent all possible topological types (where any one in the row can be selected in each case):

• *Topologically attracting orrientation preserving germs*:

$$x \mapsto \rho x, \quad 0 < \rho < 1, \quad \text{or} \quad x \mapsto x - x^{2l+1}, \quad l \in \mathbb{Z}_+;$$

• *Topologically repelling orrientation preserving germs*:

$$x \mapsto \rho x, \quad \rho > 1, \quad \text{or} \quad x \mapsto x + x^{2l+1}, \quad l \in \mathbb{Z}_+;$$

• *Mixed germs*: $x \mapsto x - x^{2l}$, $l \in \mathbb{Z}_+$ (attracting on the right);

• *Topologically attracting flips*:

$$x \mapsto \rho x, \quad -1 < \rho < 0, \quad \text{or} \quad x \mapsto -x + x^{k+1}, \quad k \in \mathbb{Z}_+;$$

• *Topologically repelling flips*:

$$x \mapsto \rho x, \quad \rho < -1, \quad \text{or} \quad x \mapsto -x - x^{k+1}, \quad k \in \mathbb{Z}_+.$$

Note that for smooth germs the fixed point 0 (and the germ f) is called *attracting* if $|\rho| < 1$, *repelling* if $|\rho| > 1$, and *parabolic* if $\rho \in \{\pm 1\}$ (compare §21.1). That is why we used "topologically" in the above definitions.

Exercise 19.28. (i) *For a smooth attracting germ $f : x \mapsto \rho x + o(x^{1+\delta})$, $0 < |\rho| < 1$, $\delta > 0$, the orbits x_n converge to zero exponentially fast: $x_n \asymp \rho^n$.*

(ii) *For a smooth parabolic germ $x \mapsto x - x^{k+1} + o(x^{k+1+\delta})$, $x > 0$, the orbits converge to zero polynomially: $x_n \asymp n^{-1/k}$.*

As we have already mentioned in §19.6, the multiplier ρ is a smooth invariant, so the smooth classification of germs differs from the topological one. Also, as we have just seen, some parabolic germs are topologically indistinguishable from attracting or repelling ones. However, smoothly they are always different:

Exercise 19.29. (i) *Any two smooth attracting germs $f_i : x \mapsto \rho x + h.o.t.$ with the same multiplier, $0 < |\rho| < 1$, are smoothly conjugate. Similarly, for smooth repelling germs.*

(ii) *Any two smooth parabolic germs of the same topological type and the same order of degeneracy are smoothly conjugate.*

(iii) *However, no parabolic germ is smoothly conjugate to an attracting or repelling one.*

Analytic classification of attracting and parabolic germs will be discussed in §§23.1.2, 23.7.3. For quasisymmetric (qs) classification, see §21.3.4.

19.8.2. *Invertible interval maps.* These are the simplest dynamical examples:

Exercise 19.30. *Let $f : I \to I$ be a continuous monotone map of an interval. Its set of fixed points,* $\mathrm{Fix}(f)$*, is a non-empty closed set.*

(i) *If f is increasing than any orbit converges to a fixed point.*

(ii) *If f is decreasing than* $\mathrm{Fix}(f)$ *is a singleton,* $\mathrm{Fix}(f) = \{\alpha\}$*, and any orbit either converges to α or it converges to a cycle of period* 2.

Here is *zigzag* pictures illustrating the above types of behavior:

19.8.3. *Circle rotations.* Let us consider a circle rotation

$$\mathrm{R}_\theta : \mathbb{T} \to \mathbb{T}, \quad z \mapsto e(\theta) z$$

by angle $\theta \in \mathbb{R}/\mathbb{Z}$. Note that in the angular coordinate $\alpha \in \mathbb{R}/\mathbb{Z}$ it becomes the translation $\alpha \mapsto \alpha + \theta \mod 1$. If the rotation number is rational, $\theta = \mathfrak{p}/\mathfrak{q}$, then the dynamics is non-interesting as $\mathrm{R}_\theta^q = \mathrm{id}$. Otherwise it exhibits several interesting features:

Exercise 19.31. *For an irrational w, the rotation R_θ is minimal and uniquely ergodic (with the Lebesgue measure m on $\mathbb{T}$ being the only invariant measure).*

Thus, for any $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$, the fractional parts $\{n\,\theta\}$ are *dense* in $[0,1]$ (*Kronecker Theorem*) and are equidistributed over there (*Weyl Equidistribution Theorem*), i.e., for any interval $I \subset [0,1]$, we have:

$$\frac{1}{N} \#\{n \in [1,N] : \ \{n\,\theta\} \in I\} \to |I| \quad \text{as } n \to \infty.$$

19.8.4. *Circle homeomorphisms.* Let S^1 be an oriented circle, and let $f : S^1 \to S^1$ be an orientation preserving circle homeomorphism. Assume it has a periodic cycle $\boldsymbol{\alpha} = (f^n \alpha)_{n=0}^{\mathfrak{q}-1}$ of period $\mathfrak{q}$. It is naturally cyclically ordered. Since f preserves the cyclic order on S^1, it does so on $\boldsymbol{\alpha}$, and hence $f|\,\boldsymbol{\alpha}$ has a well defined rotation $\mathfrak{q}/\mathfrak{p}$.

Exercise 19.32. *Under the above circumstances, any periodic point of f has period $\mathfrak{q}$ and rotation number $\mathfrak{p}/\mathfrak{q}$.*

In this situation, $\mathfrak{p}/\mathfrak{q}$ is called the *rotation number* of f.

Exercise 19.33. *For any component I of $S^1 \smallsetminus \mathrm{Per}(f)$, $f^{\mathfrak{q}}|\,I$ is an interval homeomorphism such that for any $x \in I$ we have: $f^{\mathfrak{q}n} x \to a_+$ as $n \to +\infty$ and $f^{\mathfrak{q}n} x \to a_-$ as $n \to -\infty$, where $a_\pm$ are appropriately labeled endpoints of I (with the labeling independent of x).*

Note that it can happen that $a_+ = a_-$. Then it is a fixed point of f that attracts all orbits in both forward and backward time. For instance, consider the translation $T : x \mapsto x + 1$ on $\hat{\mathbb{R}} \equiv \mathbb{R} \cup \{\infty\} \approx S^1$.

Project 19.34. *If a homeomorphism $f : S^1 \to S^1$ does not have periodic points. then it is monotonically semi-conjugate to an irrational circle rotation R_θ, $\theta \in \mathbb{R} \smallsetminus \mathbb{Q}$. Moreover,*

(i) *The preimage of any point under the conjugacy $h : S^1 \to \mathbb{R}/\mathbb{Z}$ is either a singleton or a wandering interval.*

(ii) *For any $x \in \mathbb{R}$, $\theta = \lim_{n\to+\infty} \frac{1}{n} F^n x \mod 1$, where $F : \mathbb{R} \to \mathbb{R}$ is a lift of f to the universal covering.*

Under the above circumstances, θ is called the *rotation number* of f.

Exercise 19.35. *Any orientation preserving circle homeomorphism keeping the Leb measure invariant is a rotation.*

19.9. Expanding maps.

19.9.1. *Continuous setting.* A continuous map $f : X \to X$ of a metric space is called *globally uniformly expanding* if there exists a factor $\lambda > 1$ such that

$$d(fx, fy) \geq \lambda\, d(x, y) \text{ for any two points } x, y \in X.$$

There are two natural weaker notions: f is called globally *strictly* expanding if $d(fx, fy) > d(x, y)$ for any two distinct points $x, y \in X$, and it is called globally *weakly* expanding if $d(fx, fy) \geq d(x, y)$ for any two points $x, y \in X$.

EXERCISE 19.36. *There are no globally strictly expanding maps on non-singleton compact spaces.*

So, to be useful for compact spaces this notion should be relaxed. A continuous map $f : X \to X$ is called *locally* uniformly expanding if there exists a factor $\lambda > 1$ and[26] $\varepsilon > 0$ such that

$$d(fx, fy) \geq \lambda\, d(x, y) \text{ for any } x, y \in X \text{ with } d(x, y) < \varepsilon. \tag{19.3}$$

The corresponding notions of *strict* and *weak* local expanding properties are self-explanatory.

EXERCISE 19.37. *Any locally strictly expanding map on an infinite compact space is non-injective.*

Throughout the book, *"expanding" will mean "locally expanding"*, unless otherwise is explicitly assumed.

Obviously, the iterates f^n inherit any of these local expanding properties. Moreover, in the uniform case, the expanding factor grows exponentially under the iterates. This suggests a dynamically more natural notion: A continuous map $f : X \to X$ of a compact metric space is called *dynamically* uniformly expanding if it has a uniformly expanding iterate f^n. Equivalently, there is a factor $\lambda > 1$, a constant $C > 0$ and a sequence of scales $\varepsilon_n \to 0$ such that

$$d(f^n x, f^n y) \geq C\, \lambda^n d(x, y) \quad \forall\, x, y \in X \text{ with } d(x, y) < \varepsilon_n.$$

Note that this notion of is invariant under bi-Lipschitz changes of the metric.

Let us also mention the following purely topological notion. A map $f : X \to X$ is called ε-*expansive* if there exists and $\varepsilon > 0$ such that for any two points $x, y \in X$ there exists a moment $n \in \mathbb{N}$ such that $d(f^n x, f^n y) > \varepsilon$. It is called *expansive* if it is ε-expansive for some $\varepsilon > 0$.[27]

Parabolic maps will provide interesting examples of expansive maps which are not locally uniformly expanding (see e.g., Corollary 26.8 below).

19.9.2. *Smooth setting.* In the smooth category, there is a natural infinitesimal version of the above discussion. Let $f : M \to M$ be a smooth endomorphism of a Riemannian manifold of dimension ≥ 1, and let $K \subset M$ be a closed invariant subset. The map f is called *infinitesimally strictly expanding* on K if

$$\|Df(x)v\| > \|v\| \text{ for any } x \in K \text{ and any non-vanishing tangent vector } v \in T_x M.$$

(Compare with Corollary 7.11 of the Schwarz Lemma.) It is called infinitesimally *uniformly* expanding on K if there exist a factor $\lambda > 1$ such that

$$\|Df(x)v\| \geq \lambda \|v\| \text{ for any } x \in K \text{ and any tangent vector } v \in T_x M.$$

[26]If K is non-compact then ε (end the ε_n below) shooud depend on x.

[27]For invertible maps, it is more meaningful to consider *two-sided* orbits in this definition.

Of course, if K is compact then these two notions are equivalent. The corresponding *weak* notion is

$$\|Df(x)v\| \geq \|v\| \text{ for any } x \in K \text{ and any tangent vector } v \in T_xM.$$

EXERCISE 19.38. *Check the following relations:*

(i) *The infinitesimal uniform expanding property is equivalent to the local uniform expanding property.*

(ii) *The infinitesimal strict expanding property implies the local strict expanding property, but not the other way around.*

(iii) *In case of $K = M$, the infinitesimal weal expanding property is equivalent to the local weak expanding property.*

Let us now pass to a dynamical version of this discussion. A map f as above is called *dynamically* infinitesimally uniformly expanding on K if there exist a factor $\lambda > 1$ and a constant $C > 0$ such that for any $x \in K$ and any tangent vector $v \in T_xM$ we have:

$$\|Df^n(x)v\| \geq C\lambda^n\|v\|, \quad n = 0, 1, \dots. \tag{19.4}$$

This notion is independent of the choice of a Riemannian metric.

EXERCISE 19.39. (i) *Dynamical infinitesimal and local uniform expanding properties are equivalent.*

(ii) *Under these circumstances, there exists a Riemannian metric ρ on M and a factor $\lambda_1 > 1$ such that*

$$\|Df(x)v\|_\rho \geq \lambda_1\|v\| \quad \forall\ x \in K,\ v \in T_xM.$$

Such a metric is called *Lyapunov*.

Assume now that $K = M$, i.e., f is expanding on the whole manifold. Being a local homeomorphism of a compact space, f is a covering, so it has some degree $d \in \mathbb{Z}^*$. Since f is not a homeomorphism, $|d| \geq 2$. Moreover, if f is orientation preserving then d is positive, and so $d \geq 2$.

In the smooth dynamical setting, "expanding" will usually mean "dynamically inifnitesimally uniformly expanding".

19.10. Bernoulli shift.

19.10.1. *Definition.* Consider the space $\Sigma \equiv \Sigma_2^+$ of one-sided sequences $\bar{i} = (i_0 i_1 \dots)$ of zeros and ones. Supply it with the weak topology (convergence in this topology means that each coordinate eventually stabilizes). We obtain a Cantor set. Define the shift $\sigma \equiv \sigma_2$ on this space as the map of forgetting the first coordinate,

$$\sigma : (i_0 i_1 \dots) \mapsto (i_1 i_2 \dots).$$

It is called the *(one-sided) Bernoulli shift* with two states.

EXERCISE 19.40. *Show that:*

- *σ is topologically exact and hence topologically transitive;*
- *Periodic points of σ are dense in Σ.*

EXERCISE 19.41. *Show that the only non-trivial automorphism of σ is induced by the relabeling $0 \leftrightarrow 1$.*

A *subshift* of the Bernoulli shift is the restriction of σ to a closed invariant subset $X \subset \Sigma$.

19.10.2. *Cylinders.* Given an n-string $\bar{j} \equiv (j_0, \dots j_{n-1})$ of zeros and ones, let

$$\Sigma^n_{\bar{j}} \equiv \Sigma_{j_0 \dots j_{n-1}} := \{\bar{i} \in \Sigma : \; i_k = j_k, \; k = 0, \dots, n-1\}$$

$$\equiv \{\bar{i} \in \Sigma : \; \sigma^k(\bar{i}) \in \Sigma_{j_k}, \; k = 0, \dots, n-1\}.$$

Such a set is called a *cylinder of rank* n. (We consider the whole space $\Sigma \equiv \Sigma^0$ as a "cylinder of rank 0".) Given a cylinder $\Sigma^n_{\bar{j}} \equiv \Sigma_{j_0 \dots j_{n-1}}$ of rank n, we have

(R1) $\Sigma^n_{\bar{j}}$ is the disjoint union of two cylinders of rank $n+1$, $\Sigma_{i_0 \dots i_{n-1} 0}$ and $\Sigma_{i_0 \dots i_{n-1} 1}$,

(R2) $\Sigma^n_{\bar{j}}$ is mapped by σ homeomorphically onto the cylinder $\Sigma_{i_1 \dots i_n}$ of rank $n-1$.

In particular, the whole space Σ is decomposed into 2^n disjoint cylinders of rank n each of which is mapped homeomorphically by σ^n onto the whole space Σ.

The space Σ is endowed with a natural *dyadic metric*:

$$d(\bar{i}, \bar{j}) = 2^{-n}, \; \text{where } n = \min\{k : \; i_k \neq j_k\}. \tag{19.5}$$

Thus, $\operatorname{diam} \Sigma^n_{\bar{j}} = 2^{-n}$ for any cylinder of rank n. With respect to this metric, the shift σ is locally expanding by a factor of 2:

$$d(fx, fy) = 2d(x, y) \quad \forall \; x, y \in \Sigma^1_i, \; i \in \{0, 1\}.$$

19.10.3. *Bernoulli measure.* Since the Bernoulli shift has plenty of periodic points, it has plenty of atomic invariant measures. In fact, it has a plenty of non-atomic measures as well. Among them there is one most classical that corresponds to the process of tossing of a fair coin: zeros and ones label tails and heads that appear independently in the tossing process. In other words, we consider a measure μ uniformly spread over the cylinders:

$$\mu(\Sigma_{i_0 \dots i_{n-1}}) = \frac{1}{2^n}.$$

By Kolmogorov's Theorem, this defines uniquely a Borel measure. It is called the *balanced Bernoulli measure* for σ. (Here "balanced" suggests that all the states have equal probabilities, while "Bernoulli" suggests that th events happening at different times are independent).

Exercise 19.42. *Show that μ is invariant and mixing, and hence ergodic.*

The conclusion of the Birkhoff Ergodic Theorem in this case coincides with the classical *Bernoulli Law of Large Numbers* for the coin tossing: *For a typical sequence $\bar{i} \in \Sigma$, the asymptotic frequency of appearance of zeros (or ones) is equal to* $1/2$.

In fact, the Bernoulli measure is "as mixing as one can get" as the future of the process is completely independent of the past.

Exercise 19.43. *Periodic points and iterated preimages of the Bernoulli shift are equidistributed with respect to the Belnoulli measure:*

$$\frac{1}{2^p} \sum_{\bar{i} \in \mathrm{Fix}(\sigma^p)} \delta_{\bar{i}} \to \nu, \quad \frac{1}{2^n} \sum_{\bar{i} \in \sigma^{-n}(\bar{j})} \delta_{\bar{i}} \to \nu, \quad \text{as } n \to \infty \quad \forall \; \bar{j} \in \Sigma.$$

Show that the same is valid for periodic points of minimal period p.

The balanced Bernoulli measure is also called the *measure of maximal entropy* for the shift (see §46.10.7).

REMARK 19.44. The above discussion of the Bernoulli shift admits a straightforward generalization to the case of one-sided Bernoulli shifts $\sigma_d : \Sigma_d^+ \to \Sigma_d^+$ with d states. The only subtlety is concerned with Exercise 19.41 which is in fact wrong for $d > 2$:

EXERCISE 19.45. *Give an example of a non-trivial automorphism of the one-sided Bernoulli shift σ_3.*

Along with the balanced Bernoulli measure, one can consider Bernoulli measures that assign different probabilities to different symbols. Namely, let $\mathbf{p} = (p_1, \dots, p_d) \in \mathbf{\Delta}^{d-1}$ be a probability distribution on d symbols. For a string $\bar{i} = (i_0 \dots i_{n-1})$, we define the probability of the corresponding cylinder in $\Sigma \equiv \Sigma_d$ as follows:

$$\mu_{\mathbf{p}}(\Sigma_{\bar{i}}^m) = p_{i_0} \dots p_{i_{n-1}}.$$

(So, the events at different moments are independent.)

EXERCISE 19.46. (i) *Show that this determines an invariant measure on Σ.*
(ii) *Show that each measure μ_p is ergodic.*
(ii) *Show that each of them is mixing.*

19.11. Coding: a general idea.

19.11.1. *Partitions and corresponding codings.* Let us consider a map $f : X \to X$. A *partition* (X_i) of X into *pieces* $X_i \subset X$ is a decomposition $X = \bigsqcup X_i$.

To any partition into d pieces corresponds a natural coding of the orbits of f by sequences $\bar{i} = (i_0, i_1, \dots)$ in d symbols by the rule

$$f^n x \in X_{i_n}, \ n = 0, 1, \dots$$

This gives the *coding map* $\pi : X \to \Sigma_d^+$, $x \mapsto \bar{i}(x)$. This map is equivariant, i.e., $\pi \circ f = \sigma \circ \pi$: indeed, replacement of x by fx results in forgetting the first symbol i_0 in the sequence $\bar{i}$.

A partition is called *open* if all the pieces X_i are such. For instance, the partition of the shift space Σ into cylinders of some rank n is open.

EXERCISE 19.47. *For an open partition, the coding map π is continuous. Thus, it gives a semi-conjugacy between f and a subshift of the Bernoulli shift σ.*

The sets

$$X_{\bar{i}}^n \equiv X_{i_0, \dots i_{n-1}} := \pi^{-1}(\Sigma_{i_0 \dots i_{n-1}})$$

will be referred as *cylinders of rank n* (for this coding).

Let us say that f is *expanding*[28] *with respect to a partition* (X_i) if there exist $\lambda > 1$ and $n \in \mathbb{Z}_+$ such that for any cylinder $X_{\bar{i}}^n$, we have:

$$\operatorname{dist}(f^n x, f^n y) \ge \lambda \operatorname{dist}(x, y), \quad \forall x, y \in X_{\bar{i}}^n.$$

EXERCISE 19.48. *Let (X_i) be an open partition of a compact space X. The corresponding semi-conjugacy π is injective iff* $\operatorname{diam} X_{\bar{i}}^n \to 0$ *as $n \to \infty$. In particular, this is the case when f is expanding with respect to the partition (X_i).*

[28]Here we actually deal with a "locally uniformly expanding" property, compare §19.9.

19.11.2. *Tilings and corresponding codings.* A *tile* Y in a space X is a closed subset such that $Y = \operatorname{cl} Y^\circ$. A *tiling* $\mathcal{X}$ or *tessellation* of X into tiles (also called "pieces") X_i is a decomposition $X = \bigcup X_i$ such that

$$X_i^\circ \cap X_j^\circ = \emptyset \text{ for } i \neq j.$$

We let $\partial\mathcal{X} := \bigcup \partial X_i$, $\operatorname{diam} \mathcal{X} := \max \operatorname{diam} X_i$.

Let us consider the following invariant set:

$$X^\circ = \{x \in X : \ f^n x \in \bigcup X_i^\circ, \ n = 0, 1, \dots\}$$

endowed with the open partition (X_i°). Note that if X is compact (or complete) and f is open, then by the Baire Category Theorem, X° is dense in X.

If our tessellation comprises d tiles then by the previous consideration, we obtain the coding map $\pi : X^\circ \to \Sigma_d^+$ semiconjugating $f|\,X^\circ$ to the restriction of the Bernoulli shift σ to an invariant subset Y° (which is not closed in general) .

We say that f is *expanding with respect to the tiling* (X_i) if $f|\,X^\circ$ is expanding with respect to the partition (X_i°).

Exercise 19.49. *Let X be compact, and let f be expanding with respect to the tiling (X_i). Then the inverse of the coding map $\pi : X^\circ \to Y^\circ$ extends to a surjection $\pi^{-1} : \operatorname{cl} Y^\circ \to X$ semiconjugating the subshift $\sigma|\operatorname{cl} Y^\circ$ to f.*

The *cylinders* $X_{\bar{i}}^n \equiv X_{i_0 \dots i_{n-1}}$ associated with a tiling (X_i) are defined as the closures of the corresponding cylinders $(X^\circ)_{\bar{i}}^n$ associated with the partition (X_i°). Equivalently, $X_{\bar{i}}^n = \pi^{-1}(\Sigma_{\bar{i}}^n)$. They form a nest of tilings $\mathcal{X}^n := (X_{\bar{i}}^n)$, with the natural inclusion and transformation rules: Any tile of rank n is tessellated by some tiles of rank $n+1$ (obtained by adding a symbol i_n on the right) and for $n > 0$ it is mapped to a tile of rank $n-1$ (obtained by erasing the first symbol i_0). [Compare with rules (R1)-(R2) in §19.10.2.]

Exercise 19.50. *Let $(\mathcal{X}^n)$ and $(\tilde{\mathcal{X}}^n)$ be shrinking nests of tessellations of compact spaces X and $\tilde{X}$, respectively. Let $h_n : X \to \tilde{X}$ be a sequence of homeomorphisms maping $\mathcal{X}^n$ to $\tilde{\mathcal{X}}^n$ such that $h_{n+1} \,|\, \partial\mathcal{X}^n = h_n \,|\, \partial\mathcal{X}^n$. Then the h_n uniformly converge to a homeomorphism $h : X \to \tilde{X}$.*

19.11.3. *Bernoulli generator.* A tiling $(X_i)_{i=1}^d$ of a compact space X is called (*unbranched*) *Bernoulli* if each piece X_i is mapped by f onto the whole space X and this map is injective on $\operatorname{int} X_i$. In particular, this definition can be applied in the case of an open partition.

Remark 19.51. In what follows we will also encounter situations of *branched* Bernoulli tilings when the maps $f : X_i \to X$ are branched coverings.

Exercise 19.52. (i) *If f is expanding with respect to an open Bernoulli partition then the corresponding coding $\pi : X \to \Sigma_d$ is a conjugacy between f and the Bernoulli shift σ_d.*

(ii) *If f is expanding with respect to a Bernoulli tiling then the corresponding coding $\pi^{-1} : \Sigma_d \to X$ is a semi-conjugacy between the Bernoulli shift σ_d and f. Moreover, it is one-to-one outside of a set of first Baire category.*

In this case, the tiling (X_i) is called a *Bernoulli generator*, and the map f itself is called *Bernoulli*.

FIGURE 19.1. The doubling map $\theta \mapsto 2\theta \mod 1$.

More generally, assume we have d tiles $X_i \subset X$ each of which is mapped by f homeomorphically onto the whole space X (but their union is not necessarily the whole space X). Let us call such a family (X_i) a *Bernoulli family of tiles.* Let

$$K = \{x \in X : \; f^n x \in \bigcup X_i, \; n = 0, 1, \dots\}. \tag{19.6}$$

It is a closed invariant subset of X.

EXERCISE 19.53. *If f is expanding with respect to a Bernoulli family of tiles then there is a semi-conjugacy $h : \Sigma_d \to K$ which is one-to-one over the set*

$$K^\circ := \{x \in K : \; f^n x \in \bigcup \operatorname{int} X_i, \; n = 0, 1, \dots\}.$$

If $X_i \subset \operatorname{int} X$ for all i then h is a conjugacy.

19.12. Doubling map. The *doubling map* is just the squaring map

$$f_0 : z \mapsto z^2$$

on the unit circle $\mathbb{T}$. Passing to the annular coordinate $T : \theta \in \mathbb{R}/\mathbb{Z}$, where $z = e(\theta)$, we obtain the map $\theta \mapsto 2\theta \mod \mathbb{Z}$, which justifies the term "doubling". We can also view it as a map $T : \theta \mapsto 2\theta \mod 1$ on the unit interval $I \equiv [0, 1]$, i.e.,

$$T(\theta) = 2\theta \text{ for } \theta \in [0, 1/2] \quad \text{and} \quad T(\theta) = 2\theta - 1 \text{ for } \theta \in [1/2, 1],$$

with understanding that the endpoints must be identified.

The doubling map has a unique fixed point $z = 1$, i.e., $\theta = 0$. The preimages of this point under T^n are dyadic rationals $\theta = p/2^n$, $p = 0, 1, \dots, 2^n - 1$. They tessellate the circle into (closed) *dyadic intervals*

$$I^n_{\bar{i}} \equiv I_{i_0 \dots i_{n-1}} = \left\{ \theta = i_0 + \frac{i_1}{2} + \dots + \frac{i_{n-1}}{2^{n-1}} + \frac{\theta_n}{2^n}, \quad \text{where } \theta_n \in [0, 1] \right\},$$

consisting of angles whose dyadic expansion begins with $\bar{i} \equiv (i_0, \dots, i_{n-1})$ (and may end with the infinite number of "1"'s, to make the interval closed). Note that

$$I^n_{i_0 \dots i_{n-1}} = I^{n+1}_{i_0 \dots i_{n-1}, 0} \cup I^{n+1}_{i_0 \dots i_{n-1}, 1}, \quad \text{and} \quad T(I^n_{i_0 \dots i_{n-1}}) = I^{n-1}_{i_1 \dots i_{n-1}}. \tag{19.7}$$

It follows that $T^n(I^n_{\bar{i}}) = \mathbb{T}$ for any dyadic interval, and this map is one-to-one, except that it glues the endpoints of $I^n_{\bar{i}}$ to $z = 1$.

EXERCISE 19.54. *Show that* $\operatorname{int} I^n_{i_0 \dots i_{n-1}} = \{\theta : \; T^k \theta \in \operatorname{int} I^1_{i_k}, \; k = 0, \dots, n-1\}$.

EXERCISE 19.55. *The map $\phi : \Sigma \to [0, 1]$ that associates to a dyadic sequence $\bar{i} = (i_0 i_1 \dots)$ the angle θ with this dyadic expansion, is a semiconjugacy between the Bernoulli shift $\sigma : \Sigma \to \Sigma$ and the doubling map.*

Note that this description of the doubling dynamics in terms of the usual dyadic expansions is just the coding generated by the Bernoulli tiling

$$[0, 1] = I^1_0 \cup I^1_1 \equiv [0, 1/2] \cup [1/2, 1].$$

In the angular coordinate, the doubling map is expanding by a factor of two: $T'(\theta) = 2$ for all $\theta \in \mathbb{R}/\mathbb{Z}$.

Exercise 19.56. *The Lebesgue measure m on $\mathbb{T}$ is invariant under T. The above semi-conjugacy $\phi : \Sigma \to \mathbb{T}$ induces an isomorphism between the Bernoulli shift σ with the Bernoulli measure μ and the doubling map T with m. Moreover, m is mixing and hence ergodic.*

Exercise 19.57. *Periodic points of the doubling map are rationals $\theta = p/q$ with odd denominator. Preperiodic points are rationals $\theta = p/q$ with even denominator (in the irreducible representation).*

Proposition 19.58 (Compare Exercise 19.41). *The doubling map does not have non-trivial orientation preserving automorphisms.*

Proof. Let $h : \mathbb{T} \to \mathbb{T}$ be an automorphism for the doubling map, i.e., a homeomorphism commuting with T.

Since $\theta = 0$ is the unique fixed point of T, it must be also fixed by h. Since $\theta = 1/2$ is the only T-preimage of 0 different from 0, it must be fixed by h as well. Hence the dyadic intervals $I^1_0 = [0, 1/2]$ and $I^1_1 = [1/2, 1]$ are either h-invariant or are permuted by h (fixing the endpoints). But in the latter case, h would be orientation reversing, so both intervals are invariant.

Each of them contains one T-preimage of $1/2$, respectively $\theta = 1/4$ and $\theta = 3/4$, so these points must also be fixed by h. Hence all the dyadic intervals $I^2_{i_0 i_1}$ of rank 2 are h-invariant (with the endpoints fixed).

Assume inductively that all the dyadic intervals $I^n_{i_0 \dots i_{n-1}}$ of rank n are h-invariant (with the endpoints fixed). Since each of them contains one dyadic point of next level, $\mathfrak{p}/2^{n+1} \in T^{-n+1}(0)$ with odd $\mathfrak{p}$, all these points must be fixed, and hence all the dyadic intervals of rank $n + 1$ are h-invariant.

We conclude by induction that all the dyadic points $\mathfrak{p}/2^n$, $n \in \mathbb{N}$, are fixed by h. By continuity, $h = \text{id}$. $\square$

The doubling map, and its quotients, will serve as the main dynamical model for quadratic polynomials on their Julia sets.

The above discussion can be readily generalized to the case of degree d circle maps $T_d : \mathbb{T}/\mathbb{Z} \to \mathbb{T}/\mathbb{Z}$, $\theta \mapsto d\theta$, $d \geq 2$, with one noteworthy adjustment:

Exercise 19.59. *The map T_d has $d - 1$ fixed points. Orientation preserving automorphisms of T_d are circle rotations by angles $1/(d - 1)$.*

One can find many doubling restrictions inside of the d-adic map:

Exercise 19.60. *Assume there exist two disjoint closed arcs I_1 and I_2 in $\mathbb{T}$ such that:*

(i) *One of the gaps between I_1 and I_2 is bounded by T_d-fixed points;*

(ii) *The map T_d is injective on each I_k;*

(iii) $T_d(I_k) \supset I_1 \cup I_2$, $k = 1, 2$.

Then the set

$$K := \{\theta : T^n_d(\theta) \in I_1 \cup I_2, \quad n = 0, 1, \dots \}$$

is a T_d-invariant Cantor set on which T_d is conjugate to T_2, and there exists a monotonically non-decreasing Devil K-Staircase map $h : \mathbb{T} \to \mathbb{T}$ semiconjugating $T_d | K$ to the circle doubling map. Moreover, h provides a natural period preserving one-to-one correspondence between cycles of period $p > 1$ for $T_d | K$ and for T_2.

19.13. Expanding circle maps. This theory contains germs of various important ideas that will be discussed throughout the book.

19.13.1. *Fixed point.* Let us consider a smooth orientation preserving expanding circle map $g : \mathbb{T} \to \mathbb{T}$. As we know (see §19.9.2), it is a covering of some degree $d \geq 2$. Following the spirit of our exposition, in what follows *we will assume that expanding circle maps under consideration have degree two*, unless otherwise is explicitly stated.

LEMMA 19.61. *Any expanding circle map $g : \mathbb{T} \to \mathbb{T}$ of degree two has a unique fixed point β.*

PROOF. Lifting g to the universal covering, we obtain an orientation preserving diffeomorphism $\tilde{g} : \mathbb{R} \to \mathbb{R}$ satisfying the following equivariance property

$$\tilde{g}(x+1) = \tilde{g}(x) + 2 \text{ (since } \deg g = 2).$$

It follows that $\tilde{g}(x) \sim 2x$ as $x \to \infty$, so by the Intermediate Value Theorem the equation $\tilde{g}(x) = x$ has a solution $\tilde{\beta}$.

It projects to a fixed point $\beta \in \mathbb{T}$ for g. Assume there is another fixed point $\alpha \in \mathbb{T}$. These two fixed points divide $\mathbb{T}$ into two intervals J_0 and J_1. Since g is expanding, none of these intervals can be invariant. Hence each of these intervals covers itself with degree ≥ 2 and covers the other interval with degree ≥ 1. Altogether, the degree of $g \,|\, \mathbb{T}$ would be at least 3. □

Conjugating g by a rotation, we can always *normalize* it so that the fixed point β is placed at 1. *Throughout the book, we will assume such a normalization unless otherwise is explicitly stated.*

19.13.2. *Dynamical tilings and conjugacy.* Let $\beta' \neq \beta$ be the second preimage of the fixed point $\beta = 1$. They divide $\mathbb{T}$ into two closed intervals $I^1_0 = [\beta, \beta']$ and $I^1_1 = [\beta', \beta]$, where the endpoints are labeled in the positive way with respect to the circle orientation. Moreover, the interiors of these intervals are mapped homeomorphically onto $\mathbb{T} \smallsetminus \{\beta\}$. Hence each of them contains a preimage of β', and we obtain a tiling of $\mathbb{T}$ into four closed intervals $I^2_{00}, I^2_{01}, I^2_{10}, I^2_{11}$ that appear in the listed order as we go around the circle in the positive direction. In turn, each of these intervals contains an order two preimage of 1, producing a tiling of $\mathbb{T}$ into 8 intervals $I^3_{i_0 i_1 i_2}$, where $i_k \in \{0, 1\}$.

Proceeding this way, we obtain a nest of *dynamical tilings* $\mathcal{I}^n = \mathcal{I}^n(g)$ of $\mathbb{T}$ into 2^n intervals $I^n_{i_0 \dots i_{n-1}} \equiv I^n_{i_0 \dots i_{n-1}}(g)$, where $i_k \in \{0, 1\}$, such that the lexicographic order on the set of dyadic sequences $\bar{i} = (i_0 \dots i_{n-1})$ corresponds to the order of the intervals on $\mathbb{T} \smallsetminus \{1\}$. They are nested and transformed by g in the same combinatorial way (19.7) as the standard dyadic tilings $(I^n_{\bar{i}})$ for the doubling map. We will call these intervals *dynamical (dyadic) intervals* or *tiles.*

Since g is expanding, these intervals shrink at an exponential rate:

$$C_0 \lambda_0^{-n} \leq |I^n_{\bar{i}}| \leq C \lambda^{-n}, \quad \text{where } \lambda \geq \lambda_0 > 1, \ C, C_0 > 0. \tag{19.8}$$

PROPOSITION 19.62. *Any expanding circle map g of degree two is topologically conjugate to the doubling map $T \equiv f_0$ (by an orientation preserving homeomorphism h). Moreover, the natural semi-congugacy $\phi : \Sigma \to \mathbb{T}$ between the Bernoulli shift σ and g corresponds, via h, to the the semi-conjugacy ϕ_0 for T described* Exercise 19.55*: $h \circ \phi = \phi_0$.*

FIGURE 19.2. Linearization of an expanding circle map.

PROOF. Let us consider the nest of dynamical tilings $\mathcal{I}^n(g)$ for g, along with the corresponding dyadic nest $\mathcal{I}^n(T)$ for the doubling map.

Let h_n be the piecewise linear homeomorphism $\mathbb{T} \to \mathbb{T}$ that maps $I^n_{\bar{i}}(g)$ linearly onto $I^n_{\bar{i}}(T)$. Since our tilings are transformed by g and T in the same combinatorial way, h_n conjugates g to T on the tilings' boundaries (which are equal to $g^{-n}(\beta)$ and $T^{-n}(\beta)$ respectively). Finally, (19.8) implies that the h_n uniformly converge to a homeomorphism $h : \mathbb{T} \to \mathbb{T}$ which conjugates g to T on the whole circle.

Moreover, the Bernoulli tiling $\mathcal{I}^0$ generates a Bernoulli model for g (see §19.11.3), which manifestly corresponds, via h, to the model for T. $\square$

The conjugacy h can be viewed as the *linearization* of the expanding map g.

COROLLARY 19.63. *Expanding circle maps of degree two do not admit non-trivial orientation preserving automorphisms.*

EXERCISE 19.64. *An expanding circle map $g : \mathbb{T} \to \mathbb{T}$ of degree d is topologically conjugate to $T_d : \theta \mapsto d\theta$ (by an orientation preserving homeomorphism). Hence it has $d-1$ fixed points and the group of orientation preserving automorphisms of g is the cyclic group of order $d-1$.*

19.13.3. *Balanced measure.* A Borel measure μ for an expanding circle map can be prescrbed by the masses $\nu(I^n_{\bar{i}})$ of the dynamical intervals (satisfying the natural compatibility conditoins). The easiest way of doing so it to assign to all dynamical intervals $I^n_{\bar{i}}$ of rank n the same mass 2^{-n}. We obtain an invariant measure ν called *balanced*. Note that it has obviously full support in Σ.

The balanced measure of the doubling map is the Leb measure m. Since the balanced property is preserved under conjugacies, we have $m = h_*(\nu)$, so

$$h(\theta) = \nu[0, \theta]. \tag{19.9}$$

(Note that $h(0) = 0$ due to our normalization of g.) Thus, *once we have constructed the balanced measure, we can recover the linearizing conjugacy h by means of* (19.9).

The following equidistribution properties motivate dynamical significance of the balanced measure:

EXERCISE 19.65. *Periodic points and iterated preimages of an expanding circle map are equidistributed with respect to the balanced measure:*

$$\frac{1}{2^p - 1} \sum_{x \in \mathrm{Fix}(g^p)} \delta_x \to \nu, \quad \frac{1}{2^n} \sum_{x \in g^{-n} y} \delta_x \to \nu, \quad \text{as } n \to \infty \quad \forall\, y \in \mathbb{T}.$$

Another motivation comes in the framework of the Entropy Theory where the balanced measure is also called the *measure of maximal entropy*.

19.13.4. *Distortion and quasisymmetry.*

EXERCISE 19.66. *The above conjugacy between T and g (and its inverse) is Hölder continuous.*

In fact, we can say more:

PROPOSITION 19.67. *For $g \in C^2$, the above conjugacy between T and g is quasisymmetric.*

The proof is based upon a very important *bounded distortion* property for the iterates of g:

LEMMA 19.68 (Distortion Lemma). *Let $g \in C^2$. Assume an iterate g^n is injective on an interval $J \subset \mathbb{T}$. Then for any $x, y \in J$ we have:*

$$\frac{|Dg^n(x)|}{|Dg^n(y)|} \leq C,$$

where C depends only on C^2-norm of f and on the expanding factor λ.

PROOF. Let us consider the function $\log |Dg(x)|$ written in the angular coordinate on $\mathbb{T}$. It is L-Lipschitz with

$$L = \max_{x \in \mathbb{T}} \frac{|D^2 g(x)|}{|Dg(x)|}$$

depending only on the C^2-norm of g.

Let $J_k = g^k(J)$. Since g is expanding and g^n is injective on J, we have:

$$|J_k| \leq L \frac{|J_n|}{\lambda^{n-k}} \leq \frac{L}{\lambda^{n-k}}.$$

Hence

$$|\log |Dg^n(x)| - \log |Dg^n(y)|| \leq \sum_{k=0}^{n-1} |\log |Dg(g^k x)| - \log |Dg(g^k y)||$$

$$\leq L \sum_{k=0}^{n-1} |J_k| \leq L \sum_{k=0}^{n-1} \frac{1}{\lambda^{n-k}} \leq \frac{L\lambda}{\lambda - 1}.$$

□

EXERCISE 19.69 (Denjoy Distortion Estimate). *Let $f_k : I_k \to I_{k+1}$ be a chain of C^2 diffeomorphisms between intervals I_k, $k = 0, 1, \dots, n$, and let $F = f_{n-1} \circ \cdots \circ f_1 \circ f_0$. Then*

$$\frac{|DF^n(x)|}{|DF^n(y)|} \leq \exp\left(L \cdot \sum_{k=0}^{n-1} |I_k|\right), \quad \text{where } L = \max_k \max_{x \in I_k} \frac{|D^2 f_k(x)|}{|Df_k(x)|}.$$

EXERCISE 19.70. *Under the circumstances of* Lemma 19.68*, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that*

$$\operatorname{dist}(g^n x, g^n y) < \delta \implies \frac{|Dg^n(x)|}{|Dg^n(y)|} \leq 1 + \varepsilon \quad \forall\, n \in \mathbb{N},\ x, y \in J.$$

LEMMA 19.71. *The nest of dynamical tilings $\mathcal{I}^n_{\bar{i}}$ corresponding to an expanding circle map $g \in C^2$ has a bounded geometry (see* §15.1.3*), with the bound depending only on the C^2-norm of g and the expanding factor λ.*

PROOF. Consider a nest of two intervals $I^n_{\bar{i}} \supset I^{n+1}_{\bar{i}j}$ and apply g^{n-1} to it. It will be homeomorphically mapped onto the nest $I^1_{i_n} \supset I^2_{i_n i_{n+1}}$. The latter intervals have a comparable length as there are only finitely many intervals of level 1 and 2. By the bounded distortion property of the last lemma, so do the initial intervals. □

Proof of Proposition 19.67. (Compare Exercise 15.3.) Take two adjacent intervals S and T on the circle of equal length. Select the smallest level n such that S contains a dynamical tile $S' = I_i^n$ of level n. Then it is contained in one or two adjacent dynamical tiles of level $n-1$. Calling the union of these tiles S'', we obtain

$$S' \subset S \subset S'' \quad \text{and } |S'| \asymp |S''|.$$

Since the nest of dynamical tilings has a bounded geometry, we can squeeze T in between two interval, $T' \subset T \subset T''$ of comparable length, each of which is a bounded unions of dynamical tiles of the same level $n + O(1)$.

Applying the conjugacy h to these intervals, we conclude that $h(S)$ and $h(T)$ are both squeezed in between two intervals each of which is a bounded union of dyadic intervals of level $n + O(1)$. It follows that the lengths of all six intervals in question are comparable.

19.13.5. *Analytic expanding circle maps.* Expanding circle maps $g : \mathbb{T} \to \mathbb{T}$ that appear in this book usually come from conformal annuli coverings (see e.g., §§25.3, 41.1):

LEMMA 19.72. *Let $V \subset V'$ be a nest of two $\mathbb{T}$-symmetric conformal annuli, and let $g : V \to V'$ be a holomorphic covering map of degree $d \geq 2$. Then the restriction $g|\,\mathbb{T}$ is an expanding circle map.*

PROOF. First, by symmetry, g preserves the unit circle. Second, since $\operatorname{mod} V' = d \operatorname{mod} V$, the annulus V is strictly contained in V'. By Corollary 7.11, $g|\,\mathbb{T}$ is expanding in the hyperbolic metric of V'. □

We can also control the geometry of $g|\,\mathbb{T}$ from the outside.

LEMMA 19.73. *Under the above circumstances, assume that $V \Subset V'$. Then the C^2-norm of $g : \mathbb{T} \to \mathbb{T}$ and the expanding factor λ are controlled by a lower bound μ on the modulus of the* (*external*) *fundamental annulus $A := \bar{V}' \smallsetminus (V \cup \mathbb{D})$.*

PROOF. An upper bound on the C^2-norm of $g|\,\mathbb{T}$ in terms of μ follows from the Koebe Distortion Theorem. A lower bound on λ follows from Corollary 7.11. □

In fact, any analytic expanding circle map admits such an external structure:

EXERCISE 19.74. *Let $g : \mathbb{T} \to \mathbb{T}$ be a real analytic expanding circle map. Then:*
(i) *There exists a nest of two $\mathbb{T}$-symmetric annuli $V \Subset V'$ with smooth boundary such that g admits a holomorphic extension to a covering map $V \to V'$.*
(ii) *The preimages $V^n := g^{-n}(V')$ are $\mathbb{T}$-symmetric annuli covering V' under g^n with degree d^n.*

In what follows, *we will assume without saying that an expanding analytic circle map $g : \mathbb{T} \to \mathbb{T}$ comes equipped with an annulus extension $g : V \to V'$ as above.*

19.13.6. *Ergodicity.* Let m be the normalized angular measure on the circle.

PROPOSITION 19.75. *Any expanding circle map $g \in C^2$ is ergodic with respect to m.*

PROOF. We will use background from Appendix 2 below.

Let us consider a completely invariant measurable set $X \subset \mathbb{T}$ of positive measure. We claim that X has positive (lower) density at any point $x \in \mathbb{T}$:

$$\underline{\operatorname{dens}}(X \mid x) \equiv \liminf_{I \ni x} \operatorname{dens}(X|\, I), \tag{19.10}$$

where $I \subset \mathbb{T}$ runs over all intervals containing x.

For any $n \in \mathbb{N}$, let $I^n \equiv I^n_{i(n)}$ be a dynamical dyadic interval containing x. (If there are two such intervals, select any of them.) Lemma 19.68 implies that the map $f^n : I^n \to \mathbb{T}$ has a bounded distortion:[29] there exists $C > 0$ (independent of n) such that

$$\frac{|Dg^n(y)|}{|Dg^n(y')|} \le C \quad \text{for any } y, y' \in I^n.$$

Since X is completely invariant, we obtain

$$\operatorname{dens}(X|\, I^n) \ge C^{-1} m(X).$$

By the bounded geometry of the nest of dynamical tilings (Lemma 19.71), the intervals I^n fill a dense set of scales around x. Property (19.10) follows.

Assume by contradiction that $\mathbb{T}$ is decomposed into two (completely) invariant measurable subsets of positive length: $\mathbb{T} = X_1 \sqcup X_2$. By the Lebesgue Density Points Theorem, there is a point $x \in X_1$ with $\operatorname{dens}(X_1|\, x) = 1$. Hence $\operatorname{dens}(X_2|\, x) = 0$, contradicting the above assertion. $\square$

We will see many further applications of the above method of transforming certain qualities from small scales to big scales by expanding dynamics. We will refer to it as *dynamical magnification* machinery.

19.13.7. *Absolutely continuous invariant measure* (*acim*).

THEOREM 19.76. *Any expanding circle map $g \in C^2$ has a unique acim $\rho\, dm$, where ρ is a positive continuous function. Moreover, If g is real analytic then so is the density ρ.*

PROOF. Let us push forward the angular measure under the iterates of g, $(g^n)_*(dm) = \rho_n\, dm$. By formula (19.2),

$$\rho_n(y) = \sum_{x \in g^{-n} y} \frac{1}{Dg^n(x)} = \sum_{i=1}^{d^n} Dg_i^{-n}(y), \tag{19.11}$$

where g_i^{-n} stand for the local inverse branches of g^{-n} near y, and the Dg^n, Dg_i^{-n} mean the derivatives calculated in the angular coordinate (since these derivatives are positive, we do not need to take the absolute values). By Lemma 19.68, these branches have bounded distortion, implying that the densities ρ_n have a uniformly bounded oscillation: there exists $C > 0$ such that

$$\frac{\rho_n(y)}{\rho_n(y')} \le C \quad \text{for all } y, y' \in \mathbb{T}. \tag{19.12}$$

Since $\displaystyle\int_{\mathbb{T}} \rho_n\, dm = 1$, we conclude that $\dfrac{1}{C} \le \rho_n(y) \le C$ for all $y \in \mathbb{T}$. Integration over a measurable subset $Y \subset \mathbb{T}$ implies that $m(g^{-n}(Y)) \asymp m(Y)$. By Proposition 19.18, g has an acim (with a density bounded away from 0 and ∞).

In fact, by using Exercise 19.70 instead of Lemma 19.68, the oscillation bound (19.12) can be made equicontinuos: For any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\operatorname{dist}(y, y') < \delta \implies \frac{\rho_n(y)}{\rho_n(y')} < 1 + \varepsilon,$$

[29]Of course, it does not matter that this map is slightly non-invertible, as $g^n(\partial I^n) = \{1\}$.

implying that the limiting density is continuous.

Let us show uniqueness. We have just constructed an acim $\mu \sim m$. Since m is ergodic (by Proposition 19.75), so is μ, and so would be any other acim ν. But then $\nu = \mu$ by Proposition 19.11.

If g is real analytic, it admits a holomorphic extension to an annuli map $A \to A'$ as in Exercise 19.74. Moreover, the preimages $A_n := g^{-n}(A')$ are $\mathbb{T}$-symmetric annuli covering A' under g^n with degree d^n. Hence any point $y \in \mathbb{T}$ has a neighborhood $W(y)$ where g^{-n} has d^n holomorphic inverse branches g_i^{-n}. They provide us with an analytic extension of densities (19.11) to $W(y)$, and hence to a complex neighborhood W of $\mathbb{T}$. Moreover, by the Koebe Distortion Theorem, the branches g_i^{-n} have a bounded distortion, implying that the densities ρ_n are uniformly bounded in each $W(y)$. By compactness of the circle, they are uniformly bounded in the whole neighborhood W.

Hence the average densities

$$\rho_n^{\text{ave}} := \frac{1}{n}\sum_{k=0}^{n-1} \rho_k$$

are also uniformly bounded in W. By the Little Montel Theorem, they form a normal family in W, so we can select a subsequential limit ρ. It is holomorphic in W and its restriction to $\mathbb{T}$ is the density of some invariant measure. Since the acim is unique, the conclusion follows. □

19.13.8. *Deformation space.* The *deformation space* of (degree 2) analytic expanding circle maps is the space $\mathcal{E} \equiv \mathcal{E}_2$ modulo analytic conjugacy. The term relates it to Teichmüller spaces of Reimann surfaces or to deformation spaces of Kleinian groups. Indeed. Proposiotion 19.67 asserts that all expanding circle maps $g \in \mathcal{E}$ are obtained by a qs deformation of a single map, e.g., of the square map $f_0 : z \mapsto z^2$.

Let us start with providing "best" representatives in analytic classes of $\mathcal{E}$:

Proposition 19.77. *Let $g : \mathbb{T} \to \mathbb{T}$ be a analytic expanding circle map. Then g is anlytically conjugate to a unique (normalized) analytic expanding Leb-preserving circle map $g_\circ : \mathbb{T} \to \mathbb{T}$.*

Proof. Let ρ be the density of the acim for g. Then

$$\phi(\theta) = \int_0^\theta \rho(x)\, dm = m[0,\theta],$$

is a real analytic circle diffeomorshism such that $\phi_*(\rho\, dm) = dm$. The $g_\circ := \phi \circ g \circ \phi^{-1}$ is a desired expanding map.

If there are two maps, $g_\circ$ and $\tilde{g}_\circ$, as above, then they are conjugate by an orientation preserving diffeomorphism for which m is invariant. By Exercice 19.35, it must be a rotation. □

Remark 19.78. The above Proposition is also valid in the C^r-category, $r \geq 2$, in place of the analytic one. However, the proof requires honest estimates of the derivatives of the densities ρ_n. In the analytic setting, they are not needed due to the Cauchy Estimate (Montel Theorem). It is an illustration of how analyticity makes our life easier.

The next assertion gives a sharp dichotomy for the deformation regularity:

PROPOSITION 19.79. *Let g and $\tilde g$ be two expanding analytic circle maps, and let h be the orientation preserving conjugacy between them. Then h is either analytic or singular with respect to the Lebesgue measure. Moreover, if $g \neq \tilde g$ and both maps are normalized and Leb-preserving, then h is singular.*

PROOF. As above, let $\rho\, dm$ and $\tilde\rho\, dm$ be the acim's for g and $\tilde g$ respectively. Since $\rho\, dm$ is g-ergodic, $h_*(\rho\, dm)$ is $\tilde g$-ergodic. Hence it is eather absolutely continuous or totally singular. In the latter case, the map h is singular. In the former case, $h_*(\rho\, dm)$ is an acim for $\tilde g$. By uniqueness of such a measure, $h_*(\rho\, dm) = \tilde\rho\, dm$.

Let ϕ and $\tilde\phi$ be the analytic diffeomorphisms from Proposition 19.77 that respectively bring g and $\tilde g$ to the normalized Leb-preserving models. Then the homeomorphism $\tilde\phi \circ h \circ \phi^{-1}$ is Leb-preserving as well, and hence is a rotation (see Exercise 19.35); in fact, it is the identity as $g_\circ$ and $\tilde g_\circ$ are normalized. The conclusions follow. □

Thus, *the deformation space $\mathcal{E}$ can be realized as the space of normalized Leb-preserving expanding maps $g \in \mathcal{E}$.*

As we have already mentioned (see §19.5), multipliers of periodic points remain invariant under analytic (actually, smooth) conjugacies. It terns out, that together they determine the analytic class:

THEOREM 19.80. *Two analytic expanding circle maps $g, \tilde g \in \mathcal{E}$ are analytically conjugate if and only if they have the same multipliers.*[30]

PROOF. Let us consider the dynamical tilings $\mathcal{T}^n$ and $\tilde{\mathcal{T}}^n$ for our maps. Take any two corresponding tiles $I^n_{\bar i}$ and $\tilde I^n_{\bar i}$. Since $g^n(\operatorname{int} I^n_{\bar i})$ homeomorphically covers the whole puncured circle $\mathbb{T} \smallsetminus \{\beta\}$, which contains $\operatorname{int} I^n_{\bar i}$, our tile $I^n_{\bar i}$ contains a periodic point $\alpha^n_{\bar i}$ of period n (not necessarily the smallest one). Then $\tilde I^n_{\bar i}$ contains the corresponding periodic point $\tilde\alpha^n_{\bar i}$. Let $\rho^n_{\bar i}$ and $\tilde\rho^n_{\bar i}$ be the derivatives of f^n and $\tilde f^n$ (respectively) at these points, which are equal to the appropriate powers of their multipliers. By assumption, $\rho^n_{\bar i} = \tilde\rho^n_{\bar i}$. Then the Distortion Lemma (19.68) implies $I^n_{\bar i} \asymp \tilde I^n_{\bar i}$. It follows that the conjugacy h between g and $\tilde g$ is bi-Lipschitz, and hence absolutely continuous. By Lemma 19.79, it is real analytic. □

From this point of view, *the multipliers form a complete space of coordinates ("moduli") for the space $\mathcal{E}$.*

19.13.9. *Saw-like maps.* Let $\mathcal{I} := [-1, 1]$, $\mathcal{I}^- := [-1, 0]$, $\mathcal{I}^+ := [0, 1]$. The map

$$\Lambda : \mathcal{I} \to \mathcal{I}, \quad g_0 : x \mapsto 2|x| - 1. \tag{19.13}$$

is called *(symmetric linear) saw-like map.* It is similar to the doubling map. Its non-linear (symmetric) relatives are defined as even maps $g : \mathcal{I} \to \mathcal{I}$ such that:

- Both branches $\Lambda_\pm = (\Lambda : \mathcal{I}^\pm \to \mathcal{I})$ are diffeomorphisms; one of them is oriention preserving while the other is orientation reversing (for definiteness, we assume that $f\,|\,\mathcal{I}^+$ is orientation preserving.
- Λ is expanding: $|D\Lambda_\pm(x)| \geq \lambda > 1$ for all $x \in \mathcal{I}^\pm$ (one can also understand it in the dynamical sense).

We refer to such maps as *(symmetric non-linear) saw-like maps.*[31]

[30] Here the periodic points of g and $\tilde g$ are rnaturally related by the conjugacy.

[31] Later on we will encounter more general maps

Project 19.81. *Develop a theory of saw-like maps that parallels the above theory of expanding circle maps.*

19.13.10. *Gauss map.* In many dynamical applications, one needs to generalize the above discussion to expanding maps with *infinitely many* branches. The classical Gauss map serves as a good prototype for this situation.

The map $G : (0,1] \to [0,1]$, $x \mapsto \{1/x\}$ (where $\{y\}$ stands for the fractional part of y) is called the *Gauss map.* In terms of continued fraction expansions,

$$x = \frac{1}{s_1 + \frac{1}{s_2 + \dots}} \equiv [s_1, s_2, \dots], \quad s_n \in \mathbb{N},$$

the Gauss map acts a the shift: $G(x) = [s_2, s_3, \dots]$.

Remark 19.82. Here we formally associate to 0 the "conttinued fraction" [0] and let the shift act on a single-digit fraction as $\sigma[s] = [0]$.

Exercise 19.83. *The Gauss map has the following properties:*

(i) *It is expanding:* $\exists\, \lambda > 1$ *such that* $(G^2)'(x) \geq \lambda$ *for all* $x \in (0,1]$.

(ii) *It has bounded distortion:*

$$\exists\, C > 0 \text{ such that } \forall\, n \in \mathbb{N} \text{ and } \forall\, x, y \in \left[\frac{1}{n+1}, \frac{1}{n}\right] \text{ we have : } \frac{G'(x)}{G'(y)} \leq C.$$

(iii) *The measure* $d\mu = \dfrac{1}{\log 2} \dfrac{dx}{1+x}$ *is invariant under* G.

(iv) *The measure* μ *is ergodic.*

Exercise 19.84. *Show that for a.e.* $x = [s_1, s_2, \dots, s_n] \in (0,1]$,

(i) $\dfrac{s_1 + s_2 + \dots + s_n}{n} \to \infty \quad \text{as } n \to \infty.$

(ii) *There exists a finite limit*

$$\lim_{n\to\infty} \sqrt[n]{s_1 \cdot s_2 \cdot \ldots s_n}.$$

Calculate it.

The Gauss map has many further number-theoretic and dynamical applications.

19.13.11. *Dynamical Cantor sets.* Let us start with a simple observation that the standard 1/3-Cantor K set can be dynamically generated by the tripling map

$$T_3 : [0,1] \to [0,1], \quad T_3 : x \mapsto 3x \mod 1$$

Let $I_0 = [0, 1/3]$, $I_1 = [2/3, 1]$.

Exercise 19.85. *Show that:*

(i) $K = \{x \in [0,1] : \ T_3^n x \in I_0 \cup I_1, \quad n = 0, 1, \dots\}$;

(ii) $T_3 \,|\, K$ *is naturally topologically conjugate to the Bernolli shift* σ_2*;*

(iii) $\text{length}(K) = 0$.

More generally, we can consider non-linear expaning maps of the same topological type as the tripling map. Such a map is defined on the union of two disjoint closed intervals $I_k \subset [0,1]$, $f : I_0 \cup I_1 \to [0,1]$, so that each branch $f : I_k \to [0,1]$ is an orientation preserving diffeomorphism with $f'(x) \geq \lambda > 1$. Let

$$K \equiv K_f := \{x \in [0,1] : \ f^n x \in I_0 \cup I_1, \quad n = 0, 1, \dots\}.$$

EXERCISE 19.86. *Under the above circumstances, show that:*
(i) *K is a Cantor set;*
(ii) *$f\,|\,K$ is topologically conjugate to the Bernoulli shift σ_2;*
(iii) *State and prove a bounded distortion property for K;*
(iv) *Show that K is porous* (*uniformly in all scales*) [see Appendix 2 below], *and hence* $\text{length}(K) = 0$.

One can consider a more general situation when we have d disjoint intervals I_k inside $[0,1]$, and a map $f : \bigcup I_k \to [0,1]$ such that each branch $f : I_k \to [0,1]$ is an expanding diffeomorphism (not necessarily orientation preserving). The generalization of the above Exercises to this situation is straightforward.

Such Cantor sets are called *dynamical*.

19.14. Markov shifts and maps.

19.14.1. *Markov shifts.* This section is based on the Perron-Frobenius Theory summarized in the Appendix 3 below, §19.19.

Recall that Σ_d^+ stands for the full space of one-sided sequences of symbols $1,\dots,d$. Let us consider a square matrix $A = (a_{ij})_{i,j=1}^d$ of zeros and ones, called a *transit Markov matrix*. Let

$$\Sigma_A^+ = \{\bar{i} \in \Sigma_d^+ : \ A_{i_n i_{n+1}} = 1, \ n = 0, 1, \dots\}.$$

In other words, we consider the space of all one-sided paths in the oriented graph Γ_A corresponding to A. This space is obviously closed in Σ_d^+ and invariant under the shift σ. The corresponding subshift $\sigma \equiv \sigma_A : \Sigma_A^+ \to \Sigma_A^+$ is called the *(Topological) Markov shift* or *subshifts of finite type.*

EXERCISE 19.87. *The Markov shift σ_A is topologically transitive iff the transit matrix A is irreducible. In this case:*
(i) *Periodic points are dense in Σ_A and $|\mathrm{Fix}\,\sigma_A^n| \sim r^n$, where $r = r(A)$ is the spectral radius of A;*
(ii) *Iterated preimages of any point $x \in \Sigma_A$ are dense in Σ_A, and $|f^{-n}(x)| \asymp r^n$;*
(ii) *σ_A is topologically exact (or, equivalently: topologically mixing) iff A is primitive.*

Using the Spectral Decomposition for Markov graphs, we conclude

COROLLARY 19.88. *Let A be irreducible. Then the space Σ_A^+ is decomposed into finitely many closed subsets,*

$$\Sigma_A^+ = \bigsqcup_{i=1}^{p} Y_i$$

cyclically permuted by σ and such that $\sigma^p : Y_i \to Y_i$ is topologically exact.

19.14.2. *Markov maps.* A *tiling*[32] $X = \displaystyle\bigcup_{i=1}^{d} X_i$ is called (*unbranched*) *Markov* if

(M1) the map f is injective on each X_i, and $f(X_i^\circ)$ is open;
(M2) *Markov property*: $f(X_i^\circ) \cap X_j^\circ \neq \emptyset \implies f(X_i) \supset X_j$.

[32] also referred to as a *partition*

To such a tiling, we can associate the following Markov matrix $A = (a_{ij})_{i,j=1}^d$:

(19.14) $$a_{ij} = 1 \text{ if } f(X_i) \supset X_j, \quad a_{ij} = 0 \text{ otherwise.}$$

Then
$$f(X_i) = \bigcup_{j:a_{ij}=1} X_j.$$

Exercise 19.89. *Let f be an expanding map with respect to a Markov tiling (X_i). Then the inverse coding map* $\mathrm{cl}\, Y^\circ \to X^\circ$ *from* Exercise 19.49 *has the property that* $\mathrm{cl}\, Y^\circ = \Sigma_A^+$.

Such maps f are called *Markov.* They are also called *Topological Markov Chains* (inthe Russian literature) and *subshifts of finite type* (in the Western one).

Note that Bernoulli maps introduced in §19.11.3 are exactly Markov maps with the strictly positive matrix A, i.e., $a_{ij} = 1$ for all i, j.

A Markov map is called *irreducible/primitive* if the corresponding Markov matrix A is such.

Here is an interesting class of examples:

Exercise 19.90. (i) *Let $f : I \to I$ be a piecewise monotone map of the interval, and let $I = \bigcup_{k=1}^{d} I_k$ be the tiling of I into monotonicity intervals. If the set of boundary points $\bigcup \partial I_k$ is forward invariant, then f is a Markov map.*
(ii) *Assume that f is smooth and expanding on each I_k, i.e., there is a $\lambda > 1$ such that $|f'(x)| \geq \lambda$ for any $x \in \bigcup I_k^\circ$. Then there exists a natural semi-conjugacy between the corresponding Markov shift σ_A and f. This semi-conjugacy is one-to-one except countably many points (iterated preimages of boundary points ∂I_k) over which it is two-to-one.*
(iii) *In particular, consider saw-like maps: continuous expanding maps Λ as above with two intervals of monotonicity, I_1 and I_2, on which f has opposite orientation. Such a maps is Markov iff its turning point (the common point of I_1 and I_2) is periodic or preperiodic. Show that in the periodic case, the associated matrix A is irreducible.*

Exercise 19.91. *Consider the (orientatin reversing) "anti-doubling" map of the circle*
$$T \equiv T_{-2} : \mathbb{T} \to \mathbb{T}, \quad \theta \mapsto -2\theta \mod 1.$$
Using the fixed points of T, semi-conjugate T to the Markov shift σ_A with 3×3 matrix $A = (1 - \delta_{ij})$.

More generally, we will often encounter the following situation. A family of tiles $(X_i)_{i=1}^d$ is called *Markov* if it satisfies properties (M1)–(M2). Let us consider the maximal closed invariant subset $K \subset \bigcup X_i$ (described as in (19.6)). As above (19.14), we can naturally associate to it a Markov matrix A. Similarly to the Bernoulli coding of Exercise 19.53, we now obtain a Markov coding of our dynamics:

Exercise 19.92. *If f is expanding with respect to a Markov family of tiles then there is a semi-conjugacy $h : \Sigma_A^+ \to K$ which is one-to-one over the set*
$$K^\circ := \{x \in K : \ f^n x \in \bigcup \mathrm{int}\, X_i, \ n = 0, 1, \dots\}.$$
If $f(X_i) \Supset X_j$ as long as $a_{ij} = 1$, then h is a conjugacy.

19.14.3. *Balanced Markov measures.* The Bernoulli measure for the full shift has two natural Markov versions. Let us say that a measure μ on $\Sigma_A \equiv \Sigma_A^+$ is *balanced* if any two admissible cylinders of the same rank have comparable measures:

$$\mu(\Sigma^n_{\bar{i}}) \asymp \mu(\Sigma^n_{\bar{j}}).$$

As there are $\asymp r^n$ of cylinders of rank n, where $r = r(A)$ is the spectral radius of A, for a balanced measure μ we have:

$$\mu(\Sigma^n_{\bar{i}}) \asymp r^{-n}$$

So, all balanced measures belong to the same measure class (with bounded mutual Radon-Nikodim derivatives). Note also, that any of them has full support in Σ_A.

Here is a particularly nice situation. If the shift $\sigma : \Sigma_A \to \Sigma_A$ has a a.e.-constant Jacobian $\mathrm{Jac}_m \sigma$ with respect to some quasi-invariant measure m then this Jacobian must be equal to r, and

$$m(\Sigma^n_{i_0 \dots i_{n-1}}) = r^{-n} m(\Sigma_{i_0}), \tag{19.15}$$

so such a measure is balanced.

Proposition 19.93. *Let $\sigma : \Sigma_A \to \Sigma_A$ be an irreducible (one-sided) Markov shift. Then in the balanced measure class, there is a unique measure with constant Jacobian.*

Proof. Let $m_i := m(\Sigma_i)$. As for any $\bar{i} = (i_0 \dots i_{n-1})$ we have

$$m(\Sigma_{\bar{i}}) = \bigsqcup_{a_{ik}=1} m(\Sigma_{\bar{i}\,k}), \quad \text{where } i \equiv i_{n-1},$$

the Kolmogorov compatibility condition for the measure (19.15) with constant Jacobian amounts to

$$m_i = \frac{1}{r} \sum a_{ik} m_k,$$

saying that $\mathbf{m} := (m_1 \dots m_n)$ is an r-eigenvector for A. By the Perron-Frobenius Theorem, such an eigenvector $\mathbf{m}$ exists and is unique in the probability simplex $\mathbf{\Delta}$, and moreover $\mathbf{m} > 0$. This proves the existence and uniqueness of the measure with constant Jacobian. □

Exercise 19.94. *The balanced measure class is ergodic.*

19.14.4. *Markov processes.* Let us consider a finite space $\mathcal{V} = \{1, \dots, d\}$. (Think of it as the space of *states* of some physical system or a *net* of sites.) Informally speaking, a *Markov process* is a random process on this space whose future depends only on the current moment, but not on the past history. Such a process is determined by an *initial distribution* $q = (q_1 \dots q_d)$ of states and a $d \times d$ *transit matrix* $P = (p_{ij})$ of conditional probabilities of passing from a state i to j. (Note that the vector q belongs to the probability symplex $\mathbf{\Delta}$, while the matrix P is stochastic.) This data detemines a probability measure μ on the space Σ_d^+ which gives masses

$$\mu[i_0 i_1 \dots i_{n-1} i_n] = q_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n}$$

to the cylinders. Indeed, the stochastic property of P ensures the compatibility condition for these masses:

$$\mu[i_0 i_1 \dots i_{n-1}] = \sum_j \mu[i_0 i_1 \dots . i_{n-1}\, j]$$

If μ is *shift invariant*, it is called the *stationary* measure for the Markov process. It amounts to the *dual* compatibility condition

$$\mu[i_0 i_1 \dots i_{n-1}] = \sum_j \mu[j\, i_0 i_1 \dots . i_{n-1}] \Longleftrightarrow \sum_j q_j p_{ji} = q_i \quad \forall\, i \equiv i_0.$$

meaning that q *is an invariant row for* P (or, in other words, an invariant column for the transposed matrix P^*). Now the Perron-Frobenius Theorem immediately implies:

PROPOSITION 19.95. *For any irreducible stochastic matrix P, there exists a unique stationary Markov process (i.e., a unique shift invariant measure) with the transit matrix P).*

19.14.5. *Invariant balanced measure.*

THEOREM 19.96. *Let $\sigma : \Sigma_A \to \Sigma_A$ be an irreducible (one-sided) Markov shift. Then in the balanced measure class, there is a unique invariant measure.*

This measure is naturally called the *balanced invariant measure*. Later on, it will be interpreted as the *measure of maximal entropy* (see §46.10.7).

PROOF. Let

$$p_{ij} = \frac{a_{ij} u_j}{r u_i}$$

be the stochastic matrix conjugate to A/r, where $r = r(A)$ (see Lemma 19.124), and let $(q_i) \in \mathbf{\Delta}$ be its invariant row. By Proposition 19.95, the Markov process with the initial distribution (q_i) and the transition probabilities (p_{ij}) provides us with an invariant measure μ for σ_A. This measure is balanced as for any admissible cylinder $[i_0 \dots i_n]$ we have

$$\mu[i_0 \dots i_n] = q_{i_0} \prod_{k=0}^{n-1} \frac{u_{i_{k+1}}}{r u_{i_k}} = \frac{1}{r^n} \frac{q_{i_0} u_{i_n}}{u_{i_0}} \asymp \frac{1}{r^n}.$$

Uniqueness follows from the ergodicity of the balanced measure class (Exercise 19.94) and Proposition 19.11. $\square$

19.14.6. *Equidistribution of periodic points.*

EXERCISE 19.97. *For an irreducible topoloigical Markov map $\sigma \equiv \sigma_A$, we have*
(i)

$$|\mathrm{Fix}(\sigma^p)| = \operatorname{tr} A^p \sim r^p \asymp \|A^p\|,$$

where $r \equiv r(A)$ is the spectral radius of A, and $\|\cdot\|$ is any norm in the space of matrices.
(ii) *For any $\varepsilon > 0$ there exists an N such that any orbit $(f^k x)_{k=0}^{n-1}$ of an arbitrary length $n \in \mathbb{Z}_+$ can be ε-shadowed by a piece $(f^k \alpha)_{k=0}^{n-1}$ of a periodic orbit of period $p \le n + N$:*

$$d(\sigma^k x, \sigma^k \alpha) < \varepsilon, \quad k = 0, \dots, n-1, \quad f^p \alpha = \alpha.$$

The last property is called the *Shadowing property*.

As in the expanding circle case (see Exercise 19.65), we have the following equidistribution propertty:

THEOREM 19.98. *For an irreducible topoloigical Markov map $\sigma \equiv \sigma_A$, periodic points are equidistributed with respect to the invariant Markov measure:*

$$\frac{1}{|\mathrm{Fix}(\sigma^p)|} \sum_{\alpha \in \mathrm{Fix}(\sigma^p)} \delta_\alpha \sim \frac{1}{r^p} \sum_{\alpha \in \mathrm{Fix}(\sigma^p)} \delta_\alpha \to \mu \quad \text{as } p \to \infty.$$

PROOF. Take any limit ν of the measures in the left-hand side. It is an invariant measure in the balanced measure class. The latter follows from the property that for some N any cylinder of rank n contains at least one but at most a bounded number of periodic points of period $\le n + N$ (compare with the above Shadowing property). The conclusion follows. □

EXERCISE 19.99. *State and prove a similar equidistribution result for the iterated preimages.*

19.14.7. *Markov interval maps: model with constant slope.* Using the Markov coding, the above results can be now transferred to expanding interval Markov maps (as in Exercise 19.90):

EXERCISE 19.100. *Let $f : I \to I$ be an expanding interval Markov map with and irreducible transition matrix.*

(i) *Define a quasi-invariant balanced measure for f; show that all of them belong to the same measure class with full support.*

(ii) *Show that there exists a unique balanced measure m with constant Jacobian.*

(iii) *Show that there exists a unique invariant balanced measure μ.*

(iv) *Show that periodic points are equidistributed with respect to μ.*

These results have a very interesting application. Let us say that a piecewise monotone interval map has a *constant slope* $\lambda > 1$ if it is affine on each interval I_k with slope whose absolute value is equal to λ.

EXERCISE 19.101. (i) *Let f be an expanding Markov interval map with an irreducible matrix A. Then f is topologically conjugate to a map with constnt slope $\lambda > 1$.*

(ii) *Let $l(f)$ denote the number of intervals of monotonicity of f. Then*

$$\lambda = \lim_{n\to\infty} \sqrt[n]{l(f^n)} = \lim_{n\to\infty} \sqrt[n]{\#\mathrm{Fix}(f^n)}.$$

19.15. Nielsen map associated to a Fuchsian group. In this section we will outline a classical construction of Markov circle maps associated to Fuchsian groups. This allows one to reduce many aspects of the Teichmüller Theory to the Dynamics of a single map.

Similarly to the interval case (see Exercise 19.90), being "Markov" in the circle case amounts to having an invariant finite set of pounts $X \subset \mathbb{T}$. Let us say that a map g is *orbit equivalent* to a group Γ if grand orbits of g coincide with orbits of Γ.

Let us start with Schottky groups. Consider four disjoint closed intervals, I, I', J, J', on the cirlce $\mathbb{T}$ such that the pairs I & I' and J & J' are linked. Let $\gamma : (\mathbb{D}, \partial I) \to (\mathbb{D}, \partial I')$ be a hyperbolic Möbius automorphism of $\mathbb{D}$ that stretches I to $\mathbb{T} \smallsetminus I'$, and let $\delta : (\mathbb{D}, \partial J) \to (\mathbb{D}, \partial J')$ be a similar automorphism for J. The group $\Gamma \subset \mathrm{PSL}(2, \mathbb{R})$ generated by γ and δ is an example of a *Schottky group.*

EXERCISE 19.102. (i) *The complement of the four hyperbolic half-planes based upon the intervals* I, I', J, J' *is a fundamental domain for* Γ.

(ii) Γ *is a free Fuchsian groups of a second kind.*

(iii) *The quotient* $\mathbb{D}/\Gamma$ *is a torus with a hole.*

(iv) *Define and generalize the above results to Schottky Fuchsian groups with* n *generators.*

Let us now define a map $f : I \cup I' \cup J \cup J' \to \mathbb{T}$ as follows:

$$f \,|\, I = \gamma, \;\; f \,|\, I' = \gamma^{-1}, \quad f \,|\, J = \delta, \;\; f \,|\, J' = \delta^{-1}.$$

EXERCISE 19.103. (i) f *is an expanding Markov map.*

(ii) *The corresponding dynamical Cantor set* Λ *coincides with the limit set of* Γ.

(iii) f *is orbit equivalent to* Γ.

Let us now consider the *modular group* $\Gamma = \mathrm{PSL}(2, \mathbb{Z})$ acting on the closed upper half-plane (see §2.4.12).

Define the following circle map $g : \hat{\mathbb{R}}^* \to \hat{\mathbb{R}}^*$:

$$g(x) = \begin{cases} \delta(x) = -1/x & \text{if } 0 < |x| < 1, \\ \gamma^{-1}(x) = x - 1 & \text{if } x \geq 1, \\ \gamma(x) = x + 1 & \text{if } x \leq 1. \end{cases} \tag{19.16}$$

Let T be the first return of g to the interval $[0, 1)$.

EXERCISE 19.104. *Show that*

(i) g *is orbit equivalent to* Γ, *and*

(ii) T *is the second iterate of the Gauss map;*

(iii) *Conclude that two points,* $x = [a_1 a_2 \dots]$ *and* $y = [b_1 b_2 \dots]$,[33] *in the interval* $[0, 1)$ *belong to the same* Γ*-orbit iff there exist* $k, l \in \mathbb{N}$ *such that*

$$a_{n+2l} = b_{n+2k}, \quad n = 0, 1, \dots.$$

Let us now conside groups uniformizing closed Riemann surfaces.

PROPOSITION 19.105. *Let* Γ *be a Fuchsian group acting on the disk* $\mathbb{D}$ *with a compact quotient* $S = \mathbb{D}/\Gamma$. *Then* Γ *is orbit equivalent to a (discontinuous) piecewise Möbius expanding Markov map* $g \equiv g_\Gamma : \mathbb{T} \to \mathbb{T}$ *whose Möbius branches are generators of* Γ.

Such a maps is called a *Nielsen map* associated with Γ.

EXERCISE 19.106. *Consider the congruence subgroup* Γ_2 from §2.4.13. *The vertices of the fundamebtal triangle* Δ *tessellate* $\mathbb{T}$ *into three intervals* I_k. *Define a "modular map"* $f : \mathbb{T} \to \mathbb{T}$ *by letting* $f|\, I_k$ *be the reflection in the corresponding side* γ_k *of* Δ. *Then:*

(i) f *is an orienation reversing double covering of* $\mathbb{T}$*;*

(ii) f *admits a natural semi-conjugacy with the Markov shift* σ_A from Exercise 19.91*;*

(iii) f *is conjugate to the anti-doubling map* T_{-2} *(the conjugacy is related to the "Minkowski question mark function").*

19.16. Inverse limits and Natural extensions.

[33]Recall that this notation stands for the continued fraction expansions.

19.16.1. *General construction.* Let us consider a sequence of metrizable topological spaces X_{-n} and surjective continuous maps

$$X_0 \underset{f_0}{\leftarrow} X_{-1} \underset{f_{-1}}{\leftarrow} X_{-2} \underset{f_{-2}}{\leftarrow} \dots$$

The *inverse limit* of this sequence, $\varprojlim f_n$, is the space

$$\hat{X} \equiv \varprojlim (f_{-n} : X_{-n-1} \to X_{-n}) =$$

$$= \{\hat{x} = (x_0, x_{-1}, x_{-2}, \dots) : \quad x_{-n} \in X_{-n},\ f_n(x_{-n-1}) = x_{-n}\}$$

endowed with the weak topology. It is naturally projected to all the spaces involved:

$$\pi_{-n} : \hat{X} \to X_{-n}, \quad \hat{x} \mapsto x_{-n}. \tag{19.17}$$

By general topology, $\hat{X}$ is metrizable. Moreover, if all the X_{-n} are compact then so is $\hat{X}$.

Exercise 19.107. (i) *If all the X_{-n} are finite spaces* (*with discrete topology*) *then $\hat{X}$ is a Cantor set.*

(ii) *If all the X_{-n} are topological groups, while the f_{-n} are continuous group endomorphisms endomorphisms, then $\hat{X}$ is endowed with a natural toplogical group structure.*

19.16.2. *Adding machine.* Let u snow consider a sequence $(p_n)_{n=0}^{\infty}$ of increasing natural numbers $p_n \in \mathbb{N}$, $p_n \geq 2$, such that p_{n+1} is a multiple of p_n. Let $q_0 = p_0$, and $q_n = p_n/p_{n-1}$ (for $n \geq 1$) be the corresponding *relative periods*; denote this sequence $\mathbf{q} = (q_n)_{n+0}^{\infty}$.

Letting $\mathbb{Z}/p_n\mathbb{Z} \to \mathbb{Z}/p_{n+1}\mathbb{Z}$ be the natural homomorphisms, we obtain the inverse limit

$$\mathbb{Z}_{\mathbf{q}} := \varprojlim \mathbb{Z}/p_n\mathbb{Z},$$

By Exercise 19.107, it is a compact metrizable group, and in fact, a ring with unit $e = (1, 1, \dots)$. It is called the **q**-*adic ring.* Translation by this unit $\tau_{\mathbf{q}} : \mathbb{Z}_{\mathbf{q}} \to \mathbb{Z}_{\mathbf{q}}$, $x \mapsto x + e$, is called the (**q**-*adic*) *adding machine* or *odometer.*

Exercise 19.108. *The adding machine is minimal and uniquely ergodic. Its only invariant measure is the Haar measure on $\mathbb{Z}_{\mathbf{q}}$, which is equal to the limit of the homogeneous measures on the cyclic groups $\mathbb{Z}_{p_n}$.*

In the *stationary* case, $\mathbf{q} = (q, q, q, \dots)$, the ring $\mathbb{Z}_{\mathbf{q}} \equiv \mathbb{Z}_q$ and the machine are called *q-adic.* In particular, in case $q = 2$ they are called *dyadic.*

19.16.3. *Natural extension.* This is a general useful construction that "makes" any map invertible.

Let us assume that in the inverse limit definition, all the spaces X_{-n} and all the maps f_{-n} are the same, $X_{-n} = X$ and $f_{-n} = f$ for $n \in \mathbb{N}$, so we are given just one surjective continuous map $f : X \to X$. Then $\hat{X}$ becomes the space of its *backward orbits*

$$\hat{X} = \{\hat{x} = (x \equiv x_0, x_{-1}, x_{-2}, \dots) : \quad x_{-n} \in X,\ f(x_{-n-1}) = x_{-n}\}.$$

A nice feature of this special case is that f lifts to a homeomorphism $\hat{f} : \hat{X} \to \hat{X}$,

$$\hat{f}(\hat{x}) = (fx_{-n})_{n=0}^{\infty} = (fx_0, x_0, x_1, \dots),$$

whose inverse is the "shift" given by forgetting the first coordinate: $\hat{f}^{-1}(\hat{x}) = (x_{-n})_{n=1}^{\infty}$. Moreover, the projection $\pi \equiv \pi_0 : \hat{X} \to X$ (19.17) is equivariant: $\pi \circ \hat{f} = f \circ \pi$. This lift is called the *natural extension* of f.

EXERCISE 19.109. (i) *In case when the original map f is a homeomorphism, π is also a homeomorphism conjugating $\hat{f}$ and f.*
(ii) *There is a natural one-to-one equivariant correspondence between periodic points of f and $\hat{f}$.*

EXERCISE 19.110. *There is a natural one-to-one correspondence between invariant measure for f and $\hat{f}$.*

Along with the one-sided Bernoulli shift $\sigma : \Sigma_2^+ \to \Sigma_2^+$ introduced in §19.10, we can consider the *two sided Bernoulli shift* $\hat{\sigma} : \Sigma_2 \to \Sigma_2$, where Σ_2 is the space of two-sided dyadic sequences $(i_n)_{n=-\infty}^{\infty}$ and $(\hat{\sigma})_n = i_{n+1}$ (so, a sequence is shifted by one to the left). It is a homeomorphism of Σ_2. More generally, we can consider the *two-sided Bernoully shifts $\hat{\sigma}_d$ in d symbols* and *two-sided Markov shifts $\hat{\sigma}_A : \Sigma_A \to \Sigma_A$* (see §19.14.1).

EXERCISE 19.111. (i) *These two-sided shifts are the natural extensions of the corresponding one-sided shifts.*
(ii) *Specify the construction from* Exercise 19.110 *for the shifts σ_A and σ_A^+.*

19.16.4. *Baker transformormation, horseshoe, solenoid.* Let us introduce several illuminating examples.

Let us consider the unit box $\mathbb{B} = \mathbb{I} \times \mathbb{I}$. Let us tile it into two vertical and two horizontal rectangles (respectively):

$$\Pi_0^{\text{ver}} := \{(x,y) \in \mathbb{B} :\ 0 \le x \le 1/2\}; \quad \Pi_1^{\text{ver}} : \{(x,y) \in \mathbb{B} :\ 1/2 \le x \le 1\};$$

and

$$\Pi_0^{\text{hor}} := \{(x,y) \in \mathbb{B} :\ 0 \le y \le 1/2\}; \quad \Pi_1^{\text{hor}} : \{(x,y) \in \mathbb{B} :\ 1/2 \le y \le 1\};$$

Let $\hat{T} : \mathbb{B} \to \mathbb{B}$ be the picewise affine map that affinely maps each vertical rectangle Π_k^{ver} onto the corresponding horizontal rectangle Π_k^{hor}, $k \in \{0,1\}$ (for instance it acts as $(x,y) \mapsto (2x, y/2)$ on Π_0^{ver}. It is called the *baker transformation*.

EXERCISE 19.112. (i) *The projection $\pi : (x,y) \mapsto x$ semi-conjugates $\hat{T}$ to the doubling map T turning $\hat{T}$ into the natural extension of T.*
(ii) *There is a natural semi-conjugacy between the two-sided Bernoulli shift $\sigma : \Sigma_2 \to \Sigma_2$ and $\hat{T}$. It is generically one-to-one (over all points except those whose orbits land on the vertical boundary of $\mathbb{B}$).*
(iii) *Vertical sections $L_x^{\text{ver}} := \{(x,y) :\ y \in \mathbb{I}\}$ of $\mathbb{B}$ are specified by the "future" itinerary $(i_0 i_1 \dots)$. Moreover,*

$$\operatorname{dist}(f^n x, f^n y) = \frac{1}{2^n} \operatorname{dist}(x,y) \quad \text{as } n \to +\infty.$$

For this reason, the L_x^{ver} are called "local stable manifolds" $W^s_{\text{loc}}(x)$.
(iv) *Similarly, horizontal sections L_y^{hor} of $\mathbb{B}$ are specified by the "past" itinerary $(i_{-1} i_{-2} \dots)$. Moreover,*

$$\operatorname{dist}(f^n z, f^n z) = \frac{1}{2^n} \operatorname{dist}(z,z') \quad \text{as } n \to -\infty \quad \text{for any } z, z' \in L_x.$$

They are called "local unstable manifolds" $W^u_{\mathrm{loc}}(x)$.

(vi) *Global unstable manifolds are defined as*

$$W^s(z) := \{z' : \ \mathrm{dist}(f^n z', f^n z) \to 0 \quad \text{as } n \to +\infty\}$$

(and then $\mathrm{dist}(f^n z, f^n z') \le C\dfrac{1}{2^n} \quad \text{as } n \to +\infty \quad \text{for some } C = C(z)$.*) Show that* $W^s(z) = \bigcup L_{x'}$, *where the union is taken over all* $x' \in \mathbb{I}$ *with the same "tale" as* x, *i.e.,*

$$(i'_n i'_{n+1} \dots) = (i_n i_{n+1} \dots) \quad \text{for some } n = n(z').$$

(In other words, x' *belongs to the grand fiber of* x.*)*

(vii) *Formulate the similar statement for "global unstable manifolds"* $W^u(z)$.

(viii) $\hat{T}$ *preserve the Lebesgue measure on* $\mathbb{B}$, *which corrresponds to the Bernoulli measure for* σ.

(ix) *There is a natural one-to-one correspondence between invariant measures for the doubling map and the baker transformation.*

19.17. Appendix 1: Baire category. One says that a subset $Y \subset X$ is of *first Baire category* if it is a countable union of nowhere dense subsets. According to the *Baire Category Theorem*, if X is complete, then the complement $X \smallsetminus Y$ of a set of first category is everywhere dense.

In the *topological* setting (on a complete space X), subsets $Y \subset X$ of first category are viewed to be *negligible* Accordingly, all other sets are considered to be *topologically essential*. If a subset $Y \subset X$ has a topologically negligible complement, we say that it has *full category* or it is *residual*.

Some property depending on a point $x \in X$ is called *generic* if it is satisfied on a subset of full category. (In this case, we also say that this property is satisfied *generically*, or for a *generic point* x).

A subset $Y \subset X$ has *type* G_δ if it is a countable intersection of open sets. Complementary sets are countable unions of closed subsets; such sets have *type* F_σ. Notice that any neglectable set $Y \subset X$ is contained in a neglectable set Y' of type F_σ. Dually, any residual set Y contains some residual subset Y' of type G_δ.

Exercise 19.113. *Give an example of a set* $X \subset [0,1]$ *such that* X *has zero measure but full category* (*or the other way around*).

19.18. Appendix 2: Lebesgue Density Points. Given two measureable sets, X and D in $\mathbb{R}^2$ with $m(D) > 0$, we let

$$\mathrm{dens}(X \mid D) := \frac{m(X \cap D)}{m(D)}.$$

For a point $a \in \mathbb{R}^2$, the *upper* and *lower density of* X *at* a are defined as

$$\overline{\mathrm{dens}}(X \mid a) := \limsup_{r \to 0} \mathrm{dens}(X \mid \mathbb{D}(a,r)), \quad \underline{\mathrm{dens}}(X \mid a) := \liminf_{r \to 0} \mathrm{dens}(X \mid \mathbb{D}(a,r)).$$

The *density at* a is defined as

$$\mathrm{dens}(X \mid a) := \lim_{r \to 0} \mathrm{dens}(X \mid \mathbb{D}(a,r))$$

if the limit exists. A point a is called a *Lebesgue density point of* X if $\mathrm{dens}(X \mid a) = 1$. According to the *Lebesgue Density Points Theorem*, *almost all* $a \in X$ *are density points*.

In the above definition, one can replace round disks with domains of bounded shape filling a dense set of scales (and it is important for dynamical applications):

EXERCISE 19.114. *Let $C > 1$. A point a is a density point for a measurable set $X \subset \mathbb{R}^2$ iff there is a nest of domains $D_n \ni a$ with the following properties:*
- *They shrink:* $\operatorname{diam} D_n \to 0$*;*
- *They have a C-bounded shape around a;*
- *They fill a C-dense set of scales in the sense that* $\operatorname{diam} D_{n+1} \geq C^{-1} \operatorname{diam} D_n$*;*
- $\operatorname{dens}(X \mid D_n) \to 1$ *as $n \to \infty$.*

We say that X is *porous* at $z \in X$ if there is a $\kappa \in (0,1)$ and a sequence of radii $r_n \to 0$ such that each disk $\mathbb{D}(z, \rho_n)$ contains a disk $\mathbb{D}(\zeta, \kappa\rho_n) \subset \mathbb{D}(z, \rho) \smallsetminus X$. Informally speaking, X *has definite gaps in arbitrary small scales.*

By the Lebesgue Density Points Theorem, *porous sets have zero area.*

Note that all nowhere dense compact sets are porous in moderate scales:

EXERCISE 19.115. *Let X be a nowhere dense compact subset of $\mathbb{R}^2$. Then for any $\varepsilon > 0$ there is a $\kappa = \kappa(\varepsilon) \in (0,1)$ such that any disk $\mathbb{D}(z,r)$ with $z \in X$ and $r \geq \varepsilon$ contains a gap $\mathbb{D}(\zeta, \kappa r) \subset \mathbb{D}(z,r) \smallsetminus J$.*

Let us say that a set $X \subset \mathbb{C}$ is *porous at $z \in X$ in all scales* if there is a $\kappa \in (0,1)$ such that any disk $\mathbb{D}(z,\rho)$, $\rho \in (0,1)$ contains a disk $\mathbb{D}(\zeta, \kappa\rho) \subset \mathbb{D}(z,\rho) \smallsetminus X$. Informally speaking, X *has definite gaps in all scales.*

A set J is called *uniformly porous in all scales* if the above property holds for all $z \in J$ (with the same κ).

We can also naturally define density points and porousity at z with respect to some family of domains containing x.

EXERCISE 19.116. *Let (D_n) be a shrinking nest of ovals of a bounded shape around z that fill a dense set of scales.*

(i) *If X is porous at z with respect to the nest (D_n) then X is porous at z.*

(ii) *If z is a density point of X with respect to (D_n) then z is a density point of X.*

EXERCISE 19.117. *Show that a Cantor set K in $\mathbb{R}$ or $\mathbb{C}$ with bounded geometry is uniformly porous in all scales, and hence has zero Lebesgue measure* (*in the corresponding space*).

Under maps with bounded Jac-distortion, density of sets is distorted by a bounded amount:

EXERCISE 19.118. *Let $f : D \to \Delta$ be a diffeomorphism between domains in $\mathbb{R}^2$ that has C-bounded distrotion:*
$$\frac{\operatorname{Jac} f(x)}{\operatorname{Jac} f(x')} \leq C \quad \text{for any } x, x' \in D.$$
Then for any measureable $X \subset D$ we have
$$C^{-1} \operatorname{dens}(X \mid D) \leq \operatorname{dens}(f(X) | \Delta) \leq C \operatorname{dens}(X \mid D).$$

Of course, all of the above admits a straightforward generalization to subsets of any Euclidean space $\mathbb{R}^n$, and in particular, of the real line $\mathbb{R}$. In the latter case, we can also talk about *one-sided densities at a*, $\operatorname{dens}^{\pm}(X \mid a)$ (also: upper and lower ones) defined in terms of intervals $(a, a + \varepsilon)$ and $(a - \varepsilon, a)$ respectively.

EXERCISE 19.119. *For a measurable set $X \subset \mathbb{R}$, a point $x \in X$ is a density point iff it is a one-sided density point (on both sides). Equivalently, it is a density point with respect to the family of all intervals $J \ni x$.*

19.19. Appendix 3: Perron-Frobenius Theory. *Perron-Frobenius Theory* studies spectral properties of non-negative matrices A (or, more generally: operators in linear vector spaces preserving some cone). These properties depend largely on the distribution of positive elements in A which can be described in terms of an associated graph Γ_A. So, let us start with a bit of graph theory.

19.19.1. *Spectral decomposition for graphs.* Let $\mathcal{V} = \{v_i\}_{i=1}^d$ be a finite set, and let $\Gamma \equiv \Gamma \,|\, \mathcal{V}$ be an oriented graph with vertices v_i. For a subset $\mathcal{W} \subset \mathcal{V}$, we let $\Gamma \,|\, \mathcal{W}$ be the *restriction* of Γ to $\mathcal{W}$.

Let us introduce the following relation on $\mathcal{V}$: $v_i \succeq v_j$ if there is an oriented path from v_i to v_j. It is transitive but not necessarily reflexive. A vertex v_i is called *recurrent* if $v_i \succeq v_i$, i.e., there exists an oriented loop based at v_i. Let $\mathcal{R}$ be the set of recurrent vertices. Restricted to $\mathcal{R}$, our relation becomes reflexive. Moreover, the non-recurrent vertices can be re-labeled so that $v_j \succ v_i \implies j > i$, $v_i, v_j \in \mathcal{V} \smallsetminus \mathcal{R}$.

A graph is called *recurrent* if all its vertices are recurrent, i.e., $\mathcal{V} = \mathcal{R}$. For a recurrent graph, the relation $\succeq$ is reflexive.

However, this relation is usually not anti-symmetric. So, for a recurrent graph let us consider the equivalence relation

$$v_i \sim v_j \quad \text{if } v_i \succeq v_j \text{ and } v_j \succeq v_i, \quad v_i, v_j \in \Gamma.$$

and the corresponding equivalence classes $\mathcal{E}_k$, $k = 1, \dots, l$. The restrictions $\mathcal{E}_k \equiv \Gamma \,|\, \mathcal{E}_k$ are called *irreducible components* of Γ. The relation $\succeq$ becomes a partial order $\succ$ on the space of irreducible components.

A recurrent graph Γ is called *irreducible* if it has a single irreducible component (i.e., $\mathcal{E}_1 = \mathcal{V}$). It is called *asymptotically irreducible* if it has a unique maximal irreducible component $\mathcal{E}_l$, i.e., for any vertex v_i, there is a path leading from v_i to $\mathcal{E}_l$.

Two extreme examples of irreducible graphs $\Gamma \,|\, \mathcal{E}$ are provided by a *permutation graph* generated by a cyclic permutation $\sigma : \mathcal{E} \to \mathcal{E}$ (i.e., there is an arrow from v_i to v_j iff $v_j = \sigma(v_j)$), and the *full graph* for which all (ordered) pairs of vertices are connected. We will see momentarily that any irreducible graph can be understood in terms of these two extreme cases.

For $n \in \mathbb{Z}_+$, let Γ^n be the graph with the same vertices as Γ whose arrows are (oriented) paths in Γ of length n. The graph Γ is called *primitive* if Γ^n is full for some $n \in \mathbb{Z}^+$.

EXERCISE 19.120 (Spectral Decomposition for graphs). *Any irreducible graph $\Gamma \equiv \Gamma \,|\, \mathcal{E}$ admits a decomposition into primitive components: $\mathcal{E} = \bigsqcup_{i=0}^{p} \mathcal{P}_k$, where the arrows from $\mathcal{P}_k$ go only to $\mathcal{P}_{k+1}$, $k \in \mathbb{Z}/p\mathbb{Z}$, and the blocks $\Gamma \,|\, \mathcal{P}_k$ are primitive.*

The number p of the primitive components above is called the *period* of Γ. We will also refer to the primitive components as *basic sets.*

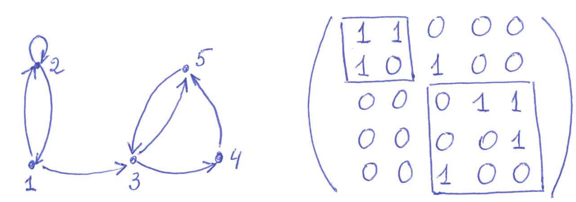

FIGURE 19.3. An irredicible Markov matrix of period 2 and its graph.

19.19.2. *Spectral Decomposition for non-negative matrices.* Let us consider a $d \times d$ matrix $A = (a_{ij})_{i,j=1}^{d}$. It is is called *non-negative* ($A \geq 0$) if $a_{ij} \geq 0$, and it is called *positive* ($A > 0$) if $a_{ij} > 0$ for all i, j. (A similar terminolody is used for vectors $u \in \mathbb{R}^d$.)

Let us asscociate to such a matrix an oriented graph $\Gamma \equiv \Gamma_A$ supported on the set $\mathcal{V} = \{v_i\}_{i=1}^{d}$ of d vertices such that v_i is connected to v_j by an arrow iff $a_{ij} > 0$. Notice that the graph Γ^n corresponds to the matrix A^n.

A non-negative matrix A is called *recurrent/irreducible/primitive* if the corresponding graph Γ_A is such. Notice that A is primitive iff $A^N > 0$ for some $N \in \mathbb{Z}_+$.

Let us label the irreducible components $\mathcal{E}_k$ so that $\mathcal{E}_k \succ \mathcal{E}_l \implies k > l$. The corresponding spaces $E_k \subset \mathbb{R}^d$ generate an *invariant flag*, i.e., the sums

$$E^k := \bigoplus_{j=0}^{k} E_j$$

are invariant under A. The operators A_k induced on the factor-spaces $E^k/E^{k-1} \approx E_k$ (where $E^0 \equiv \{0\}$) have non-negative irreducible matrices (in the bases coming from the original basis of $\mathbb{R}^d$) with graphs $\Gamma \,|\, \mathcal{E}_k$; they are called *irreducible components* of A. This brings the matrix A to a *block-triangular form*, with blocks A_k, which is called the *decomposition of A into irreducible components*.

Further decomposition of an irreducible component $\mathcal{E}_k$ into the cycles of primitive compontents $\mathcal{P}_{kl}$ leads to decomposing the corresponding space E_k into direct sum $\bigoplus_{l=0}^{p_k} P_{kl}$ so that $A_k : P_{kl} \to P_{k,l+1}$, and the operators $A_k^p : P_{kl} \to P_{kl}$, $l \in \mathbb{Z}/p_k\mathbb{Z}$, are primitive. This brings the irreducible block A_k to a *block-cyclic form*. It is called the *Spectral Decomposition* of a recurrent non-negative matrix A.

19.19.3. *Boundary spectrum.* Recall that the set of eigenvalues λ_i of A is called its *spectrum*, $\operatorname{spec} A$, and the maximum of the $|\lambda_i|$ is called the *spectral radius* $r(A)$. The set of eigenvalues λ_i on the cirle $|\lambda| = r(A)$ is called the *boundary spectrum* $\operatorname{spec}^b(A)$.

EXERCISE 19.121. *We have:* $\operatorname{spec} A \smallsetminus \operatorname{spec}(A \,|\, \mathcal{R}) \subset \{0\}$, *so*

$$\operatorname{spec}^b A = \operatorname{spec}^b(A \,|\, \mathcal{R}), \quad r(A) = r(A \,|\, \mathcal{R}).$$

This reduces the study of the boundary spectrum to the recurrent case.

The decomposition of a recurrent matrix $A \geq 0$ into irreducible components A_k brings it to the block-triangular form, implying:

$$\operatorname{spec} A = \bigcup \operatorname{spec} A_k, \quad r(A) = \max r(A_k).$$

To understand spectral properties of an irreducible matrix, let us start with the (main) primitive case:

PERRON-FROBENIUS THEOREM (Primitive case). *For any primitive matrix $A \geq 0$, we have:*

(i) *The spectral radius $r \equiv r(A)$ is a simple eigenvalue.*

(ii) *The corresponding eigenvector $u \in \mathbb{R}^d$ can be selected positive.*

(iii) *There are no other eigenvalues in the boundary spectrum.*

(iv) *There exists $\rho \in (0, r)$ such that for any $w \in \mathbb{R}^n$,*

$$A^n w = c(w) r^n u + O(\rho^n)$$

(with a uniform constant in the error term).

PROOF. The primitive case is immediately reduced to the positive one, so we can assume that $A > 0$ in the first place. Let us consider the probabilistic symplex

$$\mathbf{\Delta} := \{x = (x_i) \in \mathbb{R}^d : \ x_i \geq 0, \ \phi(x) = 1\}, \quad \text{where } \phi(x) = \sum x_i.$$

Since $A > 0$, we have a well defined continuous transformation

$$g : \mathbf{\Delta} \to \operatorname{int} \mathbf{\Delta}, \quad x \mapsto \frac{Ax}{\phi(Ax)}.$$

Let $\mathcal{L}$ be the family of straight intervals $L \subset \mathbf{\Delta}$ with $\partial L \subset \partial \mathbf{\Delta}$. Let us endow each L with the hyperbolic metric d_L. For any two points $x, y \in \mathbf{\Delta}$, we let $\delta(x, y) = d_L(c, y)$, where L is the the interval of our family $\mathcal{L}$ passing through x and y. Notice that $\delta(x, y)$ is comparable with $\|x - y\|$ on any compact subset of $\operatorname{int} \mathbf{\Delta}$ (in particular, on the image $g(\mathbf{\Delta})$).

Any $L \in \mathcal{L}$ is projectively mapped by g to the interior of some $L' \in \mathcal{L}$. Moreover, $g(L)$ has a uniformly bounded hyperbolic diameter in L'. By the 1D Schwarz Lemma, g is uniformly contracting from the hyperbolic metric of L to the hyperbolic metric of L'. Hence $\delta(g^n x, g^n y) \to 0$ uniformly exponentially, implying that

$$\|g^n x - g^n y\| \to 0 \text{ uniformly exponentially for } x, y \in \mathbf{\Delta}. \tag{19.18}$$

Hence g has a unique fixed point $u \in \operatorname{int} \mathbf{\Delta}$, and $g^n x \to u$ uniformly exponentially for $x \in \mathbf{\Delta}$. This fixed point is a positive eigenvector $u > 0$ corresponding to a positive eigenvalue $r > 0$.

Let us normalize A so that $r = 1$, (i.e., let us replace A with A/r). Then (19.18) translates to the property $A^n \to P$ exponentially fast, where P is the projection onto the line through u with $\operatorname{Ker} P = \{x \in \mathbb{R}^n : \ A^n x \to 0\}$. It follows that the spectral radius r_0 of $A \,|\, \operatorname{Ker} P$ is < 1. Moreover, $r_0 = \max\{|\lambda| : \ \lambda \in \operatorname{spec} A \smallsetminus \{r\}\}$.

Now all the conclusions follow (with any $\rho > r_0$ and $c(w)u = Pw$). □

The difference $r - r_0$ in the above proof is called the *spectral gap*: it controls the rate of convergence of the (normalized) A^n to the projection P.

Perron-Frobenius Theorem (Irreducible case). *For an irreducible matrix A of period p, we have:*

(i) *The spectral radius $r \equiv r(A)$ is a simple eigenvalue.*

(ii) *The corresponding eigenvector u can be selected positive.*

(iii) *The boundary spectrum is obtained from r by rotating it by the cyclic group*

$$\lambda \mapsto e(l/p)\,\lambda, \quad l \in \mathbb{Z}/p\mathbb{Z}.$$

(iii) *There exists $\rho \in (0, r)$ such that for any $w \in \mathbb{R}^n$,*

$$\frac{1}{n} \sum_{k=0}^{n-1} A^k w = c(w) r^n u + O(\rho^n)$$

Proof. The proof illustrates the *averaging method* widely used in the Representation Theory.

By rescaling A, we reduce the result to the case $r = 1$. Let us consider the Spectral Decomposition into the primitive components,

$$\mathbb{R}^d = \bigcup_{l=0}^{p-1} \mathcal{P}_l,$$

and let $u \in \mathcal{P}_0$ be the positive invariant vector for A^p constructed in the primitive case. Then for any $\varepsilon := e(m/p)$, $m \in \mathbb{Z}/p\mathbb{Z}$, the vector

$$\mathbf{u}_m := \sum_{l=0}^{p-1} \varepsilon^{-l} A^l u$$

is a positive eigenvector corresponding to the eigenvalue ε.

We leave the reader to supply the rest of the argument. □

Exercise 19.122. *Deduce from the above the general form of the Perron-Frobenius Theorem: For any non-negative matrix $A \geq 0$, the spectral radius $r \equiv r(A)$ itself is an eigenvalue of some multiplicity m. The corresponding basis of* *eigenvectors* $u_k \in \mathbb{R}^d$, $k = 1, \dots, m$, *can be selected positive. The boundary spectrum is the union of* k r*-scaled cyclic groups of some orders* p_k,

$$\operatorname{spec}^b(A) = \bigcup_{k=0}^{m} \{e(l/p_k)\, r : \; l \in \mathbb{Z}/p_k\mathbb{Z}\}.$$

19.19.4. *Stochastic matrices.* A non-negative matrix $P = (p_{ij})$ is called *stochastic* if all its rows $\mathbf{p}_i := (p_{ij})_j$ are probability distributions:

$$\sum_j p_{ij} = 1 \quad \forall\, i,$$

i.e., all the $\mathbf{p}_i$ belong to the probability simplex $\mathbf{\Delta}$. Equivalently, *the unit vector* $\mathbf{1} := (1 \dots 1)$ *is invariant under* P.

Exercise 19.123. *The spectral radius of a stochastic matrix is equial to* 1.

Lemma 19.124. *Any irreducible non-negative matrix* A *with spectral radius* 1 *is conjugate to a stochastic matrix.*

Proof. By the Perron-Frobenius Theorem, A has a positive invariant vector $(u_i) \in \mathbb{R}^d_+$. Let U be the diagonal matrix with the diagonal elements u_i. Then

$$U^{-1} A\, U = (u_i^{-1} a_{ij} u_j)$$

is the desired stochastic matrix. □

19.19.5. *Weighted permutations.* Let us introduce an important example of an irreducible matrix. A matrix $A \geq 0$ is called a *weighted permutation* if its graph Γ_A is a permutation σ of vertices: $a_{ij} \neq 0$ iff $j = \sigma(i)$. Such a matrix has a block structure with the blocks corresponding to the cycles of σ. The *multiplier* of a cyclic weighted permutation is defines as

$$\lambda(A) = \prod_{j=\sigma(i)} a_{ij}$$

Exercise 19.125. *If* A *is a* $p \times p$ *weighted permutation matrix then*

$$r(A) = \sqrt[p]{\lambda(A)}.$$

A Markov chain is called a *permutation* iff its matrix A is a weighted permutation (with all the weights being equal to 1).

Notes

Section "General Dynamics and Ergodic Theory" of the bibliography lists several introductory text books in these areas: Halmos (for the first reading in the Ergodic Theory) [**Ha2**] Brin-Stuck [**BS**] (for first reading in the Hyperbolic Theory), Katok-Hasselbatt [**KaH**] (for a broad variety of topics), supplemented with more advanced sources: Kornfeld-Sinai-Fomin [**KSF**], Pollicott [**Po**], and Walters [**Wa**].

The Denjoy Distortion Estimate (Exercise 19.69) appeared in the classical work [**Den**] (see Notes to §30 for more comments).

A general exploration of expanding maps was initiated in Shub's thesis [**Shu**] (though of course, Fatou studied "expanding Julia sets"). Theorem 19.79 on the analytic/singular dichotomy is due to Shub and Sullivan [**ShuS**]. It is related to the last step of the proof of the Mostow Rigidity Theorem.

Markov partitions for general expanding endomorphisms were constructed by Krzyzewski [**Krz**] (following earlier work by Adler et al, Sinai, and Bowen for hyperbolic automorphisms).

The balanced measures for toplogical Markov shifts appeared in Parry [**Par**].

The general notion of measure-theoretic/topological attractor of §19.7 was introduced by Milnor [**M1**].

CHAPTER 3

Dynamical plane I: basic objects

20. Holomorphic dynamics: Fatou and Julia sets

Below

$$f \equiv f_c : z \mapsto z^2 + c$$

unless otherwise is stated. Dynamical objects will be labeled by either f or c, whatever is more convenient in a particular situation (for instance, $\mathcal{D}_f(\infty) \equiv \mathcal{D}_c(\infty)$, $\mathcal{J}(f_c) \equiv \mathcal{J}_c$ by default). We will also use interchangably notation like $\mathcal{J}_f \equiv \mathcal{J}(f)$. Moreover, the label can be skipped altogether if f is not varied.

In this parametrization of the quadratic family, the variable "c" plays two different roles: the *parameter* and the *critical value*. Sometimes, when we want to emphasize the particular meaning, we use notation $v := f(0)$ for the critical value.

20.1. Critical points and values. First note that f^n is a branched covering of $\mathbb{C}$ over itself of degree 2^n. Its critical points and values have a good dynamical meaning:

EXERCISE 20.1. *The set of finite critical points of f^n is $\bigcup_{k=0}^{n-1} f^{-k}(0)$. The set of critical values of f^n is $\{0_k \equiv f^k 0\}_{k=1}^n$.*

Note that there are much fewer critical values than critical points!

We let

$$\mathrm{Crit}^n_f \equiv \mathrm{Crit}^n(f) := \bigcup_{k=0}^{n-1} f^{-k}(0), \quad n = 0, 1, \dots, \infty \tag{20.1}$$

be the set of *critical points of iterated f*. Let

$$\mathcal{P}_f := \mathrm{orb}\, v, \quad \overline{\mathcal{P}}_f \equiv \mathrm{cl}\, \mathcal{P}_f = \overline{\mathrm{orb}\, v}$$

be the *post-valuable sets*[1] of f. They are forward invariant and contain the critical value of f. (Of course, we will often skip the label f in the notation.) The corresponding *postcritical sets* are defined respectively as

$$\mathrm{orb}\, 0 = \mathcal{P} \cup \{0\}, \quad \overline{\mathrm{orb}\, 0} = \overline{\mathcal{P}} \cup \{0\}.$$

The iterate f^n is an unbranched covering over the complement of $\{f^k 0\}_{k=1}^n$, so *all the iterates are unbranched over $\mathbb{C} \smallsetminus \overline{\mathcal{P}}$.*

[1] If confusion can arise, then we call $\overline{\mathcal{P}}_f$ the *closed* post-valuable set (and similarly for the postcritical sets).

Corollary 20.2. *Let V be a topological disk which does not contain points $0_k \equiv f^k(0)$, $k = 1, 2, \dots, n$. Then the inverse function f^{-n} has 2^n single-values branches f_i^{-n} that univalently map V onto pairwise disjoint topological disks U_i, $i = 1, 2, \dots, 2^n$.*

These simple remarks explain why the forward orbit of 0 plays a very special role. We will have many occasions to see that this single orbit is responsible for the complexity and variety of the global dynamics of f.

However, f has one more critical point overlooked so far:

20.2. Looking from infinity. Extend f to an endomorphism of the Riemann sphere $\hat{\mathbb{C}}$. This extension has a critical point at ∞ fixed under f. We will start exploring the dynamics of f from there. The first observation is that $\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_R$ is f-invariant for a sufficiently big R, and moreover $f^n z \to \infty$ as $n \to \infty$ for $z \in \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_R$. This can be expressed by saying that $\mathbb{C} \smallsetminus \mathbb{D}_R$ belongs to the *basin of infinity* defined as the set of all *escaping* points:

$$\mathcal{D}_f(\infty) = \{z \in \hat{\mathbb{C}} : f^n z \to \infty \text{ as } n \to \infty\} = \bigcup_{n=0}^{\infty} f^{-n}(\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_R).$$

Proposition 20.3. *The basin of infinity $\mathcal{D}_f(\infty)$ is a completely invariant domain containing ∞.*

Proof. The only non-obvious statement to check is connectivity of $\mathcal{D}_f(\infty)$. To this end let us show inductively that the sets $U_n = f^{-n}(\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_R)$ are connected. Indeed, assume that U_n is connected while U_{n+1} is not. Consider a bounded component V of U_{n+1}. Then the restriction $f : V \to U_n$ is proper and hence surjective (Corollary 1.105). In particular f would have a pole in V – contradiction. □

20.3. Basic Dichotomy for Julia sets. We can now introduce the fundamental dynamical object, the *filled Julia set*

$$\mathcal{K}_f \equiv \mathcal{K}(f) := \hat{\mathbb{C}} \smallsetminus \mathcal{D}_f(\infty).$$

Proposition 20.3 implies that $\mathcal{K}(f)$ is a completely invariant compact subset of $\mathbb{C}$. Moreover, it is *full*, i.e., it does not separate the plane (since $\mathcal{D}_f(\infty)$ is connected).

Exercise 20.4. (i) *The filled Julia set consists of more than one point.*
(ii) *Each component D of* $\operatorname{int} \mathcal{K}(f)$ *is simply connected.*

The filled Julia set and the basin of infinity have a common boundary, which is called the *Julia set*,

$$\mathcal{J}_f \equiv \mathcal{J}(f) := \partial \mathcal{K}(f) = \partial \mathcal{D}_f(\infty).$$

Figures in this section show several representative pictures of the Julia sets $\mathcal{J}_c \equiv \mathcal{J}(f_c)$ for different parameter values c.

Generally, topology and geometry of the Julia set is intricate, and it takes an effort to put a hold on it. However, the following rough classification will give us some guiding principle:

Theorem 20.5 (Basic Dichotomy). *The Julia set* (*and the filled Julia set*) *is either connected or Cantor. The latter happens if and only if the critical point escapes to infinity: $f^n(0) \to \infty$ as $n \to \infty$.*

PROOF. As in the proof of Proposition 20.3, let us consider the increasing sequence of domains $U_n = f^{-n}(\hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_R)$ exhausting the basin of infinity. Assume first that the critical point does not escape to ∞. Then $f : U_{n+1} \to U_n$ is a branched double covering with the only branched point at ∞. By the Riemann-Hurwitz formula, if U_n is simply connected then U_{n+1} is simply connected as well. We conclude inductively that all the domains U_n are simply connected. Hence their union, $\mathcal{D}_f(\infty)$, is simply connected as well, and its complement, $K(f)$, is connected. But the boundary of a full connected compact set is connected. Hence $\mathcal{J}(f)$ is connected, too.

Assume now that the critical point escapes to infinity. Then 0 belongs to some domain U_n. Take the smallest n with this property. Adjust the radius R in such a way that the orbit of 0 does not pass through $\mathbb{T}_R = \partial U_0$. Then $0 \notin \partial U_{n-1}$, and hence ∂U_{n-1} is a Jordan curve. Let us consider the complementary Jordan disk $D \equiv D^0 = \mathbb{C} \smallsetminus \overline{U}_{n-1}$. Since $f(0) \in U_{n-1}$, f is unbranched over D. Hence $f^{-1}(D) = D_0^1 \cup D_1^1$, where the $D_i^1 \Subset D$ are disjoint topological disks conformally mapped onto D. The disks D_0^1 and D_0^1 form a Bernoulli family of tiles in the sense of §19.92.

Take now the f-preimages of $D_0^1 \cup D_1^1$ in D_0^1. We obtain two Jordan disks D_{00}^2 and D_{01}^2 with disjoint closures conformally mapped by f onto D_0^1 and D_1^1 respectively. Similar disks, D_{10}^2 and D_{11}^2, we find in D_1^1 (see Figure 20.1).

Iterating this procedure, we will find that $f^{-n}(D)$ is the union of 2^n Jordan disks $D^n_{i_0 i_1 \dots i_{n-1}}$ such that $D^n_{i_0 \dots i_{n-1}}$ is compactly contained in $D^{n-1}_{i_0 \dots i_{n-2}}$ and is conformally mapped by f onto $D^{n-1}_{i_1 \dots i_{n-1}}$.

Since $D_0^1 \cup D_1^1$ is compactly contained in D, the branches of the inverse map, $f^{-1} : D_j^1 \to D_{ij}^2$, are uniformly contracting in the hyperbolic metric of D (by the Schwarz Lemma). Since the domains $D^n_{i_0 i_1 \dots i_{n-1}}$ are obtained by iterating these branches, they uniformly exponentially shrink as $n \to \infty$. Hence the filled Julia set $\mathcal{K}(f) = \bigcap f^{-n}(D)$ is a Cantor set. Of course, the Julia set $\mathcal{J}(f)$ coincides with $\mathcal{K}(f)$ in this case. $\square$

REMARK 20.6. The above proof also shows that in the case when $\mathcal{J}(f)$ is a Cantor set, the map f is *expanding* on it: $\|Df(z)\| \geq \rho > 1$ for any $z \in \mathcal{J}(f)$, with respect to the hyperbolic metric on the disk $D \supset \mathcal{J}(f)$.

The Basic Dichotomy gives us the first glimpse of how the behavior of the critical point may influence the global dynamics.

Since neither Cantor nor connected sets (containing more than one point) can have isolated points, we conclude:

COROLLARY 20.7. *Both $\mathcal{K}(f)$ and $\mathcal{J}(f)$ are perfect sets.*

When the Julia set is Cantor, there is an explicit *symbolic model* for the dynamics of f on it. Namely, the partition $\mathcal{J} = (\mathcal{J} \cap D_0^1) \sqcup (\mathcal{J} \cap D_1^1)$ constructed in the proof of Theorem 20.5 is a Bernoulli generator for $f|\mathcal{J}$. So by Exercise 19.52 we conclude:

COROLLARY 20.8. *If the Julia set $\mathcal{J}(f)$ is Cantor then the dynamics on it is topologically conjugate to the one-sided Bernoulli shift $\sigma_2 : \Sigma_2^+ \to \Sigma_2^+$ with two states.*

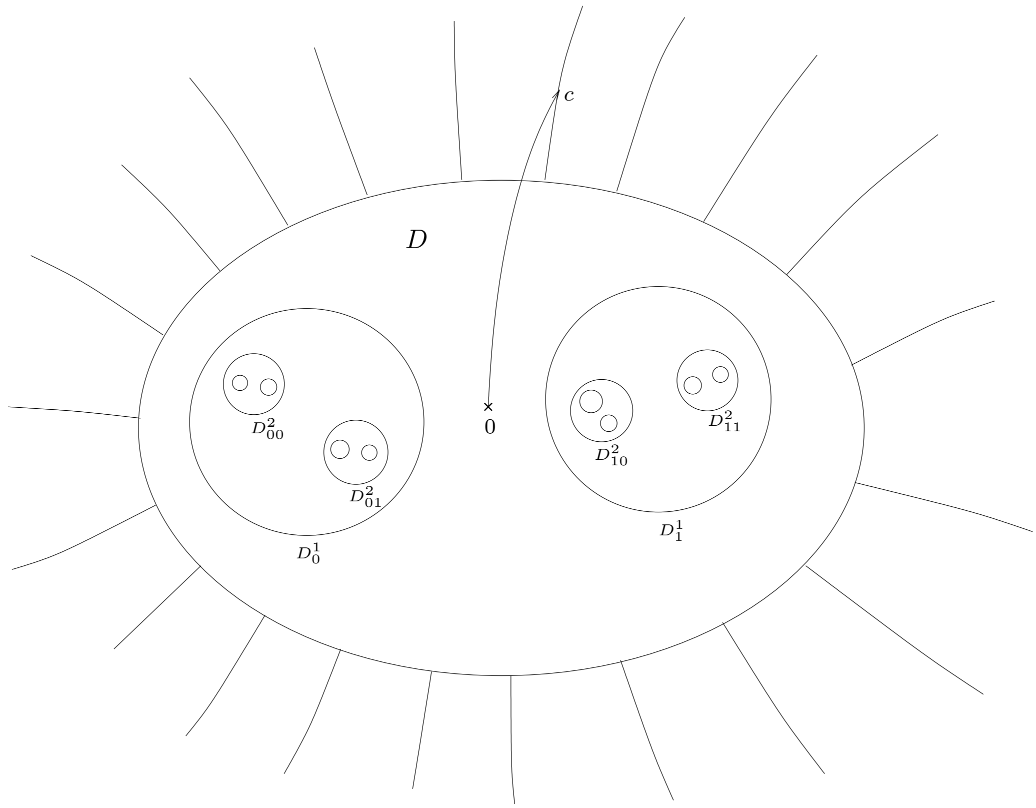

FIGURE 20.1. Generating a Cantor Julia set.

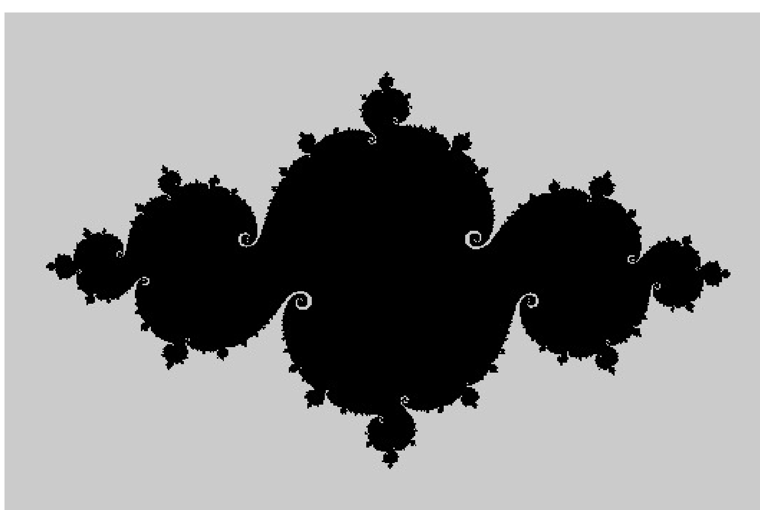

FIGURE 20.2. The filled Julia set of $z^2+\varepsilon$ for small ε is a quasidisk. It contains an attracting fixed point inside.

FIGURE 20.3. Cauliflower: the filled Julia set of $z^2 + 1/4$. It contains a parabolic fixed point $1/2$: the prominent cusp on the right.

For this reason, the set of parameters $c \in \mathbb{C}$ for which the Julia set $\mathcal{J}_c$ is Cantor is called the *shift locus*. The complementary set of parameters, $\mathcal{M}$, is called the *connectedness locus* or the *Mandebrot set*. It will be a central object of study in this book (see Ch. 5).

20.4. Real quadratic family.

20.4.1. *Two real shift loci.* In the case of real parameter values c, the Bernoulli coding of $\mathcal{J}_c$ becomes particularly nice:

EXERCISE 20.9. *Consider a quadratic polynomial $f_c : z \mapsto z^2 + c$ with a real c. Let $\mathcal{J} \equiv \mathcal{J}(f_c)$.*

(i) *If $c < -2$ then $\mathcal{J}$ is a Cantor set on the real line. In this case the Bernoulli generator for f_c consists of*
$$\mathcal{J}^0 = \mathcal{J} \cap \{z : \operatorname{Re} z < 0\} \text{ and } \mathcal{J}^1 = \mathcal{J} \cap \{z : \operatorname{Re} z > 0\}.$$

picture

(ii) *If $c > 1/4$ then J is a Cantor set disjoint from the real line. In this case the Bernoulli generator for f_c consists of*
$$\mathcal{J}^0 = \mathcal{J} \cap \{z : \operatorname{Im} z > 0\} \text{ and } \mathcal{J}^1 = \mathcal{J} \cap \{z : \operatorname{Im} z < 0\}.$$

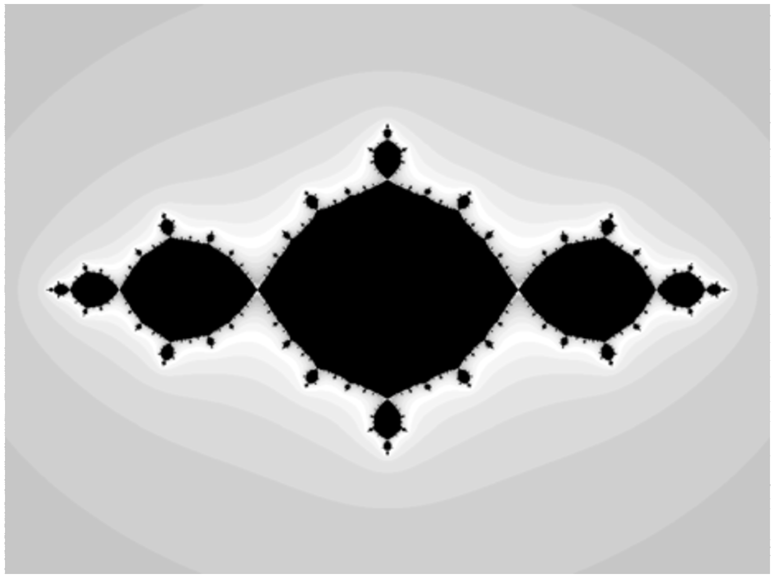

FIGURE 20.4. Basilica: the filled Julia set of $z^2 - 1$. It has a superattracting cycle $\{0, -1\}$ of period 2.

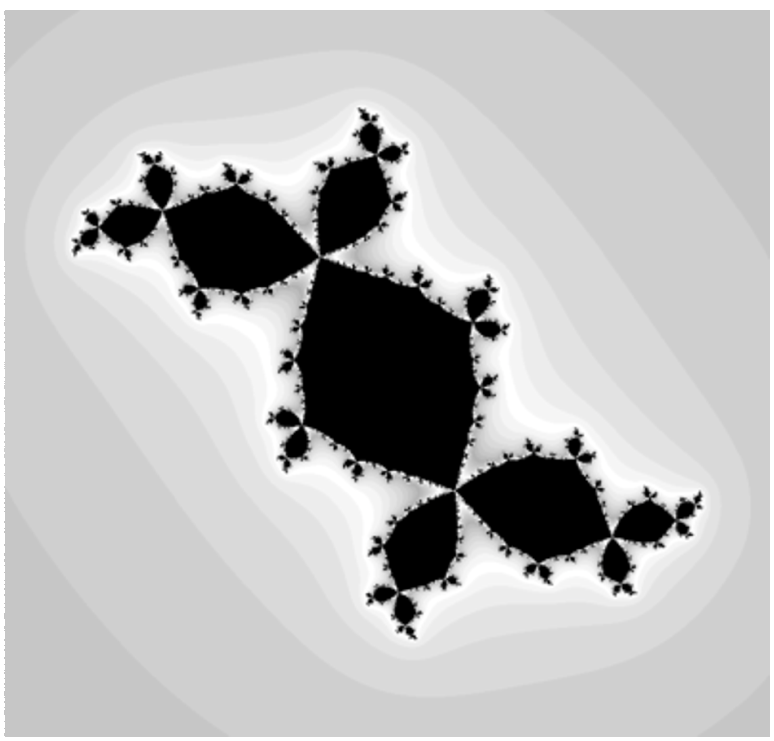

FIGURE 20.5. Douady rabbit: his head and two ears form the immediate basin for a superattracting cycle of period 3.

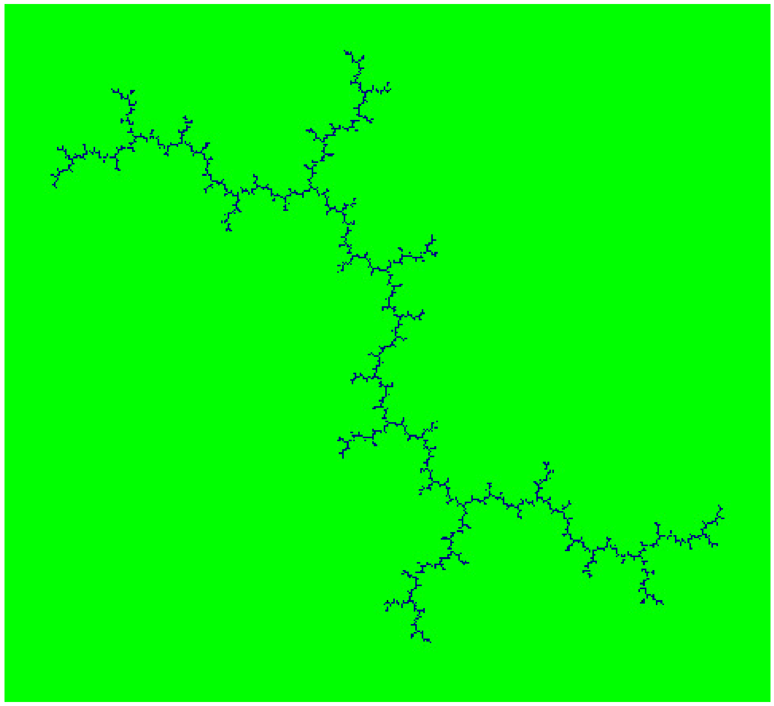

FIGURE 20.6. Dendrite: the Julia set of $z^2 + i$. Here the critical point is preperiodic with period 2.

We see, in particular, that for $c \in \mathbb{R} \smallsetminus [-2, 1/4]$, the map f_c has no invariant intervals on the real line.

20.4.2. *Maximal invariant interval.* The following material relies on the notions from §19.8.1.

EXERCISE 20.10. *For $c \in [-2, 1/4]$, the map f_c has an invariant interval. The maximal invariant interval has a form $\mathcal{I}_c = [-\beta_c, \beta_c]$, where β_c is an orientation preserving fixed point of f_c. Moreover, this point is either repelling or parabolic, i.e. $f_c'(\beta_c) \geq 1$ (it is parabolic only for $c = 1/4$). The orbits of $x \in \mathbb{R} \smallsetminus \mathcal{I}_c$ escape to ∞. Finally, the Julia set $\mathcal{J}_c$ is connected, and $\mathcal{I}_c$ is its real slice:*

$$\mathcal{I}_c = \mathcal{J}_c \cap \mathbb{R} = \mathcal{K}_c \cap \mathbb{R}.$$

Putting Exercises 20.9 and 20.10 together, we obtain:

COROLLARY 20.11. *For real c, the Julia set $\mathcal{J}_c$ is connected if and only if $c \in [-2, 1/4]$.*

REMARK 20.12. In Exercise 20.18 we will identify the smallest invariant interval $\mathcal{T}_c$ for $c < -1$.

EXERCISE 20.13. *For a quadratic polynomial f, assume* $\operatorname{int} \mathcal{K} \neq \emptyset$. *If a component D of* $\operatorname{int} \mathcal{K}$ *meets $\mathbb{R}$ then its real slice $D \cap \mathbb{R}$ is an interval.*

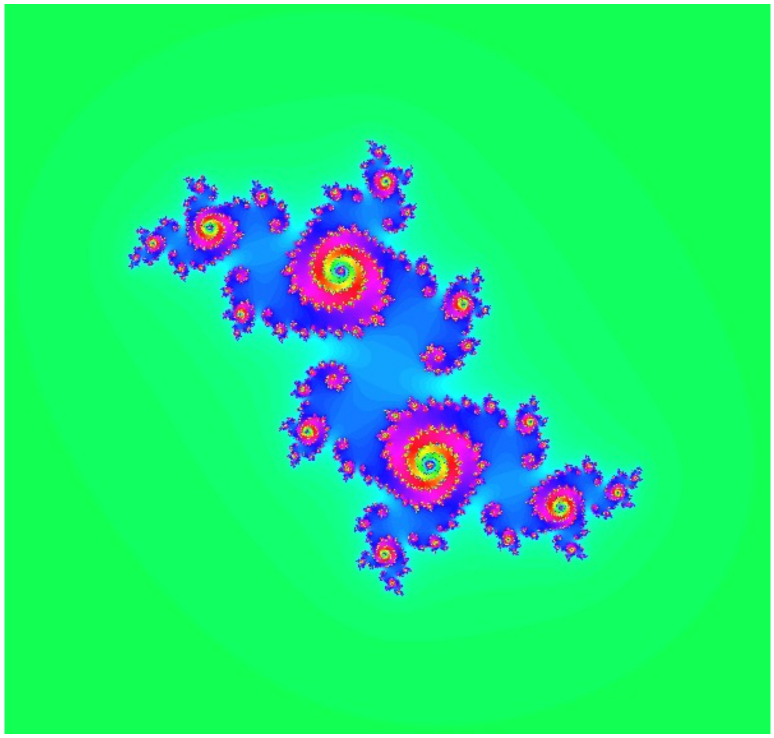

FIGURE 20.7. Cantor dust.

20.4.3. *Unimodal maps.* The quadratic maps $f_c : \mathcal{I}_c \to \mathcal{I}_c$, $c \in [-2, 1/4]$, give us examples of unimodal interval maps. A *unimodal* interval map $f : \mathcal{I} \to \mathcal{I}$ is a continuous map that has exactly two intervals of monotonicity, and hence, f has exactly one extremum in $\operatorname{int} I$. Let us make a few *default conventions/assumptions* (that will be assumed unless otherwise is explicitly stated):

– f is real analytic in a neighborhood of $\mathcal{I}$;

– The extremum of f is located at the origin, which is the only critical point of f;

– 0 is non-degenerate: $f''(0) \neq 0$;

– Any *initial* unimodal map[2] under consideration is normalized so that 0 is the *minimum* point of f. So the critical value $v := f(0)$ is the global minimum of f.

Slightly abusing terminology, we say that a unimodal map f is *proper* if $f(\partial\mathcal{I}) \subset \partial\mathcal{I}$, i.e., one of the boundary points, $\beta \in \mathcal{I}$, is fixed, while the other is its preimage. Note that the β-fixed point must have a positive multiplier ρ, and under our convention $\beta > 0$.

We say that a proper unimodal map (and its β-fixed point) is *repulsive* if β is locally repelling in the exterior of $\mathcal{I}$,[3] i.e, there exists $\varepsilon > 0$ such that

$$\forall\ x \in (\beta, \beta + \varepsilon)\ \exists\ n \in \mathbb{Z}_+ \text{ s.t. } f^n x > \beta + \varepsilon.$$

[2]A typical example of a "non-initial" unimodal map is a restricted iterate of another map.

[3]The role of this assumption will become clear when we study quadratic-like renormalization.

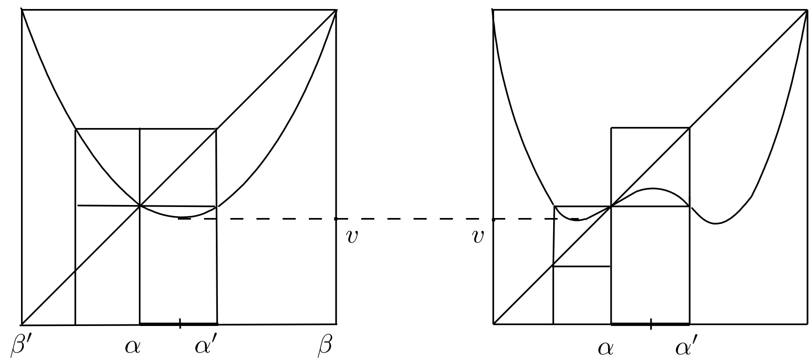

FIGURE 20.8. Orientation reversing attracting fixed point α, with f^2 depicted on the right-hand side.

Then necessarily $\rho \geq 1$, and sufficiently $\rho > 1$ or $\rho = 1$ and the Taylor expansion of f near β looks like

$$f(x) = x + a(x-\beta)^2 + \dots \text{ with } a > 0. \tag{20.2}$$

A non-repulsive proper unimodal map will also be referred to as *attractive*.

A proper unimodal map has a *dynamical symmetry*: $\sigma : x \mapsto x'$, where $f(x) = f(x')$, so under our conventions, $\mathcal{I} = [\beta', \beta]$. We will be mostly dealing with *even* unimodal maps, so σ is the usual central symmetry $x \mapsto -x$. In fact, in what follows *we will assume that f is even* unless otherwise is explicitly stated. In this case, $\mathcal{I} = [-\beta, \beta]$.

20.4.4. *Fixed points.* Let us now return to the quadratic family. The boundary parameter values $c = 1/4$ and $c = -2$ play a special role in the one-dimensional dynamics (both real and complex).

EXERCISE 20.14. *The map $f \equiv f_{1/4} : z \mapsto z^2 + 1/4$ is singled out among the quadratic maps f_c, $c \in \mathbb{C}$, by the property that it has a double fixed point $\alpha = \beta = 1/2$, i.e., $f(\beta) = \beta$, $f'(\beta) = 1$. In this case, $f^n x \to \beta$ for all $x \in \mathcal{I}$.*

The Julia set of $f_{1/4}$ is a Jordan curve depicted on Figure 20.3 (see §26.1 for an explanation of some features of this picture). It is called the *cauliflower*, and the map $f_c : z \mapsto z^2 + 1/4$ itself is sometimes called the *cauliflower map*.

Let us take a look at what happens as c crosses $1/4$:

EXERCISE 20.15. *For any $c < 1/4$, the map f_c has two fixed points $\alpha_c < \beta_c$.*
(i) *The point α_c is attracting for $c \in (-3/4, 1/4)$, and repelling for $c < -3/4$. It is orientation preserving for $c \in (0, 1/4)$, and it is a flip for $c < 0$.*
(ii) *For $c \in (-3/4, 1/4)$, $f_c^n x \to \alpha_c$ for all $x \in \operatorname{int} \mathcal{I}$* (see Figure 20.8).
(iii) *The multiplier $\rho : c \mapsto f_c'(\alpha)$ is an orientation preserving diffeomorphism*

$$\rho : (-3/4, 1/4) \to (-1, 1).$$

We say that the *saddle-node* bifurcation occurred at $c = 1/4$ and the *superattracting bifurcation* occurred at $c = 0$.

The above fixed points, α_c and β_c, will be called α- and β-fixed points respectively. They play quite a different dynamical role. In §24.4.2 a similar classification of the fixed points will be given for any quadratic polynomial with connected Julia set.

20.4.5. *Period doubling bifurcation.* This section will give us the first glimpse of a fundamental phenomenon that will play a central role throughout this book.

EXERCISE 20.16. (i) *For parameters $c < -3/4$ near $-3/4$, the map f_c has an attracting cycle $\boldsymbol{\gamma}_c$ of period* 2.

(ii) *This attracting cycle persists on the parameter interval $(-5/4, -3/4)$ and its multiplier $\rho_c \equiv \rho(\boldsymbol{\gamma}_c)$ is an orientation preserving diffeomorphism from this interval onto $(-1, 1)$.*

(iii) $\rho_c < 0$ *for $c < -1$ (so the cycle $\boldsymbol{\gamma}_c$ is a flip) and $\rho_c < -1$ for $c < -5/4$.*

EXERCISE 20.17. *For $c \in [-5/4, 0)$, we have:*

(i) *The interval $I \equiv I_c = [-\alpha_c, \alpha_c]$ is a periodic interval of period* 2*, i.e., $f^2(I) \subset I$, and moreover, $f(I) \cap I = \{\alpha\}$* (see Figure 20.9).

(ii) *The restriction $f^2 \,|\, I$ is a unimodal map; this map is repulsive iff $c < -3/4$.*

(iii) *For $c \in (-5/4, -3/4)$, all the orbits in $\operatorname{int} I$ converge to the attracting cycle $\boldsymbol{\gamma}$.*

(iv) *Any orbit in $\operatorname{int} \mathcal{I}$ eventually lands in I, i.e., for any $x \neq \pm\beta$, we have $f^n x \in I$ for some $n \in \mathbb{N}$.*

Moreover, I_c remains periodic until much smaller parameter $c_* < -5/4$ – which one?

We will call the map $f_{-3/4}$ that has a flip parabolic fixed point (and the corresponding doubling bifurcation parameter $c = -3/4$) *Myrberg*.

Let us conclude with a useful remark:

EXERCISE 20.18. *Let us consider a real quadratic map $f \equiv f_c$, $c < -1$, that has a flip cycle of period two. Then the critical point* 0 *lies in between v and $f(v)$, and the interval $\mathcal{T} \equiv \mathcal{T}_c := [v, f(v)]$ is invariant. Thus, this is the smallest invariant interval containing* 0*. Moreover, any orbit in $\operatorname{int} \mathcal{I}$ eventually lands in $\mathcal{T}$.*

20.4.6. *Chebyshev map and its saw-like model.* The parameter $c = -2$ is specified by the property that the second iterate of the critical point is fixed under f_c: $0 \mapsto -2 \mapsto 2 \mapsto 2$ (see Figure 20.10). The corresponding map $f_{-2} \equiv \text{Ч}$ is called *Chebyshev* or *Ulam-Neumann* (in the dynamical setting). The Julia set of this map is unusually simple:

EXERCISE 20.19 (Chebyshev map). *Let $\text{Ч} \equiv f_{-2} : z \mapsto z^2 - 2$.*

(i) *Zhukovsky map $\text{Ж} : (\mathbb{C}^*, \mathbb{T}) \to (\mathbb{C}, [-2, 2])$, $z \mapsto z + 1/z$ semiconjugates $f_0 : z \mapsto z^2$ to Ч.*

(ii) *The interval $\mathcal{I} = [-2, 2]$ is completely invariant under Ч, i.e., $\text{Ч}^{-1}(\mathcal{I}) = \mathcal{I}$.*

(iii) $\mathcal{J}(\text{Ч}) = \mathcal{I}$.

(iv) *The map $h : \omega \mapsto 2 \sin \dfrac{\pi\omega}{2}$ conjugates the saw-like map Λ (19.13) to $\text{Ч} \,|\, \mathcal{I}$.*

(v) *The map $\text{Ч}|\mathcal{I}$ is nicely semi-conjugated to the one-sided Bernoulli shift $\sigma : \Sigma \to \Sigma$. Namely, there exists a natural semi-conjugacy $h : \Sigma \to \mathcal{I}$ such that* $\operatorname{card} h^{-1}x = 1$ *for all $x \in \mathcal{I}$ except countable many points (preimages of the fixed point $\beta = 2$ under iterates of f). For these special points,* $\operatorname{card} h^{-1}(x) = 2$.

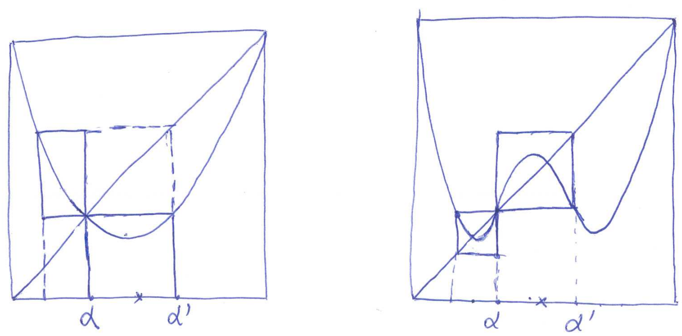

FIGURE 20.9. Periodic interval of period 2, with f^2 depicted on the right-hand side.

(vi) *The measure*

$$d\mu = \frac{dx}{\sqrt{4 - x^2}}$$

is Ч*-invariant and ergodic.*

(vii) *Generalize the above discussion to higher degree Chebyshev polynomials* (*see* §2.10.1) .

20.5. Fatou set. The Fatou set is defined as the complement of the Julia set:

$$\mathcal{F}_f \equiv \mathcal{F}(f) := \hat{\mathbb{C}} \setminus \mathcal{J}(f) = \mathcal{D}_f(\infty) \cup \operatorname{int} \mathcal{K}(f).$$

Since $\mathcal{K}(f)$ is full, all components of $\operatorname{int} \mathcal{K}(f)$ are simply connected. Only one of them can contain the critical point 0. Such a component D_0 (if exists) is called *critical* or *central*. Its image $f(D_0)$ contains the critical value $c = f(0)$; it is called *valuable*.

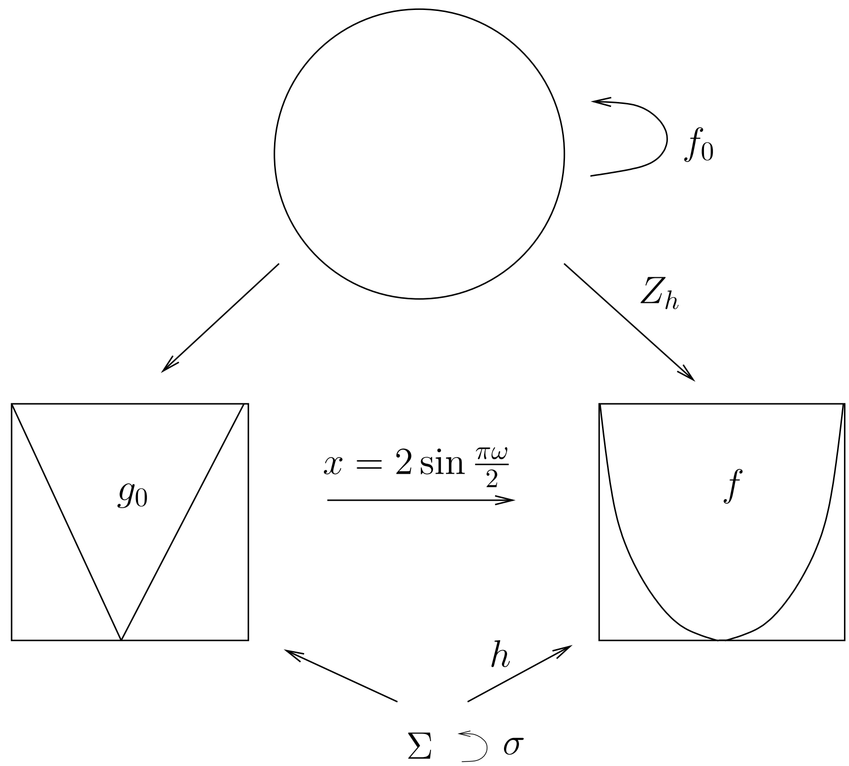

FIGURE 20.10. Doubling, Chebyshev, saw-like, and shift maps.

Let U be one of the components of $\operatorname{int}\mathcal{K}$. Since $\operatorname{int}\mathcal{K}$ is invariant, it is mapped by f to some other component V. Moreover, $f(\partial U) \subset \partial V$ since the Julia set is also invariant. Hence $f : U \to V$ is proper, and thus surjective. Moreover, since V is simply connected, $f : U \to V$ is either a conformal isomorphism (if U is not critical), or is a double branched covering (if U is critical).

The Fatou set can be also characterized as *the set of normality* (and was actually classically defined in this way):

PROPOSITION 20.20. *The Fatou set $\mathcal{F}(f)$ is the maximal set on which the family of iterates f^n is normal.*

PROOF. On $\mathcal{D}_f(\infty)$, the iterates of f locally uniformly converge to ∞, while on $\operatorname{int}\mathcal{K}(f)$ they are uniformly bounded. Hence they form a normal family on $\mathcal{F}(f)$. On the other hand, if $z \in \mathcal{J}(f)$, then the orbit of z is bounded while there are nearby points escaping to ∞. Hence the family of iterates is not normal near z. $\square$

Thus, the family of iterates $(f^n)_{n\in\mathbb{N}}$ is locally equicontinuous on the Fatou set, implying that *points $x \in \mathcal{F}(f)$ are Lyapunov stable.* We will see in §21.5 that this characterizes the Fatou set. This gives a good sense of the Fatou set as the *regular* ($\equiv$ *non-chaotic*) part of the dynamical plane.

20.6. Preimages of points.

PROPOSITION 20.21. *Let $f : z \mapsto z^2 + c$. If $c \neq 0$, then for any neighborhood U intersecting $\mathcal{J}(f)$ we have:*

$$\operatorname{orb} U := \bigcup_{n=0}^{\infty} f^n(U) = \mathbb{C}.$$

If $c = 0$ and $U \not\ni 0$ then $\operatorname{orb} U = \mathbb{C}^*$.

PROOF. By the Montel Theorem, $\mathbb{C} \smallsetminus \operatorname{orb} U$ contains at most one point. If there is one, a, then $f^{-1}a = \{a\}$. Hence a is the critical point of f, i.e., $a = 0$. Moreover, $f(a) = a$, so $c = 0$. □

This result immediately yields:

COROLLARY 20.22. *For any point $z \in \mathbb{C}$, except $z = 0$ in case $f : z \mapsto z^2$, we have:*

$$\operatorname{cl} \bigcup_{n=0}^{\infty} f^{-n} z \supset \mathcal{J}(f).$$

COROLLARY 20.23. *If $J \subset \mathcal{J}$ is a non-empty backward invariant closed subset of $\mathcal{J}$ then $J = \mathcal{J}$. If $K \subset \mathcal{K}$ is a non-empty full backward invariant closed subset of $\mathcal{K}$ then $K = \mathcal{K}$.*

Together with Exercise 19.4, the above Proposition also yields:

COROLLARY 20.24. *Any polynomial f is topologically transitive on its Julia set.*

20.7. Inverse branches.

20.7.1. *Normality and Koebe control.*

LEMMA 20.25 (Normality Lemma). *Given a domain $U \subset \mathbb{C}$, the family of inverse branches $f_i^{-n}|U$ that are well defined on U is normal.*[4]

PROOF. Since normality is a local property, we can assume that U is bounded. Take R so big that $U \subset \mathbb{D}_R$ and $f^{-1}(\mathbb{D}_R) \subset \mathbb{D}_R$. Then $f^{-n}(U) \subset \mathbb{D}_R$ for all $n \in \mathbb{N}$ and in particular, $f_i^{-n}(U) \subset \mathbb{D}_R$ for all the inverse branches under consideration. So, this family is normal by the Little Montel Theorem. □

This allows us to control the distortion of the inverse branches (which also follows directly from the Koebe Distortion Theorem):

COROLLARY 20.26. *Under the above circumstances, let $\mathbb{D}(z,r) \Subset U$ and*

$$\operatorname{mod}(U \smallsetminus \mathbb{D}(z,r)) \geq \mu > 0.$$

Then the inverse branches f_i^{-n} have a bounded distortion on $\mathbb{D}(z,r)$ and map it onto ovals of bounded shape (around the $f_i^{-n}(z)$). The bounds depend only μ.

[4]Here U is allowed to intersect the postcritical set.

20.7.2. *Shrinking Lemma.*

LEMMA 20.27. *Let $U \subset \mathcal{D}(\infty)$ be a domain in the basin of ∞, and let $f_i^{-n}|U$ be an infinite family of inverse branches that are well defined on U.*[5] *Then for any set $V \Subset U$, $\operatorname{diam}(f_i^{-n}(V)) \to 0$ as $n \to \infty$.*

PROOF. Select a base point $z \in V$. By the Koebe control of the inverse branches, the pullbacks $f_i^{-n}(V)$ have a bounded shape around points $f_i^{-n}(z)$. So it is sufficient to show that the inner radii around these points go to 0. But this is obvious since these points go to the boundary of $\mathcal{D}(\infty)$ (the Julia set). □

SHRINKING LEMMA. Let $U \subset \mathbb{C}$ be a domain intersecting the Julia set $\mathcal{J}(f)$, and let $f_i^{-n}|U$ be an infinite family of inverse branches that are well defined on U. Then for any set $V \Subset U$, $\operatorname{diam}(f_i^{-n}(V)) \to 0$ as $n \to \infty$.

PROOF. Since the family of inverse branches f_i^{-n} is normal on U, it is sufficient to show that $\operatorname{diam}(f_i^{-n}(V)) \to 0$ for *some* domain $V \Subset U$. But since U intersects $\mathcal{J}$, it intersects the basin $\mathcal{D}(\infty)$ as well, and the conclusion follows from Lemma 20.27. □

20.7.3. *Formation of monotonicity intervals.* Let us consider a real quadratic polynomial $f \equiv f_c$ with $c \in [-2, 1/4]$, and let $\mathcal{I} \equiv \mathcal{I}_c$ be it invariant interval from Exercise 20.10. We let $\operatorname{Crit}^n_{\mathbb{R}} \equiv \operatorname{Crit}^n_{\mathbb{R}}(f) := \operatorname{Crit}^n(f) \cap \mathbb{R}$ be the set of real critical points of f^n.

For $n \in \mathbb{Z}_+$ and $x \in \mathcal{I} \smallsetminus \operatorname{Crit}^n_{\mathbb{R}}$, let $L_n(x) \subset \mathbb{R}$ be the maximal interval containing x on which f^n is monotone. The boundary points of $L_n(x)$ belong to $\operatorname{Crit}^n_{\mathbb{R}} \cup \partial\mathcal{I}$. By (20.1), for each endpoint $a \in \partial L_n(x)$ that does not belong to $\partial\mathcal{I}$, there exists an integer $k \in [0, n-1]$ such that $f^k a = 0$, so the interval $f^k(L)$ "grabs" the critical point 0 and "carries it forward" to the image $M_n(x) := f^n(L_n(x))$.

20.7.4. *Inverse branches for real maps.* For an interval $M \subset \mathbb{R}$ we let M° be its interior rel the real line. By Corollary 20.2, there is a well defined inverse branch

$$f^{-n} : \mathbb{C}(M_n^\circ(x)) \to \mathbb{C}$$

that maps $M_n(x)$ to $L_n(x)$.

LEMMA 20.28. *The image of the half-plane $\mathbb{H}_+$ under the above branch of f^{-n} is contained in one of the half planes $\mathbb{H}_+$ or $\mathbb{H}_-$ (depending on whether $f^n : L_n(x) \to M_n(x)$ is orientation preserving or reversing). Similarly for the half-plane $\mathbb{H}_-$.*

PROOF. Since $f^n(\mathbb{R}) \subset \mathbb{R}$, we have $f^{-n}(\mathbb{C} \smallsetminus \mathbb{R}) \subset \mathbb{C} \smallsetminus \mathbb{R}$. All the more, any inverse branch of f^{-n} maps the half-plane $\mathbb{H}_+$ inside $\mathbb{C} \smallsetminus \mathbb{R}$. The orientation rule comes from the fact that f^n preserves orientation of $\mathbb{C}$. □

There is a nice way to visualize these branches. Let us color the half-plane $\mathbb{H}_+$ in black while keeping $\mathbb{H}_-$ white. Then the complement $\mathbb{C} \smallsetminus f^{-n}(\mathbb{C} \smallsetminus \mathbb{R})$ assumes the checker-board coloring illuminating the corresponding branches.

[5] If $\mathcal{J}$ is connected then any topological disk in $\mathcal{D}(\infty)$ would serve for *all* inverse branches.

20.7.5. *Pullback of the real space.* Let now consider an $\mathbb{R}$-symmetric polynomial $f \equiv f_c$. Assume we have two pairs of real intervals, (M, L) and (M', L'), such that f^n monotonically maps (M', L') onto (M, L). Then the inverse branch $f^{-n} : (M, L) \to (M', L')$ admits an extension to a conformal embedding $f^{-n} : \mathbb{C}(M) \to \mathbb{C}(M')$. Applying Lemma 6.18, we obtain:

Corollary 20.29. *Under the above circumstances,*

$$\operatorname{mod}_{\mathbb{R}}(M, L) \geq \delta > 0 \implies \operatorname{mod}_{\mathbb{R}}(M', L') \geq \varepsilon(\delta) > 0.$$

Furthermore, Corollary 7.2 of the Symmetric Schwarz Lemma and Koebe Theorem yields:

Lemma 20.30. *Let $f^n : L' \to L$ be a monotonic map between two real open intervals viewed as hyperbolic lines* (see §2.4.5). *Then f^n expands the hyperbolic metric:*

$$\|Df^n(x)\|_{\text{hyp}} > 1 \quad \forall\, x \in L'.$$

If $f^n(L') \Subset \operatorname{int} L$ with $\operatorname{mod}_{\mathbb{R}}(L : f^n(L')) \geq \delta > 0$, then the expansion is uniform, i.e., the above norm is $\geq \rho(\delta) > 1$, and the map f^n has a uniformly bounded distortion:

$$\frac{Df^n(x)}{Df^n(y)}| \leq C(\delta) \quad \forall\, x, y \in L'.$$

20.8. Expanding (hyperbolic) sets. The map f is tremendously contracting near the critical point 0, and under iteration this contraction propagates through the postcritical set. The following lemma is the first indication that otherwise the map f tends to be expanding:

Lemma 20.31. *Let $c \neq 0$. Then any component of $\mathbb{C} \setminus \overline{\mathcal{P}}_f$ is a hyperbolic domain. Let $z, f(z) \in \mathbb{C} \setminus \overline{\mathcal{P}}_f$, and let $\Omega \ni z$, $V \ni f(z)$ be the components of $\mathbb{C} \setminus \overline{\mathcal{P}}_f$ containing the corresponding points. Assume Ω intersects $f^{-1}(\overline{\mathcal{P}}_f)$. Then f is strictly expanding with respect to the hyperbolic metrics of Ω and V:*

$$\|Df(z)\|_{\text{hyp}} > 1.$$

Proof. If $\mathbb{C} \setminus \overline{\mathcal{P}}_f$ is not hyperbolic, then $\overline{\mathcal{P}}_f$ consists of a single point, c. But then $f(c) = c$ and hence $c = 0$.

Let Ω' be the component of $f^{-1}(V)$ containing z. Since $\overline{\mathcal{P}}_f$ is forward invariant, $\Omega' \subset \Omega$. Moreover, $\Omega' \subset \Omega \setminus f^{-1}(\overline{\mathcal{P}}_f)$, so by assumption, Ω' is strictly smaller than Ω. Since V does not contain the critical value c, the map $f : \Omega' \to V$ is a covering. The conclusion now follows from Corollary 7.1. ☐

The notion of expanding circle map introduced in §19.13 admits the following natural generalization. A compact f-invariant set $Z \subset \mathbb{C}$ is called *expanding* or *hyperbolic*[6] (and also, f is called expanding/hyperbolic on Z) if there exist constants $C > 0$ and $\rho > 1$ such that

$$|Df^n(z)| \geq C\rho^n \text{ for any } z \in Z,\ n \in \mathbb{N}. \tag{20.3}$$

Of course, we can define the expanding property with respect to another Riemannian metric $\|\cdot\|$ on Z. Since all such metrics are equivalent, the expanding property is independent of a particular choice of the metric.

For instance, a *Cantor* Julia set $\mathcal{J}(f)$ of a quadratic polynomial f is always expanding (see Remark 20.6). Lemma 20.31 yields a useful general criterion:

[6]Should not be confused with hyperbolicity of a Riemann surface.

COROLLARY 20.32. *Let $\mathcal{O} \supset \overline{\mathcal{P}}_f$ be a forward invariant compact set (e.g., $\overline{\mathcal{P}}_f$ itself). Assume that a component Ω of $\mathbb{C} \smallsetminus \mathcal{O}$ intersects $f^{-1}(\mathcal{O})$. Then any f-invariant compact set $Z \subset \Omega$ is expanding.*

20.9. Higher degree polynomials. The above basic definitions and results admit a straightforward extension to higher degree polynomials

$$f : z \mapsto a_0 z^d + a_1 z^{d-1} + \cdots + a_d, \quad d \geq 2, \ a_0 \neq 0,$$

with obvious adjustments. For instance, the exceptional cases in Proposition 20.21 are polynomials affinely conjugate to $z \mapsto z^d$.

The following point should be kept in mind though: the Basic Dichotomy is not valid any more in the higher degree case. Instead, there is the following partial description of the topology of the Julia set:

- The Julia set $\mathcal{J}(f)$ (and the filled Julia set $\mathcal{K}(f)$) is connected if and only all the critical points c_i are non-escaping, i.e., $c_i \in \mathcal{K}(f)$.

- If all the critical points escape to ∞, then $\mathcal{J}(f)$ is a Cantor set on which the dynamics is conjugate to the Bernoulli shift with d symbols.[7] However, the reverse is not true anymore, e.g., there are cubic polynomials with one non-escaping critical point whose Julia sets are Cantor. The dynamics is not expanding on these Julia sets.

Note that the Basic Dichotomy is still valid in the case of *unicritical* polynomials, that is, the ones that have a single critical point. Any such polynomial is affinely conjugate to $z \mapsto z^d + c$, where c is defined uniquely up to multiplication by $e(1/(d-1))$.

PROJECT 20.33. *Work out the basic dynamical definitions and results in the case of higher degree polynomials.*

In the theory of quadratic maps f_c, higher degree polynomials still appear as the iterates of f_c. It is useful to know that they have the same Julia set:

EXERCISE 20.34. *Show that $\mathcal{K}(f^n) = \mathcal{K}(f)$ for any polynomial f.*

Let Crit_f be the set of critical points of f. Then similarly to the quadratic case, we will use notation

$$\mathrm{Crit}_f^n := \bigcup_{k=0}^{n} f^{-k}(\mathrm{Crit}_f), \quad \mathcal{P}_f := \bigcup_{k=1}^{\infty} f^k(\mathrm{Crit}_f), \quad \overline{\mathcal{P}}_f \equiv \mathrm{cl}\,\mathcal{P}_f \tag{20.4}$$

for the set of critical points of the iterates of f (including $n = \infty$) and for the post-valuable sets, respectively.

In the discussion below, we will often refer to polynomials of degree two as just "polynomials". If we want to emphasize that some results hold for arbitrary degree, we will do it explicitly.

[7]The latter result is non-trivial, but it is easy to show that *some iterate of f* is conjugate to the Bernoulli shift.

Notes. The Chebyshev map (§20.4.6) was first considered in the dynamical context by Ulam and von Neumann [**UvN**] who observed numerically that it exhibits a chaotic behavior on its invariant interval $\mathcal{I}$ (which is related to the existence of the absolutely continuous invariant measure μ on $\mathcal{I}$). This was probably the first occasion of using computers in dynamics.

The doubling bifurcation (§20.4.5) was discovered by Myrberg in the 1970s [**Myr1**].

21. Periodic motions

"Periodic solutions is the only opening through which we can try to penetrate to the domain that was viewed unaccessible" (Poincaré).

21.1. Rough classification of periodic points by the multiplier. Consider a periodic point α of period p. The local dynamics near its cycle $\boldsymbol{\alpha} = (f^n\alpha)_{n=0}^{p-1}$ depends first of all on its *multiplier*

$$\rho = (f^p)'(z) = \prod_{n=0}^{p-1} f'(f^n\alpha) \quad \text{(compare §19.5).}$$

The point (and its cycle)[8] is called *attracting* if $|\rho| < 1$ (compare §19.7). A particular case of an attracting point is a *superattracting* one when $\rho = 0$. In this case, the critical point 0 belongs to the cycle, and we sometimes call it

$$\mathbf{0} \equiv \{f^n(0)\}_{n=0}^{p-1} \equiv \{0_n\}_{n=0}^{p-1}.$$

When we want to emphasize that an attracting periodic point is not superattracting, we call it *simply* attracting.

A periodic point is called *repelling* if $|\rho| > 1$, and *neutral* if $\rho = e(\theta)$, $\theta \in \mathbb{R}/\mathbb{Z}$. In latter case, θ is called the *rotation number* of α. Local dynamics near a neutral cycle depends delicately on the arithmetic of the rotation number. A neutral point is called *parabolic* if the rotation number is rational, $\theta = \mathfrak{p}/\mathfrak{q}$, and is called *irrational* otherwise. An irrational periodic point can be of *Siegel* and *Cremer* type, to be defined below. Compare §19.8.1.

We will consider these cases one by one. We first analyze the dynamics locally, near the periodic point in question, and then globalize it to the basin (such a globalization will be called *semi-local*). For the local analysis, it is convenient to put the fixed point at the origin and consider a holomorphic map

$$f : z \mapsto \rho z + a_2 z^2 + \dots \tag{21.1}$$

nearby. We will often refer to this object as a *holomorphic germ* near 0 since we are interested in its local properties, which allows us to restrict its domain of definition (compare §50.1.)

Further semi-local analysis of periodic motion will be carried in §23.

21.2. Attracting cycles.

[8]All the terminology introduced for periodic points applies to their cycles, and vice versa.

21.2.1. *Fundamental annulus.* Let $\boldsymbol{\alpha}$ be an attracting cycle. The orbits of all nearby points uniformly converge to $\boldsymbol{\alpha}$ and, in particular, are bounded. It follows that attracting cycles belong to the Fatou set $\mathcal{F}(f)$. The rate of convergence is exponential in the simply attracting case and superexponential in the superattracting case.

For a simply attracting periodic point α, we say that a piecewise smooth (open) disk $P \ni \alpha$ is a (*attracting*) *petal* of α if $f|P$ is univalent and $f(P) \Subset P$. (For instance, one can take a small round disk $\mathbb{D}(\alpha, \varepsilon)$ as a petal.) Then the annulus $A = \overline{P} \smallsetminus f^p(P)$ is called a *fundamental annulus* of α.

In the superattracting case, a *petal* is a smooth disk $P \ni \alpha$ such that $f^p : P \to f^p(P)$ is a branched covering of degree d (with a single critical point at α), and $f(P) \Subset P$. (For instance, one can let P be the component of $f^{-p}(\mathbb{D}(\alpha, \varepsilon))$ containing α.) The corresponding *fundamental annulus* is $\overline{P} \smallsetminus f^p(P)$.

Exercise 21.1. *Being (super-)attracting/repelling/neutral is a topologically invariant property of a germ: If two germs f and $\tilde{f}$ are topologically conjugate and f is (super-)attracting (resp., neutral or repelling) then so is $\tilde{f}$.*

The corresponding statement is false over the reals: e.g., real germs $x \mapsto x/2$ and $x \mapsto x - x^3$ are topologically conjugate near the origin: see Exercise 19.25 (iii).

21.2.2. *Basin of attraction.* The *basin of attraction* of an attracting cycle $\boldsymbol{\alpha}$ is the set of all points whose orbits converge to $\boldsymbol{\alpha}$:

$$\mathcal{D}(\boldsymbol{\alpha}) = \mathcal{D}_f(\boldsymbol{\alpha}) = \{z : \ f^n z \to \boldsymbol{\alpha} \text{ as } n \to \infty.\}$$

Exercise 21.2. *Show that the basin $\mathcal{D}(\boldsymbol{\alpha})$ is a completely invariant union of components of* $\operatorname{int} \mathcal{K}(f)$.

The union of components of $\mathcal{D}(\boldsymbol{\alpha})$ containing the points of $\boldsymbol{\alpha}$ is called the *immediate basin* of attraction of the cycle $\boldsymbol{\alpha}$. We will denote it by $\mathcal{D}^\bullet(\boldsymbol{\alpha}) \equiv \mathcal{D}^\bullet_f(\boldsymbol{\alpha})$. The component of $\mathcal{D}^\bullet(\boldsymbol{\alpha})$ containing α will be denoted $\mathcal{D}^\bullet(\alpha) \equiv \mathcal{D}^\bullet_f(\alpha)$.

Exercise 21.3. (i) *The immediate basin of an attracting cycle $\boldsymbol{\alpha}$ consists of exactly p components, where p is the period of $\boldsymbol{\alpha}$.*
(ii) *Show that it can be constructed as follows. Let P_0 be a petal of a point $\alpha \in \boldsymbol{\alpha}$ and let P_n be defined inductively as the component of $f^{-p}(P_{n-1})$ containing α. Then*

$$P_0 \subset P_1 \subset P_2 \subset \dots \quad \text{and} \quad \mathcal{D}^\bullet(\alpha) = \bigcup_{n=0}^{\infty} P_n.$$

21.2.3. *Critical point in the basin.* We will now state one of the most important facts of the classical holomorphic dynamics:

Theorem 21.4. *The immediate basin of attraction $\mathcal{D}^\bullet_f(\boldsymbol{\alpha})$ of an attracting cycle $\boldsymbol{\alpha}$ contains the critical point* 0. *Moreover, if $\boldsymbol{\alpha}$ is simply attracting then* $\operatorname{orb}(0)$ *crosses any fundamental annulus A.*

Remark 21.5. Of course, the assertion is trivial when $\boldsymbol{\alpha}$ is superattracting as $0 \in \boldsymbol{\alpha}$ in this case.

Proof. Otherwise f^p would conformally map each component D of the immediate basin onto itself. Hence it would be a hyperbolic isometry of D, despite the fact that $|(f^p)'(\alpha)| < 1$.

To prove the second assertion (which would also give another proof of the first one), let us consider a petal P_0 containing some point $\alpha \in \boldsymbol{\alpha}$, and let us define P_n inductively as the component of $f^{-p}(P_{n-1})$ containing α (compare with Exercise 21.3 above). Then $P_0 \subset P_1 \subset P_2 \subset \dots$. If none of these domains contains a critical point of f^p, then the all the maps $f^p : P_n \to P_{n-1}$ are isomorphisms and all the P_n are topological disks. Hence their union, P_∞, is a topological disk as well, and $f^p : P_\infty \to P_\infty$ is an automorphism. Hence it is a hyperbolic isometry contradicting the fact that α is attracting.

Hence some P_n contains a critical point of f^p. Take the first such n (obviously, $n \geq 1$). Then $P_{n-1} \smallsetminus f^p(P_{n-1})$ contains a critical value of f^p, which is contained in orb(0). Applying further iterates of f^p, we will bring it to the fundamental annulus $\overline{P}_0 \smallsetminus f^p(P_0)$. □

Remark 21.6. The above argument proves a more general statement for *proper attracting basins. Let $f : (D, \alpha) \to (D, \alpha)$ be a holomorphic endomorphism of a hyperbolic Riemann surface which has an attracting fixed point α. Then f has a critical point.* (The second assertion concerning the fundamental annulus also holds.)

Corollary 21.7. *A quadratic polynomial can have at most one attracting cycle. If it has one, all other cycles are repelling.*

Proof. The first assertion is immediate. For the second one, notice that under the circumstances, the postcritical set $\overline{\mathcal{P}}_f$ is a discrete set accumulating on the attracting cycle $\boldsymbol{\alpha}$. Hence it does not divide the complex plane, and $0 \in \mathbb{C} \smallsetminus \overline{\mathcal{P}}_f$. Applying Lemma 20.31, we conclude that $|\rho(\beta)| = \|Df^q(\beta)\|_{\text{hyp}} > 1$ for any other periodic point β of period q. □

Of course, the period of the attracting cycle can be arbitrary big. A quadratic polynomial is called *hyperbolic* if it either has an attracting cycle, or if its Julia set is Cantor. (The unifying property is that for hyperbolic maps, orb(0) converges to an attracting cycle in the Riemann sphere.) For instance, the doubling map $z \mapsto z^2$, the basilica map $z \mapsto z^2 - 1$ (see Figure 20.4), and the rabbit map (see Figure 20.5) are all hyperbolic. Though dynamically non-trivial, it is a well understood class of quadratic polynomials (see §25).

In what follows, we will usually mark an attracting cycle

$$\boldsymbol{\alpha} \equiv (f^n \alpha)_{n=0}^{p-1} \equiv (\alpha_n)_{n=0}^{p-1}$$

so that the immediate basin $\mathcal{D}_0 \equiv \mathcal{D}^\bullet(\alpha)$ contains the critical point 0. This domain and various associated objects will be called *central*. For instance, the attracting periodic point $\alpha \equiv \alpha_0 \in \mathcal{D}_0$ is "central".

The immediate basin $\mathcal{D}^v \equiv \mathcal{D}_1 \equiv \mathcal{D}(\alpha_1)$ containing the critical value v, and associated objects (e.g., the periodic point $\alpha^v \equiv \alpha_1 \equiv f(\alpha) \in \mathcal{D}^v$) will be called *valuable*.

21.2.4. *Real basin of attraction.* Let us now consider a real quadratic polynomial $f = f_c$, $c \in [-2, 1/4]$, that has an attracting cycle $\boldsymbol{\alpha} = (\alpha_k)_{k=0}^{p-1} \in \mathbb{C}$. Then in fact $\boldsymbol{\alpha}$ is real since $f^n(0) \to \boldsymbol{\alpha}$.

Let $\mathcal{D}_\mathbb{R}(\boldsymbol{\alpha}) \subset I$ be the real attracting basin of $\boldsymbol{\alpha}$. Of course, it is open, so it is the disjoint union of open intervals. The *real immediate basin* $\mathcal{D}^\bullet_\mathbb{R}(\alpha_k)$ of α_k is the component of $\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ containing α_k, and the real immediate basin $\mathcal{D}^{\bullet}_{\mathbb{R}}(\boldsymbol{\alpha})$ is the union of the $\mathcal{D}^{\bullet}_{\mathbb{R}}(\alpha_k)$.

COROLLARY 21.8. (i) *We have:* $\mathcal{D}^{\bullet}_{\mathbb{R}}(\alpha_k) = \mathcal{D}^{\bullet}(\alpha_k) \cap \mathbb{R}$.

(ii) *The real immediate basin* $\mathcal{D}^{\bullet}_{\mathbb{R}}(\boldsymbol{\alpha})$ *contains the critical point* 0.

PROOF. Assertion (i) follows from Exercise 20.13. Together with Theorem 21.4, it implies (ii). $\square$

21.3. Parabolic cycles.

21.3.1. *Local expansions.* Let us consider a parabolic germ

$$f : z \mapsto e(\mathfrak{p}/\mathfrak{q}) \cdot z + a_{k+1} z^{k+1} + \dots, \quad k \geq 1,\ a_{k+1} \neq 0, \tag{21.2}$$

with rotation number $\theta = \mathfrak{p}/\mathfrak{q} \in \mathbb{R}/\mathbb{Z}$ near the origin. If $\theta = 0$, we call f (and the corresponding parabolic point) *primitive*; otherwise we call it *satellite*.

EXERCISE 21.9. *If* $f^{\mathfrak{q}} = \mathrm{id}$ *then* f *is locally conformally conjugate to the rotation* $z \mapsto e(\mathfrak{p}/\mathfrak{q}) \cdot z$. *Otherwise,* $f^{\mathfrak{q}}$ *admits a local expansion*

$$f^{\mathfrak{q}}(z) = z + b_{k+1} z^{k+1} + \dots, \quad k \geq 1,\ b_{k+1} \neq 0,$$

with $k = \mathfrak{q} l$ *for some* $l \in \mathbb{Z}_+$. *Moreover,*

$$f^{\mathfrak{q} n} z = z + n\, b_{k+1} z^{k+1} + \dots .$$

We call l the *order of degeneracy* of f at 0. In the case $l = 1$, the parabolic germ f is called *non-degenerate*. In case when $f^{\mathfrak{q}} = \mathrm{id}$ we can formally let $l = \infty$ (of course, this is impossible if f is a polynomial of degree $d \geq 2$).

21.3.2. *Leau-Fatou Flowers.* An open Jordan disk P is called an *attracting petal* for f if:

- $0 \in \partial P$;
- $f^{\mathfrak{q}}(P) \subset P$ and $f^{\mathfrak{q}} | P$ is univalent;
- 0 is the only point where ∂P and $\partial(f^{\mathfrak{q}} P)$ touch;
- $f^{\mathfrak{q} n} z \to 0$ as $n \to \infty$ locally uniformly in P.

Given such a petal, the set $\overline{P} \smallsetminus f(P)$ is called an attracting *fundamental crescent* $\mathcal{C}^{\mathrm{a}}$.

EXERCISE 21.10. *Show that the last condition in the definition of the petal follows from the first three.*

We say that a petal P has a γ-*wedge* at 0 if both local branches of the boundary $\partial P \smallsetminus \{0\}$ have tangent lines at 0 that meet at angle $2\pi\gamma$. A *bisector* L_i of such a petal is a smooth curve landing at 0 that divides the γ-wedge into two $(\gamma/2)$-subwedges. We say that an orbit (z_n) is *asymptotic* to the bisectors of the petals if there is the union of $\mathfrak{q}$ bisectors L_i cyclically permuted by f such that for any $\varepsilon > 0$, the points z_n are eventually trapped in the union of the ε-wedges around the L_i.

Two attracting petals are called *equivalent* if they overlap[9].

[9]At the moment, it is not evident that this is an equivalence relation, but the next theorem shows that it is.

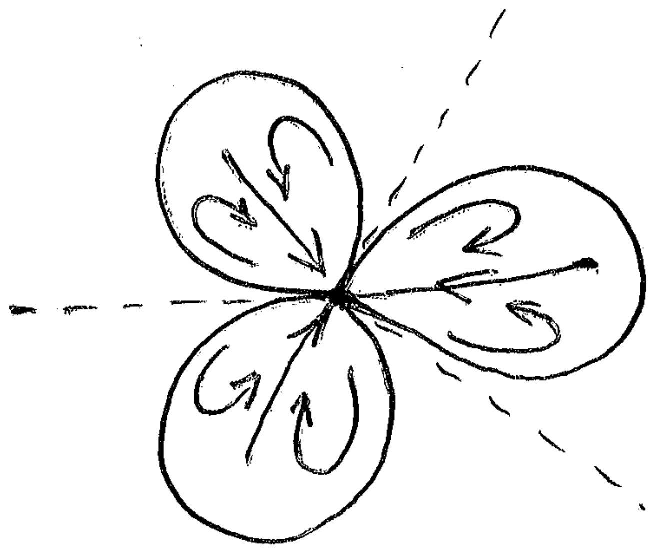

FIGURE 21.1. Attracting flower of a parabolic point.

THEOREM 21.11 (Flower Theorem). *For a parabolic germ* (21.2) *which is not conformally conjugate to a rotation, there is a choice of disjoint* $l\mathfrak{q}$ *petals* P_i *with wedge* $1/\mathfrak{q}l$ *at* 0 *such that the flower* $\Phi \equiv \Phi^{\mathrm{a}} := \bigcup P_i$ *is invariant under rotation by* $1/\mathfrak{q}l$ *and under* f. *The orbits of points* $z \in \Phi$ *converge to* 0 *locally uniformly and asymptotically to the bisectors of the petals. Moreover, for any* $\gamma' \in (0, 1/\mathfrak{q}l)$, *there is a similar flower* $\Phi' \subset \Phi$ *with wedge* γ' *at* 0 *in which convergence to* 0 *is uniform.*

Vice versa, if some $\operatorname{orb} Z$ *converges to* 0 *without direct landing at* 0 *then eventually it lands in any flower* Φ'.

The smaller flowers Φ' and the corresponding petals will be called *uniformly attracting*.

PROOF. The proof will be split in several cases. The main analysis happens in the following one:

The germ f *is primitive and non-degenerate*:

$$f : z \mapsto z + az^2 + \dots,\ a \neq 0. \tag{21.3}$$

Conjugating f by complex scaling $\zeta = az$ we make $a = 1$.

Let us move the fixed point to ∞ by inversion $Z = -\dfrac{1}{z}$. It brings f to the form

$$F : Z \mapsto Z + 1 + O(\frac{1}{Z}) \tag{21.4}$$

near ∞. It is obvious from this asymptotical expression that any right half-plane

$$Q_t = \{Z :\ \operatorname{Re} Z > t\} \tag{21.5}$$

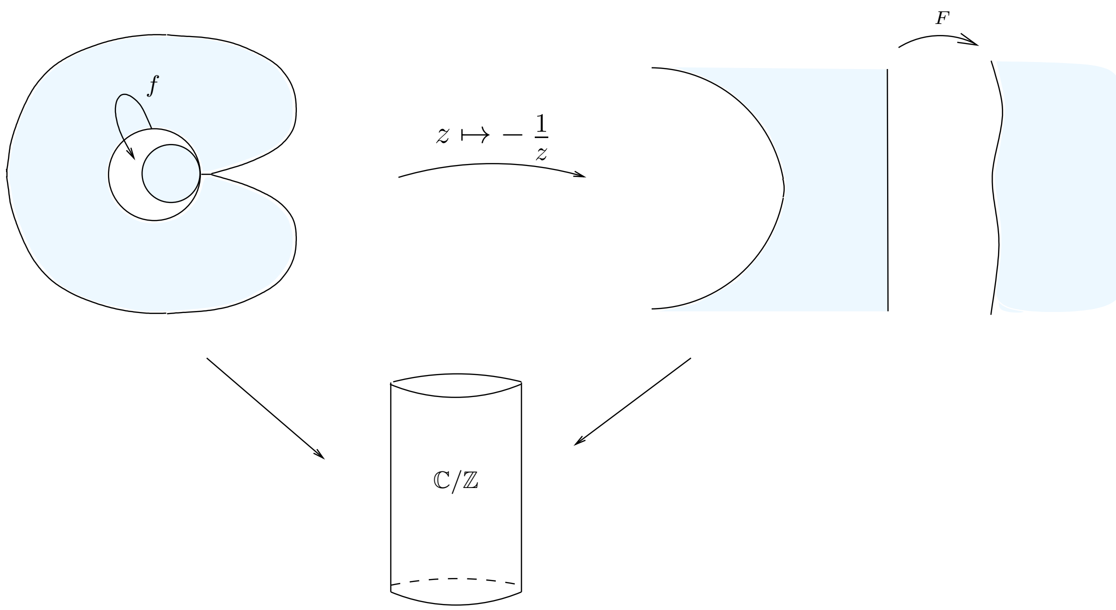

Figure 21.2. Parabolic petal in the non-degenerate case and the corresponding Écalle-Voronin cylinder.

with $t > 0$ sufficiently big is invariant under F, and in fact

$$F(Q_t) \subset Q_{t+1-\varepsilon}, \tag{21.6}$$

where $\varepsilon = \varepsilon(t) \to 0$ as $t \to \infty$. So, such a half-plane provides us with a petal with wedge $1/2$ at ∞. Moreover, for any $\varepsilon > 0$ eventually

$$\operatorname{Re} Z_n \geq \operatorname{Re} Z + (1-\varepsilon)n, \quad \operatorname{Im} Z_n \leq \varepsilon n, \text{ where } Z_n \equiv F^n Z, \tag{21.7}$$

so the orbits in Q_t converge to ∞ locally uniformly and asymptotically horizontally. (Note that the horizontal direction is the bisector of the petal P at ∞).

Vice versa, if $Z_n \to \infty$ without direct landing at ∞, then due to asymptotical expression (21.4) we eventually have $\operatorname{Re} Z_{n+1} \geq \operatorname{Re} Z_n + 1 - \varepsilon$. Hence $\operatorname{Re} Z_n \to +\infty$ and $\operatorname{orb} Z$ eventually lands in the half-plane Q_t.

Now we would like to enlarge Q_t to a petal P with the full wedge 1 at ∞. To this end let us consider two logarithmic curves

$$\Gamma_\pm = \{Y = \pm C \log(t - X + 1) + R)\}, \quad X \leq t, \text{ where } Z = X + iY.$$

If R is big enough then $\Gamma_\pm$ lie in the domain where the asymptotics (21.4) applies. If C is big enough then the half-slope of these curves is bigger (in absolute value) than the slope of the displacement vector $F(Z) - Z$. It follows that F moves the curves $\Gamma_\pm$ to the right, and the region $P \supset Q_t$ bounded by these curves and the segment of the vertical line $\operatorname{Re} Z = t$ in between is mapped univalently into itself. This is the desired petal with wedge 1 at ∞.

A petal Q' with an intermediate angles $\gamma' \in (0, 1)$ can be obtained as the wedge $\{|\arg(Z - t')| < 2\pi\gamma'\}$ centered at a sufficiently big real point $t' \in \mathbb{R}_+$. It is easy to see (as for the half-planes Q_t above) that convergence is uniform in such a wedge, and that all orbits converging to ∞ eventually land in it.

Let f *be a general primitive parabolic germ*:

$$f : z \mapsto z + bz^{k+1} + \dots \text{ with } k \geq 1,\ b \neq 0. \tag{21.8}$$

Again, conjugating f by a complex scaling $\zeta = \lambda z$, where $\lambda^k = b$, we make $b = 1$.

Let us now use a non-invertible change of variable $\zeta = z^k$. A formal calculation shows that it conjugates f to a multi-valued germ

$$g : \zeta \mapsto \zeta + \zeta^2 + O(|\zeta|^{2+1/k})$$

where the residual term is given by a power series in $\zeta^{1/k}$. (Such an expression is called *Puiseux series*.) Making now a change of variable $Z = -1/z$, we come up with a multi-valued germ

$$G : Z \mapsto Z + 1 + O(1/|Z|^{1/k}) \tag{21.9}$$

near ∞. Let us consider any single-valued branch of this germ on the slit plane $\mathbb{C} \smallsetminus \mathbb{R}_-$. Then the same considerations as in the non-degenerate case show that G has desired petals. Lifting these petals to the z-plane provides us with desired attracting flowers Φ^{a} for f.

The *satellite parabolic case with rotation number* $\mathfrak{p}/\mathfrak{q}$ is reduced to the above by considering $f^{\mathfrak{q}}$. $\square$

Notice that in the half-plane model (21.9), the attracting fundamental crescent $\mathcal{C}^{\mathrm{a}}$ becomes a vertical topological strip $\mathrm{cl}(G(Q_t)) \smallsetminus Q_t$.

21.3.3. *More on the local dynamics.* The above picture provides us with a simple description of *small orbits* near 0:

Exercise 21.12. *Let f be a parabolic germ near* 0. *There is an $\varepsilon > 0$ such that if* $\operatorname{orb} z$ *of some point $z \neq 0$ stays in the ε-neighborhood of* 0 *then this orbit is eventually trapped in the attracting flower Φ^{a}, and thus, $f^n z \to 0$ locally uniformly near* 0.

Exercise 21.13. *Let f be a primitive non-degenerate parabolic germ* (21.3). *Then for any orbit $z_n \to 0$, except for the stationary one, we have:*

$$z_n \asymp \frac{1}{n}, \quad |z_{n+1} - z_n| \asymp \frac{1}{n^2}.$$

What is the convergence rate in the degenerate case (21.8)? *For a satellite parabolic germ* (21.2)?

Exercise 21.14. (i) *For a parabolic germ with zero rotation number* (21.8), *there is a smooth invariant curve γ landing at* 0, *and any such curve is a bisector of an attracting petal.*

(ii) *For a parabolic germ with rotation number $\mathfrak{p}/\mathfrak{q}$, there is a smooth periodic curve γ landing at* 0; *moreover, any such curve has period $\mathfrak{q}$ and is a bisector of an attracting petal.*

Applying the above discussion to the local inverse f^{-1}, we obtain *repelling petals*, *repelling Leau-Fatou flowers* Φ^{r}, *repelling fundamental crescents* $\mathcal{C}^{\mathrm{r}}$, etc.

Exercise 21.15. *Show that Φ^{r} can be obtained from Φ^{a} by rotating through angle* $1/(2\mathfrak{q})$ (*measured in revolutions*).

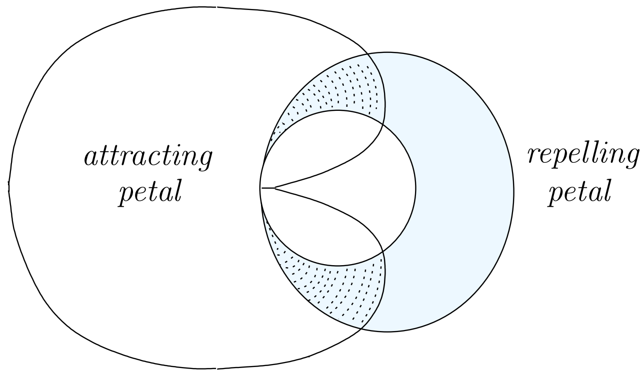

FIGURE 21.3. Repelling crescent and rectangle for a non-degenerate parabolic point with multiplier 1.

It follows that the union $\Phi^{\mathrm{a}} \cup \Phi^{\mathrm{r}} \cup \{0\}$ is a neighborhood of 0, which gives us full understanding of the local dynamics near a parabolic point. In particular, it provides us with *two-sided small orbits* that converge to 0 in both positive and negative time.

Removing from repelling fundamental crescent $\mathcal{C}^{\mathrm{r}}$ an attracting flower Φ^{a}, we obtain a *repelling fundamental rectangle* Δ^{r} (see Figure 21.3). It is a nice fundamental domain for the space of backward orbits converging to the parabolic point. We will refer to the boundary intervals of Δ^{r} contained in $\partial\mathcal{C}^{\mathrm{r}}$ as *vertical sides* of Δ^{r}.

EXERCISE 21.16. (i) *A parabolic germ is never topologically conjugate to a hyperbolic one (attracting or repelling).*
(ii) *Let $Q \ni 0$ be a locally completely invariant compact set for a parabolic germ f. Then there is no qc map $h : (\mathbb{C}, Q, 0) \to (\mathbb{C}, \tilde{Q}, 0)$ that conjugates $f\,|\,Q$ to a hyperbolic germ $\tilde{f}$ restricted to $\tilde{Q}$.*

The above discussion leads to a complete topological (and in fact, quasiconformal) classification of parabolic germs:

PROBLEM 21.17. *Show that two parabolic germs* (21.2) *are locally topologically (and in fact, qc) conjugate if and only if they have the same rotation number $\mathfrak{p}/\mathfrak{q}$ and the same order of degeneracy l.*

21.3.4. *Real parabolic germs.* In conclusion, let us take a look at a real parabolic germ

$$f : (\mathbb{R}, 0) \to (\mathbb{R}, 0), \quad z \mapsto \rho z + b z^{k+1} + \dots, \quad \rho \in \{\pm 1\},\ b \in \mathbb{R}^*. \tag{21.10}$$

It is *primitive* or *satellite* depending on whether $\rho = 1$ or $\rho = -1$. In the latter case, the points *flip* form one side of 0 to the other under the dynamics.

EXERCISE 21.18. *For a real parabolic germ, the Leau-Fatou flower of f is $\mathbb{R}$-symmetric. Each ray $\mathbb{R}_\pm$ is a bisector of an attracting or repelling petal $P_\pm$.*

If both of the above petals $P_\pm$ are attracting then f (and its parabolic point) is called *parabolic-attracting* (on the real line). In this case, $f^n x \to 0$ for all real x near 0. Similarly, if both petals are repelling then f is called *parabolic-repelling.* If one of the petals is attracting while the other is repelling then f is called *parabolic-semi-attracting*. In this case, orbits are attracted to 0 on one side and are repelled on the other. It is possible only in the primitive case.

EXERCISE 21.19. (i) *Classify the local topological* (*quasisymmetric*) *dynamics for real germs* (21.10) *according to* ρ, k, *and* $\operatorname{sign} b$.
(ii) *Show that a real parabolic germ is never qs conjugate to a repelling or attracting germ* (*even on one side of* 0).

21.3.5. *Écalle-Voronin cylinders.* Let us take an attracting petal P. The return map $g := f^{p\mathfrak{q}}$ transforms one boundary component of the fundamental crescent $\overline{P} \smallsetminus f(P)$ to the other, so the quotient $\mathrm{Cyl} \equiv \mathrm{Cyl}_f := P/g$ is a topological cylinder called *Écalle-Voronin cylinder*. *A priori*, there are several options for the conformal type of Cyl_f: it can be isomorphic to an annulus $\mathbb{A}(r, R)$, or to the punctured disc $\mathbb{H}/ < z+1 > \approx \mathbb{D}^*$, or to the bi-infinite cylinder $\mathbb{C}/ < z+1 > \approx \mathbb{C}^*$. In fact, the latter happens:

LEMMA 21.20. *The Écalle-Voronin cylinder* Cyl_f *is isomorphic to the bi-infinite cylinder.*

PROOF. Notice first that the cylinder $\mathrm{Cyl} = P/ < g >$ is independent of the petal P in the equivalence class, so we can make any convenient choice. Let us use for this purpose the half-plane $Q \equiv Q_t$ (21.5) near ∞. Then the fundamental crescent $\overline{Q} \smallsetminus F(Q)$ becomes a vertical strip S whose boundary curves stay distance ~ 1 apart, by (21.4). Moreover, the right-hand boundary curve $Y \mapsto F(t+iY)$ is almost vertical, so each straight interval $I_Y := [t+iY, F(t+iY)]$ cuts S into two half-strips. Projecting these intervals to the cylinder $\mathrm{Cyl} \approx S/F$, we obtain a horizontal foliation Γ on S by circles that we also denote I_Y.

Let $\mathrm{Cyl}^{\pm}$ be the half-cylinders obtained by cutting Cyl by the circle I_0. It is enough to show that

$$\operatorname{mod} \mathrm{Cyl}^{\pm} \equiv \mathcal{W}(\Gamma|\,\mathrm{Cyl}^{\pm}) = \infty. \tag{21.11}$$

Let us deal with Cyl^{+} for definiteness. Let us further cut the cylinder on some big height $H > 0$ by the circle I_H, and call the corresponding finite cylinder Cyl^{+}_{H}. Put any conformal metric $\rho = \rho(z)|dz|$ on Cyl^{+}_{H} with

$$\int_{I_Y} \rho\, dl_Y = l_\rho(I_Y) \geq \mathcal{L}_\rho(\Gamma|\,\mathrm{Cyl}^{+}_{H}) \geq 1 \quad \text{for any } Y \in [0, H], \tag{21.12}$$

where dl_Y is the Euclidean length element along I_Y. Since the circles I_Y are almost horizontal, we have for the Euclidean area form $dm \geq (1/2)\, dl_Y\, dY$. Hence, integrating (21.12) over dY gives us:

$$\int_{C^{+}_{H}} \rho\, dm \geq \frac{1}{2} H.$$

By the Cauchy-Schwarz Inequality (compare (6.1)), we obtain:

$$H\, m_\rho(\mathrm{Cyl}^{+}_{H}) \asymp \operatorname{area}(\mathrm{Cyl}^{+}_{H}) \int \rho^2\, dm \geq \left(\int \rho\, dm\right)^2 \geq \frac{1}{4} H^2.$$

Taking the infimum over all admissible ρ, we obtain $\operatorname{mod}(\mathrm{Cyl}^{+}_{H}) \geq cH$ with $c > 0$ independent of ρ. So, $\operatorname{mod}(\mathrm{Cyl}^{+}_{H}) \to \infty$ as $H \to \infty$, and we are done. $\square$

EXERCISE 21.21. *Show that the return map* $f^{p\mathfrak{q}} : \mathcal{D} \to \mathcal{D}$ *to the immediate parabolic basin* $\mathcal{D} \supset P$ *has zero displacement:* $\inf_{z\in\mathcal{D}} \operatorname{dist}_{\mathrm{hyp}}(z, f^{p\mathfrak{q}}z) = 0$.

EXERCISE 21.22. *Let us glue the boundary components of the strip* $\overline{\mathbb{S}} = \{z : \ 0 \le \operatorname{Im} z \le 1\}$ *by the map* $\mathbb{R} \to \mathbb{R} + i$, $x \mapsto \lambda x + i$, *where* $\lambda > 1$. *Show that the quotient* $\mathbb{S}/\sim$ *is a cylinder of finite modulus.*

21.3.6. *Polynomial case: parabolic basin and the critical point.* Let us now go back to quadratic polynomials, so $f \equiv f_c$. Consider a parabolic periodic point α with period p and rotation number $\mathfrak{p}/\mathfrak{q}$.

PROPOSITION 21.23. *Any parabolic point* α *belongs to the Julia set* $\mathcal{J}$.

PROOF. Let us give two proofs for this simple assertion.

As we have seen (Exercise 21.9), there is a $k \ge 2$ such that the kth Taylor coefficient of the iterates $f^{p\mathfrak{q}n}$ at α grows as $n \to \infty$. By the Cauchy estimate, it is impossible if the family of iterates is bounded near α.

Another way is to apply the local dynamical picture near α: existence of a repelling flower implies Lyapunov instability near α. $\square$

Let us now apply general notions from §19.7 to our setting. The *parabolic realm* and the *parabolic basin* of a parabolic cycle $\boldsymbol{\alpha}$ are defined as follows:

$$\mathcal{R}(\boldsymbol{\alpha}) \equiv \mathcal{R}_f(\boldsymbol{\alpha}) := \{z : f^n z \to \boldsymbol{\alpha} \text{ as } n \to \infty\}, \quad \mathcal{D}(\boldsymbol{\alpha}) \equiv \mathcal{D}_f(\boldsymbol{\alpha}) := \operatorname{int} \mathcal{R}(\boldsymbol{\alpha}).$$

EXERCISE 21.24. *Let* $\boldsymbol{\alpha}$ *be a parabolic cycle of a polynomial* f. *Then:*

(i) $\mathcal{D}(\boldsymbol{\alpha}) = \mathcal{R}(\boldsymbol{\alpha}) \smallsetminus \operatorname{Orb}_-(\boldsymbol{\alpha}) \equiv \{z : f^n z \to \boldsymbol{\alpha} \text{ but } f^n z \notin \boldsymbol{\alpha} \ \forall\ n \in \mathbb{N}\}$;

(ii) *The basin* $\mathcal{D}(\boldsymbol{\alpha})$ *is a completely invariant union of components of* $\operatorname{int}\mathcal{K}$*;*

(iii) *Among these components there are* $p\mathfrak{q}l$ *components attached to* $\boldsymbol{\alpha}$ *and permuted by* f, *while all others are preimages of these.*

The union of these $p\mathfrak{q}l$ components is called the *immediate parabolic basin* of $\boldsymbol{\alpha}$. It will be denoted as $\mathcal{D}^\bullet(\boldsymbol{\alpha}) \equiv \mathcal{D}^\bullet_f(\boldsymbol{\alpha})$. Each of these components is periodic with period $p\mathfrak{q}$. So, the immediate basin comprises l cycles of periodic components.

As in the attracting case, we have:

THEOREM 21.25. *The immediate parabolic basin* $\mathcal{D}^\bullet(\boldsymbol{\alpha})$ *of a parabolic cycle contains a critical point. In fact, each cycle of components of* $\mathcal{D}^\bullet(\boldsymbol{\alpha})$ *contains a critical point.*

PROOF. Let D be a component of $\mathcal{D}^\bullet(\boldsymbol{\alpha})$. If it does not contain critical points of $g := f^{p\mathfrak{q}}$, then $g : D \to D$ is an (unbranched) covering, and hence an automorphism of D (since D is simply connected). Since the orbits of g in D escape to infinity (of D) and $D \approx \mathbb{D}$ is hyperbolic, the quotient $D/ < g >$ is isomorphic to either an annulus $\mathbb{A}(r, R)$ (if g is hyperbolic) or to the punctured disc $\mathbb{D}^*$ (if g is parabolic), contradicting Lemma 21.20. $\square$

As in the hyperbolic case, we now conclude:

COROLLARY 21.26. *A quadratic polynomial* f *can have at most one parabolic cycle* $\boldsymbol{\alpha}$. *Moreover, this cycle is non-degenerate: there is a unique cycle of petals attached to it.* (*Thus, there are* $p\mathfrak{q}$ *petals attached to* $\boldsymbol{\alpha}$, *where* p *is the period of* $\boldsymbol{\alpha}$ *and* $\mathfrak{p}/\mathfrak{q}$ *is its rotation number.*)

If f *has a parabolic cycle, then all other cycles are repelling.*

Such a quadratic polynomial is naturally called *parabolic*.

As in the hyperbolic case, we call the immediate basin $\mathcal{D}_0 \equiv \mathcal{D}(\alpha) \ni 0$ (and all associated object) *central*, while we call the immediate basin $\mathcal{D}_1 \equiv f(\mathcal{D}_0) \ni c$ (and all associated objects) *valuable*. In particular, the parabolic point $\alpha \in \mathcal{D}_0$ is *central*, while $\alpha^v \equiv \alpha_1 \in \mathcal{D}_1$ is *valuable*.

In conclusion, let us mention one consequence of Exercise 21.12:

PROPOSITION 21.27. *Let f be a parabolic quadratic polynomial with a parabolic cycle $\boldsymbol{\alpha}$. There exists an $\varepsilon > 0$ such that for any $\operatorname{orb} z$ in the Julia set $\mathcal{J}$ that does not land in $\boldsymbol{\alpha}$, there exists an infinite sequence of moments n_k such that $\operatorname{dist}(f^{n_k} z, \boldsymbol{\alpha}) \geq \varepsilon$.*

In other words, the *transit map through the ε-neighborhood of $\boldsymbol{\alpha}$ is well defined for all orbits in the Julia set that do not land in $\boldsymbol{\alpha}$.*

21.3.7. *Real parabolic basin.* Now, the discussion of attracting cycles for real maps (§21.2.4) can be carried to the parabolic case. Let $f = f_c$, $c \in [-2, 1/4]$, be a real polynomial that has a parabolic cycle $\boldsymbol{\alpha} = \{\alpha_k\}_{k=0}^{p-1} \in \mathbb{C}$. Then $\boldsymbol{\alpha}$ is real since $f^n(0) \to \boldsymbol{\alpha}$.

The multiplier of $\boldsymbol{\alpha}$ is either 1 or -1, and accordingly, we call $\boldsymbol{\alpha}$ (and f itself) *primitive* or *satellite* (≡*flip*) *parabolic* (see §§19.8.1, 21.3.1). These cases are dynamically quite different:

EXERCISE 21.28. *In the primitive parabolic case, $\boldsymbol{\alpha}$ is locally topologically semi-attracting, while it is locally topologically attracting in the satellite case.*

Similarly to §21.3.6, let us consider the *real parabolic realm* and the *real parabolic basin* of $\boldsymbol{\alpha}$:

$$\mathcal{R}_{\mathbb{R}}(\boldsymbol{\alpha}) = \{x \in \mathcal{I} : f^n x \to \boldsymbol{\alpha}\}, \quad \mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha}) = \operatorname{int} \mathcal{R}_{\mathbb{R}}(\boldsymbol{\alpha}).$$

Obviously, $\mathcal{R}_{\mathbb{R}}(\boldsymbol{\alpha}) = \mathcal{R}(\boldsymbol{\alpha}) \cap \mathbb{R}$, but the situation with the real basin is more subtle.

Being open, $\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha}) \subset \mathcal{I}$ is the disjoint union of open intervals. The *real immediate parabolic basin* $\mathcal{D}^{\bullet}_{\mathbb{R}}(\alpha_k)$ of α_k is the component of $\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ such that $\operatorname{cl} \mathcal{D}^{\bullet}_{\mathbb{R}}(\alpha_k)$ contains α_k, and the *real immediate parabolic basin* $\mathcal{D}^{\bullet}_{\mathbb{R}}(\boldsymbol{\alpha})$ is the union of the $\mathcal{D}^{\bullet}_{\mathbb{R}}(\alpha_k)$.

EXERCISE 21.29. (i) *In the primitive case, $\boldsymbol{\alpha} \in \mathcal{R}_{\mathbb{R}}(\boldsymbol{\alpha}) \smallsetminus \mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ and*

$$\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha}) = \mathcal{D}(\boldsymbol{\alpha}) \cap \mathbb{R}.$$

(ii) *In the satellite case, $\boldsymbol{\alpha} \in \mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ and*

$$\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha}) = \mathcal{R}_{\mathbb{R}}(\boldsymbol{\alpha}) = (\mathcal{D}(\boldsymbol{\alpha}) \cap \mathbb{R}) \cup \operatorname{Orb}_{-}(\boldsymbol{\alpha}).$$

(iii) *The real immediate basin $\mathcal{D}^{\bullet}_{\mathbb{R}}(\boldsymbol{\alpha})$ contains the critical point 0. Moreover,*

$$0 \notin \operatorname{Orb}_{-}(\boldsymbol{\alpha}).$$

21.4. Repelling cycles. Let us now consider a repelling cycle $\boldsymbol{\alpha} = (f^k \alpha)_{k=0}^{p-1}$. Nearby points escape (exponentially fast) from a small neighborhood of $\boldsymbol{\alpha}$, which implies that the family of iterates f^n is not normal near $\boldsymbol{\alpha}$. Hence repelling periodic points belong to the Julia set. In fact, as we are about to demonstrate, they are dense in the Julia set, so that the Julia can be alternatively defined as the closure of repelling cycles. It gives us a view of the Julia set "from inside".

But first, let us now show that almost all cycles are repelling:

LEMMA 21.30. *A quadratic polynomial may have at most two non-repelling cycles.*

PROOF. Let $\alpha_\circ$ be a neutral periodic point of period p with multiplier $\rho_\circ$ of a quadratic polynomial $f_\circ : z \mapsto z^2 + c_\circ$. Due to Corollary 21.26, we can assume that $\rho_\circ \neq 1$. Then by the Implicit Function Theorem, the equation $f^p(z) = z$ has a local holomorphic solution $z = \alpha_c$ assuming value $\alpha_\circ$ at $c_\circ$. The multiplier of this periodic point, $\rho_c = (f_c^p)'(\alpha_c)$ is also a local holomorphic function of c. In fact, it is a global algebraic function. So, if it was locally constant then it would be globally constant, and the map $f_0 : z \mapsto z^2$ would have a neutral cycle. Since this is not the case, the multiplier is not constant, and hence near $c_\circ$ it assumes all values in some neighborhood of $\rho_\circ$. In particular, it assumes values with $|\rho| < 1$. Moreover, if near $c_\circ$

$$\rho(c) = \rho_\circ + a(c - c_\circ)^k + \dots, \quad a \neq 0,$$

then the set $\{c : \ |\rho(c)| < 1\}$ is the union of k sectors that asymptotically occupy $1/2$ of the area of a small disk $\mathbb{D}(c_\circ, \varepsilon)$. It follows that if we take three of such multiplier functions, then two of them must have overlapping sectors, so that the corresponding two cycles can be made simultaneously attracting, contradicting Corollary 21.7. □

THEOREM 21.31. *The Julia set is the closure of repelling cycles.*

PROOF. Let us first show that any point of the Julia set can be approximated by a periodic point. Let $z \in \mathcal{J}(f)$ be a point we want to approximate. Since the Julia set does not have isolated points (see Corollary 20.7), we can assume that z is not the critical value. Then in a small neighborhood $U \ni z$, there exist two branches of the inverse function, $\phi_1 = f_1^{-1}$ and $\phi_2 = f_2^{-1}$. Since the family of iterates is not normal in U, one of the equations, $f^n z = z$, $f^n z = \phi_1(z)$, or $f^n z = \phi_2(z)$, has a solution in U for some $n \geq 1$ (by the Refined Montel Theorem). If it is an equation of the first series, we find in U a periodic point of period n (maybe, not the least one). Otherwise, we find a periodic point of period $n + 1$.

Since by Lemma 21.30, almost all periodic points are repelling, we come to the desired conclusion. □

A quadratic polynomial is called *periodically repelling* if all of its cycles (in $\mathbb{C}$) are repelling.

21.5. Topological exactness (leo property). We know from Corollary 20.24 that any polynomial is topologically transitive on its Julia set. In fact, it enjoys stronger mixing properties:

PROPOSITION 21.32. *Any polynomial f is topologically exact, and hence topologically mixing, on its Julia set.*

This follows from a slightly sharper statement:

LEMMA 21.33. *Let $f : z \mapsto z^2 + c$, and let Q be a compact subset of $\mathbb{C}$ for which we assume that in case $c = 0$, $Q \not\ni 0$. Then for any neighborhood $U \subset \mathbb{C}$ intersecting the Julia set $\mathcal{J}(f)$, we have $f^n(U) \supset Q$ for all n sufficiently big.*

PROOF. Enlarging Q if needed, we can assume without loss of generality that $Q = \mathbb{D}_R$ for $c \neq 0$ and $Q = \mathbb{A}[R^{-1}, R]$ for $c = 0$, where R is so big that $f(Q) \supset Q$. Then it is enough to show that $f^n(U) \supset Q$ for some $n \in \mathbb{N}$.

By Theorem 21.31, U contains a repelling periodic point α of some period p. Hence, U contains a little disk $D \ni \alpha$ such that $f^p(D) \supset D$.

Let $D_k = f^k(D)$. By the higher degree version of Proposition 20.21 applied to f^p (see the discussion in §20.9), we have: $\bigcup_{m\in\mathbb{N}} D_{pm} \supset Q$. Since

$$D \subset D_p \subset D_{2p} \subset \dots,$$

one of these sets, say D_{pm}, contains Q. All the more, $f^{pm}(U) \supset Q$. $\square$

As we have noticed in §20.5, the dynamics on the Fatou set is Lyapunov stable. We see now that the dynamics on the Julia set is :yapunov unstable. Thus, we have:

PROPOSITION 21.34. *The Fatou set coincides with the set of Lyapunov stable points.*

The above lemma also implies a refined version of the Shrinking Lemma:

EXERCISE 21.35. *Assume there is a family of inverse branches $f^{-n_k} : \mathbb{D}_k \to \mathbb{C}$ with bounded dilatation, where $n_k \to \infty$ and $\mathbb{D}_k$ are round disks with bounded radii centered at some points of the Julia set. Then* $\operatorname{diam}(f^{-n_k}(\mathbb{D}_k)) \to 0$ *as $k \to \infty$.*

21.6. Siegel and Cremer cycles.

21.6.1. *Lyapunov criterion.* Let us start with a local situation. Let us consider a germ $f : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$

$$f : z \mapsto e(\theta)z + b_2 z^2 + \dots \tag{21.13}$$

with irrational rotation number, $\theta \in (\mathbb{R}/\mathbb{Q})/\mathbb{Z}$.

The germ f (and its fixed point 0) is called *linearizable* or *Siegel* if it is conformally conjugate to the rotation by θ, i.e., there exists a conformal map $\phi : (U, 0) \to (\mathbb{D}_r, 0)$ from a neighborhood of 0 to a disk $\mathbb{D}_r$, normalizaed so that $\phi'(0) = 1$, such that

$$\phi(fz) = e(\theta)\,\phi(z). \tag{21.14}$$

The variable $\zeta = \phi(z)$ is called the *linearizing coordinate* for f.

PROPOSITION 21.36. *A fixed point is Siegel iff it is Lyapunov stable.*

PROOF. Obviously, Siegel points are Lyapunov stable.

Vice versa, if 0 is stable then it has an arbitrary small invariant neighborhood. Filling it in, we obtain a small simply connected invariant neighborhood $U \ni 0$. It can be conformally mapped onto a disk $\mathbb{D}_r$ by the Riemann map $\phi : (U, 0) \to (\mathbb{D}_r, 0)$, normalized so that $\phi'(0) = 1$. Then $g := \phi \circ f \circ \phi^{-1}$ is a holomorphic endomorphism of the unit disk fixing 0, with $|g'(0)| = |e(\theta)| = 1$. By the Schwarz Lemma, $g(z) = e(\theta)z$. $\square$

In the polynomial case, linearizability can be characterized in terms of the Fatou-Julia Dichotomy:

PROPOSITION 21.37. (i) *For a polynomial f, a neutral periodic point α is Siegel iff $\alpha \in \mathcal{F}$.*
(ii) *In the Siegel case, let D be the component of $\mathcal{F}$ containing α. Then $f^p|U$ is conformally conjugate to the rotation of $\mathbb{D}$ by θ (where p is the period of α and θ is its rotation number).*

PROOF. The first assertion follows from Propositions 21.34 and 21.36. To prove the second one, consider the Riemann mapping $\phi : (D, \alpha) \to (\mathbb{D}, 0)$. Applying the Schwarz Lemma in the same way as in the proof of Proposition 21.36, we concude that $\phi \circ f^p \circ \phi^{-1}$ is the rotation of $\mathbb{D}$ by θ. □

The component D of $\mathcal{F}(f)$ containing a Siegel point is called a *Siegel disk*.

We will see later on that *a quadratic polynomial can have at most one non-repelling cycle* (see Theorem 28.13). If it has one, it can be non-contradictory classified as either hyperbolic, or parabolic, or Siegel, or Cremer.

21.6.2. *Existence of Cremer points.* Let us show that Cremer cycles indeed exist:

PROPOSITION 21.38. *In the family $f_\theta : z \mapsto e(\theta)z + z^2$, $\theta \in \mathbb{R}/\mathbb{Z}$, the origin* 0 *is the Cremer fixed point for a generic rotation number θ.*

PROOF. Let us consider the set $\Lambda \subset \mathbb{R}/\mathbb{Z}$ of rotation numbers θ for which $0 \in \mathcal{J}(f_\theta)$. We have $\Lambda = \Lambda_p \sqcup \Lambda_C$, where Λ_p is the set of parabolic (i.e., is rational), rotation numbers, while Λ_C is the set of Cremer (i.e., irrational) numbers.

We will show that Λ *is of type G_δ*. Since Λ_p is dense, Λ is a dense G_δ, so rotation numbers $\theta \in \Lambda$ are generic by definition. Of course, removing a countable subset preserves genericity, so the conclusion would follow.

To prove that Λ is G_δ, let us consider a function

$$d : \mathbb{R}/\mathbb{Z} \to \mathbb{R}_{\geq 0}, \quad d(\theta) = \operatorname{dist}(0, \mathcal{J}(f_\theta)).$$

Then Λ is the set of zeros of d. We will show that d is upper-semicontinuous. Since the set of zeros of a non-negative upper semicontinuous function is of type G_δ, it will complete the proof.

So, take an $\varepsilon > 0$ and let $d(\theta_\circ) = d_\circ$. Then by Theorem 21.31, there is a repelling periodic point $\alpha_\circ \in \mathcal{J}_\circ$ such that $|\alpha_\circ| < d_\circ + \varepsilon$. But repelling periodic points persist under perturbations: for any θ near $\theta_\circ$, the map f_θ has a repelling periodic point α_θ continuously depending on θ (see Lemma 33.3 below for this easy property). Hence $|\alpha_\theta| < d_\circ + 2\varepsilon$ for all nearby θ's. As $d(\theta) \leq |\alpha_\theta|$, the semicontinuity follows. □

Existence of Siegel points will be shown in §23.2.

21.6.3. *Postcritical set.*

PROPOSITION 21.39. (i) *If D is a Siegel disk then $\partial D \subset \omega(0)$.*
(ii) *If α is a Cremer point then $\alpha \in \omega(0)$.*

PROOF. (i) If $D \not\subset \omega(0)$ then there is a disk W intersecting ∂D but disjoint from $\omega(0)$. Then all inverse branches of f^{-n} are well defined in W. Take a point $z \in W \cap D$, let $z \equiv z_0, z_{-1}, z_{-2}, \dots$ be its backward orbit that stays in D, and let $f^{-n} : W \to W_{-n}$ be the corresponding inverse branches. By the Shrinking Lemma, $|Df^{-n}(z)| \to 0$ as $n \to \infty$. However, this is impossible as $f^{-p}|\,D$ is conformally equivalent to the disk rotation (where p is the period of D).

(ii) The proof is similar. If there were a disk $W \ni \alpha$ disjoint from $\omega(0)$, then the Shrinking Lemma would imply that $|Df^{-n}(\alpha)| \to 0$ as $n \to \infty$. □

It immediately follows that $\operatorname{orb} 0$ is infinite in the Siegel case.

EXERCISE 21.40. *Show that* $\operatorname{orb} 0$ *is infinite in the Cremer case as well.*

21.6.4. *Perez-Marco hedgehog.* An ε*-small orbit* near a fixed point α is an orbit contained in $\overline{\mathbb{D}}_\varepsilon$. We can talk about *two-sided* and *one-sided* small orbits.) If an ε-small orbit exists for any $\varepsilon > 0$, we say that "there are small orbits near α". Obviously, there are small two-sided orbits near any Siegel point, and by the Flower Theorem, there are such orbits near any parabolic point. In this section we will construct small orbits near any Cremer point.

Let us begin with the one-sided case. Let

$$f : z \mapsto \rho z + a z^2 + \dots \tag{21.15}$$

be a holomorphic germ near 0 univalent on $\mathbb{D}_r$. For $0 < \varepsilon < r$, let $K_+^\varepsilon \equiv K_+^\varepsilon(f)$ be the connected component of $\{z : \ f^n z \in \overline{\mathbb{D}}_\varepsilon, \ n = 0, 1, \dots\}$ containing 0. In other words, K_+^ε is the maximal forward invariant connected subset of $\overline{\mathbb{D}}_\varepsilon$ containing 0.

Let $\mathcal{D}^\varepsilon \equiv \mathcal{D}_f^\varepsilon$ be the component of $\operatorname{int} K_+^\varepsilon(f)$ containing 0.

EXERCISE 21.41. (i) *For an attracting germ* f *as above, all forward orbits in* $\mathcal{D}^\varepsilon$ *converge to* 0 (*in this case* $\mathcal{D}^\varepsilon$ *is called the immediate basin of attraction of* 0 *for the germ* f *in* $\overline{\mathbb{D}}_\varepsilon$.)

(ii) *For an attracting germ* f, $\mathcal{D}^\varepsilon$ *stretches all the way to the boundary circle* $\mathbb{T}_\varepsilon$ (i.e., $\mathbb{T}_\varepsilon \cap \operatorname{cl} \mathcal{D}^\varepsilon \neq \emptyset$).

(iii) *For a neutral germ* f, *the set* K_+^ε *stretches all the way to the boundary circle* $\mathbb{T}_\varepsilon$.

REMARK 21.42. Taking the union of the hulls K_+^ε over all $\varepsilon < r$, we obtain an invariant connected subset $K_+^r \subset \mathbb{D}_r$ that stretches all the way to $\mathbb{T}_r$. However, this set is not closed, so formally speaking, it is not a hull.

Before going to the proof of the existence of the two-sided small orbits, recall from §19.8.4 the structure of circle homeomorphisms with rational rotation number.

LEMMA 21.43. *Let* $g : \mathbb{T} \to \mathbb{T}$ *be a real analytic orientation preserving circle diffeomorphism with a rational rotation number* $\mathfrak{p}/\mathfrak{q}$. *Then for any* $\rho > 1$ *there exists a completely invariant open set* $\mathcal{U} \subset \mathbb{A}(1/\rho, \rho)$ *such that* $\mathcal{U} \cap \mathbb{T} = \mathbb{T} \smallsetminus \operatorname{Per}(g)$ *and* $\mathcal{U} \cup \mathbb{T}$ *is connected. Moreover, all orbits in* $\mathcal{V}$ *converge to periodic cycles in* $\mathbb{T}$ *in both negative and positive times.*

PROOF. Let us consider a component $I = (a_-, a_+)$ of $\mathbb{T} \smallsetminus \operatorname{Per}(g)$ (see Exercise 19.33). Then for any $x \in I$, we have $g^{\mathfrak{q}n} x \to a_+$ as $n \to +\infty$. Hence a_+ is either attracting or parabolic. In either case, x belongs to the attracting basin of a_+, which is open in $\mathbb{C}$. Hence for any point $x \in I$ there is a neighborhood V whose forward orbit is contained in $\mathbb{A}(1/\rho, \rho)$ and the domains $g^{\mathfrak{q}n}(V)$ uniformly converge to a_+ as $n \to +\infty$.

Replacing f with f^{-1}, we obtain a similar statement for a_- in the negative time. In this way we obtain a neighborhood $V \ni x$ whose full orbit is contained in $\mathbb{A}(1/\rho, \rho)$ and the $g^{\mathfrak{q}n}(V)$ uniformly converge to $a_\pm$ as $n \to \pm\infty$.

Covering a fundamental interval $J \subset I$ with finitely many such neighborhoods V_k, we construct a completely invariant open set $\mathcal{U}_I$ in $\mathbb{A}(1/\rho, \rho)$ whose trace by $\mathbb{T}$ is equal to the orbit of I and $\mathbb{T} \cup \mathcal{U}_I$ is connected (and all orbits in $\mathcal{U}_I$ converge to the boundary periodic cycles of $\bigcup g^n I$ in both negative and positive times). Doing this for all components I of $\mathbb{T} \smallsetminus \operatorname{Per}(g)$, we obtain a desired open set. $\square$

Let $K^\varepsilon \equiv K^\varepsilon(f)$ be the connected component of the *two-sided* non-escaping set $\{z : \ f^n z \in \overline{\mathbb{D}}_\varepsilon, \quad \forall\, n \in \mathbb{Z}\}$ containing 0. In other words, K^ε is the maximal completely invariant connected subset of $\overline{\mathbb{D}}_\varepsilon$ containing 0.

THEOREM 21.44. *Let $f : z \mapsto e(\theta)z + az^2 + \dots$ be a neutral holomorphic germ near* 0 *univalent on $\mathbb{D}_r$. Then for any $0 < \varepsilon < r$, the set K^ε is a completely invariant hull that stretches all the way to the boundary circle $\mathbb{T}_\varepsilon$.*

PROOF. We fix an $\varepsilon \in (0, r)$ and will not necessarily emphasize dependence of various objects on it, e.g., $K \equiv K^\varepsilon$.

Obviously, K is compact. By the Maximal Principle, K is full, so it is a hull or a singleton. We need to show that it stretches all the way to $\mathbb{T}_\varepsilon$.

Let us first consider the parabolic case, $\theta = \mathfrak{p}/\mathfrak{q}$. In this case, K contains the intersection $\Phi^a \cap \Phi^r$ of local attracting and repelling Leau-Fatou flowers (see §21.3.2), so it is a hull (rather than a singleton). Assume by contradiction that $K \subset \mathbb{D}_\varepsilon$.

Let us consider the Riemann mapping $\phi : \mathbb{C} \smallsetminus K \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ and define the *external map* $g := \phi \circ f \circ \phi^{-1}$ in an outer neighborhood $\Omega^+ := \phi(\mathbb{D}_\varepsilon \smallsetminus K)$ of $\mathbb{T}$. It is a conformal map, proper near $\mathbb{T}$. Hence it extends continuously to $\mathbb{T}$ (see Proposition 2.60), and then it extends by reflection to a continuous map in a symmetric neighborhood Ω of $\mathbb{T}$ which is conformal off $\mathbb{T}$. Since $\mathbb{T}$ is removable (see §16), g is conformal on the whole Ω, so it restricts to an analytic diffeomorphism of $\mathbb{T}$.

As we know from Exercise 21.14 (applied to f^{-1}), there is a curve $\gamma \subset \Phi^r$ landing at 0 which is f^{-1}–periodic with period $\mathfrak{q}$, $f^{-\mathfrak{q}}(\gamma) \subset \gamma$. Moreover, the curves $f^{-n}(\gamma)$, $n = 0, \dots, \mathfrak{q}-1$, represent different homotopy classes in $\mathbb{C} \smallsetminus K$ rel 0 (since attracting petals separate any two of these curves). Let $\delta = \phi(\gamma)$. Then δ is g^{-1}–periodic with period $\mathfrak{q}$, implying by Lemma 8.18 that it lands at a $\mathfrak{q}$-periodic point $\alpha \in \mathbb{T}$. So, g has a rational rotation number.

By Lemma 21.43, there is a completely invariant open set $\mathcal{U} \subset \Omega^+$ such that $\mathbb{T} \cup \mathcal{U}$ is connected. Let $\mathcal{V} := \phi^{-1}(\mathcal{U})$. The set $K \cup \mathcal{V}$ is connected. Indeed, $\overline{U}_i \cap \mathbb{T} \neq \emptyset$ for each component U_i of $\mathcal{U}$. Hence $\overline{V}_i \cap K \neq \emptyset$ for each component V_i of $\mathcal{V}$. It follows that all the sets $V_i \cup K$, and hence their union, $\mathcal{V} \cup K$, are connected.

Thus, $K \cup \mathcal{V}$ is a completely invariant connected set contained in $\mathbb{D}_\varepsilon$, contradicting the definition of K (as the maximal connected completely invariant subset of $\overline{\mathbb{D}}_\varepsilon$). This completes the proof in the case of rational θ.

In general, consider any sequence of parabolic maps $f_n : \mathbb{D}_r \to \mathbb{C}$ converging to f (uniformly on compact subsets of $\mathbb{D}_r$) and the corresponding sequence of completely invariant compact connected sets $0 \in K_n \subset \overline{\mathbb{D}}_\varepsilon$ stretching all the way to $\mathbb{T}_\varepsilon$. Any Hausdorff limit K' of this sequence inherits all of these properties. Filling in its holes, if necessary, we turn K' into a hull K, supplying us with a desired hedgehog for f.

$\square$

Though it is not obvious, the hedgehog K^ε is uniquely determined by being a completely invariant hull in the disk $\overline{\mathbb{D}_\varepsilon}$ touching its boundary [**PM2**].

REMARK 21.45. Similarly, one can associate a canonical hedgehog K^U to any Jordan disk $U \ni 0$ such that f is univalent in some neighborhood of $\overline{U}$.

21.6.5. *Appendix: associated circle diffeomorphism.* Let us uniformize the complement of the hedgehog by the complement of the unit disk, $\phi : \mathbb{C} \smallsetminus K \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$. Then the map $f : e(\theta)z + z^2 + \dots$ in an outer neighborhood of K is conjugate to a proper map g in an outer neighborhood of $\mathbb{T}$. The latter extends to an analytic diffeomorphism of $\mathbb{T}$, for which we keep the same notation g (compare §41.1.2). As Perez-Marco has demonstrated, this map is a very useful tool for understanding local dynamics near a Cremer point. Let us mention one application.

LEMMA 21.46. *The map $g : \mathbb{T} \to \mathbb{T}$ has rotation number θ.*

PROOF. Approximation of f_θ with parabolic maps $f_{\mathfrak{p}_n/\mathfrak{q}_n}$ leads to approximation of f with circle diffeomorphisms with rotation numbers $\mathfrak{p}_n/\mathfrak{q}_n$. □

It led Perez-Marco to a new proof of the topological invariance of the rotation number of a Cremer point (see **[PM1]**):

NAISHUL'S THEOREM. Topologically conjugate neutral germs have the same rotation number.

21.7. Periodic components. The notions of a periodic component of $\mathcal{F}(f)$ and its cycle are self-explanatory. It is classically known that such a component is always associated with a non-repelling periodic point:

THEOREM 21.47. *Let $\mathbf{U} = (U_i)_{i=1}^p$ be a cycle of periodic components of* $\operatorname{int} \mathcal{K}(f)$. *Then one of the following three possibilities can happen:*

- $\mathbf{U}$ *is the immediate basin of an attracting cycle;*
- $\mathbf{U}$ *is the immediate basin of a parabolic cycle $\boldsymbol{\alpha} \subset \partial\mathbf{U}$ of some period $q|p$;*
- $\mathbf{U}$ *is the cycle of Siegel disks.*

Proof. Take a component U of the cycle $\mathbf{U}$, and let $g = f^p$. By the Schwarz Lemma, $g|U$ is either a conformal automorphism of U, or it strictly contracts the hyperbolic metric $\operatorname{dist}_{\text{hyp}}$ on U. In the former case, it is either elliptic, or otherwise. If g is elliptic then U is a Siegel disk. Otherwise the orbits of g converge to the boundary of U.

Let us show that if an orbit $(z_n = g^n z)$, $z \in U$, converges to ∂U, then it converges to a g-fixed point $\beta \in \partial U$. Join z and $g(z)$ with a smooth arc γ, and let $\gamma_n = f^n\gamma$. By the Schwarz Lemma, the hyperbolic length of the arcs γ_n stays bounded. Hence they uniformly escape to the boundary of U. Moreover, by the relation between the hyperbolic and Euclidean metrics (Lemma 7.7), the Euclidean length of the γ_n shrinks to 0. In particular,

$$|g(z_n) - z_n| = |z_{n+1} - z_n| \to 0 \quad \text{as } n \to \infty. \tag{21.16}$$

By continuity, all limit points of the orbit (z_n) are fixed under g. But g, being a polynomial, has only finitely many fixed points. On the other hand, (21.16) implies the ω-limit set of the orbit (z_n) is connected. Hence it consists of a single fixed point β.

Moreover, the orbit (ζ_n) of any other point $\zeta \in U$ must converge to the same fixed point β. Indeed, the hyperbolic distance between z_n and ζ_n stays bounded and hence the Euclidean distance between these points shrink to 0.

Thus either U is a Siegel disk, or the g-orbits in U converge to a g-fixed point β, or the map $g : U \to U$ strictly contracts the hyperbolic metric and its orbits do not escape to the boundary ∂U. Let us show that in the latter case, g has an attracting fixed point α in U.

Take a g-orbit (z_n), and let $d_n = \mathrm{dist}_{\mathrm{hyp}}(z_0, z_n)$. Since g is strictly contracting,

$$\mathrm{dist}_{\mathrm{hyp}}(z_{n+1}, z_n) \le \rho(d_n)\, \mathrm{dist}_{\mathrm{hyp}}(z_n, z_{n-1}),$$

where the contraction factor $\rho(d_n) < 1$ depends only on $\mathrm{dist}_{\mathrm{hyp}}(z_n, z_0)$. Since the orbit (z_n) does not escape to ∂U, this contraction factor is bounded away from 1 for infinitely many moments n, and hence $\mathrm{dist}_{\mathrm{hyp}}(z_{n+1}, z_n) \to 0$. It follows that any ω-limit point of this orbit in U is fixed under g.

By the strict contraction, g can have only one fixed point in U, and hence any orbit must converge to this point. Strict contraction also implies that this point is attracting.

We still need to prove the most delicate property: in the case when the orbits escape to the boundary point $\beta \in \partial U$, this point is parabolic. In fact, we will show that $g'(\beta) = 1$. Of course, this point cannot be either repelling (since it attracts some orbits) or attracting (since it lies on the Julia set). So it is a neutral point with some rotation number $\theta \in [0, 1)$. The following lemma will complete the proof.

Lemma 21.48 (Necklace Lemma). *Let $f : z \mapsto \rho z + a_2 z^2 + \dots$ be a holomorphic map near the origin, and let $|\rho| = 1$. Assume that there exists a domain $\Omega \subset \mathbb{C}^*$ such that all iterates f^n are well-defined on Ω, $f(\Omega) \cap \Omega \neq \emptyset$, and $f^n(\Omega) \to 0$ as $n \to \infty$. Then $\rho = 1$.*

Proof. Consider a chain of domains $\Omega_n = f^n \Omega$ converging to 0. Without loss of generality we can assume that all the domains lie in a small neighborhood of 0 and hence the iterates $f^n|\Omega$ are univalent. Fix a base point $a \in \Omega$ such that $f(a) \in \Omega$, and let

$$\phi_n(z) = \frac{f^n(z)}{f^n(a)}.$$

These functions are univalent, normalized by $\phi_n(a) = 1$, and do not have zeros. By the Koebe Distortion Theorem (the version given in Exercise 4.14(ii)), they form a normal family. Moreover, if $\rho \neq 1$ then any limit function ϕ of this family is non-constant since $\phi(f(a)) = \rho \neq 1 = \phi(a)$. Therefore, the derivatives $\phi_n' \,|\, \Omega$ are bounded away from 0, and hence $\mathrm{dist}(1, \partial\Omega_n) \ge \varepsilon > 0$ for all $n \in \mathbb{N}$. It follows that

$$\mathrm{dist}(f^n(a), \partial\Omega_n) \ge \varepsilon\, r_n, \quad n \in \mathbb{N},$$

where $r_n = |f^n(a)|$.

On the other hand, f acts almost as the rotation by θ near 0, where $\theta = \arg \rho \in (0, 1)$. Since this rotation is recurrent (see Exercise 19.31) and $\theta \neq 0$, there exists an $l > 0$ such that

$$\mathrm{dist}(f^{n+l}(a), f^n(a)) = o(r_n) \quad \text{as } n \to \infty.$$

The last two estimates imply that $\Omega_{n+l} \cap \Omega_n \neq \emptyset$ for all sufficiently big n.

Thus, the chain of domains $\Omega_n, \dots, \Omega_{n+l}$ closes up, and their union form a "necklace" around 0. Take a Jordan curve γ in this necklace, and let D be the disk bounded by γ. Then $f^n(\gamma) \to 0$ as $n \to \infty$. By the Maximum Principle, $f^N(D) \Subset D$ for some N. By the Schwarz Lemma, $|\rho| < 1$ – contradiction. □

Notes. Classical Theorem 21.4 on existence of a critical point in the attracting basin (due to Fatou and Julia) plays a fundamental role in the field. It is valid for a general rational function f of degree d and implies that f may have at most $2d-2$ (the number of critical points) attracting cycles. In particular, a polynomials of degree d may have at most $d-1$ finite (in $\mathbb{C}$) attracting cycles.

Lemma 21.30 on the number of non-repelling cycles is due to Fatou: it gives the twice bigger bound than was anticipated.

Existence of non-linearizable neutral points ("Cremer maps") was demonstrated by Pfeiffer in 1917 [**Pf**]. The work of Cremer [**Cr**] appeared two decades later. The Perez-Marco hedgehogs appeared in [**PM1**]. A more topological proof was recently given in [**FLRT**]: instead of the Riemann Mapping Theorem, it makes use of the Brouwer Translation Theorem for plane homeomorphisms (see [**Fr**]).

Description of the dynamics in an invariant Fatou component (Theorem 21.47) is a version of the *Denjoy-Wolff Theorem* on the holomorphic dynamics in the disk (see [**Va**, §43]). The Necklace Lemma (and its proof given in this book) is due to Fatou [**F3**, §54]. There is a more topological proof of this lemma that suggested the name *Snail Lemma* commonly used these days (see [**M2**, §16]).

22. Postcritical set as the global attractor

22.1. Remark on wandering domains. Consistently with the general terminology of §19.1, a component D of the Fatou set $\mathcal{F}(f)$ is called *wandering* if $f^n D \cap f^m D = \emptyset$ for any natural $n < m$. Such components will also be referred to as "wandering domains".[10]

In §29.2, we will prove that wandering domains do not exist. Here we will make an observation that implies this in some special cases (but it will not be used in the general argument.)

Proposition 22.1. *If D is a wandering domain then, $\omega(z) \subset \omega(0)$ for any $z \in D$.*

Proof. Let r_n be the inner radius of the $f^n(D)$ around $f^n z$.

Since area $f^n(D) \to 0$, we have $r_n \to 0$.

By the Koebe 1/4−Theorem this implies that

$$|Df^n(z)| \to 0 \quad \text{as } n \to \infty. \tag{22.1}$$

Assume now that $\omega(z) \not\subset \omega(0)$. Then there is an open disk W intersecting $\omega(z)$ but disjoint from $\omega(0)$. Let f^{-n} be the iterated inverse branches on W that send $f^n z$ back to z. By the Normality Lemma, 20.25, they form a normal family, implying boundedness of the derivatives $|Df^{-n}(f^n z)|$, which contradicts (22.1). □

[10]One can consider more general "wandering domains", not necessarily full components of $\mathcal{F}(f)$. Such domains can certainly exist (in the basins of attracting and parabolic points). We hope this slight terminological inconsistency will not cause a problem.

22.2. Global measure-theoretic attractor. The following result is a manifestation of the leading role of the critical point in global holomorphic dynamics:

THEOREM 22.2. *If the Julia set $\mathcal{J}(f)$ has positive area then $f^n z \to \omega(0)$ for a.e. $z \in \mathcal{J}(f)$.*

PROOF. For any point $\zeta \in \mathcal{J} \smallsetminus \mathrm{cl}(\mathrm{orb}\, v))$, there is a disk $\mathbb{D}(\zeta, 2\varepsilon)$ disjoint from $\mathrm{orb}\, v$. Then all the inverse branches f_i^{-n} are well defined on this disk. By the Koebe Distortion Theorem, they have an absolutely bounded distortion on the twice smaller disk:

$$\left| \frac{Df_i^{-n}(\zeta')}{Df_i^{-n}(\zeta)} \right| \asymp 1 \quad \forall\, \zeta' \in \mathbb{D}(\zeta, \varepsilon).$$

Moreover, by the Shrinking Lemma,

$$\mathrm{diam}(f_i^{-n}(\mathbb{D}(\zeta, \varepsilon)) \to 0 \quad \text{as } n \to \infty$$

uniformly in i.

Thus, the pullbacks $\Delta_i^n := f_i^{-n}(\mathbb{D}(\zeta, \varepsilon))$ are shrinking ovals of bounded shape mapped by f^n onto $\mathbb{D}(\zeta, \varepsilon)$ with bounded distortion. Let $\mathbb{D}(\zeta', \delta) \subset \mathbb{D}(\zeta, \varepsilon) \smallsetminus \mathcal{J}$ be a gap in the Julia set. Then its pullbacks $f_i^{-n}(\mathbb{D}(\zeta', \delta)) \subset \Delta_i^n$ are gaps in $\mathcal{J}$ of definite size (i.e., their inner radii are comparable with $\mathrm{diam}\, \Delta_i^n$, uniformly in (n, i)).

Let now $X := \{z \in \mathcal{J} :\ \omega(z) \not\subset \omega(0)\}$. Then for any $z \in X$, $\omega(z) \not\subset \mathrm{cl}(\mathrm{orb}\, v)$, so there is a point $\zeta \in \omega(z) \smallsetminus \mathrm{cl}(\mathrm{orb}\, v)$. Hence there is a sequence $n_k \to \infty$ such that $f^{n_k} z \in \mathbb{D}(\zeta, \varepsilon/2)$.

Let Δ^{n_k} be the $f^{n_k}-$pullback of $\mathbb{D}(\zeta, \varepsilon)$ around z. By the previous discussion, it is a shrinking sequence of ovals of bounded shape around z that contain definite gaps in $\mathcal{J}$. Hence $\mathcal{J}$ is porous at z (as defined in §19.18), so z is not a density point for $\mathcal{J}$. By the Lebesgue Density Theorem, $\mathrm{area}\, X = 0$. □

With the notion of global measure-theoretic attractor in hands (§19.7), we obtain:

COROLLARY 22.3. *Under the above circumstances, the global measure-theoretic attractor of $f|\mathcal{J}$ is contained in $\omega(0)$.*

REMARK 22.4. This result is a strating point for exploring the area of the Julia set. In particular, it implies that $\mathrm{area}\, \mathcal{J}(f) = 0$ for hyperbolic, parabolic and postrictically preperiodic maps (compare §§25.5, 26.5, 27.1.1).

Notes. Proposition 22.1 on wandering domains is due to Fatou [**F2**, §30].

In Holomorphic Dynamics, the global measure-theoretic attractor (§22.2) appeared in [**L6**]. It was one of the first applications of the Koebe Distortion Theorem to measurable dynamics. Besides having immediate consequences in the area problem (see §§25.5, 26.5, 27.1.1) it was the starting point for the study of real 1D attractors by A. Blokh and the author (see §46).

23. Remarkable functional equations

Study of certain functional equations was one of the main motivations for the classical work in holomorphic dynamics. By means of these equations the local dynamics near periodic points of different types can be reduced to the simplest normal form. But it turns out that the role of the equations goes far beyond local issues: global solutions of the equations play a crucial role in understanding the dynamics.

We will start with the local analysis and then globalize it (though sometimes one can go the other way around). For the local analysis we put the fixed point at the origin and consider a holomorphic germ (21.1) near the origin. The key question is whether f can be locally conjugated to its linear part

$$L_\rho : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \rho z.$$

If so, tha *map* f is called *locally linearizable*, or else: the corresponding *germ* is called *linearizable* (compare §21.6). However, in the parabolic case, the question assumes a more subtle form...

23.1. Linearizing coordinate in the attracting case. Let us start with the easiest case of a (simply) attracting (i.e., with $0 < |\rho| < 1$) fixed point. The topological picture is particularly simple as all these maps are topologically, and in fact quasiconformally, equivalent.

23.1.1. *Local qc classification.*

EXERCISE 23.1. *Show that any two attracting linear maps L_ρ and $L_{\rho'}$ with $0 < |\rho|, |\rho'| < 1$ are globally quasiconformally conjugate. However, if $\rho \neq \rho'$ then the conjugacy is not differentiable at* 0.

PROPOSITION 23.2. *Any two* (*simply*) *attracting holomorphic maps* (21.1) *are locally quasiconformally conjugate.*

PROOF. Let f, $\tilde{f}$ be two attracting holomorphic maps in question, both defined in some disk $\mathbb{D}_{2r}$. Let us consider fundamental annuli $A = \overline{\mathbb{D}}^r \smallsetminus f(\mathbb{D}_r)$ and $\tilde{A} = \overline{\mathbb{D}}^r \smallsetminus \tilde{f}(\mathbb{D}_r)$.

Start with taking any diffeomorphism $h : A \to \tilde{A}$ equivariant on $\mathbb{T}_r$, i.e., such that $h(fz) = \tilde{f}(h(z))$ for any $z \in \mathbb{T}_r$.

By means of the dynamics, lift h to a diffeomorphism $h_1 : f(A) \to \tilde{f}(A)$, i.e., let $h_1(f(z)) = \tilde{f}(h(z))$ (which is not a problem as f is invertible). Since h was selected equivariant on ∂A, h_1 matches with h on $\partial(f(A))$. By qc removability of smooth curves (Gluing Lemma) we obtain a qc homeomorphism $H_1 : A \cup f(A) \to \tilde{A} \cup \tilde{f}(\tilde{A})$ that extends h and is equivariant on A. Moreover, since f and $\tilde{f}$ are conformal, h_1, and hence H_1, have the same dilatation as h.

Let $A^n = \displaystyle\bigcup_{k=0}^{n} f^k(A)$, and let $\tilde{A}^n$ be the similarly annulus for $\tilde{f}$. Note that

$$\bigcup A^n = \mathbb{D}_r^* = \bigcup \tilde{A}^n.$$

Proceeding as above, we consecutively construct a sequence of qc homeomorphisms $H_n : A^n \to \tilde{A}^n$ equivariant on A^{n-1} and extending H_{n-1} from the latter without increasing dilatation. Taking the union of these annuli, we obtain an equivariant qc homeomorphism $H : \mathbb{D}_r^* \to \mathbb{D}_r^*$.

By qc removability of isolated points, we extend it to a desired qc conjugacy $\mathbb{D}_r \to \mathbb{D}_r$ between f and $\tilde{f}$. $\square$

Thus, in the attracting case, the value of the multiplier ρ does not have any significance form the point of view of topological dynamics.

EXERCISE 23.3. *Show that the above conjugacy can be selected smooth in a punctured neighborhood of* 0.

However, there is an obstruction for smoothness at the origin, as the multiplier is a smooth invariant: *If a conjugacy between two maps* (21.1) *is differentiable at* 0*, then the maps have the same multiplier* ρ (see Proposition 19.6; compare with §19.8.1).

23.1.2. *Local conformal linearization.* In fact, the multiplier is the only obstruction for smooth, and in fact, for conformal, conjugacy. Indeed, any map (21.1) can be conformally linearized near the origin (compare with Exercise 19.29):

THEOREM 23.4 (Linearization Theorem). *Consider a holomorphic map* (21.1) *near the origin. Assume* $0 < |\rho| < 1$*. Then there exists an* f*-invariant Jordan disk* $V \ni 0$*,* $r > 0$*, and a conformal map* $\phi : (V, 0) \to \mathbb{D}_r$ *with* $\phi'(0) = 1$ *satisfying the equation:*

$$\phi(fz) = \rho\, \phi(z) \tag{23.1}$$

The above properties determine uniquely the germ of ϕ *at the origin.*

The above function ϕ is called the *linearizing coordinate* for f near 0 or the *Königs function*. The linearizing equation (23.1) is also called the *Schröder equation*. It locally conjugates f to its linear part $z \mapsto \rho z$.

PROOF. The linearizer ϕ can be given by the following explicit formula:

$$\phi(z) = \lim_{n\to\infty} \rho^{-n} f^n z. \tag{23.2}$$

To see that the limit exists (uniformly near the origin), let $z_n = f^n z$, $z_0 \equiv z$, notice that $z_n = O(|z\rho|^n)$ uniformly near the origin, and take the ratio of the two consecutive terms in (23.2):

$$\frac{\rho^{-n-1} z_{n+1}}{\rho^{-n} z_n} = \rho^{-1} \frac{\rho z_n (1 + O(|z_n|))}{z_n} = 1 + O(|z\rho^n|).$$

Hence

$$\phi(z) = z \prod_{n=0}^{\infty} \frac{\rho^{-n-1} z_{n+1}}{\rho^{-n} z_n} = z(1 + O(|z|))$$

uniformly near the origin, and the conclusion follows.

Obviously, ϕ is a linearizer. Its uniqueness follows from the exercise below. □

EXERCISE 23.5. *Show that if a holomorphic germ* f *near the origin commutes with the linear germ* $z \mapsto \rho z$*,* $0 < |\rho| < 1$*, then* f *is itself linear.*

REMARK 23.6. We see that the conjugacy ϕ is constructed by going forward by the iterates of f and then returning back by the iterates of the corresponding linear map. This method of constructing a conjugacy between two maps will be used on several other occasions: see (23.7) and Project 23.41.

Let us note in conclusion that the Königs function $\phi = \phi_f$ depends holomorphically on f:

LEMMA 23.7. *Let* $f_\lambda(z) : z \mapsto \rho(\lambda) z + a_2(\lambda) z^2 + \dots$ *be a holomorphic family of local maps with attracting fixed point* 0*. Then the Königs function* $\phi_\lambda(z)$ *depends holomorphically on* λ.

Proof. The above proof shows that convergence in Königs formula (23.2) is locally uniform over λ. Hence the limit is holomorphic in (λ, z). $\square$

23.1.3. *Fundamental torus.* Take a little disk $D := \mathbb{D}_\varepsilon$ around 0. It is invariant under f, and the quotient of D under the action of f is a conformal torus $\mathbb{T}^2_f$. It can be obtained by taking a fundamental annulus $A_f := \overline{D} \smallsetminus f(D)$ and gluing its boundary components by the dynamical relation [identifying a point $z \in \partial D$ to $f(z) \in \partial(f(D))$].

The torus $\mathbb{T}^2_f$ is naturally partially marked: its fundamental group $\Gamma = \pi_1(\mathbb{T}^2_f)$ has a marked generator corresponding to a little circle around 0, compare §2.6.3. The second generator of Γ is represented by a proper arc γ in the fundamental annulus A_f connecting two dynamically related points. Given the endpoints, such an arc is defined up to a twist by $n \in \mathbb{Z}$ revolutions.

By the Linearization Theorem, 23.4, the action of f on D is conformally equivalent to the linear action of $L_\rho : \zeta \mapsto \rho\,\zeta$ on $\mathbb{D}^*$. Hence $\mathbb{T}^2_f$ is conformally equivalent (as partially marked torus) to $\mathbb{T}^2_\rho = \mathbb{D}^*/L_\rho$, so ρ is the modulus of $\mathbb{T}^2_f$, see §2.6.3.

By means of the universal covering map $z \mapsto e(z)$, the torus $\mathbb{T}^2_\rho$ is identified with the quotient $\mathbb{C}$ by the lattice generated by two translations $\alpha : z \mapsto z+1$ and $\beta : z \mapsto z + \dfrac{2\pi i}{\log \rho}$. Moreover, the $2\pi i$-ambiguity in the choice of $\log \rho$ corresponds to the twist ambiguity in the choice of the second generator.

In case of an attracting periodic point α of period p, we can apply the above discussion to f^p near α to obtain the corresponding fundamental torus $\mathbb{T}^2_{f,\alpha}$.

23.1.4. *Extension to the immediate basin.* Next, we will extend the Königs function to the immediate basin of attraction:

Proposition 23.8. *Let f be a polynomial with attracting periodic point α. Then the Königs function ϕ analytically extends to the immediate basin $D \equiv D^\bullet(\alpha)$, and it satisfies there Schröder functional equation* (23.1)*. Moreover, the map $\phi : D \to \mathbb{C}$ is a branched covering of infinite degree, branched on $D \cap \mathrm{Crit}^\infty_f$. The fibers of ϕ are petit orbits of $f|\,D$.*

Proof. We can assume without loss of generality that α is fixed, $f(\alpha) = \alpha$. The immediate basin D is exhausted by an increasing nest of domains

$$P_0 \subset P_1 \subset P_2 \subset \dots$$

where P_0 is a domain for the local solution of (23.1) and P_{n+1} is the component of $f^{-1}(P_n)$ containing α (compare proof of Theorem 21.4). Then we can consecutively extend ϕ from P_n to P_{n+1} by means of the Schröder equation:

$$\phi_{n+1}(z) = \rho^{-1}\phi_n(fz), \quad z \in P_{n+1}.$$

Since the maps $f : P_{n+1} \to P_n$ are branched coverings, all the extensions $\phi_n : P_n \to \rho^{-n}\phi(P_0)$ are branched coverings, and hence the limiting map $\phi : D \to \mathbb{C}$ is a branched covering as well. As $\deg(f|\,P_n) > 1$ eventually for all n (once the P_n contain a critical point of f), we have $\deg \phi_n \to \infty$.

Moreover, any critical point of ϕ_{n+1} is either a critical point of f or else an f-preimage of a critical point ϕ_n.

The last assertion is also easily supplied, consecutively for the maps ϕ_n. $\square$

So, in the quadratic case, $f : z \mapsto z^2 + c$, the map ϕ branches on $(f \mid D)^{-pm}(0)$, where p is the period of α. Moreover, the critical points of ϕ are simple in this case, and its critical values are $\rho^{-n}\,\phi(0)$, $n = 0, 1, \dots$.

23.1.5. *Cycles of curves.* Periodic curves landing at periodic points play a key role in the Polynomial Dynamics. Let us take a first glance at this phenomenon.

EXERCISE 23.9. *For any attracting linear map $L_\rho : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$, there is a proper invariant ray $\mathcal{R} : \mathbb{R}_+ \to \mathbb{C}^*$ (going from 0 to ∞) .*

As simple as it is, this fact is somewhat counterintuitive when $\arg\rho \neq 0 \mod 2\pi$, as it may look like the twist by $\arg\rho$ is not compatible with the invariance of the ray. What it really implies that an invariant curve *spirals* around the origin. This phenomenon is in the heart of the spiralling nature of Julia sets, clearly visible on the pictures.

LEMMA 23.10. *For any attracting germ f (21.1) there exists an invariant curve landing at 0.*

PROOF. Let $D := \mathbb{D}_\varepsilon$ for a small ε, and let $A := \overline{D} \smallsetminus f(D)$ be the corresponding fundamental annulus. Connect a point $z \in \partial D$ to its image $f(z) \in \partial(fD)$ by an arc γ properly embedded into A, and let $\mathcal{R} := \bigcup_{n=0}^{\infty} f^n(\gamma)$.

Alternatively, the lemma follows from the previous Exercise by linearization (Proposition 23.2). ☐

An arc γ landing at 0 is called *periodic* with period $\mathfrak{q}$ (under an attracting map (21.1)) if the curves $f^n(\gamma)$, $n = 0, 1, \dots, \mathfrak{q} - 1$, meet only at 0, and $f^{\mathfrak{q}}(\gamma) \subset \gamma$. Let us call these curves γ_k, $k \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$, labeling them in a cyclic order around 0 (see §1.3.3). Being an orientation preserving local homeomorphism, f preserves this order, so it induces a rotation of this cycle of curves: $f(\gamma_k) \subset \gamma_{k+\mathfrak{p}}$. Under these circumstances, $\mathfrak{p}/\mathfrak{q}$ is called the *combinatorial rotation number* of the cycle of curves.

LEMMA 23.11. *For any attracting map f (21.1), $0 < |\rho| < 1$, and any rational $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}/\mathbb{Z}$, there exists a cycle of curves γ_k landing at 0 with combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$.*

PROOF. For the linear map $L \equiv L_{\mathfrak{p}/\mathfrak{q}} : z \mapsto e(\mathfrak{p}/\mathfrak{q})\, z$, we can let $\gamma = \mathbb{R}_+$. Since any attracting map f is locally conjugate to L (Proposition 23.2), the conclusion follows. ☐

EXERCISE 23.12. *Construct such a cycle of curves directly, without using* Proposition 23.2.

EXERCISE 23.13. (i) *Two disjoint invariant curves, γ_0 and γ_1, can be connected by an arc $\delta : [0, 1] \to \mathbb{C}$ with $\delta(i) \in \gamma_i$ for $i = 0, 1$, whose interior is disjoint from the γ_i and such that $f(\delta) \cap \delta = \emptyset$ (a "bridge" between γ_1 and γ_2).*
(ii) *For a cycle of curves γ_k, there exists a Jordan disk D such that ∂D intersects each γ_k at a single point and $f(D) \Subset D$.*

A cycle of curves $(f^k\gamma)$ as above naturally projects to a simple closed curve $\boldsymbol{\gamma}$ in the fundamental torus $\mathbb{T}^2_f$. It represents some homology class $[\boldsymbol{\gamma}] \in H_1(\mathbb{T}^2_f)$. This homology group is naturally realized as the lattice in $\mathbb{C}$ spanned by $\alpha = 2\pi i$ and $\beta = \log\rho$ (or equivalently: $\alpha = 1$, $\beta = \log\rho/2\pi i$).

EXERCISE 23.14. *If a cycle of curves* $(f^k\gamma)$ *has combinatorial rotation number* $\mathfrak{p}/\mathfrak{q}$ *then* $[\gamma] = \mathfrak{q}\beta - \mathfrak{p}\alpha$ *(for an appropriate choice of* $\log\rho$*).*

23.2. Existence of Siegel disks. We will now give a simple proof of existence of Siegel disks in the quadratic family. Here it will be convenient to put a fixed point at the origin and to normalize the quadratic term so that $f_\lambda(z) = \lambda z + z^2$.

PROPOSITION 23.15. *In the quadratic family* $f_\lambda(z) = \lambda z + z^2$, $\lambda = e(\theta)$ *with* $\theta \in \mathbb{R}/\mathbb{Z}$*, the map* f_λ *is linearizable for Lebesgue almost all rotation numbers* θ*.*

PROOF. The idea is to construct Siegel disks as limits of attracting petals. To this end we need to control the size of the latter. By Proposition 23.8, the Königs map ϕ_λ is unbranched over the disk $\mathbb{D}_r$, where $r \equiv r_\lambda = |\phi_\lambda(-\lambda/2)|$. Hence there exists a petal $D_\lambda \ni 0$ containing the critical point $-\lambda/2$ on its boundary which is univalently mapped by ϕ_λ onto $\mathbb{D}_r$.

By Lemma 23.7, the function $\lambda \mapsto \phi_\lambda(-\lambda/2)$ is holomorphic on the unit disc $\mathbb{D}$. Let us show that it is also bounded, and in fact $r_\lambda < 2$. Indeed, it is trivial to check that the filled Julia set $\mathcal{K}(f_\lambda)$ is contained in the disc $\overline{\mathbb{D}}_2$. Hence

$$D_\lambda \subset \operatorname{int}\mathcal{K}(f_\lambda) \subset \mathbb{D}_2.$$

But then $r_\lambda < 2$ by the Schwarz Lemma applied to the inverse function

$$\psi_\lambda = \phi_\lambda^{-1} : (\mathbb{D}_{r_\lambda}, 0) \to (D_\lambda, 0), \quad \psi_\lambda'(0) = 1. \tag{23.3}$$

By the Fatou and Riesz'-Privalov Theorems (see Appendix §8.4) the function $g(\lambda) := \phi_\lambda(-\lambda/2)$ has non-vanishing radial limits

$$\bar{g}(\theta) = \lim_{\rho\to 1} g(\rho\, e(\theta)) \quad \text{for almost all } \theta \in \mathbb{R}/\mathbb{Z}.$$

Let us finally show that for such a θ, the map f_λ with $\lambda = e(\theta)$ is linearizable on the disk of radius $\bar{r} := |\bar{g}(\theta)|/2 > 0$. Indeed, the family of functions ψ_λ (23.3) with $\lambda = \rho\, e(\theta)$ is well defined, normalized, and normal (by the Little Montel) on the disk of radius $\bar{r}$ (as long as ρ is sufficiently close to 1). Then any limit function ψ conjugates the θ-rotation of $\mathbb{D}_{\bar{r}}$ to $f_\lambda \,|\, \psi(\mathbb{D}_{\bar{r}})$. □

So, in the family $f_\theta : z \mapsto e(\theta)\, z + z^2$, $\theta \in \mathbb{R}/\mathbb{Z}$, *almost all germs are renormalizable* (by the above Proposition), while *generic germs are not* (by Proposition 21.38). It is a striking illustration of the difference between probabilistic and topological viewpoints!

23.3. Natural extension. In this section, we will apply the natural extension construction from §19.16.3 to polynomial maps (not necessarily of degree two). It will give us an extra insight into the idea of *global linearization*.

Let us consider a polynomial $f : \mathbb{C} \to \mathbb{C}$, and let $\hat{f} : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be its natural extension.[11] For a neighborhood U of z, let $\hat{U} = \hat{U}(\hat{z}) = (U_{-n})_{n=0}^\infty$ be the *pullback* of U along $\hat{z}$, i.e., U_{-n-1} is defined inductively as the component of $f^{-1}(U_{-n})$ containing z_{-n-1}. Let us call the pullback $\hat{U}$ *regular* if the maps $f : U_{-n-1} \to U_{-n}$ are eventually univalent. In this case $\hat{U}$ is called a *local leaf* of $\hat{z}$.

A point $\hat{z} \in \hat{X}$ is called *regular* if it is contained in some local leaf. Let $\mathcal{R}_f$ be the space of regular points in $\hat{\mathbb{C}}$. Path connected components of $\mathcal{R}_f$ are called *global leaves.* We let $\mathcal{L}(\hat{z})$ be the global leaf though a point $\hat{z} \in \hat{\mathbb{C}}$.

[11] Note that in this section, $\hat{\mathbb{C}}$ stands for the *natural extension space* , rather than the *Riemann sphere.*

We define the *intrinsic* topology on a global leaf $\mathcal{L} \subset \mathcal{R}_f$ by letting all the regular pullbacks $\hat{U}(\hat{z})$ be the basis of neighborhoods of points $\hat{z} \in \mathcal{L}$. Moreover, if $f : U_{-n-1} \to U_{-n}$ are univalent for $n \geq N$, then the projection

$$\pi_{-N} : \hat{U} \to U_{-N}, \quad \hat{z} \mapsto z_{-N}$$

is homeomorphic, and we take it as a local chart on $\mathcal{L}$. Transition maps between such local charts are given by iterates of f, so that, they turn $\mathcal{L}$ into a Riemann surface.

EXERCISE 23.16. *Show that:*

(i) *the projections* $\pi_{-n} : \mathcal{L} \to \mathbb{C}$ (19.17) *are holomorphic;*

(ii) *the critical points of* $\pi : \mathcal{L} \to \mathcal{C}$ *are the backward orbits* $\hat{z} = (z_{-n})_{n=0}^{\infty}$ *passing through critical points of* f *(such orbits are called "critical"); find the degree of branching of* π *at* $\hat{z}$*;*

(iii) *the set of critical values of* $\pi : \mathcal{L} \to \mathbb{C}$ *is contained in the post-valuable set* $\mathcal{P}_f$*.*

(iv) *For any* $\hat{z}$*,* $\hat{f}$ *restricts to a conformal isomorphism* $\mathcal{L}(\hat{z}) \to \mathcal{L}(\hat{f}\hat{z})$*.*

We call $\mathcal{R}_f$ the *regular leaf space* for f. The map $\hat{f} : \mathcal{R}_f \to \mathcal{R}_f$ is an *isomorphism* of the regular leaf space (i.e., it is a homeomorphism that conformally maps leaves to leaves).

LEMMA 23.17. *Let* $\hat{\mathcal{C}}_f = \pi^{-1}(\overline{\mathcal{P}}_f)$*. Then the map* $\mathcal{L} \smallsetminus \hat{\mathcal{C}}_f \to \mathcal{L} \smallsetminus \overline{\mathcal{P}}_f$ *is a covering.*

PROOF. Let $z \in \mathbb{C} \smallsetminus \hat{\mathcal{C}}_f$ and let $U \subset \mathbb{C} \smallsetminus \overline{\mathcal{P}}_f$ be a little disk around z. Then

$$\pi^{-1}(U) = \bigsqcup_{\hat{z} \in \pi^{-1} z} \hat{U}(\hat{z}),$$

where each local leaf $\hat{U}(\hat{z})$ projects univalently onto U. $\square$

23.4. Global leaf of a repelling point. Local theory near a repelling periodic point α can be immediately reduced to the above local theory near an attracting point by taking the local inverse of f^p (where p is the period of α). In particular, repelling germs are also locally linearizable:

COROLLARY 23.18. *Consider a holomorphic germ* (21.1)*. Then there exist Jordan disks* $V \Supset V' \ni 0$ *such that* $f(V') = V$*,* $r > 0$*, and a conformal map* $\phi : (V, 0) \to \mathbb{D}_r$ *with* $\phi'(0) = 1$ *satisfying the equation:*

$$\phi(fz) = \rho\, \phi(z), \quad z \in V'. \tag{23.4}$$

The above properties determine uniquely the germ of ϕ *at the origin.*

As in §23.1.3, we can proceed by defining the fundamental torus $\mathbb{T}^2_{f,a} \approx \mathbb{T}^2_\rho = \mathbb{C}^*/L_\rho$ for f^p near α, where $L_\rho : z \mapsto \rho z$. (Note that $\mathbb{T}^2_\rho \approx \mathbb{T}^2_{1/\rho}$.) We can also consider "invariant curves" γ landing at 0, though strictly speaking γ is not invariant but rather $f(\gamma) \supset \gamma$. When we need to emphasize this detail, we call γ *essentially invariant*. With a similar convention, we can consider *cycles of curves* $\boldsymbol{\gamma} \equiv (f^k(\gamma))_{k=0}^{\mathfrak{q}-1}$ for f. Each cycle $\boldsymbol{\gamma}$ is endowed with its *combinatorial rotation number* $\mathfrak{p}/\mathfrak{q}$ (which is equal to the negative of the combinatorial rotation number of $\boldsymbol{\gamma}$ under f^{-1}).

However, as we have already seen, the global effect of repelling points on the dynamics is completely different from that of attracting points. In particular, as we will show momentarily, the inverse of the linearizing coordinate for a repelling point admits a global extension to an entire function.

Assume that $f : \mathbb{C} \to \mathbb{C}$ is a polynomial with a repelling fixed point α. Let us consider the inverse linearizing function $\psi : (\mathbb{D}_r, 0) \to (V, \alpha)$, $\psi = \phi^{-1}$. It satisfies the functional equation

$$\psi(\rho z) = f(\psi(z)), \quad z \in \mathbb{D}_{r/|\rho|}. \tag{23.5}$$

It allows us to extend ψ holomorphically to the disk $\mathbb{D}_{|\rho|r}$ by letting $\psi(\zeta) = f(\psi(\zeta/\rho))$ for $\zeta \in \mathbb{D}_{|\rho|r}$. Repeating this procedure, we can consecutively extend f to the disks $\mathbb{D}_{|\rho|^n r}$, $n = 1, 2, \dots$, so that in the end we obtain an entire function $\psi : \mathbb{C} \to \mathbb{C}$ satisfying (23.5). This funcion is called *Poincaré*.

The natural extension gives a nice dynamical way to construction the Riemann surface $\mathcal{L}$ of the inverse (multivalued) function $\phi = \psi^{-1}$. Namely, it can be interpreted as the global leaf $\mathcal{L}(\hat{\alpha})$ of the fixed point $\hat{\alpha} = (\alpha, \alpha, \alpha, \dots) \in \hat{\mathbb{C}}$ of $\hat{f}$. More precisely, let us consider the space of inverse orbits of f *converging to the fixed point* α:

$$\mathcal{L} = \{\hat{z} = (z_{-n})_{n=0}^{\infty} \in \hat{\mathbb{C}} : \quad z_{-n} \to \alpha\}.$$

The map f lifts to an invertible map $\hat{f} : \mathcal{L} \to \mathcal{L}$,

Since $z_{-n} \to 0$, $z_{-n} \in V$ for all $n \geq N$. Selecting U so small that $U_{-N} \subset V$, we see that $U_{-n} \subset V$ for all $n \geq N$, and hence all the maps $f : U_{-n-1} \to U_{-n}$ are univalent for $n \geq N$. Thus, $\hat{U}$ is regular for a sufficiently small U. Hence $\mathcal{L}$ is contained in a leaf of $\mathcal{R}_f$.

EXERCISE 23.19. *Show that $\mathcal{L}$ is a full leaf of $\mathcal{R}_f$.*

The following statement shows that $\mathcal{L}$ is the indeed the Riemann surface for ϕ:

PROPOSITION 23.20. *The maps ψ and ϕ lift to mutually inverse conformal isomorphisms $\hat{\psi} : (\mathbb{C}, 0) \to (\mathcal{L}, \hat{\alpha})$ and $\hat{\phi} : (\mathcal{L}, \hat{\alpha}) \to (\mathbb{C}, 0)$ conjugating $z \mapsto \rho z$ to $\hat{f}$ and such that $\pi \circ \hat{\psi} = \psi$.*

PROOF. For $u \in \mathbb{C}$, we let $\hat{\psi}(u) = (\psi(u/\rho^n))_{n=0}^{\infty} \in \mathcal{L}$.

Vice versa, if $\hat{z} = (z_{-n})_{n=0}^{\infty} \in \mathcal{L}$ then eventually $z_{-n} \in V$, so that the local linearizer ϕ is well defined on all z_{-n}, $n \geq N$. Let now $\hat{\phi}(\hat{z}) = \rho^n \phi(z_{-n})$ for any $n \geq N$. It does not depend on the choice of n since $\phi|V$ conjugates f to $z \to \rho z$.

We leave to the reader to check the stated properties of these maps. □

Let $\hat{\mathcal{K}}(f) = \pi^{-1}(\mathcal{K}(f))$.

LEMMA 23.21. *Assume $\mathcal{K}(f)$ is connected. Let U be a component of $\mathcal{L} \smallsetminus \hat{\mathcal{K}}(f)$. Then U is simply connected, so that, the projection $\pi : U \to \mathcal{D}_f(\infty)$ is a universal covering.*

PROOF. Since $\mathcal{K}(f)$ is connected, $\overline{\mathcal{P}}_f \subset \mathcal{K}(f)$. By Lemma 23.17, $U \to \mathcal{D}_f(\infty)$ is a covering map. Since $D_f(\infty)$ is conformally equivalent to $\mathbb{D}^*$, U is either conformally equivalent to $\mathbb{D}^*$ or is simply connected. But in the former case U would be a neighborhood of ∞ in $\mathcal{L} \approx \mathbb{C}$, so that, $\hat{\mathcal{K}}(f)$ would be bounded in $\mathcal{L}$. It is impossible since $\hat{\mathcal{K}}(f)$ is $\hat{f}$-invariant, and by Proposition 23.20 $\hat{f} \,|\, \mathcal{L}$ is conjugate to $z \mapsto \rho z$ with $|\rho| > 1$. □

COROLLARY 23.22. *Under the circumstance of the previous lemma, assume U is periodic with period $\mathfrak{q}$, i.e., $\hat{f}^{\mathfrak{q}}(U) = U$. Then $\hat{f}^{\mathfrak{q}} : U \to U$ is conformally conjugate to a hyperbolic automorphism $T : z \mapsto 2^{\mathfrak{q}} z$ of $\mathbb{H}$.*

PROOF. The Böttcher isomorphism $B : \mathcal{D}_f(\infty) \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ (see §23.5.2 below) lifts to an isomorphism $\hat{B} : U \to \mathbb{H}$ between the universal coverings. It conjugates $\hat{f}^{\mathfrak{q}}|U$ to a lift of $z \mapsto z^{2^{\mathfrak{q}}}$ by the exponential $\exp : \mathbb{H} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$, which is the desired hyperbolic automorphism. □

Thus, the quotient

$$A := U/\hat{f}^{\mathfrak{q}} = \left(\bigcup_{k=0}^{\mathfrak{q}-1} \hat{f}^k(U) \right) / \hat{f}$$

is an annulus of modulus $\pi/(\mathfrak{q} \log 2)$. Moreover, it is naturally embedded into the quotient torus $\mathbb{T}^2_{f,a} = \mathcal{L}^*/\hat{f} \approx \mathbb{T}^2_\rho$. The equator E of A is represented by a periodic curve γ landing at 0 with period $\mathfrak{q}$. This curve has some combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$. By Exercise 23.14, E represents a cycle in $H_1(\mathbb{T}^2_\rho)$ with coordinates $(\mathfrak{q}, -\mathfrak{p})$ in the basis $(\log \rho, 2\pi i)$. We will make use of these remarks in §24.6.

23.5. Superattracting points and Böttcher coordinates.

23.5.1. *Böttcher equation.*

THEOREM 23.23. *Let $f : z \mapsto z^d + a_{d+1} z^{d+1} + \dots$ be a holomorphic map near the origin, $d \geq 2$. Then there exists an f-invariant Jordan disk $V \ni 0$, $r \in (0,1)$, and a conformal map $B : (V, 0) \to (\mathbb{D}_r, 0)$ satisfying the equation:*

$$B(fz) = B(z)^d. \tag{23.6}$$

The above properties determine uniquely the germ of B at the origin, up to post-composition with rotation $z \mapsto e(1/(d-1)) \cdot z$ (so, it is unique in the quadratic case $d = 2$). Moreover, it can be normalized so that $B'(0) = 1$.

The map B is called the *Böttcher function*, or the *Böttcher coordinate* near 0. Equation (23.6) is called the *Böttcher equation.* In the Böttcher coordinate the map f assumes the normal form $z \mapsto z^d$.

PROOF. The Böttcher function can be given by the following explicit formula:

$$B(z) = \lim_{n \to \infty} \sqrt[d^n]{f^n z}, \tag{23.7}$$

where the value of the d^nth root is selected so that it is tangent to the id at ∞. Obviously, this function, if exists, satisfied the Böttcher equation. So, we only need to check that the limit exists.

Let $z_n = f^n z$, where $z_0 \equiv z$. Then

$$\frac{\sqrt[d^{n+1}]{z_{n+1}}}{\sqrt[d^n]{z_n}} = \frac{\sqrt[d^{n+1}]{z_n^d(1 + O(z_n))}}{\sqrt[d^n]{z_n}} = \sqrt[d^{n+1}]{(1 + O(z_n)} = 1 + O\left(\frac{z_n}{d^{n+1}}\right).$$

Hence

$$B(z) = \lim_{n \to \infty} \sqrt[d^n]{z_n} = z \prod_{n=0}^{\infty} \frac{\sqrt[d^{n+1}]{z_{n+1}}}{\sqrt[d^n]{z_n}} = z \prod_{n=0}^{\infty} \left(1 + O\left(\frac{z_n}{d^{n+1}}\right)\right) = z(1 + O(z)),$$

where the last product is convergent uniformly at a superexponential rate.

Finally, uniqueness of the Böttcher function follows from the exercise below. □

EXERCISE 23.24. *Let $d \geq 2$. Show that there are no conformal germs commuting with $g : z \mapsto z^d$ near the origin, except rotations*

$$z \mapsto e(k/(d-1)) \cdot z, \quad k \in \mathbb{Z}/(d-1)\mathbb{Z}.$$

Let us now consider a quadratic polynomial f_c near ∞. Since ∞ is a superattracting fixed point of f of degree 2, the map f_c near ∞ can be reduced in the Böttcher coordinate to the map $z \mapsto z^2$ (Theorem 23.23). Thus, there is a Jordan disk $V = V_c \subset \mathbb{C}$ whose complement $\mathbb{C} \smallsetminus V$ is f_c–invariant, some $R > 1$, and a conformal map $B_c : \mathbb{C} \smallsetminus V \to \mathbb{C} \smallsetminus \mathbb{D}_R$ satisfying the Böttcher equation:

$$B_c(f_c z) = B_c(z)^2. \tag{23.8}$$

Moreover, $B_c(z) \sim z$ as $z \to \infty$.

In this situation, explicit formula (23.7) assumes the

$$B_c(z) = \lim_{n\to\infty} (f_c^n(z))^{1/2^n}, \ |z| > R, \tag{23.9}$$

where R is sufficiently big and the root in the right-hand side is selected in such a way that it is tangent to the identity at ∞.

Next, we will globalize the Böttcher function B_c.

23.5.2. *Connected case: Böttcher vs Riemann.*

THEOREM 23.25. *Let $f_c : z \mapsto z^2 + c$ be a quadratic polynomial with connected Julia set. Then the Böttcher function admits an analytic extension to the whole basin of ∞. Moreover, it conformally maps $\mathcal{D}_c(\infty)$ onto the complement of the unit disk and globally satisfies* (23.8).

PROOF. We will skip label c from the notations. Let, as usually, $f_0(z) = z^2$.

Let $U^n = \hat{\mathbb{C}} \smallsetminus f^{-n}\overline{V}$. Then $U^0 \subset U^1 \subset U^2 \subset \dots$ and $\bigcup U^n = \mathcal{D}_c(\infty)$. Since the filled Julia set $\mathcal{K}(f)$ is connected, the domains U^n are topological disks and the maps $f : U^{n+1} \to U^n$ are double coverings branched point at ∞ (recall the proof of Theorem 20.5).

Let $\Delta^n = \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}_{R^{1/2^n}}$. By Lemma 3.11, the Böttcher function $B : U^0 \to \Delta^0$ admits a lift $\tilde{B} : U^1 \to \Delta^1$ such that $f_0 \circ \tilde{B} = B \circ f$ (see Figure 23.5.2). But the Böttcher equation tells us that $B : U^0 \to \Delta^0$ is a lift of its restriction $B : f(U^0) \to f_0(\Delta^0)$. If we select $\tilde{B}$ so that $\tilde{B}(z) = B(z)$ at some finite point $z \in U^0$, then these two lifts must coincide on U^0, i.e., $\tilde{B}\,|\,U^0 = B$. Thus, $\tilde{B}$ is the analytic extension of B to U^1. Obviously, it satisfies the Böttcher equation as well.

In the same way, the Böttcher function can be consecutively extended to all the domains U^n and hence to their union, $\mathcal{D}_c(\infty)$. □

Thus, the Böttcher function provides us with the Riemann mapping from $\mathbb{C} \smallsetminus \mathcal{K}(f)$ to the complement of the unit disk. Given the intricate fractal structure of the Julia set, this is quite remarkable that its complement can be uniformized in this explicit way!

One can also go the other way around and construct the Böttcher function by means of uniformization:

EXERCISE 23.26. *Let $f \equiv f_c$ be a quadratic polynomial with connected Julia set. Then the basin of infinity $\mathcal{D}_f(\infty)$ is a conformal disk. Uniformize it by the complement of the unit disk, $\psi : (\mathbb{C} \smallsetminus \overline{\mathbb{D}}, \infty) \to (\mathcal{D}_f(\infty), \infty)$, normalized at ∞ so*

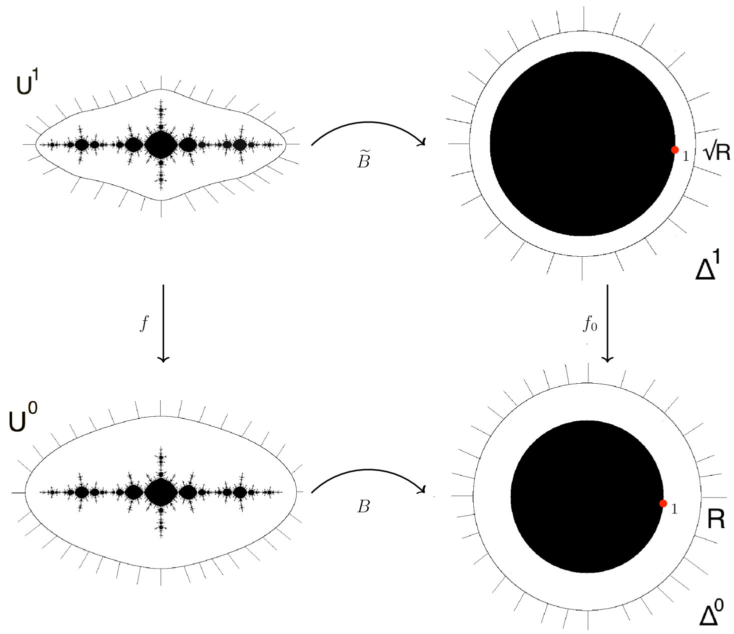

FIGURE 23.1. Lift of the Böttcher function.

that $\psi(z) \sim \rho z$ *with* $\rho > 0$. *Prove* (*without using the* Böttcher Theorem) *that* ψ *conjugates* $f_0 : z \mapsto z^2$ *on* $\mathbb{C} \smallsetminus \overline{\mathbb{D}}$ *to* f *on* $\mathcal{D}_f(\infty)$ (*and that* $\rho = 1$).

EXERCISE 23.27. *Prove that in the connected case,* $\mathcal{D}(\infty)$ *is the maximal domain of analyticity of the Böttcher function* B.

Given two quadratic polynomials, f_c and f_b, with connected Julia sets, the composition

$$h : \mathcal{D}_c \to \mathcal{D}_b, \quad h \equiv h_{cb} := B_b^{-1} \circ B_c, \tag{23.10}$$

provides us with a conformal conjugacy between f_c and f_b on their basins of infinity. We call it the *Böttcher conjugacy.* In particular, for $b = 0$, the Böttcher conjugacy $h_{c0} : \mathcal{D}_c \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ can be identified with the Böttcher function B_c itself.

Let us finish with a curious consequence of Theorem 23.25. The *capacity* of a connected compact set $K \subset \mathbb{C}$ is defined as the radius R of the disk $\mathbb{D}_R$ such that the domain $\mathbb{C} \smallsetminus K$ can be conformally mapped onto $\mathbb{C} \smallsetminus \overline{\mathbb{D}}_R$ by a map tangent to id at ∞.

COROLLARY 23.28. *Let* $f_c : z \mapsto z^2 + c$. *Then the capacity of the filled Julia set* $\mathcal{K}_c$ *is equal to* 1.

23.5.3. *Cantor case: Böttcher position of the critical value.* In the disconnected case the Böttcher function $B \equiv B_c$ cannot be any more extended to the whole basin of ∞, as it branches at the critical point 0. However, B can still be extended to a big invariant region $\Omega \equiv \Omega_c$ containing 0 on its boundary.

THEOREM 23.29. *Let* $f \equiv f_c : z \mapsto z^2 + c$ *be a quadratic polynomial with disconnected Julia set. Then the Böttcher function* $B \equiv B_c$ *admits the analytic* *extension to a domain* $\Omega \equiv \Omega_c$ *bounded by a "figure eight" curve branched at the critical point 0. Moreover,* B *maps* Ω *conformally onto the complement of some disk* $\overline{\mathbb{D}}_R$ *with* $R > 1$*. The inverse map extends continuously to a map* $\mathbb{C} \setminus \mathbb{D}_R \to \overline{\Omega}$ *which is one-to-one except that it maps two antipodal points* $\pm \operatorname{Re}(\theta) \in \mathbb{T}_R$ *to* 0*.*

PROOF. Since $0 \in \mathcal{D}_f(\infty)$, the orb(0) lands at the domain V of the Böttcher function near ∞. By shrinking V, we can make $f^n 0 \in \partial V$ for some $n > 0$. Then there are no obstructions for consecutive extensions of B to the domains

$$U^k = \hat{\mathbb{C}} \setminus f^{-k}\overline{V}, \quad k = 0, 1, \dots, n$$

(in the same way as in the connected case). All these domains are bounded by real analytic curves except the last one, U^n, which is bounded by a figure eight curve branched at 0. This is the desired domain Ω.

The last statement follows from the theory of prime ends (see Exercise 8.14) and the symmetry of the figure-eight with respect to the origin. ☐

Since the critical value c belongs to $\partial U^{n-1} \subset \Omega$ (in the notation of the above proof), the expression $B_c(c)$ is well-defined (provided the Julia set $\mathcal{J}_c$ is disconnected). It gives the *Böttcher position of the critical value* as a function of the parameter c. This function will play a crucial role in what follows.

In case when $\mathcal{K}_c$ is connected, we let $\Omega \equiv \Omega_c$ be the whole basin of infinity, $\Omega_c := \mathcal{D}_c(\infty)$. In either case, formula (23.9) remains valid globally (by construction of the extension):

$$B_c(z) = \lim_{n\to\infty} (f_c^n(z))^{1/2^n}, \ z \in \Omega_c, \tag{23.11}$$

where the root in the right-hand side is selected to be tangent to the identity at ∞.

For a point $z \in \Omega$, the polar coordinates (r, θ) of $B(z)$ are called the *external coordinates* of z.

Let us consider the logarithmic differential of B in Ω,

$$\omega \equiv \omega_c := i\frac{dB}{B}, \quad \text{i.e.,} \quad \omega(z)\, dz := i\frac{B'(z)}{B(z)}\, dz.$$

It is an Abelian differential which is nicely transformed under the dynamics:

$$f^*\omega = 2\omega, \quad \text{i.e.,} \quad f'(z)\, \omega(fz) = 2\, \omega(z) \tag{23.12}$$

(since $f^*(\log B) = 2 \log B$ by the Böttcher equation). By means of this functional equation, ω can be extended to an Abelian differential on the whole basin $\mathcal{D}(\infty)$ (denoted in the same way). It has simple zeros at 0 and all its iterated preimages, i.e., its zero divisor is equal to $\operatorname{Crit}_f^\infty$. It supplies $\mathcal{D}(\infty) \setminus \operatorname{Crit}_f^\infty$ with a (Euc(1)) flat structure, turning it to a translation surface (see §2.7.5).

23.5.4. *External rays and equipotentials.* Foliations by external rays and equipotentials, introduced in the general context (see §8.3) have a good dynamical meaning on the basin of ∞.

The map $f_0 : z \mapsto z^2$ on $\mathbb{C} \setminus \overline{\mathbb{D}}$ has two invariant foliations, by the straight rays going to ∞ and by round circles centered at the origin. (Note that the latter foliation is dynamically defined: see the hint to Exercise 23.24.) We will label the rays by their angles $\theta \in \mathbb{R}/\mathbb{Z}$ and the circles by their radii $r > 1$ or by their "heights" $t = \log r \in \mathbb{R}_+$. So,

$$\mathcal{R}_0^\theta = \{re(\theta) : \ r \in \mathbb{R}_+\}, \quad \mathcal{E}_0^r \equiv \mathcal{E}_0^t = \{re(\theta) : \ \theta \in \mathbb{R}/\mathbb{Z}\}, \ t = \log r,$$

where the subscript 0 suggests affiliation to the map f_0. Note that

$$f_0(\mathcal{R}_0^\theta) = \mathcal{R}_0^{2\theta} \quad \text{and} \quad f_0(\mathcal{E}_0^t) = \mathcal{E}_0^{2t}.$$

If we now take an arbitrary quadratic polynomial $f \equiv f_c$, then by means of the Böttcher function $B \equiv B_c$, the above two foliations can be transferred to the domain $\Omega \equiv \Omega_c \subset \mathcal{D}_c(\infty)$, supplying it with the foliation by *external rays* and *equipotentials*. The rays naturally labeled by the corresponding *external angles* θ, while the equipotentials are labeled by the *equipotential radii* r *or heights* t. Let $\mathcal{R}^\theta \equiv \mathcal{R}_c^\theta$ stand for the external ray of angle θ and let $\mathcal{E}^r \equiv \mathcal{E}_c^r$ or $\mathcal{E}^t \equiv \mathcal{E}_c^t$ stand for the equipotential of height $t = \log r$. (We will also use notation $\mathcal{R}^\theta(t) \equiv \mathcal{R}^\theta(r)$ for the point on the ray $\mathcal{R}^\theta$ whose equipotential level is equal to $t = \log r$.)

Note that these foliations can also be interpreted as *vertical* and *horizontal* foliations of the Abelian differential ω (23.12). (The factor "i" in the definition of ω was introduced in order to make the rays "vertical".)

If $\mathcal{K}(f)$ is connected then $\Omega = \mathcal{D}(\infty)$, so *the whole basin of infinity is foliated by the external rays and equipotentials.*

In the disconnected case, we can pull the two foliations in Ω back by the iterates of f to obtain *singular foliations* on the whole basin of ∞. Equivalently, we can consider the vertical and horizontal foliations of the Abelian differential ω. (See §2.7.5.)

These foliations have simple singularities located at the zeros of ω, i.e., on the set Crit_f^∞. Horizontal leaves are relatively compact in $\mathcal{D}(\infty)$, and are either simple closed curves or loops of figure-eights. Vertical leaves can either go to ∞ in the positive direction or else *crash* at a singular point. The former will still be called *external rays*, the latter will be referred to as *separatrices*. In the negative direction, any vertical leaf either goes to the Julia set or crashes at a singular point.

Note that any external ray is the maximal non-singular extension of a ray defined earlier in Ω. There are at most countably many rays that *crash* at iterated preimages of the 0. Two rays landing at the critical point 0 will be called the *critical rays*. The particularly important ray going through the critical value will be called the *valuable ray* (its external angle will be also called "valuable"). Of course, it contains the (coinciding) images of the critical rays.

The union of all external rays is a simply connected domain $\hat{\Omega} \equiv \hat{\Omega}_c \supset \Omega_c$ obtained by removing from $\mathcal{D}_c(\infty)$ the closures of all separatrices.

Problem 23.30. *The Böttcher function B analytically extends to a single valued function on the whole domain $\hat{\Omega}$ that univalently maps it onto a the complement of a "Levin-Sodin hedgehog" obtained by attaching to $\overline{\mathbb{D}}$ countably many "needles"*

$$N_k := \{re^{i\theta_k}: \ 1 \le r \le R_k\}.$$

of shrinking length (i.e., $R_k \to 0$) attached to a dense set of points $e(\theta_k) \in \mathbb{T}$. What are the values of θ_k and R_k?

The figure-eight that bounds Ω will be called the *critical figure-eight* or the *critical equipotential*.

For $r = e^t > B_c(0)$, we let

$$\Omega_c(r) \equiv \Omega_c(t) = \{z : \ |B_c(z)| > r\},$$

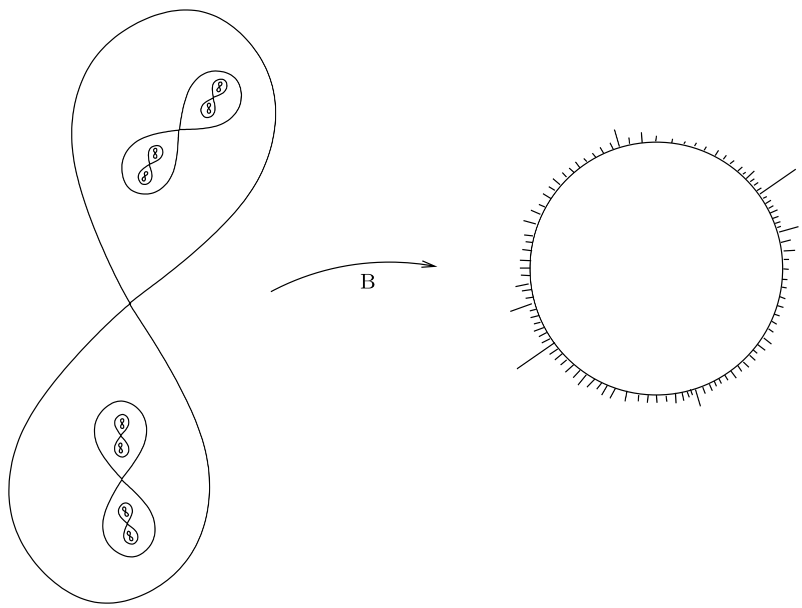

FIGURE 23.2. Levin-Sodin hedgehog.

and call it a *superpotential disk* of radius r, or of height t. The complementary Jordan disk $\Sigma_c(r) \equiv \Sigma_c(t) = \mathbb{C} \smallsetminus \Omega_c(r)$ will be called a *subpotential* disk of radius r, or of height t.

23.6. Dynamical Green function.

23.6.1. *Brolin Formula.* The *Green function* of a quadratic polynomial $f = f_c$ near ∞ can be introduced as follows:

$$G(z) \equiv G_c(z) = \log |B_c(z)|, \tag{23.13}$$

where B_c is the Böttcher function of f_c. The Green function is harmonic wherever the Böttcher function is defined (since the Böttcher function never vanishes) and has a logarithmic singularity at ∞ [compare §10.9]:

$$G(z) = \log |z| + o(1).$$

In the connected case, (23.13) defines the Green function in the whole basin $\mathcal{D}_c(\infty)$. In the disconnected case definition (23.13) can be used only on a subdomain of $\mathcal{D}_c(\infty)$, say on the complement of the Levin-Sodin hedgehog [which contains the critical superpotential domain Ω_c, see (23.11)]. However, in either case the Green function satisfies the equation:

$$G(fz) = 2G(z). \tag{23.14}$$

This equation can be obviously used in order to extend the Green function harmonically to the *whole basin of* ∞. Let us summarize simple properties of this extension:

EXERCISE 23.31. (i) *In the connected case the Green function does not have critical points. In the disconnected case, its critical points are simple saddles located at the critical points of the iterated f (i.e., on the set* Crit_f^∞*).*

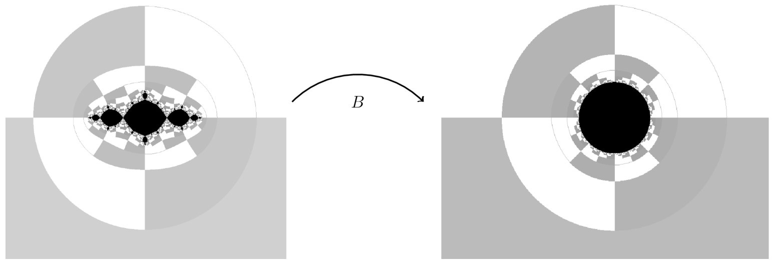

FIGURE 23.3. Dyadic grid.

(ii) *Equipotentials are the level sets of the Green function, while external rays* (*and their preimages*) *are its gradient curves.*

(iii) *The Brolin formula holds:*

$$G(z) = \lim_{n\to\infty} \frac{1}{2^n} \log |f^n z|, \quad z \in \mathcal{D}(\infty). \tag{23.15}$$

(iv) *Extension of the Green function by* 0 *through the filled Julia set* $\mathcal{K}(f)$ *gives a continuous subharmonic function on the whole complex plane.*

(v) *The Julia set is Dirichlet regular.*

These properties show that the dynamical Green function G is indeed the Green function of $\mathcal{D}(\infty)$ with the pole at ∞ as was defined in the general context in §10.9. Moreover, the dynamical notion of external rays and equipotentials matches with the general one.

We also see that though the Böttcher function B is not globally defined, its absolute value $|B| = \exp G$ admits a continuous extension to the whole plane satisfying the functional equation $|B(fz)| = |B(z)|^2$. In what follows, we assume that $|B|$ is extended this way.

23.6.2. *Dyadic grid.* Let us fix some $r = e^t > 1$, and consider the annulus $\mathbb{A}(1, r]$ cut along the real line, $\Delta^0 = \mathbb{A}(1, r] \smallsetminus (1, r]$. Let us pull it back by the dynamics of $f_0 : z \mapsto z^2$; let $\Delta^n_{\bar{i}} \equiv \Delta^n_{i_0\dots i_{n-1}}$ be the pullback under the branch of $f_0^{-n} : \mathbb{C} \smallsetminus \mathbb{R}_+ \to \mathbb{C}$ that maps $\mathbb{T} \smallsetminus \{1\}$ to the dyadic interval $I^n_{\bar{i}}$ (see §19.12).

It provides us with the tiling of each annulus $\mathbb{A}(1, r^{1/2^n})$ by 2^n rectangles $\Delta^n_{\bar{i}}$ such that

$$\Delta^{n+1}_{i_0\dots i_n} \subset \Delta^n_{i_0\dots i_{n-1}} \quad \text{and} \quad f_0(\Delta^{n+1}_{i_0\dots i_n}) = \Delta^n_{i_1\dots i_n}. \tag{23.16}$$

Let now f be a quadratic polynomial with connected Julia set. Taking the pullback of the above grid under the Böttcher map, $D^n_{\bar{i}} := B^{-1}(\Delta^n_{\bar{i}})$, we obtain the corresponding tilings of external annuli neighborhoods of the Julia set. Since B is equivariant, the behavior of this grid under the dynamics and the inclusion is the same as in (23.16).

This grid gives a useful dynamical picture for f in the external neighborhood of the Julia set.

23.6.3. *Holomorphic dependence on parameters.* Let Λ be a domain in $\mathbb{C}$. Let us consider a holomorphic family of superattracting germs over Λ,

$$f_\lambda(z) = z^d + a_{d+1}(\lambda)\, z + \dots, \quad \lambda \in \Lambda,$$

meaning that $(\lambda, z) \mapsto f_\lambda(z)$ is a holomorphic function in two variables on some domain in $\Lambda \times \mathbb{C}$. For each $\lambda \in \Lambda$, let us normalize the corresponding Böttcher functions B_λ so that $B'_\lambda(0) = 1$.

PROPOSITION 23.32. *Under the above circumstances,* $\forall\ \lambda_0 \in \Lambda\ \exists\ \delta > 0$ *such that* $(\lambda, z) \mapsto B_\lambda(z)$ *is well defined and holomorphic in the bidisk* $\mathbb{D}_\delta(\lambda_0) \times \mathbb{D}_\delta$.

PROOF. Let us go back to the explicit formula (23.7) for the Böttcher function. All the estimates that prove convergence in this formula are locally uniform in (λ, z). Hence the limit is holomorphic in both variables. □

REMARK 23.33. The parameter domain Λ can be an arbitrary complex manifold, even infinite-dimensional. The proof is the same. (Alternatively, one can use that holomorphicity can be detected by one-dimensional slices.)

Let us apply the above Proposition to the quadratic family. Let

$$\mathbf{D} \equiv \mathbf{D}(\infty) = \{(c, z) \in \mathbb{C}^2 :\ z \in \mathcal{D}_c(\infty)\}. \tag{23.17}$$

EXERCISE 23.34. *The set* $\mathbf{D}(\infty)$ *is a domain in* $\mathbb{C}^2$.

LEMMA 23.35. *The fibered Green function* $\mathbf{G} : (c, z) \mapsto G_c(z)$ *is continuous on* $\mathbf{D}$.[12]

PROOF. The orbits of $(f_c^n z)_{n\in\mathbb{N}}$, $(c, z) \in \mathbf{D}$, escape to ∞ at a locally uniform rate, which implies that convergence in the Brolin formula (23.15) is locally uniform on $\mathbf{D}$. □

Let

$$\mathbf{\Omega} = \{(c, z) \in \mathbb{C}^2 :\ z \in \Omega_c\} = \{(c, z) \in \mathbf{D} :\ G_c(z) > G_c(0)\}.$$

COROLLARY 23.36. *The set* $\mathbf{\Omega}$ *is open in* $\mathbb{C}^2$.

COROLLARY 23.37. *The Böttcher function* $B_c(z)$ *is holomorphic on* $\mathbf{\Omega}$.

PROOF. By Proposition 23.32, $B_c(z)$ is holomorphic in some neighborhood of the line at infinity, $\mathbb{C} \times \{\infty\} \subset \mathbb{C} \times \hat{\mathbb{C}}$. Its extension to $\mathbf{\Omega}$ is obtained by several liftings by the *fibered dynamics*

$$\mathbf{f} : \mathbb{C}^2 \to \mathbb{C}^2, \quad (c, z) \mapsto (c, f_c(z)). \tag{23.18}$$

Since $\mathbf{f}$ is holomorphic on $\mathbf{\Omega}$, these liftings are holomorphic as well. □

23.7. Fatou coordinates in parabolic petals.

23.7.1. *Linearization.* A parabolic germ f cannot be generally linearized in the whole neighborhood of a parabolic point, but it can be linearized in the petals. Below we let T be the translation $z \mapsto z + 1$.

THEOREM 23.38. *Let* $f : z \mapsto z + a_{\mathfrak{q}+1} z^{\mathfrak{q}+1} + \dots$ *be a parabolic germ near* 0. *Then in any attracting direction there exists an attracting petal* P *and a conformal isomorphism* $\phi : P \to \Pi$ *from* P *onto a* T*-invariant domain* Π (*which can be selected as the right half-plane*) *satisfying the Abel functional equation*

$$\phi(fz) = \phi(z) + 1. \tag{23.19}$$

This function is unique up to a translation $\phi + a$ (*and up to restricting/enlarging its domain*).

---

[12]In fact, this function is *pluriharmonic* on $\mathbf{D}$, i.e., its restrictions to one-dimensional holomorphic curves in $\mathbf{D}$ are harmonic.

In other words, ϕ *conformally conjugates* $f|\,P$ *to the translation* $z \mapsto z+1$ *on the right half-plane.*

We will outline three proofs for this result that give different insights into its nature.

Proof 1: Conformal viewpoint. Let us consider the projection of the petal onto the Écalle-Voronin cylinder, $p : P \to \mathrm{Cyl} = P/ < f >$. As was shown in Lemma 21.20, the latter is conformally isomorphic to the $\mathbb{C}/T$. Let $i : \mathrm{Cyl} \to \mathbb{C}/\mathbb{Z}$ be an isomorphism.

Let us consider the universal covering $\pi : \mathbb{C} \to \mathrm{Cyl} = \mathbb{C}/T$ of the cylinder, with the generating deck transformation T. Then i lifts to a map $\phi : P \to \mathbb{C}$ conjugating $f|\,P$ to $T|\,\phi(P)$. $\square$

Proof 2: Quasiconformal viewpoint. Let us go back to the proof of Lemma 21.20. Consider the fundamental strip S for the F-action in the right half-plane Q. Its boundary components are related by F which is close to the horizontal translation $T : z \mapsto z+1$. Let us first straighten this strip:

Exercise 23.39. *The strip S can be mapped onto the straight strip $\{0 \le \operatorname{Re} z \le 1\}$ by a qc homeomorphism h_0 which is equivariant on ∂S, i.e., $h_0(Fz) = h_0(z) + 1$ for z in the left-hand boundary component of S.*

Now h_0 can be equivariantly extended to the whole half-plane Q by letting $h(F^n z) = h_0(z) + n$ for $z \in S$. We obtain a qc homeomorphism $h : Q \to \Pi \equiv \{\operatorname{Re} z \ge 0\}$ conjugating $F|\,Q$ to the translation $T|\,\Pi$. Since F is conformal, the push-forward conformal structure $\mu = h_*(\sigma)$ is T-invariant. Pulling it back by the translations T^{-n}, $n \in \mathbb{N}$, we extend μ to a T-invariant conformal structure on the whole plane $\mathbb{C}$. (We will keep the same notation for it.)

By MRMT (see §14), there is a qc homeomorphism $\psi : \mathbb{C} \to \mathbb{C}$ that solves the Beltrami equation $\mu = \psi^*(\sigma)$. Conjugating T by ψ, we obtain a conformal automorphism acting freely on $\mathbb{C}$. Hence it is a translation. Normalizing ψ appropriately, we make this translation equal to T, so $\psi \circ T \circ \psi^{-1} = T$.

The composition $\psi \circ h$ is the desired solution of the Abel equation. $\square$

Exercise 23.40. *Show that the solution of the Abel equation depends continuously on f.* (*Defining topologies is part of the exercise.*)

23.7.2. *Proof 3: Analytic viewpoint.* In the non-degenerate case, consider the conjugate map $F : Z \mapsto Z + 1 + O(1/|Z|)$ near ∞ (21.4).

Project 23.41. (i) *Show that the forward orbits of F have the following assymptotics in the right-half plane $\{\operatorname{Re} Z > R\}$:*

$$F^n(Z) \sim n + a \log n + \Phi(Z), \quad n \to +\infty,$$

where $\Phi(Z) \sim Z$ is univalent.

(ii) *Show that Φ is a solution of the Abel equation*

$$\Phi(F(Z)) = \Phi(Z) + 1.$$

(iii) *Derive the degenerate case from the non-degenerate one.*

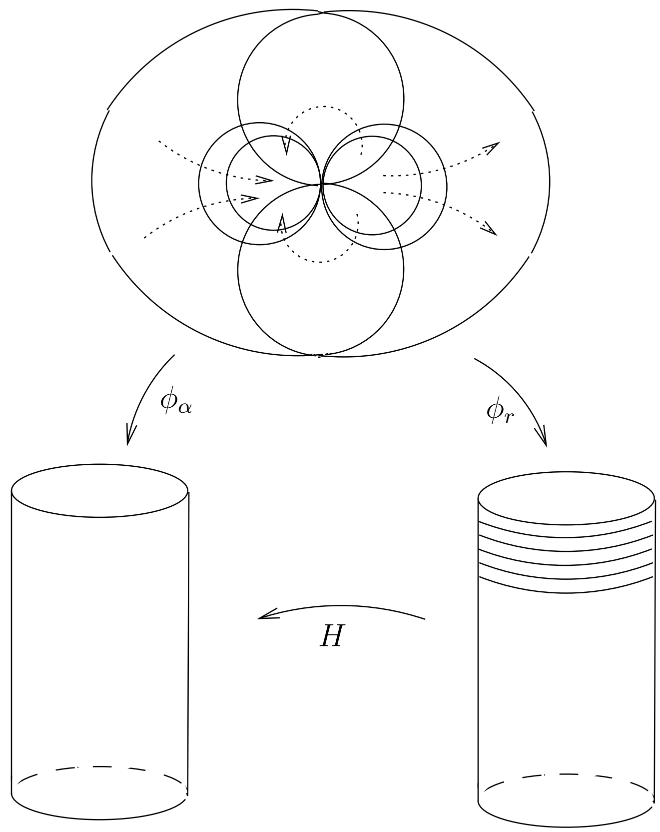

FIGURE 23.4. The "upper" horn map for a non-degenerate parabolic germ.

23.7.3. *Horn map and Écalle-Voronin moduli.* Since a parabolic germ is locally invertible, we can conside, along with the linearizing coordinate $\phi_{\rm a}$ on the attracting flower $\Phi^{\rm a}$, a linearizing coordinate $\phi_{\rm r}$ on the repelling one, satisfying the same functional equation

$$\phi_{\rm r}(f^{\mathfrak{q}}z) = \phi_{\rm r}(z) + 1.$$

Since neigboring petals in the flowers $\Phi^{\rm a}$ and $\Phi^{\rm r}$ overlap, we can consider $2\mathfrak{q}l$ compositions $\hat{H}_i = \phi_{\rm a} \circ \phi_{\rm r}^{-1}$ defined on the components of $\Phi^{\rm a} \cap \Phi^{\rm r}$. As the $\hat{H}_i$ commute with the translation by 1, they descend to partial maps between the cylinders,

$$H_i : \mathbb{C}/\mathbb{Z} \approx \mathrm{Cyl}^i_{\rm r} \to \mathrm{Cyl}^i_{\rm a} \approx \mathbb{C}/\mathbb{Z} \quad \text{(see Figure 23.4).}$$

As each H_i is defined near an appropriate end of the corresponding repelling cylinder, it is called the *horn map.* Mapping the cylinders to $\mathbb{C}^*$ by the exponential map e, we conjugate the H_i to $2\mathfrak{q}l$ germs h_i near 0.

If f is replaced with a conformally conjugate germ then the germs h_i do not change (up to complex scalings), so they provide us with *functional invariants* for the conformal conjugacy. Taking their Taylor coefficients, we obtain *Écalle-Voronin moduli* for conformal classes of parabolic germs. In fact, *they provide us with a complete conformal classification of parabolic germs.*

PROJECT 23.42. *Complete the theory of conformal classification of parabolic germs, including the realization theorem for a parabolic map with given horn maps.*

EXERCISE 23.43. *Show that a parabolic germ can be embedded into a local holomorphic flow iff the gluing maps are* id (*up to scaling*). *Such a germ is conformally equivalent to the Möbius one,* $z \mapsto z/(1-z)$. *What is the corresponding flow?*

23.7.4. *Globalization.* As in the attracting-repelling cases, in the polynomial case, the Fatou coordinates (and hence the horn map) can be globalized leading to a nice transcendental map (see Figure 23.5):

PROBLEM 23.44. *Let* $f \equiv f_c$ *be a quadratic polynomial with a parabolic periodic point* α. *Let* $\mathcal{D}^\bullet(\alpha)$ *be the immediate basin of* α *containing the critical point* 0. *Then:*

(i) *The attracting Fatou coordinate* $\phi_{\rm a}$ *for* α *extends to a branched covering* $\mathcal{D}^\bullet(\alpha) \to \mathbb{C}$ *of infinite degree, branched with degree two over* $\phi_{\rm a}(0) - \mathbb{N}$.

(ii) *The inverse repelling Fatou coordinate* $\phi_{\rm r}$ *extends to a branched covering* $\phi_{\rm r}^{-1} : \mathbb{C} \to \mathbb{C}$, *branched with degree two over the post-valuable set* $\mathcal{P} \equiv \operatorname{orb} v$.

(iii) *The horn map* H *extends to a branched covering* $\operatorname{Dom} H \to \mathbb{C}$ *of infinite degree branched, with degree two, over a single point* $\phi_{\rm a}(v) = \phi_{\rm a}(0) \mod \mathbb{Z}$. *Here* $\operatorname{Dom} H$ *is a punctured disk representing the upper end of the cylinder. The corresponding map* h *extends to a holomorphic map on a topological disk* $\operatorname{Dom} h = e(\operatorname{Dom} H) \cup \{0\}$ *fixing* 0.

Notes. The Linearization Theorem for simply attracting and repelling fixed points is due to Schröder [**Schr**] and Koenigs [**Ko**]. This result inuagurated the beginning of Holomorphic Dynamics.

Existence of linearizable neutral points (and hence, "Siegel disks") was demonstrated by Siegel in 1942 [**Si**]. A simple proof in the quadratic case (Proposition 23.15) is due to Yoccoz.

As the name adequately suggests, the local theory for superattracting germs was developed by Böttcher [**Bot**]. The global extension of the Böttcher funcion and the associated external objects (coordinates, rays and equipotentials) appeared in Douady & Hubbard's Orsay Notes [**DH2**].

The linearizing coordinates for parabolic maps were constructed by Leau [**Leau**] and Fatou [**F3**, Ch. VII]. The local conformal classification of parabolic gemrs was carried by Écalle [**Ec**] and Voronin[**Vo**]. This theory proved to be of great importance for understanding the phenomenon of *parabolic implosion* see [**D4, Lav, Sh1**].

The regular leaf space of the natural extension (§23.4) was introduced in [**LMin**].

24. Periodic ray configurations

24.1. Motivating problems. Consider a quadratic polynomial $f \equiv f_c$ with connected Julia set. As we know (Theorem 23.25), its basin of infinity is uniformized by the Böttcher map $B : \mathcal{D}_f(\infty) \to \mathbb{C} \smallsetminus \mathbb{D}$, which conjugates f to $z \mapsto z^2$. If the Julia set was locally connected then by the Carathéodory-Torhorst Theorem the inverse map $\phi := B^{-1}$ would extend continuously to the unit circle $\mathbb{T}$. This would

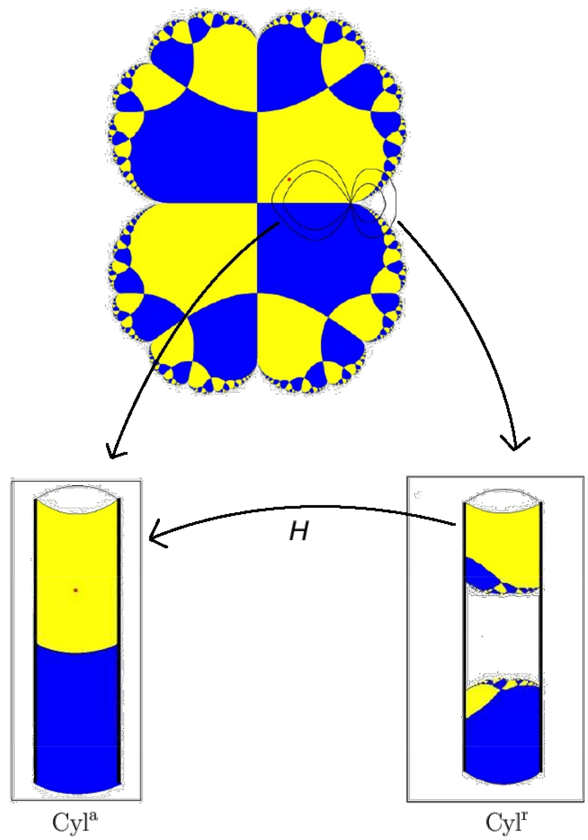

FIGURE 23.5. The cauliflower horn map.

give a representation of $f|\,\mathcal{J}_f$ as a quotient of the the doubling map $\theta \mapsto 2\theta \mod 1$ of the circle $\mathbb{R}/\mathbb{Z} \approx \mathbb{T}$. This observation immediately leads to the following problems:

1) Describe explicitly equivalence relations on the circle corresponding to all possible Julia sets;

2) Study the problem of local connectivity of the Julia sets.

It turns out that the first problem can be addressed in a comprehensive way. The second problem is very delicate. However, even non-locally connected examples can be partially treated due to the fact that many external rays always land at some points of the Julia set. This is the main theme of the following discussion.

24.2. Landing of rational rays. We say that an external ray $\mathcal{R}^\theta$ *lands* at some point z of the Julia set if $\mathcal{R}^\theta(t) \to z$ as $t \to 0$. Two rays $\mathcal{R}^{\theta/2}$ and $\mathcal{R}^{\theta/2+1/2}$ will be called "preimages" of the ray $\mathcal{R}^\theta$. Obviously, if some ray lands, then its image and both its preimages land as well.

An external ray $\mathcal{R}^\theta$ is called *rational* if $\theta \in \mathbb{Q}/\mathbb{Z}$, and *irrational* otherwise. Dynamically the rational rays are characterized by the property of being either periodic or preperiodic:

EXERCISE 24.1. *Let* $\mathcal{R} \equiv \mathcal{R}^\theta$.

If θ *is irrational then the rays* $f^n(\mathcal{R})$, $n = 0, 1, \dots,$ *are all distinct.*

Assume θ *is rational:* $\theta = \mathfrak{q}/\mathfrak{p}$, *where* $\mathfrak{q}$ *and* $\mathfrak{p}$ *are mutually prime. Then*

(i) *If* $\mathfrak{p}$ *is odd then* $\mathcal{R}$ *is periodic: there exists a* p *such that* $f^p(\mathcal{R}) = \mathcal{R}$.

(ii) *If* $\mathfrak{p}$ *is even then* $\mathcal{R}$ *is preperiodic: there are* p *and* $r > 0$ *such that* $f^r(\mathcal{R})$ *is a periodic ray of period* p, *while the rays* $f^k(\mathcal{R})$, $k = 0, 1, \dots, r-1$, *are not periodic.*

How to calculate p *and* r*?*

Let $\mathcal{R}^\theta[t_1, t_2] = \{R^\theta(t) : \ t_1 \le t \le t_2\}$ be the arc of the ray $\mathcal{R}^\theta$ between equipotentials of level t_1 and t_2.

LEMMA 24.2. *Let* f *be a quadratic polynomial with connected Julia set. Then the Euclidean length of any arc* $\mathcal{R}^\theta[t, 2t]$ *goes to 0 as* $t \to 0$, *uniformly in* θ.

PROOF. Endow the basin $\mathcal{D} \equiv \mathcal{D}_f(\infty)$ with the hyperbolic metric ρ. By compactness, the hyperbolic length of any arc $\mathcal{R}[t_0, 2t_0]$ with $1 \le t_0 \le 2$ is squeezed in between some constants $0 < l < L$.

For any $t \in (0,1)$, select $n \sim \log 1/t \in \mathbb{N}$ so that $2^n t \in [1,2)$. Since $g : \mathcal{D} \to \mathcal{D}$ is a covering map, it locally preserves ρ. Hence the hyperbolic length of any ray arc $\mathcal{R}^\theta[t, 2t]$ is also squeezed in between l and L.

But all these arcs accumulate on the Julia set as $t \to 0$, uniformly in θ, and the conclusion follows from the relation between the hyperbolic and Euclidean metrics (Lemma 7.7). $\square$

THEOREM 24.3. *Let* f *be a polynomial with connected Julia set. Then any periodic ray* $\mathcal{R} = \mathcal{R}_f^{\mathfrak{p}/\mathfrak{q}}$, $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}_{\mathrm{odd}}/\mathbb{Z}$, *of some period* p *lands at some repelling or parabolic point of* f *of a period dividing* p.

PROOF. As the ray $\mathcal{R}$ is periodic, it is invariant under some iterate $g = f^p$. Let $d = 2^p$. Consider a sequence of points $z_n = \mathcal{R}(1/d^n)$, and let γ_n be the sequence of arcs on $\mathcal{R}$ bounded by the points z_n and z_{n+1}. Then $g(\gamma_n) = \gamma_{n-1}$.

By Lemma 24.2, the Euclidean length of these arcs goes to 0 as $n \to \infty$. Hence the limit set of the sequence $\{z_n\}$ is a connected set consisting of the fixed points of g. Since g has only finitely many fixed points, this limit set consists of a single g-fixed point β (which is f-periodic with a period dividing p). It follows that the ray $\mathcal{R}$ lands at $\beta \in \mathcal{J}(f)$ (compare with the proof of Theorem 21.47).

Since $\beta \in \mathcal{J}(f)$, it can be either repelling, or parabolic, or Cremer. But the latter case is excluded by the Necklace Lemma, 21.48. $\square$

COROLLARY 24.4. *For a polynomial with connected Julia set, any preperiodic ray* $\mathcal{R} = \mathcal{R}_f^{\mathfrak{p}/2^r\mathfrak{q}}$, *where* $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}_{\mathrm{odd}}/\mathbb{Z}$ *and* $r > 0$, *lands at some repelling or parabolic preperiodic point of* f. *Moreover, its preperiod is equal to* r.

PROOF. To justify the last assertion, it is enough to show that for $r = 1$, our ray cannot land at any periodic point α. Otherwise the map f^p (fixing α) would not be locally injective near α. $\square$

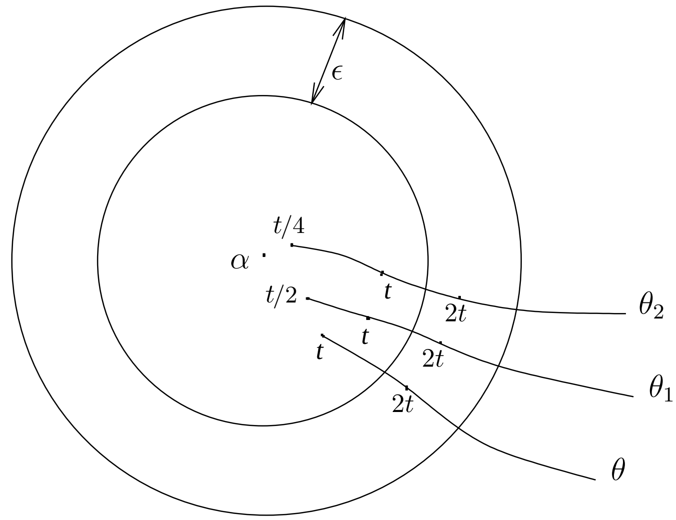

FIGURE 24.1. Construction of a landing ray.

24.3. Inverse Theorem: periodic points are landing points. It is less obvious that, vice versa, any repelling or parabolic point is a landing point of at least one ray:

24.3.1. *Repelling case.*

THEOREM 24.5. *Let f be a polynomial (of any degree $d \geq 2$) with connected Julia set. Then any repelling periodic point α is the landing point of at least one periodic ray. Moreover, there are only finitely many rays $\mathcal{R}^\theta$ landing at α, and all of them are periodic with the same period. (In particular, in the quadratic case, all angles θ are rational with odd denominator.)*

PROOF. Replacing f with its iterate, we can assume without loss of generality that α is a fixed point (albeit, f is not quadratic anymore but some polynomial of degree $d \geq 2$).

Let us consider a small disk $D := \mathbb{D}(\alpha, r)$ around α which is univalently mapped by f onto a strictly bigger disk $D' := f(D) \Supset D$ so that $|f'(z)| \geq \rho > 1$ for all $z \in D$. We let f^{-1} be the inverse branch in D' fixing α. The disks $D^n := f^{-n}(D)$, $n = 0, 1, \dots$ form a nest shrinking to α.

Let $\varepsilon = \operatorname{dist}(D, \partial D')$. By Lemma 24.2, there is a $t_0 > 0$ such that $\mathcal{R}^\theta[t, 2t] < \varepsilon$ for any $\theta \in \mathbb{R}/\mathbb{Z}$ and any $t \leq t_0$.

Let us now consider any point $\mathcal{R}^\theta(t) \in D$ with $t \leq t_0$. By our choices, $\mathcal{R}^\theta[t, 2t] \subset D'$. Then $\mathcal{R}^{\theta_1}[t/2, t] := f^{-1}(\mathcal{R}^\theta[t, 2t]) \subset D$. Let us extend this ray arc to a ray arc $\mathcal{R}^{\theta_1}[t/2, 2t]$. For the same reason as above, $\mathcal{R}^{\theta_1}[t, 2t] \subset D'$, and hence $\mathcal{R}^{\theta_1}[t/2, 2t] \subset D'$.

Let us now take the preimage of this ray arc, $\mathcal{R}^{\theta_2}[t/4, t] \subset D$, and extend it to a ray arc $\mathcal{R}^{\theta_2}[t/4, 2t]$. For the same reason, it is contained in D'. Proceeding in the same way, we inductively construct a sequence of ray arcs $\mathcal{R}^{\theta_n}[t/2^n, 2t] \subset D'$ such that

$$\mathcal{R}^{\theta_n}[t/2^n, t/2^k] \subset D^k, \quad k = 0, 1 \dots, n-1, \tag{24.1}$$

and $\mathcal{R}^{\theta_{n+1}}[t/2^{n+1}, t] = f^{-1}(\mathcal{R}^{\theta_n}[t/2^n, 2t])$. Let $\theta \in \mathbb{R}/\mathbb{Z}$ be the limit for a subsequence $\theta_{n(j)}$. By continuity of the Böttcher coordinate, for any $\tau \in (0, 2t)$, the ray arcs $\mathcal{R}^{\theta_{n(j)}}[\tau, 2t] \subset \overline{D}'$ uniformly converge to $\mathcal{R}^{\theta}[\tau, 2t]$. By (24.1),

$$\mathcal{R}^{\theta}(0, t/2^k] \subset D^k, \quad k = 0, 1, \dots$$

Hence the ray $\mathcal{R}^{\theta}$ lands at α.

Finiteness of the number of landing rays comes from the interplay between the doubling dynamics at infinity and an invertible local dynamics near α.

Namely, let us consider the set $\Theta \subset \mathbb{R}/\mathbb{Z}$ of the angles θ of all the rays $\mathcal{R}^{\theta}$ landing at α. It is invariant under the d-adic circle map $T \equiv T_d : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, where $d = \deg f$. Let $\Theta_0 \subset \Theta$ is the subset of angles θ constructed above (i.e., the set of all limit angles of the sequence (θ_n)). By Exercise 19.3, Θ_0 is a closed $T-$invariant subset of $\mathbb{R}/\mathbb{Z}$, on which T is surjective. (Note that at this point it is unclear whether Θ itself is a closed.)

Since f is a local diffeomorphism near α, the map $T : \Theta \to \Theta$ is injective. Hence its restriction to Θ_0 is bijective. But this map is expanding. By Exercise 19.37, Θ_0 is finite. Thus, $T : \Theta_0 \to \Theta_0$ is a permutation of a finite set, so all its points are periodic.

Finally, let us show that all other angles $\theta \in \Theta$ are periodic with the same period. Let us first consider any finite T-invariant subset $\Theta' \subset \Theta$. As a subset of $\mathbb{R}/\mathbb{Z}$, it is cyclically ordered. Moreover, this cyclic order coincides with the natural cyclic order of the corresponding rays $\mathcal{R}^{\theta}$, $\theta \in \Theta'$ (see §1.3.3). But since f is an orientation preserving local homeomorphism near α, its restriction to this set of rays preserves its cyclic order. By Exercise 1.127, all angles $\theta \in \Theta'$ have the same period $\mathfrak{q}$.

It follows that any T-periodic angle $\theta \subset \Theta$ has period $\mathfrak{q}$. Since the d-adic map has only finitely many periodic points of a given period, Θ contains only finitely many periodic angles. Let Θ_{per} be this finite set.

Assume that some $\theta \in \Theta$ is not periodic. Let I be the component of $(\mathbb{R}/\mathbb{Z}) \smallsetminus \Theta_{\text{per}}$ containing θ. Then $I \cap \Theta$ is an open interval in Θ in the sense of §1.11. Let $g := T^{\mathfrak{q}}$. Since $g|\,\Theta$ preserves the cyclic order and fixes $\partial I \in \Theta_{\text{per}}$, the set $I \cap \Theta$ is invariant under g, and the restriction of g to this set is monotonically increasing.

Assume for definiteness that $\theta < g(\theta)$. Then by monotonicity,

$$\theta < g(\theta) < g^2(\theta) < \dots,$$

and the whole orbit $(g^n(\theta))$ is contained in $I \cap \Theta$. Then it converges to some fixed point θ_0, which is impossible since θ_0 is repelling for g. The Theorem is proved. $\square$

24.3.2. *Parabolic case.* A similar result is valid in the parabolic case:

THEOREM 24.6. *Let f be a polynomial (of any degree $d \geq 2$) with connected Julia set. Then any parabolic periodic point α is the landing point of at least one periodic ray. Moreover, there are only finitely many rays $\mathcal{R}^{\theta}$ landing at α, and all of them are periodic with the same period. They are permuted with combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$ equal to the rotation number of α.*

PROOF. Replace f with its iterate $f^{p\mathfrak{q}}$ that turns α into a fixed point with multiplier 1. Let us consider a repelling petal P', and let $f^{-1} : P' \to P$ be the inverse branch of f that maps it to a smaller petal P. Let us also consider an attracting flower $\Phi \equiv \Phi^{\mathrm{a}}$ overlapping with these petals so that $P' \smallsetminus \Phi$ is a repelling fundamental rectangle $\Delta \equiv \Delta^{\mathrm{r}}$ defined in §21.3.3 (see Figure 21.3). Let $\varepsilon > 0$ be the distance between the vertical sides of this rectangle, and let a level t_0 be selected as in the repelling case.

We can now follow the argument of the repelling case replacing the fundamental annulus $\overline{D}' \smallsetminus D$ with the repelling fundamental rectangle Δ. The key observation is that if we have a point $\mathcal{R}^\theta(t) \in P$ with $t \geq t_0$, then the ray arc $\mathcal{R}^\theta[t, 2t]$ is contained in P'. Indeed, $\mathcal{R}^\theta[t, 2t] \cap P'$ is contained in $P' \smallsetminus \Phi$ (since $\Phi \subset \mathcal{K}_f$). Hence, if the $\mathcal{R}^\theta[t, 2t]$ escaped from P' then it would cross both vertical sides of Δ, which is impossible since it has length $< \varepsilon$.

Hence we can apply f^{-1} to the arc $\mathcal{R}^\theta[t, 2t]$, to obtain a ray arc $\mathcal{R}^{\theta_1}[t/2, t]$; then extend it to a ray arc $\mathcal{R}^{\theta_1}[t/2, 2t] \subset P'$. The argument now proceeds as in the repelling case.

The only special remark is that the combinatorial rotation number of α is equal to its rotation number $\mathfrak{p}/\mathfrak{q}$ as a parabolic point. Indeed, the rays land at α with certain slopes (see Exercise 21.14). The combinatorial rotation number of the rays is equal to the rotation number of the slopes, and the latter is equal to $\mathfrak{p}/\mathfrak{q}$. □

24.3.3. *Cantor case.*

Proposition 24.7. *Let $f \equiv f_c : z \mapsto z^2 + c$ be a quadratic polynomial with Cantor Julia set, i.e., $c \in \mathbb{C} \smallsetminus \mathcal{M}$. Then any external ray $\mathcal{R}^\theta$ that does not crash at a precritical point lands at some point of $\mathcal{J}(f)$.*

Proof. Let $\mathcal{E}$ be the critical equipotential, i.e., the figure-eight centered at 0. There are two critical rays that crash at 0. All other rays cross $\mathcal{E}$ at regular points and get trapped in the body Ω of this figure-eight. It contains two figure-eights of the next level, $\mathcal{E}_0$ and $\mathcal{E}_1$, centered at the first preimages of 0, z_0 and z_1. If a ray enters Ω but does not crash at z_i, it crosses one of these figure eights, say $\mathcal{E}_i$, at its regular point, and gets trapped in its body Ω_i. In turn, each Ω_i contains two figure-eights of the second level, $\mathcal{E}_{i0}$ and $\mathcal{E}_{i1}$, and so on. (See Figure 24.2.)

If a ray does not crash at any of the pre-critical points, then it consecutively crosses figure-eights $\mathcal{E}_{i_0 \dots i_{n-1}}$ of all levels at their regular points and gets trapped in the corresponding bodies. Since the dynamics on the Julia set is expanding,

$$\operatorname{diam} \mathcal{E}_{i_0 \dots i_{n-1}} = O(\lambda^{-n}) \quad \text{with some } \lambda > 1,$$

implying that such a ray lands at some point of $\mathcal{J}$. □

24.4. Fixed points and their combinatorial rotation number.

24.4.1. *Combinatorial rotation number.* Let us now consider a polynomial f of degree d with connected Julia set. Let a be its repelling or parabolic fixed point, and let $\mathcal{R}_i \equiv \mathcal{R}^{\theta_i}$ be the rays landing at a. The set of angles $\Theta(a) = \{\theta_i\} \subset \mathbb{T}$ is called the *ray portrait* of a. It is invariant under the map $T_d : \theta \mapsto d\,\theta$.

Lemma 24.8. (i) *The ray portrait $\Theta(a)$ is rotated by T_d with some rotation number $\mathfrak{p}/\mathfrak{q}$, called the combinatorial rotation number of a.*

(ii) *If a is parabolic with rotation number γ then its combinatorial rotation number coinsides with γ.*

Proof. (i) The map f, being an orientation preserving local homeomorphism near a, preserves the cyclic order of the rays $\mathcal{R}_i$ (see §1.3.3). Hence T_d preserves the cyclic order of the corresponding angles θ_i, implying the conclusion.

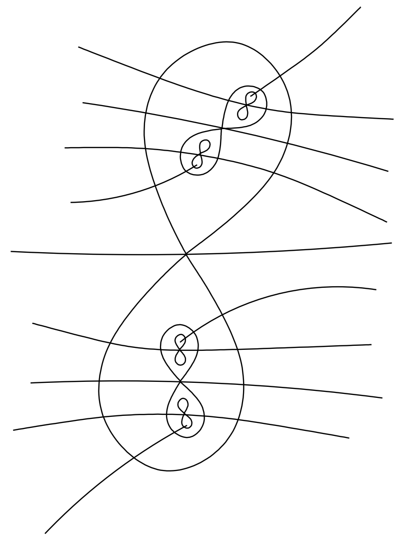

FIGURE 24.2. Crashing and landing rays for a Cantor Julia set.

(ii) In the parabolic case, the rays have distinct asymptotic slopes η_i at a which are rotated with rotation number γ. □

For a repelling or parabolic cycle $\mathbf{a} = (a_k)_{k=0}^{p-1}$ the *combinatorial rotation number* of $\mathbf{a}$ (and of each periodic point a_k) is defined by considering the a_k as fixed points of f^p (the answer is independent of k).

24.4.2. *The $\alpha-$ and $\beta-$ fixed points of a quadratic polynomial.* Let us now assume that $f = f_c$ is a quadratic polynomial $z \mapsto z^2 + c$ with connected Julia set. It turns out that the two fixed points of f (which are statically indistinguishable) play very different dynamical role.

The polynomial f has only one invariant ray, $\mathcal{R}^0$. By Theorem 24.3, this ray lands at some fixed point called β; moreover, this point is either repelling or parabolic with multiplier 1 (the last property follows from Lemma 24.8 (ii)). The ray $\mathcal{R}^0$ is the only ray landing at β (for any other ray would be also invariant by Lemma 24.8 (i)).

REMARK 24.9. In the locally connected case, it immediately follows that β is *non-dividing*, i.e., removing it from $\mathcal{J}$ does not disconnect the latter. In fact, it is always true (compare with Corollaries 32.7 and 9.8).

For $c \neq 1/4$, f_c has the second fixed point called α. It is either attracting, or neutral, or repelling. If α is repelling or parabolic, then by Theorems 24.5 and 24.6 it is a landing point of some periodic ray $\mathcal{R} = \mathcal{R}^\theta$. Since $\theta \neq 0 \mod 1$, the period $\mathfrak{q}$ of this ray is greater than 1. Of course, all the rays $\mathcal{R}_i = f^i(\mathcal{R})$, $n = 0, 1, \dots, \mathfrak{q} - 1$, also land at α, so α is the *dividing* fixed point.

By Lemma 24.8 the ray portrait $\Theta(\alpha) \subset \mathbb{T}$ is a *rotation cycle* for the doubling map $\theta \mapsto 2\theta$. By Proposition 24.27 (from the Appendix), it is in fact, a single rotation cycle. Hence *the rays $\mathcal{R}_i$ are cyclically permuted by f with a combinatorial rotation number* $\mathfrak{p}/\mathfrak{q}$. This rotation number, $\rho_{\rm com}(f_c) \equiv \rho_{\rm com}(c)$, is also called the *combinatorial rotation number of f* (or of the corresponding parameter c).

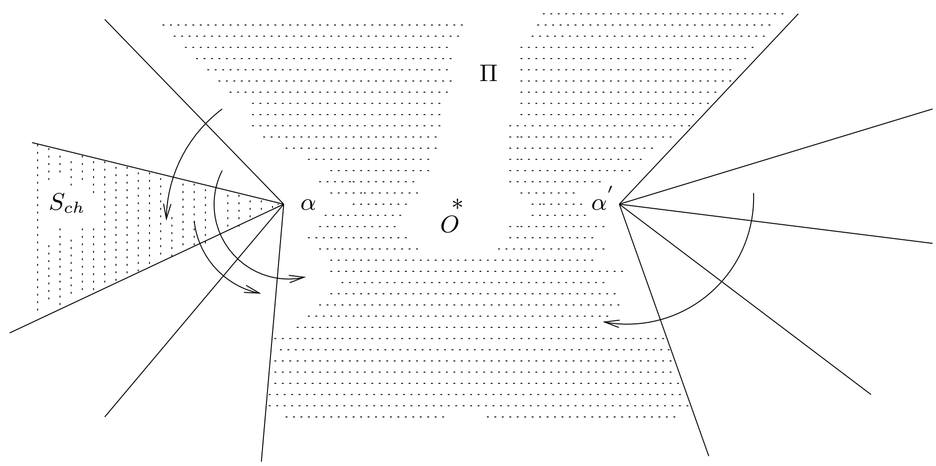

FIGURE 24.3. Configuration of α-sectors with rotation number 2/5.

In §37.4 we will describe the set of parameters with a given combinatorial rotation number.

24.4.3. *Configuration of α-sectors.* The rays $\mathcal{R}_i$ divide the plane into $\mathfrak{q}$ sectors S_i, $i = 0.\dots, q-1$, which cut off arcs ω_i at the circle at infinity. We study these arcs in Lemma 24.25 from the Appendix. Recall that the longest of these arcs, labeled $\omega_0 \equiv \omega_q$, is called *critical*, while the shortest, $\omega_{\rm ch} \equiv \omega_1$, is called *characteristic* or *valuable*. The corresponding sectors, $S_0 \equiv S_{\mathfrak{q}}$ and $S_{\rm ch} \equiv S_1$, will be called in the same way.

LEMMA 24.10. *For $i = 1, \dots, \mathfrak{q}-1$, the map f univalently maps the sectors S_i onto S_{i+1}. The critical sector S_0 contains the critical point* 0*, while the characteristic sector $S_{\rm ch}$ contains the critical value $c = f(0)$.*

PROOF. Let $\overline{S}_i$ be the compactification of the sector S_i at infinity obtained by attaching the arc ω_i to S_i. This is a topological triangle. For $i = 1, \dots, \mathfrak{q}-1$, the boundary of S_i is homeomorphically mapped onto the boundary of S_{i+1}. By the Argument Principle, the whole triangle $\overline{S}_i$ is homeomorphically mapped onto $\overline{S}_{i+1}$. Hence there are no critical points in these S_i, so that, $0 \in S_0$. □

Let $\alpha' = -\alpha$; this is the second preimage of the fixed point α. There are $\mathfrak{q}$ rays $\mathcal{R}'_i$ landing at α' symmetric to the rays $\mathcal{R}_i$, so that, $f(\mathcal{R}'_i) = \mathcal{R}_{i+1}$, $i \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$. Altogether, the rays $\mathcal{R}_i$ and $\mathcal{R}'_i$ partition the plane into $\mathfrak{q}-1$ pairs of symmetric sectors S_i, S'_i, $i = 1, \dots, \mathfrak{q}-1$ (bounded by two rays each) and a central strip $\Pi \ni 0$ bounded by two pairs of symmetric rays.

LEMMA 24.11. *The central strip Π is mapped onto the characteristic sector $S_{\rm ch}$ as a double branched covering.*

PROOF. Each pair of symmetric rays that bound Π is mapped homeomorphically onto a characteristic ray that bound $S_{\rm ch}$, so we have a 2-to-1 map $\partial\Pi \to \partial S_{\rm ch}$.

Let $\overline{\Pi}$ be the compactification of Π by two symmetric arcs η and η' at infinity (where the arc η appears in the proof of Lemma 24.25). Each of these arcs is mapped homeomorphically onto the characteristic arc ω_1.

We see that the boundary of $\overline{\Pi}$ is mapped to the boundary of $\overline{S}_{\mathrm{ch}}$ as a double covering, and the conclusion follows. $\square$

The above lemmas could also be derived from the following assertion:

EXERCISE 24.12. *Let a point $a \in \mathcal{J}$ be the landing point of two rays $\mathcal{R}^{\theta_1}$ and $\mathcal{R}^{\theta_2}$. Assume that the critical value v does not lie on the corresponding cut-line $L := \mathcal{R}^{\theta_1} \cup \mathcal{R}^{\theta_2} \cup \{a\}$, and let S_{ch} be the component of $\mathbb{C} \smallsetminus L$ containing v. Then the preimage $f^{-1}(L)$ comprises two cut-lines that divide $\mathbb{C}$ into three domains. The central one (containing 0) is mapped onto S_{ch} with degree two. The other two are mapped univalently onto the other component of $\mathbb{C} \smallsetminus L$.*

Generalize this assertion to a configuration of $\mathfrak{q}$ rays landing at a.

Finally, let us slightly *accelerate* the dynamics on the sector S_0, by letting

$$F|\,\Pi = f^{\mathfrak{q}}, \quad F|\,S_i' = f^{\mathfrak{q}-i},\ i = 1, \dots, \mathfrak{q}-1. \tag{24.2}$$

This map is a double branched covering of the strip Π over the sector S_0 and is a conformal isomorphism of each lateral sector S_i' onto S_0. It is *Bernoulli* in the sense that its range $\overline{S}_0$ is tiled by its domains $\overline{\Pi}$ and the $\overline{S}_i'$ (compare §19.11.3).

24.4.4. *Spine and skeleton.* In this section we assume that the Julia set $\mathcal{J} \equiv \mathcal{J}(f)$ is connected and locally connected. Then by the Carathéodory-Torhorst Theorem the inverse Böttcher function $\phi \equiv B^{-1} : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathcal{D}(\infty)$ admits a continuous extension to a function $\mathbb{C} \smallsetminus \mathbb{D} \to \mathcal{D}(\infty) \cup \mathcal{J}$ (denoted in the same way). It follows that every ray $\mathcal{R}^\theta$ lands at some point $z^\theta := \phi(e(\theta)) \in \mathcal{J}$.

Let us consider the *co-fixed* point $\beta' \equiv -\beta$. The legal arc $\sigma \equiv \sigma_f := [\beta, \beta']$ is called the *spine* of $\mathcal{K}$. As usual, we let $\sigma^\circ \equiv (\beta, \beta')$ be the intrinsic interior of the spine. It turns out that the spine and its preimages capture all cut-points in $\mathcal{K}$.

REMARK 24.13. For this discussion, the choice of the spine in the interior components of $\mathcal{K}$ can be made in an arbitrary way.

LEMMA 24.14. *Assume the Julia set $\mathcal{J} \equiv \mathcal{J}(f)$ is connected and locally connected. Then a point $a \in \mathcal{J}$ belongs to σ° if and only of there are two rays, $\mathcal{R}^{\theta_0}$ and $\mathcal{R}^{\theta_1}$, landing at a such that the dyadic expansions for θ_0 and θ_1 begin with different digits* (0 *and* 1 *respectively*).

PROOF. Let us consider the arc Γ composed of the spine and two rays, $\mathcal{R}^0$ and $\mathcal{R}^{1/2}$. It is the line $\mathbb{R}$ properly embedded into $\mathbb{C}$, so by the Jordan Theorem, its complement is the union of two open topological half planes, $\Pi_\pm$.

Let us also consider the following set: $X := \overline{\mathbb{D}} \cup \mathbb{R}$. By Theorem 9.28, the inverse Böttcher map admits a continuous extension to a map

$$\hat{\phi} : (\mathbb{C}, X) \to (\mathbb{C}, \mathcal{K} \cup \mathcal{R}^0 \cup \mathcal{R}^{1/2}).$$

Since $K \cup \mathcal{R}^0 \cup \mathcal{R}^{1/2} \supset \sigma$, the images of the half-planes, $\hat{\phi}(\mathbb{H}_\pm)$, are contained in the half-planes $\Pi_\pm$. But $\hat{\phi}(\mathbb{H}_+)$ is the union of rays $\mathcal{R}^\theta$ with $\theta = (0\,\varepsilon_2 \dots) \in (0, 1/2)$, while $\hat{\phi}(\mathbb{H}_-)$ is the union of rays $\mathcal{R}^\theta$ with $\theta = (1\,\varepsilon_2 \dots) \in (1/2, 1)$,

It follows that $\mathcal{R}^{\theta_0}$ lands in $\overline{\Pi}_+$, while $\mathcal{R}^{\theta_1}$ lands in $\overline{\Pi}_-$. Hence the common landing point must lie in the intersection $(\overline{\Pi}_+ \cap \overline{\Pi}_-) \cap \mathcal{K}$, which is the spine.

Vice versa, Exercise 9.24 implies that any point of $\mathcal{J} \cap \sigma^\circ$ is a cut-point that can be accessed from both Π_+ and Π_-. $\square$

Taking preimages of the spine σ, we obtain the *skeleton* of f,

$$\mathcal{S}k \equiv \mathcal{S}k_f := \bigcup_{n=0}^{\infty} f^{-n}(\sigma); \quad \mathcal{S}k^\circ := \bigcup_{n=0}^{\infty} f^{-n}(\sigma^\circ).$$

It is an infinite "tree-like" set in the Julia set containing all cut-points of $\mathcal{J}$:

PROPOSITION 24.15. *Assume the Julia set $\mathcal{J} \equiv \mathcal{J}(f)$ is connected and locally connected. A point $a \in \mathcal{J}$ is a cut-point if and only if it belongs to the skeleton $\mathcal{S}k^\circ$, i.e., iff $f^n a \in \sigma^\circ$ for some $n \in \mathbb{N}$.*

PROOF. The "if" part follows directly from Lemma 24.14.

Vice versa, if $a \in J$ is a cut-point then by definition there are at least two rays, $\mathcal{R}^{\theta^\pm}$, landing at a. The dyadic expansions of the corresponding angles, $\theta^\pm$, differ at some place $n \in \mathbb{N}$. Then the dyadic expansions for $2^n\theta^+$ and $2^n\theta^- \mod 1$ differ at the first place. By Lemma 24.14, $f^n a \in \sigma^\circ$. □

24.4.5. *Real case.* In the case of $f \equiv f_c$ with real c, we have (compare §20.4.2):
(i) The zero-ray $\mathcal{R}^0$ is the ray $(\beta, 0)$ in $\mathbb{R}$ landing at the β-fixed point;
(ii) The spine σ is the maximal invariant interval $\mathcal{I} = [\beta, \beta]$;
(iii) The skeleton $\mathcal{S}k$ is the set of points in $\mathcal{J}$ eventually landing in $\mathbb{R}$.

Note that all these notions make sense even if the Julia set is not locally connected. In the locally connected case, we have:

PROPOSITION 24.16. *Assume that f is real-symmetric and $\mathcal{J}(f)$ is locally connected.*[13] *Then all cut points of $\mathcal{J}$ land in $\mathcal{I}^\circ \equiv (\beta, \beta)$.*

24.5. General periodic ray configurations.

24.5.1. *Sectors and strips.* More generally, let us consider a *repelling or parabolic periodic cut-point* α and its cycle $\boldsymbol{\alpha} = (\alpha_n \equiv f^n\alpha)_{n=0}^{p-1}$. Then there is at least two (and at most finitely many, by Theorems 24.5 and 24.6) rays $\mathcal{R}_i(\alpha_n)$, $i = 0, \dots, r-1$, landing at each α_n. Each configuration

$$\mathfrak{R}(\alpha_n) := \{\alpha_n\} \cup \bigcup_i \mathcal{R}_i(\alpha_n)$$

is an open star properly embedded into $\mathbb{C}$. Under f^p, the rays $\mathcal{R}_i(\alpha_n)$ are permuted with some rotation number $\mathfrak{p}/\mathfrak{q}$.

The union

$$\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\alpha}) := \bigcup_n \mathfrak{R}(\alpha_n)$$

is called the *ray configuration* of $\boldsymbol{\alpha}$. Clearly, it is f-invariant.

Let $S_i(\alpha_n)$, $i = 0, 1, \dots, r-1$, be the complementary sectors to the star $\mathfrak{R}(\alpha_n)$, labeled so that

- $S_0(\alpha_n) \ni 0$; this sector is called *critical*;
- The ray $\mathcal{R}_i(\alpha_n)$ oriented from α_n to ∞ is the positively oriented boundary ray of $S_i(\alpha_n)$.

Note that the angular size of $S_0(\alpha_n)$ at infinity is $> 1/2$, while all other sectors have angular size $< 1/2$ (compare with Lemma 24.10). Let $S_v(\alpha_n)$ be the sector $S_i(\alpha_n)$ containing the critical value $f(0)$. It is called *valuable*.

[13] In fact, it is known that $\mathcal{J} \equiv \mathcal{J}(f_c)$ is locally connected for all $c \in [-2, 1/4]$.

Exercise 24.12 implies that each non-critical sector $S_i(\alpha_n)$ is mapped univalently by f onto some non-valuable sector $S_j(\alpha_{n+1})$ of twice bigger angular size (and any non-valuable sector can be obtained in this way). The image of any critical sector $S_0(\alpha_n)$ covers twice the valuable sector $S_v(\alpha_{n+1})$ and covers once its complement. The preimage $\Pi(\alpha_n) := f^{-1}(S_v(\alpha_{n+1})) \ni 0$ is a 0-symmetric topological strip bounded by the rays landing at α_n and two symmetric rays landing at $\alpha'_n \equiv -\alpha_n$.

The sector S_{ch} of the smallest angular size is called *characteristic*. We will see momentarily that it is unique, but for the moment let us select one and label the periodic points α_n so that α_1 is the vertex of S_{ch}.

LEMMA 24.17. (i) *The characteristic sector S_{ch} is the innermost valuable sector among the $S_v(\alpha_n)$ (in particular, it is uniquely defined).*
(ii) S_{ch} *is disjoint from* $\mathfrak{R}$.
(iii) *There is a topological strip Π_{ch} contained in S_{ch} such that $f^p : \Pi_{\mathrm{ch}} \to S_{\mathrm{ch}}$ is a double branched covering.*
(iv) *The images $f^m(\Pi_{\mathrm{ch}})$, $m = 1, \dots, p\mathfrak{q}$, are pairwise disjoint, and the maps*

$$f : f^m(\Pi_{\mathrm{ch}}) \to f^{m+1}(\Pi_{\mathrm{ch}}), \quad m = 0, 1, \dots, p\mathfrak{q} - 2,$$

are univalent, while $f : f^{\mathbf{p}\mathfrak{q}-1}(\Pi_{\mathrm{ch}}) \to S_{\mathrm{ch}}$ is a double branched covering. (*Note that $f^{\mathbf{p}\mathfrak{q}-1}(\Pi_{\mathrm{ch}}) = \Pi(\alpha_0)$.*)

PROOF. (i) If $S_{\mathrm{ch}} = S_j(\alpha_1)$ were not valuable then it would be the univalent image of a sector $S_i(\alpha_0)$ of twice smaller angular size, contradicting the definition of S_{ch}.

Furthermore, any two valuable sectors $S_v(\alpha_n)$ are nested. Hence S_{ch} is the innermost valuable sector, which determines it uniquely.

(ii) If $S_{\mathrm{ch}} \cap \mathfrak{R}(\boldsymbol{\alpha}) \neq \emptyset$ then S_{ch} would contain some sector $S_i(\alpha_n)$, $n \neq 1$, which would necessarily have a smaller angular size.

(iii)–(iv) Let $\Pi_1 \equiv S_{\mathrm{ch}}$, $L_1 := \partial S_{\mathrm{ch}}$, and let $L_{-m} \subset \mathfrak{R}(\alpha_{-m})$ be the lift of L_{ch} under f^{m+1} passing through $\alpha_{-(m+1)}$, $m = 0, 1, \dots, p\mathfrak{q} - 2$ (where $-m$ in α_{-m} is taken mod p). Let Π_{-m} be the pullback of $\Pi_0 \equiv \Pi(\alpha_0)$ by f^m attached to L_{-m}. In other words, it is the appropriate component of $\mathbb{C} \smallsetminus f^{-(m+1)}(\mathfrak{R})$ attached to L_{-m}.

Since $\mathfrak{R}$ is forwards invariant, the configurations $f^{-m}(\mathfrak{R})$ form an increasing sequence of sets. Hence any two components $\Pi_1, \Pi_0, \dots, \Pi_{-(p\mathfrak{q}-2)}$ are either disjoint or nested. Assume there are two nested components: $\Pi_{-m} \supset \Pi_{-n}$ for some $-1 \le m < n \le p\mathfrak{q} - 2$. Since $L_{-m} \neq L_{-n}$, the domain Π_{-m} would contain $L_{-n} \subset \mathfrak{R}$. Applying f^{m+1}, we conclude that S_{ch} would intersect $\mathfrak{R}$, contradicting (ii).

Let us take one more pullback, $\Pi_{\mathrm{ch}} := \Pi_{-(p\mathfrak{q}-1)}$. It is contained in S_{ch} since it is mapped onto S_{ch} under $f^{p\mathfrak{q}}$, which preserves orientation of $L_{\mathrm{ch}} = \partial S_{\mathrm{ch}} = \partial \Pi_{\mathrm{ch}} \cap \partial S_{\mathrm{ch}}$. Hence Π_{ch} is disjoint from all the Π_{-m}, $m = 0, 1, \dots, p\mathfrak{q} - 2$ as well.

It follows that all the domains Π_{-m}, $m = 1, \dots, p\mathfrak{q}-1$, are disjoint from $\Pi_0 \ni 0$, so they do not contain 0. Hence the corresponding maps $f : \Pi_{-m} \to \Pi_{-(m-1)}$ are all univalent, and so is their composition $f^{p\mathfrak{q}-1} : \Pi_{\mathrm{ch}} \to \Pi_0$. As $f : \Pi_0 \to S_{\mathrm{ch}}$ is a double branched covering, we conclude that $f^{p\mathfrak{q}} : \Pi_{\mathrm{ch}} \to S_{\mathrm{ch}}$ is a double branched covering, too.

Note finally that Π_{ch} is a topological strip as it is bounded by two properly embedded lines, L_{ch} and L'_{ch}, where L'_{ch} is the other lift of L_{ch} under $f^{p\mathfrak{q}} : \overline{\Pi}_{\mathrm{ch}} \to \overline{S}_{\mathrm{ch}}$. $\square$

The rays $\mathcal{R}^{\pm}_{\rm ch}$ on boundary of the sector $S_{\rm ch}$, their landing point $\alpha_{\rm ch} \in \boldsymbol{\alpha}$, the topological line $L_{\rm ch} = \partial S_{\rm ch} = \mathcal{R}^{+}_{\rm ch} \cup \mathcal{R}^{-}_{\rm ch} \cup \{\alpha_{\rm ch}\}$, and the strip $\Pi_{\rm ch}$, are all called *characteristic* or *valuable*. We label the periodic points α_n so that $\alpha_1 = \alpha_{\rm ch}$.

The strip $\Pi_0 = f^{-1}(S_{\rm ch}) \ni 0$ is called *critical* or *central*. Let us also consider a set $\Upsilon := f^{-1}(\Pi_{\rm ch}) \subset \Pi_0$. If $f(0) \in \Pi_{\rm ch}$ then Υ is bounded by four cut-lines that are pairwise asymptotic at ∞. We call it a *central ameba* .

Exercise 24.18. *If $f(0) \in \Pi_{\rm ch}$ then the map $f^{p\mathfrak{q}} : \Upsilon \to \Pi_0$ is a double branched covering.*

24.5.2. *Periodic ray portraits.* A *(periodic) ray portrait* $\Theta \subset \mathbb{T}$ is a finite set of periodic angles with the same period decomposed into unlinked subsets Θ_n, $n = 0, \dots, p-1$, of cardinality $r \geq 2$ each (so $|\Theta| = pr$) that are rotated under the doubling map $T : \mathbb{T} \to \mathbb{T}$.

To each cut-cycle $\boldsymbol{\alpha}$ corresponds a ray portrait $\Theta(\boldsymbol{\alpha})$ by taking the set of angles of the ray configuration $\mathfrak{R}(\boldsymbol{\alpha})$. The unlinked subsets $\Theta_n(\boldsymbol{\alpha})$ correspond to the rays landing at the same point $\alpha_n \in \boldsymbol{\alpha}$.

We will show later on that vice versa, any perioic ray portrait can be realized as a portrait of a ray configuration for some quadratic polynomials f_c, and will describe the corresponding set of parameters c (see the Wake Theorem in §37.3).

24.5.3. *QC geometry of ray configurations.* Given a ray configuration $\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\alpha})$ as above, let us truncate it by some equipotential $\mathcal{E}(t)$. We obtain a cell decomposition $\mathcal{C}$ of the subpotential domain $\Sigma \equiv \Sigma(t)$.

Given two maps f and $\tilde{f}$, we say that the respective cell decompositions, $\mathcal{C}$ and $\mathcal{C}'$ have *the same combinatorics* if there exists a homeomorphism $h : \Sigma \to \tilde{\Sigma}$ that maps cells to cells respecting the boundary Böttcher marking. We say that they are *qc equivalent* if this map can be selected to be quasiconformal.

Exercise 24.19. (i) *If the cycle $\boldsymbol{\alpha}$ is repelling then the cells of the corresponding cell decomposition are quasidisks.*
(ii) *If two cell decompositions $\mathcal{C}$ and $\tilde{\mathcal{C}}$ (associated with repelling cycles) are combinatorially equivalent then they are qc equivalent.*

24.6. Yoccoz Inequality. Let us consider a holomorphic repelling germ

$$f : z \mapsto \rho z + a z^2 + \dots, \quad |\rho| > 1,$$

Assume it has a cycle of (open) topological sectors S_i centered at 0, $i \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$, which is rotated under f with combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$. Precisely, this means:

- The sectors S_i are pairwise disjoint.
- If they are labeled according to their cyclic order around 0, then $f(S_i) = S_{i+\mathfrak{p}}$ near 0 (where the equality between sectors is understood in the sense of "sectorial germs").

Let $S \equiv S_0$. Then the quotient $S/f^{\mathfrak{q}}$ is a conformal annulus. Call it A.

Lemma 24.20. *Under the above circumstances, we have:*

$$\operatorname{mod} A \leq \frac{2\pi \log |\rho|}{|\mathfrak{q} \log \rho - 2\pi i \mathfrak{p}|^2},$$

for an appropriate branch of $\log \rho$.

PROOF. Let $r > 0$ be so small that f maps the disk $\mathbb{D}_r$ biholomorphically onto the image, and $f(\mathbb{D}_r) \Supset \mathbb{D}_r$. Then the quotient annulus A is naturally embedded into the quotient torus $\mathbb{T}^2 := \mathbb{D}_r^*/f$.

By means of the exponential map $\exp : \mathbb{C} \to \mathbb{C}^*$, this torus is represented as $\mathbb{C}/\Gamma$, where Γ is the lattice generated by translations by $\beta := \log \rho$ and $\alpha := 2\pi i$. Moreover, β and α represent a basis in the first homology group $H_1(\mathbb{T}^2)$. In this basis, horizontal curves of the annulus A represent the class $\mathfrak{q}\beta - \mathfrak{p}\alpha \in H_1(\mathbb{T}^2)$ (with an appropriate choice of $\log \rho$), see Exercise 23.14.

Now Proposition 6.25 implies the desired estimate. □

COROLLARY 24.21. *There is a choice of* $\log \rho$ *that belongs to the round disk* $\Delta \subset \{\operatorname{Re} \mu \geq 0\}$ *of diameter* $d = \dfrac{2\pi}{\mathfrak{q}^2 \operatorname{mod} A}$ *which is tangent to the imaginary axis at point* $2\pi i\, \mathfrak{p}/\mathfrak{q}$.

PROOF. The estimate of the above lemma can be re-written as

$$\frac{|\mu - 2\pi i\, \mathfrak{p}/\mathfrak{q}|}{\cos \theta} \leq d,$$

where $\mu = \log \rho$ and $\theta = \arg(\mu - 2\pi i\, \mathfrak{p}/\mathfrak{q})$. The conclusion follows from the following elementary fact: □

EXERCISE 24.22. *The equation* $r = d \cos \theta$ *in the polar coordinates describes the circle based upon* $[0, d]$ *as a diameter.*

Let us now go back to the quadratic case.

LEMMA 24.23. *Let* f *be a quadratic polynomial with connected Julia set, and let* α *be a repelling periodic point of period* p *with combinatorial rotation number* $\mathfrak{p}/\mathfrak{q}$. *Then the corresponding fundamental torus* $\mathbb{T}^2_{f,\alpha}$ *contains a* $\mathfrak{p}/\mathfrak{q}$*-annulus* A *with*

$$\operatorname{mod} A = \frac{\pi}{p\, \mathfrak{q} \log 2}.$$

PROOF. See Corollary 23.22 and remarks afterwards. □

Putting this together with Corollary 24.21, we obtain:

COROLLARY 24.24. *Under the circumstances of the above lemma, there is a choice of* $\log \rho$ *that belongs to the round disk* $\Delta \subset \{\operatorname{Re} \mu \geq 0\}$ *of diameter* $d = \dfrac{p \log 4}{\mathfrak{q}}$ *which is tangent to the imaginary axis at the point* $2\pi i\, \mathfrak{p}/\mathfrak{q}$.

24.7. Appendix: Rotation sets for the doubling map. Let us analyze rotation cycles for the doubling map $T : \theta \mapsto 2\theta \mod 1$.

LEMMA 24.25. *Let* Θ *be a rotation cycle for* T *with rotation number* $\mathfrak{p}/\mathfrak{q}$. *Then complementary arcs* ω_i *to* Θ (*counted according to the action of* T *starting with the shortest one*) *have lengths* $2^{i-1}/(2^{\mathfrak{q}} - 1)$, $i = 1, \dots \mathfrak{q}$.

PROOF. Recall from §1.11 that a rotation of a cyclically ordered finite set Θ preserves the relation "κ is next to θ". In case when the order on Θ is induced from the ambient oriented circle $\mathbb{T}$, this relation amounts to saying that the positive arc[14] (θ, κ) is complementary to Θ. Thus, for any complementary arc (θ, κ), the

[14]There are two arcs in $\mathbb{T}$ with endpoints θ and κ. "Positive" is the one whose orientation from θ to κ coincides with the orientation induced from $\mathbb{T}$.

arc $(T(\theta), T(\kappa))$ is also complementary to Θ. It follows that $T|\,\Theta$ can be extended, arc by arc, to an orientation preserving homeomorphism $\tilde{T} : \mathbb{T} \to \mathbb{T}$. Moreover, if a complementary arc $\omega = [\theta, \kappa]$ is less than the half-circle then $T|\,\omega$ itself is an orientation preserving homeomorphism, so we let $\tilde{T}|\,\omega = T$ in this case.

Since T is not a homeomorphism of $\mathbb{T}$, this surgery should be non-trivial on some complementary arc ω_i. Hence, among the complementary arcs ω_i there should be one (and of course, only one) which is longer than the half-circle. Let us call it $\omega_0 = [\theta_0, \kappa_0]$.

This arc is the union of the half-circle $\xi = (\theta_0, \theta_0']$ and the arc $\eta = (\theta_0', \kappa_0)$ of length $\varepsilon/2$, where $\theta_0' = \theta_0 + 1/2$ is the point symmetric with θ_0. Moreover, the arc ξ is bijectively mapped under T onto the whole circle $\mathbb{T}$, while η is homeomorphically mapped onto a complementary arc $(\theta_1, \kappa_1) = \omega_1$. We see that $|\omega_1| = \varepsilon$.

Since each complementary arc ω_i, $i = 1, \dots, \mathfrak{q}-1$, is shorter than the half-circle, it is mapped under T homeomorphically onto the arc ω_{i+1}, and $|\omega_{i+1}| = 2|\omega_i|$. Hence $|\omega_i| = 2^{i-1}\varepsilon$, $i = 1, \dots, \mathfrak{q}$. Since $\mathfrak{q} = 0$ in $\mathbb{Z}/\mathfrak{q}\mathbb{Z}$, we obtain an equation

$$\frac{1+\varepsilon}{2} = |\omega_0| = |\omega_{\mathfrak{q}}| = 2^{\mathfrak{q}-1}\varepsilon,$$

which gives us the desired value of ε. □

The arc ω_0 which is longer than the half-circle is called *major*. The shortest arc ω_1 is called *characteristic*.

EXERCISE 24.26. *Show that the major arc ω_0 contains the fixed point $\theta = 0$.*

PROPOSITION 24.27. *For the doubling map $T : \theta \mapsto 2\theta$ on $\mathbb{T}$ and any rational $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}/\mathbb{Z}$, there exists a unique rotation cycle $\Theta_{\mathfrak{p}/\mathfrak{q}} \subset \mathbb{T}$ with rotation number $\mathfrak{p}/\mathfrak{q}$.*

PROOF. Let Θ be a rotation cycle with rotation number $\mathfrak{p}/\mathfrak{q}$. Let us consider its characteristic arc $\omega_1 = (\theta, \kappa)$. Since κ is next to θ in Θ, we have: $\kappa = 2^l\theta \mod 1$ with

$$l\mathfrak{p} = 1 \quad \text{in} \quad \mathbb{Z}/\mathfrak{q}\mathbb{Z}. \tag{24.3}$$

On the other hand, $\kappa = \theta + 1/(2^{\mathfrak{q}} - 1)$ by Lemma 24.25. Hence

$$(2^l - 1)\,\theta = 1/(2^{\mathfrak{q}} - 1) \mod 1. \tag{24.4}$$

Since θ is T-periodic with period $\mathfrak{q}$, $2^{\mathfrak{q}}\theta = \theta \mod 1$, so

$$\theta = t/(2^{\mathfrak{q}} - 1) \text{ with some } t \in \mathbb{Z}. \tag{24.5}$$

Plugging this into (24.4), we come up with an equation

$$(2^l - 1)\,t = 1 \mod 2^{\mathfrak{q}} - 1. \tag{24.6}$$

Since l and $\mathfrak{q}$ are mutually prime by (24.3), so are $2^l - 1$ and $2^{\mathfrak{q}} - 1$. Hence (24.6) has a unique solution mod $2^{\mathfrak{q}} - 1$. This proves uniqueness of the rotation cycle.

To prove existence, let us go backwards as follows:
– first find l satisfying (24.3);
– then take the solution $t \in \mathbb{Z}$ of (24.6);
– define the angle θ by (24.5), so that, (24.4) is satified;
– and finally let $\kappa = 2^l\theta$, $\omega_1 = (\theta, \kappa)$, so that, we have: $|\omega_1| = 1/(2^{\mathfrak{q}} - 1)$.

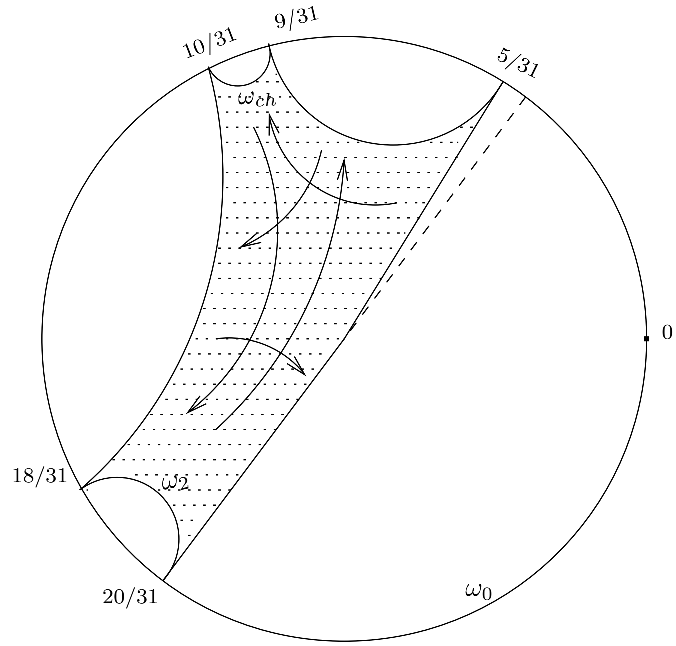

FIGURE 24.4. Rotation set with rotation number 2/5.

Then $\omega_2 = T^l(\omega_1)$ is the arc of length $2/(2^{\mathfrak{q}} - 1)$ adjacent to ω_1; $\omega_3 = T^{2l}(\omega_1)$ is the arc of length $4/(2^{\mathfrak{q}}-1)$ adjacent to ω_2, etc., up to the arc $\omega_{\mathfrak{q}} = T^{l(\mathfrak{q}-1)}(\omega_1)$ of length $2^{\mathfrak{q}-1}/(2^{\mathfrak{q}} - 1) > 1/2$. Since the total length of these arcs is equal to 1, their closures tessellate the whole circle $\mathbb{T}$, so $\Theta = \operatorname{orb}\theta$ is a rotation cycle for T^l with rotation number $1/\mathfrak{q}$. By (24.3) we have: $T|\,\Theta = (T^l)^{\mathfrak{p}}|\,\Theta$, so $T|\,\Theta$ has rotation number $\mathfrak{p}/\mathfrak{q}$. $\square$

As we know from Exercise 19.37, there are no infinite closed T-invariant subsets $\Theta \subset \mathbb{T}$ such that the restriction $T|\,\Theta$ is invertible. However, an "almost invertible" scenario can happen:

PROBLEM 24.28. *Prove that for any irrational $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$, there exists a unique closed T-invariant set $\Theta \equiv \Theta_\theta \subset \mathbb{T}$ such that for any $\eta \in \Theta$, the map*

$$\operatorname{orb}_T \eta \to \mathbb{T}, \quad T^n\eta \mapsto n\,\theta \mod 1$$

is cyclic order preserving. Moreover:

- $T|\,\Theta$ *is is one-to-one except for two opposite points γ and $\gamma' = \gamma + 1/2$;*
- Θ *is contained in one of the half-circles bounded by γ and γ'; the other half-circle contains the fixed point $\theta = 0$;*
- Θ *is a Cantor set and there exists a Devil Staircase map $\mathbb{T} \to \mathbb{T}$ (collapsing complementary to Θ intervals to points) that semi-conjugates $T|\,\Theta$ to the circle rotation $\eta \mapsto \eta + \theta \mod 1$.*

Thus, for any rotation number $\theta \in \mathbb{R}/\mathbb{Z}$, we have one *rotation set* $\Theta_\theta \subset \mathbb{T}$ such that $T\,|\,\Theta_\theta$ is almost a rotation by angle θ. Moreover, the doubling map preserves the cyclic order on Θ_θ.

Exercise 24.29. *Show that an invariant set $\Theta \subset \mathbb{T}$ is a rotation set under the doubling map iff one of the gaps in Θ has angular size $\geq 1/2$.*

Let us finish with introducing special rotation sets for iterates of the doubling map (which appear naturally in the renormalization context). Assume there exist $p \geq 2$ and two disjoint closed arcs I_1 and I_2 in $\mathbb{T}$ satisfying assumptions of Exercise 19.60 for the iterate T^p. Then we obtain a T^p-invariant Cantor set $K \subset I_1 \cup I_2$ and the associated Devil K-Staircase map $h : \mathbb{T} \to \mathbb{T}$ semiconjugating $T^p| K$ to the circle doubling map. It provides us with a natural one-to-one correspondence between cycles of period > 1 for $T^p| K$ and for T. Since h is monotonically non-decreasing, we obtain a one-to-one correspondence between rotation cycles for $T^p| K$ of period > 1 and those for the doubling map. We call such cycles *tuned rotation cycles.* We conclude:

Corollary 24.30. *Under the above circumstances, for any non-zero $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, there exist a unique tuned rotation cycle $\Theta_{\mathfrak{p}/\mathfrak{q}} \subset K$ for T^p with rotation number $\mathfrak{p}/\mathfrak{q}$.*

Exercise 24.31. *Extend the above discussion of irrational rotation sets (Problem 24.28) to tuned ones.*

Notes

According to Milnor [**M2**], Theorems 24.5 and 24.6 (on rays landing at periodic points) are due to Douady. The proof given here is more recent: it appeared in [**BeL**].

Description of the ray portraits (§24.4 and §24.5) is due to Douady & Hubbard [**DH2**] and Lavaurs [**Lav**]. See also [**M5**].

The Yoccoz Inequality (§24.6) appeared in [**H2**]. Its Kleinian group counterpart had been already known to Bers. See also Pommerenke and Levin.

Rotation sets for the doubling map (§24.7) was a folklore in the early 1980s. A written account (in the general higher degree case) appeared in [**G, GM**]. See [**Z**] for recent developments.

Notes to Chapter 3

Local Holomorphic Dynamics was founded in the second half of the XIXth–early XXth century by Schröder [**Schr**] and Koenigs [**Ko**] (attracting case), Leau [**Leau**] (parabolic case), and Böttcher [**Bot**] (superattracting case). It was focused on solving the corresponding functional equations. See [**Al**] for a comprehensive account of the early history of the field.

The founder of the Global Holomorphic Dynamics is Pierre Fatou. His first note in this field appeared as early as 1906 [**F:CR06**]. He observed there that a Cantor set can appear as the boundary of the basin of attraction of a fixed point. A comprehensive theory started to emerge one decade later, announced in two 1917 Comptes Rendus notes [**F:CR17**] (in particular, the "Julia set" appeared in the second of these notes under the name of the *set $\mathcal{F}$ of irregular points*). They followed in 1919–20 by fundamental Memoires [**F1, F2, F3**] . Solving of functional equations was considered to be so central at that time that the whole set of memoirs was called accordingly.

The main tool of the theory was Montel's theory of normal families developed in the turn of the 20th century [**Mo1**]. *Irregular* points had been introduced by Montel in this general context, as points where a sequence of meromorphic functions fails to be normal.

Meanwhile, some pieces of the theory, accompanied with a detailed study of some examples, were developed by Julia [**J:CR17, J:CR18, J**]. This work won the Grand Prix de l'Academie des Sciences (1918).

See [**Au**] for a historical account of the above events.

Note on the names. It seems that Douady is responsible for a significant portion of the names for various objects in contemporary Holomorphic Dynamics, including *basilica, cauliflower, rabbit, Mandelbrot set and its limbs, Misiurewicz point, Hubbard tree,...* A few exceptions include *Julia set* (see [**Au**] for the history), *Fatou set* [**Bl**], *kneading sequence* [**MT**], *Baker domain* [**EL1**], *witch's broom* [**M2**], *parabolic wake* [**At**],.... In this book we mostly use commonly accepted terminology, though some adjustments would be historically well justified (like *Leau-Fatou coordinates*, *Pfeifer-Cremer points* or *Arnold-Herman rings*).

CHAPTER 4

Dynamical plane II: fine structures and models

25. Hyperbolic maps

In the next several sections, we will take a closer look at some special important classes of quadratic polynomials: hyperbolic, parabolic, and preperiodic.

25.1. Definition revisited.

EXERCISE 25.1. *If the Julia set $\mathcal{J}$ is expanding then it is repelling in the following sense: there exists an $\varepsilon > 0$ such that for any point z in the Fatou set $\mathcal{F}$ there exists an $n \in \mathbb{N}$ such that* $\operatorname{dist}(f^n z, \mathcal{J}) > \varepsilon$.

THEOREM 25.2. *Let f be a quadratic polynomial with connected Julia set $\mathcal{J}$. Then f is expanding on $\mathcal{J}$ if and only if f has an attracting cycle $\boldsymbol{\alpha}$. Moreover, in this case all points $z \in \operatorname{int} \mathcal{K}$ are attracted to the cycle $\boldsymbol{\alpha}$.*

Taking into account Remark 20.6, we see that a quadratic polynomial f is hyperbolic in the sense of §21.2.3 if and only if its Julia set $\mathcal{J}$ is hyperbolic ($\equiv$ expanding) – so, the terminology is consistent.

PROOF. Assume f has an attracting cycle $\boldsymbol{\alpha} = (\alpha_k)_{k=0}^{p-1}$. Take a small invariant neighborhood $U = \bigcup U_k \Subset \mathcal{F}$ of $\boldsymbol{\alpha}$. Let n be the first moment when $f^n(0)$ lands in U, and moreover, let $f^n(0) \in U_k$. Let V_i be the pullback (see §19.1) of U_k containing $f^i(0)$, $k = 0, 1, \dots, n-1$, $V = \bigcup V_k$, and let

$$\Omega' = \mathbb{C} \setminus (\overline{U} \cup \overline{V}), \quad \Omega = f^{-1}(\Omega').$$

Then $\Omega \subset \Omega'$, $\Omega \neq \Omega'$, and $f : \Omega \to \Omega'$ is a covering map. By Corollary 7.1, $\|Df(z)\| > 1$ for any $z \in \Omega$, in the hyperbolic metric of Ω'. Since $\mathcal{J}$ is compactly contained in Ω', there exists $\lambda > 1$ such that $\|Df(z)\| \geq \lambda$, $z \in \mathcal{J}$, so f is expanding on $\mathcal{J}$ with respect to this hyperbolic metric. Since the hyperbolic and the Euclidean metrics over the Julia set $\mathcal{J} \Subset \Omega'$ are equivalent, f is expanding with respect to the latter as well.

REMARK 25.3. This assertion also follows from the Shrinking Lemma (§20.7.2).

Vice versa, assume $\mathcal{J}$ is connected and expanding. The former implies $0 \in \mathcal{K}$ while the latter implies $0 \notin \mathcal{J}$. Thus, $0 \in \operatorname{int} \mathcal{K}$, and in particular, $\operatorname{int} \mathcal{K} \neq \emptyset$.

Let D be a component of $\operatorname{int} \mathcal{K}$. Then we have two possibilities: either D is wandering, i.e., $f^n(D) \cap f^m(D) = \emptyset$ for $n > m \geq 0$, or it is eventually periodic, i.e., there exist $n \geq 0$ and $p \geq 1$ such that $f^n(D) = f^{n+p}(D)$.

If $\mathcal{J}$ is expanding then there exists an $\varepsilon > 0$ such that

$$|Df^n(z)| > 1 \text{ when } \operatorname{dist}(z, \mathcal{J}) < \varepsilon. \tag{25.1}$$

In the wandering case, the iterates $f^n(D)$ are eventually contained in an ε-neighborhood of $\mathcal{J}$. Then by (25.1), $\operatorname{area}(f^{n+1}D) > \operatorname{area}(f^n D)$ for n big enough, which is impossible as the total area of $\mathcal{K}$ is finite.

If D is eventually periodic, we can assume without loss of generality that it is actually periodic. Then by (25.1), there is a domain $D' \Subset D$ such that $f^p(D') \Subset D'$. By the Schwarz Lemma, D' contains an attracting periodic point. Since f can have only one attracting cycle, the conclusion follows. $\square$

25.2. Local connectivity of the Julia set.

THEOREM 25.4. *Let f be a hyperbolic quadratic polynomial. If the Julia set $\mathcal{J}$ is connected then it is locally connected. Moreover, the Böttcher uniformization $B^{-1} : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus \mathcal{K}$ admits a Hölder continuous extension to the boundary.*

PROOF. We will prove directly that the inverse Böttcher function B^{-1} extends Hölder continuously to $\mathbb{T}$.

It is convenient to lift $f_0 : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ to the doubling map $T : \theta \mapsto 2\theta$ on $\mathbb{H}_+$ by means of the exponential $e(z) = e^{2\pi i z}$. Let $\phi := B^{-1} \circ e$, so

$$\phi \circ T = f \circ \phi.$$

For $z \in \mathbb{R}$, let $\mathbb{D}_+(z, \varepsilon) := \mathbb{D}(z, \varepsilon) \cap \mathbb{H}_+$.

Since f is hyperbolic, we can select $\varepsilon_0 < 1/4$ so that the strip

$$\Pi := \{0 < \operatorname{Im} z < \varepsilon_0\}$$

is mapped by ϕ to an annulus $A := \phi(\Pi)$ contained in the neighborhood of $\mathcal{J}$ where $|Df^{-1}(z)| \leq 1/\lambda < 1$ for any branch of f^{-1}. Note also that $f^{-1}(A) \subset A$ since $T^{-1}(\Pi) \subset \Pi$.

Let n_0 be such that $2^{-n_0} < \varepsilon_0$. For any $z \in \mathbb{R}$ and $n \in \mathbb{N}$, let us consider the half-disk $D^+ := \mathbb{D}_+(z, 2^{-(n+n_0)})$. Then the blown-up half-disk $T^n(D_+) = \mathbb{D}_+(T^n z, 2^{-n_0})$ is contained in Π, so the image $\phi(T^n(D_+))$ is contained in A. It follows that any branch of f^{-n} contracts $\phi(T^n(D_+))$ by factor λ^{-n}. But for an appropriate inverse branch, we have:

$$f^{-n}(\phi(T^n(D_+))) = \phi(D_+).$$

Thus,

$$\operatorname{diam}(\phi(D_+)) = O(\lambda^{-n}) = O((\operatorname{diam} D_+)^\kappa), \quad \text{where } \kappa = \frac{\log \lambda}{\log 2},$$

and the conclusion follows. $\square$

We can now refer to general properties of lc hulls (see Exercise 1.33 and Proposition 9.21) to conclude:

COROLLARY 25.5. *Given a hyperbolic map f with connected Julia set, let D_i be the components on $\operatorname{int} \mathcal{K}$ (arbitrary labeled). Then each D_i is a Jordan disk, and $\operatorname{diam} D_i \to 0$.*

Since f acts as the doubling map $T : \theta \mapsto 2\theta \mod 1$ on the external angles, we also obtain:

COROLLARY 25.6. *For a hyperbolic map with connected Julia set, the Böttcher uniformization extends to a semi-conjugacy $B^{-1} : \mathbb{T} \to \mathcal{J}$ between the doubling map T and $f \mid \mathcal{J}$. Moreover, $B^{-1}(e(\theta)) \in \mathcal{J}$ is the landing point of the external ray $\mathcal{R}^\theta$.*

25.3. Blaschke model for the immediate basin.

25.3.1. *Conjugacy.* As above, let $\boldsymbol{\alpha} = (\alpha_k)_{k=0}^{p-1}$ be an attracting cycle of f with the central periodic point α_0. By Corollary 25.5, the immediate basin $\mathcal{D}_0 \ni 0$ of α_0 is a Jordan disk. Let us uniformize it by the unit disk, $\psi : (\mathcal{D}_0, \alpha_0) \to (\mathbb{D}, 0)$. By the Conformal Schönflies Theorem, ψ extends to a homeomorphism $\psi : \overline{\mathcal{D}}_0 \to \overline{\mathbb{D}}$. Let

$$g = \psi \circ f^p \circ \psi^{-1} : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$$

PROPOSITION 25.7. *If the Riemann mapping ψ is appropriately normalized, then*

$$g(z) = z\frac{z+\rho}{1+\bar{\rho}z},$$

where $\rho \in \mathbb{D}$ is the multiplier of the attracting cycle $\boldsymbol{\alpha}$.

PROOF. Consider first an arbitrary uniformization $\psi : (\mathcal{D}_0, \alpha_0) \to (\mathbb{D}, 0)$. The map $g : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ is a double branched covering of the disk fixing 0 and preserving the unit circle $\mathbb{T}$. A general form of such a map is

$$g(z) = \lambda z \frac{z-a}{1-\bar{a}z}, \text{ with } |\lambda| = 1 \tag{25.2}$$

(see Exercise 3.3). Replacing ψ with $\lambda\psi$ results in replacing $g(z)$ with $\lambda g(z/\lambda)$ killing the coefficient λ in front of the Blaschke product (and replacing a with λa).

Since the multiplier is invariant under conformal conjugacies, we have:

$$\rho = g'(0) = -a \equiv \rho.$$

□

REMARK 25.8. In fact, we did not need to know that $\mathcal{D}_0$ is a Jordan disk to conclude that g is the above Blaschke map. It would follow from the property that $g : \mathbb{D} \to \mathbb{D}$ is proper.

COROLLARY 25.9. *In the superattracting case* (*when* 0 *is periodic*), *the return map $f^p : \overline{\mathcal{D}}_0 \to \overline{\mathcal{D}}_0$ is conformally conjugate to $f_0 : z \mapsto z^2$ on $\overline{\mathbb{D}}$.*

In this case, the punctured disk $\mathcal{D}_0^*$ is foliated by the *internal rays* and *equipotentials* corresponding, under the Riemann map $\psi : \mathcal{D}_0 \to \mathbb{D}$, to the radii and circles centered at 0 (compare §9.2.3). They form two orthogonal invariant foliations of $\mathcal{D}_0^*$.

Let us finally mention that the above discussion carries through for other immediate basins $\mathcal{D}_k \equiv \mathcal{D}(\alpha_k)$, $k = 1, \dots, p-1$, with the critical point, $f^{-(p-k)}(0)$, playing the role of 0.

25.3.2. *Hyperbolicity of the Blaschke map.* The Blaschke map g is an example of a *hyperbolic rational map* (with the Julia set $\mathbb{T}$).

LEMMA 25.10. *Blaschke map g* (25.2) *has the following properties:*

(i) $g^n z \to 0$ *in* $\mathbb{D}$; $g^n z \to \infty$ *in* $\mathbb{C} \smallsetminus \overline{\mathbb{D}}$;

(ii) $g : \mathbb{T} \to \mathbb{T}$ *is expanding.*

PROOF. Since g preserves the unit disk $\mathbb{D}$, the family of iterates g^n is normal on $\mathbb{D}$ (by the Little Montel Theorem). Since 0 is an attracting cycle, $g^n z \to 0$ near 0. By normality, $g^n z \to 0$ on the whole disk $\mathbb{D}$.

As $g : \mathbb{D} \to \mathbb{D}$ is a double branched covering, by the Riemann-Hurwitz Formula, it has a unique critical point in $\mathbb{D}$. Let $v \in \mathbb{D}$ be the corresponding critical value; $g^n v \to 0$.

By $\mathbb{T}$-symmetry, g has a critical value, $1/\bar{v} \in \mathbb{C} \setminus \overline{\mathbb{D}}$ and $g^n(1/\bar{v}) \to \infty$. Since $g : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a degree two rational function, it has at most two critical values. Thus, v and $1/\bar{v}$ are the only critical values of g.

Let us consider the post-valuable set $\overline{\mathcal{P}}_g = \operatorname{orb} v \cup \operatorname{orb}(1/\bar{v}) \cup \{0, \infty\}$. Let Ω' be its complement, and $\Omega = g^{-1}(\Omega')$. Since $\overline{\mathcal{P}}_g$ is forward invariant, $\Omega \subset \Omega'$. Since Ω' does not contain the critical values, $g : \Omega \to \Omega'$ is an unramified degree two covering. By Corollary 7.1 of the Schwarz Lemma, $g : \mathbb{T} \to \mathbb{T}$ is an expanding circle map. □

Putting this together with Proposition 19.67, we conclude:

COROLLARY 25.11. *Any Blaschke map $g : \mathbb{T} \to \mathbb{T}$ (25.2) is quasisymmetrically conjugate to the doubling map $f_0|\,\mathbb{T}$.*

EXERCISE 25.12. *Any Blaschke product (25.2) admits a restriction to a double covering $V \to V'$ between two $\mathbb{T}$-symmetric annuli $V \Subset V'$.*

Like any expanding circle map of degree two (see Lemma 19.61), the Blaschke map g has a unique fixed point on $\mathbb{T}$ (of course, this is obvious algebraicly, too). Coming back to our hyperbolic quadratic polynomial f, we conclude that the return map $f^p : \overline{\mathcal{D}}_0 \to \overline{\mathcal{D}}_0$ has a unique fixed point $\beta_0 \in \partial\mathcal{D}_0$. This point is called the *root* of $\mathcal{D}_0$, or the *central root* of f.

Any other component D of the full basin $\mathcal{D}(\boldsymbol{\alpha})$, is conformally mapped under some iterate f^n onto $\mathcal{D}_0$ inducing a homeomorphism $f^n : \partial D \to \partial\mathcal{D}_0$. The point $\beta_D \in \partial D$ corresponding to β_0 under this homeomorphism is called the *root of* D. According to our general convention (see §21.2.3), $\beta^v = f(\beta_0) \in \partial\mathcal{D}^v$ is called the *valuable root*.

25.3.3. *A couple of remarks.* The conjugacy between the Blaschke maps, albeit quasisymmetric, has low regularity. Indeed, by Proposition 19.79, we have:

COROLLARY 25.13. *The (orientation preserving) conjugacy $h : \mathbb{T} \to \mathbb{T}$ between two different Blaschke maps g and $\tilde{g}$ (25.2) is singular.*

EXERCISE 25.14. *Let us consider a more general Blaschke map*

$$g : \overline{\mathbb{D}} \to \overline{\mathbb{D}}, \quad z \mapsto e(\theta)\,\frac{z-a}{1-\bar{a}z}\,\frac{z-b}{1-\bar{b}z}, \quad |a|<1,\ |b|<1.$$

If g has a fixed point $\alpha \in \mathbb{D}$ (e.g., for small b) then:

(i) *It is (dynamically) expanding on the unit circle $\mathbb{T}$;*

(ii) *It preserves the Poisson measure $dm_\alpha(\zeta) = P(\alpha,\zeta)\,dm(\zeta)$ on $\mathbb{T}$, see (10.5);*

(iii) *It is conjugate to a Blaschke map (25.2) by a Möbius automorphism $h : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$.*

25.4. Bounded distortion, quasi-self-similarity, and qc geometry.

25.4.1. *Bounded distortion and dynamical quasi-self-similarity.* The following property provides a fundamental tool in the study of hyperbolic maps:

LEMMA 25.15. *For any hyperbolic polynomial f, there exists an $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$, any point $z \in \mathcal{J}(f)$, and any $n \in \mathbb{N}$, all the inverse branches*

f_i^{-n} are well defined in the disk $\mathbb{D}(z,\varepsilon)$, have an absolutely bounded distortion:

$$\left|\frac{Df_i^{-n}(\zeta)}{Df_i^{-n}(z)}\right| \asymp 1 \quad \forall\, \zeta \in \mathbb{D}(z,\varepsilon),$$

and shrink at a uniformly exponential rate:

$$\operatorname{diam} f_i^{-n}(\mathbb{D}(z,\varepsilon)) = O(\lambda^{-n})$$

(with the constant depending on f only).

Proof. Follows immediately from the fact that $\mathrm{cl}(\operatorname{orb} v) \cap \mathcal{J} = \emptyset$, the Koebe Distortion Theorem, and the expanding property of $f|\,\mathcal{J}$ (compare with the proof of Theorem 22.2). □

Exercise 25.16. *Prove the above distortion bounds following the method of* Lemma 19.68 *instead of the* Koebe Distortion Theorem.

A slight modification of Lemma 25.15 shows that hyperbolic Julia sets in small scales look roughly the same as they do in moderate scales:

Exercise 25.17. *Let f be a hyperbolic quadratic map. Then there exist $\varepsilon > 0$ such that for any $z \in \mathcal{J}(f)$ and $\rho \in (0,\varepsilon)$ there is $n \in \mathbb{N}$ such that f^n univalently and with bounded distortion maps the disk $\mathbb{D}(z,\rho)$ onto an oval of size of order ε and of bounded shape around $f^n z$. (All constants depend on f.)*

Corollary 25.18. *Connected hyperbolic Julia sets $\mathcal{J}(f_c)$ do not have cusps neither from outside (i.e., from $\mathcal{D}(\infty)$) no from inside (i.e., from components of* $\operatorname{int}\mathcal{K}$*).*

25.4.2. *QC geometry.* As the next application of distortion bounds, let us treat the geometry of the components of $\operatorname{int}\mathcal{K}$:

Proposition 25.19. *For a hyperbolic map f, all components D_i of* $\operatorname{int}\mathcal{K}(f)$ *are K-quasidisks, with a uniform K.*

Proof. Let us begin with the central component $\mathcal{D}_0 \ni 0$ of the immediate basin. By Proposition 25.7, the Riemann mapping $\psi : \overline{\mathcal{D}}_0 \to \overline{\mathbb{D}}$ conjugates $f|\,\partial\mathcal{D}_0$ to the restricted Blaschke product $g|\,\mathbb{T}$. Both of these maps are expanding (by Definition (20.3) and Lemma 25.10). By repeating the argument of Lemma 19.67, we see that $\psi : \partial\mathcal{D}_0 \to \mathbb{T}$ is quasisymmetric (Exercise!), so $\partial\mathcal{D}_0$ is a quasicircle.

Any other component D_i of $\operatorname{int}\mathcal{K}(f)$ is univalently mapped onto $\mathcal{D}_0$ by some iterate f^n. By Lemma 25.15, $f^n : \partial D_i \to \partial\mathcal{D}_0$ is a homeomorphism with a bounded distortion, so the D_i are all K-quasicircles with a uniform K. □

Corollary 25.20. *There is a qc homeomorphism $h_0 : (\mathbb{C}, \overline{\mathcal{D}}_0, 0) \to \overline{\mathbb{C}}, \overline{\mathbb{D}}, 0)$ conjugating $f^p|\,\partial\mathcal{D}_0$ to the doubling map $f_0|\,\mathbb{T}$.*

Proof. By Proposition 25.19, the Riemann mapping $\phi : (\mathcal{D}_0, 0) \to (\mathbb{D}, 0)$ admits a qc extension to the whole plane. By Proposition 25.7, it conjugates the return map $f^p|\,\overline{\mathcal{D}}_0$ to a Blaschke map $g|\,\overline{\mathbb{D}}$. By Corollary 25.11, $g|\,\mathbb{T}$ is quasisymmetrically conjugate to the doubling map $f_0|\,\mathbb{T}$. Since the latter conjugacy admits a qc extension to the whole plane (by the Ahlfors-Beurling Theorem), the conclusion follows. □

25.4.3. *Structure of maps with attracting fixed point.* Let us consider, in particular, a hyperbolic map $f \equiv f_c$ that has an attracting fixed point. (Such a c beloings to the main hyperbolic component of the Mandelbrot set: see Exercise 33.2 below).

Recall the notion of qc welding from §15.4.

PROPOSITION 25.21. *Let f has an attracting fixed point, and let $\mathcal{D}_0$ be its immediate basin. Then the Julia set $\mathcal{J} \equiv \mathcal{J}(f)$ coincides with $\partial\mathcal{D}_0$. Moreover, it is a quasicircle obtained by qc welding by means of a homeomorphism $h : \mathbb{T} \to \mathbb{T}$ conjugating a Blaschke map g to f_0.*

PROOF. Since $f : \mathcal{D}_0 \to \mathcal{D}_0$ is a double branched covering, $f^{-1}(\mathcal{D}_0) = \mathcal{D}_0$. Hence $f^{-1}(\partial\mathcal{D}_0) = \partial\mathcal{D}_0$, implying (by Corollary 20.23) that $\partial\mathcal{D}_0 = \mathcal{J}$. Now Proposition 25.19 implies that $\mathcal{J}$ is a quasicircle.

By Proposition 25.7, $f \,|\, \mathcal{D}_0$ is conformally conjugate to a Blaschke map g, while by Theorem 23.25, $f \,|\, \hat{\mathbb{C}} \smallsetminus \overline{\mathcal{D}}_0$ is conformally conjugate to $f_0 \,|\, \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$. Moreover, since $\partial\mathcal{D}_0$ is a Jordan curve, all the conjugacies extend continuously to the boundary. Hence f is the qc welding by means of the map $h : \mathbb{T} \to \mathbb{T}$ conjugating $g \,|\, \mathbb{T}$ to $f_0 \,|\, \mathbb{T}$. □

Under the above circumstances, f is called the (*qc*) *mating* between the Blaschke map $g : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ and the square map $f_0 : \hat{\mathbb{C}} \smallsetminus \mathbb{D} \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}$.

REMARK 25.22. Proposition 25.21 can be seen is several other ways. One follows from the Attracting-Superatttracting QC Surgery turning f into f_0 (see §25.8 below). Another one follows from the Structural Stability Theory (developed in §36 below) implying that the Julia set $\mathcal{J}_c$ moves holomorphically over the main hyperbolic component of the Mandelbrot set.

These approaches would bipass Proposition 25.19, and in fact, the latter can be then derived from Proposition 25.21 by means of the Renormalization (see §28.4.3) and Straightening (see §40). (Of course, it would be much more sophisticated approach than the one we have carried above, but it would also provide us with an instructive insight into the matter.)

25.5. Porosity and area of the Julia set. Theorem 22.2 on the gobal measure-theoretic attractor immediately implies that area $\mathcal{J}(f) = 0$ for any hyperbolic f. Here is a sharper statement:

PROPOSITION 25.23. *Any hyperbolic Julia set $\mathcal{J}(f)$ is uniformly porous in all scales. Hence it has zero area. Moreover, if $\mathcal{J}(f) \subset \mathbb{R}$ then it has zero length.*

PROOF. By Lemma 25.17, any small scale disk $\mathbb{D}(z, \rho)$ with $z \in \mathcal{J}$ can be mapped with bounded distortion onto a moderate scale oval of bounded shape. Since the latter contains a definite gap, the former contains one as well. □

EXERCISE 25.24. *Generalize* Proposition 25.23 (*along with* Lemma 25.15 *and* Exercise 25.17) *to arbitrary invariant hyperbolic sets $Q \subset \mathcal{J}$.*

25.6. Hubbard tree for a superattracting map. We will now attach to any superattracting quadratic polynomial a topological object called a (*geometric*) *Hubbard tree*. It will be eventually shown (see Theorem 35.1) that these polynomials (and there are countably many of them) are fully classified by their Hubbard trees.

25.6.1. *Definition and first properties.* Let $f = f_c$ be a superattracting quadratic polynomial, so its critical point 0 is periodic with some period $p \in \mathbb{Z}_+$. Let $0_k = f^k(0)$, $k = 0, 1, \dots, p-1$, $\mathbf{0} = \{0_k\}_{k=0}^{p-1}$, and let $\mathcal{D}_k$ be the component of $\operatorname{int}\mathcal{K} \equiv \operatorname{int}\mathcal{K}(f)$ containing 0_k. By the last part of Theorem 25.2, $\operatorname{int}\mathcal{K}$ coincides with the basin of $\boldsymbol{\alpha}$, so for any component D of $\operatorname{int}\mathcal{K}$ there exists a unique $n = n(D) \in \mathbb{N}$ such that f^n univalently maps D onto $\mathcal{D}_0$. Let us mark in D the preimage of 0 under this map and call it the *center* 0_D of D (in particular, 0_k is the center of $\mathcal{D}_k$). This makes $\mathcal{K}$ a *pointed lc hull*.

By Exercise 9.23, any two points z, ζ in $\mathcal{K}$ can be connected by a unique legal arc $[z, \zeta]$. Let us consider the legal hull $\mathcal{T} \equiv \mathcal{T}_f$ of the points 0_k. When 0 is a fixed point (i.e., $p = 1$), then $\mathcal{T} = \{0\}$ is trivial, so *in what follows we will assume that* $p > 1$. Then by Exercise 9.26, $\mathcal{T}$ is a topological tree called the (*geometric*) *Hubbard tree*. Let $\mathbf{b} = \{b_j\}$ be the set of its branch points. Let us also mark on $\mathcal{T}$ the points 0_k.

Proposition 25.25. *We have:*

(i) *The marked Hubbard tree* $(\mathcal{T}, \mathbf{0} \cup \mathbf{b})$ *is invariant under* f; *hence all branch points of* $\mathcal{T}$ *are* (*pre-*)*periodic; moreover,* $f : \mathcal{T} \to \mathcal{T}$ *is surjective.*

(ii) *The critical value* $v = c$ *is a tip of* $\mathcal{T}$; *the critical point* 0 *is not a branch point of* $\mathcal{T}$.

Proof. (i) Let us take the legal path $\gamma_k = [0, 0_k] \subset \mathcal{T}$ connecting 0 to another 0_k, $k = 1, \dots, p-1$. Since $\operatorname{int}\gamma_k \not\ni 0$, γ_k does not contain symmetric points (Exercise 9.25). Hence f is injective on γ_k, which implies that $f(\gamma_k)$ is the legal arc connecting 0_1 to 0_{k+1} (where $0_p \equiv 0$). Since $\mathcal{T}$ is legally convex, $f(\gamma_k) \subset \mathcal{T}$. Since $\mathcal{T} = \bigcup \gamma_k$ (being a tree), and similarly, $f(\mathcal{T}) = \bigcup\, [0_1, 0_k] = \bigcup f(\gamma_k)$, the first assertion follows.

(ii) Since $\mathcal{T} = \bigcup [0_k, 0_j]$, all the tips of $\mathcal{T}$ are contained in the cycle $\mathbf{0}$. So, one of the points 0_k must be a tip. But if 0_k with $k \neq 0 \,(\mathrm{mod}\, p)$ is not a tip then 0_{k+1} is not either, since the map $f : \mathcal{T} \to \mathcal{T}$ near any non-critical point is a local emebdding. Thus, if 0_1 is not a tip then none of the 0_k are – contradiction. □

Remark 25.26. The above argument shows that there is an $l \in [1, p]$ such that the tips of $\mathcal{T}$ are exactly the points 0_k, $k = 1, \dots, l$.

25.6.2. *Attracting basin.*

Lemma 25.27. *The basin* $\mathcal{D}(\mathbf{0}) \cap \mathcal{T}$ *is dense in the Hubbard tree* $\mathcal{T}$.

Proof. Otherwise, there is an arc $L \subset \mathcal{T}$ such that all the images $f^n(L)$ are disjoint from the immediate basin $\mathcal{D}_0$, and in particular, they do not contain 0. Then all the iterates $f^n : L \to \mathcal{T}$ are embeddings.

Let us consider the hyperbolic metric on $\mathbb{C} \smallsetminus \mathbf{0}$, and define the corresponding "hyperbolic distance" on each component J of $\mathcal{T} \smallsetminus \mathbf{0}$ as follows. For $x, y \in J$, a smooth path γ connecting x to y in $\mathbb{C} \smallsetminus \mathbf{0}$ is called *admissible* if it can be retracted to $[x, y] \subset \mathcal{T}$ in $\mathbb{C} \smallsetminus \mathbf{0}$ rel the endpoints. Then we let

$$d_{\mathrm{hyp}}(x, y) = \inf_{\gamma} l_{\mathrm{hyp}}(\gamma), \tag{25.3}$$

where γ runs over all admissible paths connecting x to y.

Note that if the images $x' = f(x)$ and $y' = f(y)$ also lie in the same component J' of $\mathcal{T} \smallsetminus \mathbf{0}$, then any admissible path γ' connecting x' to y' lifts by f to an admissible path γ connecting x to y. By Corollary 7.1, $l_{\text{hyp}}(\gamma') > l_{\text{hyp}}(\gamma)$, which implies that

$$d_{\text{hyp}}(x', y') \geq \lambda\, d_{\text{hyp}}(x, y), \tag{25.4}$$

with $\lambda \geq \lambda(\varepsilon) > 1$, provided x' and y' stay ε-away from $\mathbf{0}$.

It follows that for the above arc $L = [a, b] \subset \mathcal{T}$, we have $d_{\text{hyp}}(f^n a, f^n b) \to \infty$ (exponentially fast), which is of course impossible. $\square$

Lemma 25.28. *If two components of the immediate basin $\mathcal{D}^\bullet(\mathbf{0})$ do not have disjoint closures then they touch at the root points.*

Proof. By invariance of the immediate basin, we can assume that one of the components in question is valuable, $\mathcal{D}_1 \ni v$. Proposition 25.25(ii) implies that $\overline{\mathcal{D}}_1 \cap \mathcal{T}$ is a closed internal ray with endpoints v and $b \in \partial \mathcal{D}_1$. Since it is invariant under f^p, we conclude that $f^p(b) = b$, so b is the root $\mathcal{D}_1$. $\square$

Let us consider the cycle $(\beta_k)_{k=0}^{l-1}$ of the root $\beta_0 \in \partial \mathcal{D}_0$. We see that the components $\mathcal{D}_k$ of the immediate basin $\mathcal{D}^\bullet(\boldsymbol{\alpha})$ are organized into l *bouquets* $\mathcal{B}_k$ sharing the root β_k. Each bouquet comprises $\mathfrak{q} = p/l$ components cyclically permuted by f^l with some combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$ (independent of the bouquet). Moreover, the bouquets have disjoint closures.

If all the components $\mathcal{D}_k$ have disjoint closures (i.e., the bouquets $\mathcal{B}_k$ comprise one component each: $\mathfrak{q} = 1$) then the hyperbolic map f is called *primitive*. Otherwise, it is called *satellite*.

25.6.3. *Markov tiling.* If we puncture out all the points 0_k and b_j, from the Hubbard tree $\mathcal{T}$, the rest will be disjoint union of topological intervals J_s°. Their closures J_s form a tiling $\mathfrak{J}$ of $\mathcal{T}$.

Proposition 25.29. *The tiling $\mathfrak{J}$ is Markov. It generates a natural semi-conjugacy $h : \Sigma_A \to \mathcal{T} \smallsetminus \mathcal{D}(\mathbf{0})$. This semi-conjugacy is one-to-one over all points of $\mathcal{T} \smallsetminus \mathcal{D}(\mathbf{0})$ except for the branch points $b \in \mathbf{b}$ and their dynamical preimages.*

Proof. Let $J_s, J_t \in \mathfrak{J}$ and $f(J_s^\circ) \cap J_t \neq \emptyset$. Since $J_s^\circ \not\ni 0$, f homeomorphically maps J_s° onto an interval in the tree $\mathcal{T}$. Since the tiling boundary $\mathbf{0} \cup \mathbf{b}$ is forward invariant, $\partial(f(J_s)) \cap J_t^\circ = \emptyset$. Since J_t° does not contain branch points, $f(J_s) \supset J_t$. It proves the Markov property.

The map f is expanding on the tiles with respect to the hyperbolic metric, see (25.4). Hence there exists a natural semi-conjugacy $h : \Sigma_A \to \mathcal{T} \smallsetminus \mathcal{D}(\mathbf{0})$ (see Exercise 19.92). The only points over which h may fail to be one-to-one are common points of different tiles (which are, in this case, only branch points of $\mathcal{T}$), and their dynamical preimages. $\square$

Exercise 25.30. *Let $J_v \in \mathfrak{J}$ be the valuable interval ending at the critical value v. Then it can be reached in the Markov graph Γ_A* (see Appendix, §19.19) *from any other interval $J_s \in \mathfrak{J}$.*

See §25.6.13 for more properties of this Markov chain.

25.6.4. *The $\alpha-$fixed point.*

LEMMA 25.31. *The legal path γ connecting the critical point 0 to the critical value $0_1 = c$ contains the $\alpha-$fixed point of f.*

PROOF. The image $\delta := f(\gamma)$ is the legal path connecting c to $0_2 \equiv f(c)$. Let us orient γ from 0 to c, and respectively, orient δ from c to 0_2. Since the critical value c is a tip of $\mathcal{T}$ (Proposition 25.25) these orientations are opposite near c.

Note that the inclusion $\delta \subset \gamma$ is impossible, since in this case a topological interval, γ, is mapped by f to itself, so it would contain a fixed point which is attracting at least on one side (as long as we know that the fixed points are isolated, which is certainly the case for polynomials).

If $\delta \supset \gamma$ then the inverse branch $f^{-1} : \delta \to \gamma$ maps a topological interval, δ, to a smaller interval γ, so it has a fixed point $\alpha \in \gamma$.

Otherwise, $\delta \cap \gamma = [c, \alpha]$, where α is a branch point of $\mathcal{T}$. Let us show that this point is fixed under f. If this is not the case, then γ contains a point $\alpha_{-1} \neq \alpha$ such that $f(\alpha_{-1}) = \alpha$. Let us consider two cases:

Case a): $\alpha_{-1} \in (0, \alpha)$. Then the image $\alpha_1 := f(\alpha) \neq \alpha$ is a point of the path $(\alpha, 0_2)$ that branches off γ. Let us consider the topological interval $I_0 = [\alpha, \alpha_1] \subset [\alpha, 0_2]$. It is oriented *away from* the critical point 0 (in the sense that the path $[0, \alpha]$ is disjoint from I_0). Hence I_0 extends the intervals $I_{-1} := [\alpha_{-1}, \alpha]$ beyond α to form the interval $[\alpha_{-1}, \alpha_1]$.

Then the interval $I_1 \equiv [\alpha_1, \alpha_2] := f(I_0) \subset \mathcal{T}$ extends $I_0 = f(I_{-1})$ beyond α_1 to form the interval $[\alpha, \alpha_2]$ oriented away from 0 as well. Attaching to it the next iterate $I_3 \equiv [\alpha_2, \alpha_3] := f^2[\alpha, \alpha_1]$, we obtain a bigger interval $[\alpha, \alpha_3]$ with the same property, etc. Note that we will never hit the critical point (since the intervals $[\alpha, \alpha_n]$ grow away from it in the Hubbard tree), so the intervals in question will never be folded under f. In this way, we obtain infinitely many branch points $\alpha_n = f^n(\alpha)$ of $\mathcal{T}$, which is impossible.

Case b): $\alpha_{-1} \in (\alpha, c)$. Then the interval $I_0 = [\alpha, \alpha_1]$ is contained in $[\alpha, c]$. However, it cannot be contained in $[\alpha, \alpha_{-1}]$, for otherwise the latter interval would be invariant under f; and then it would contain a semi-attracting fixed point or a semi-attracting cylce of period two. Hence $\alpha_1 \in (c, \alpha_{-1})$ and then $[\alpha, \alpha_2]$ is an interval that branches off γ. Moreover, the interval $[\alpha_{-1}, \alpha]$ contains a preimage α_{-2} of α_{-1}. Thus, the interval $[\alpha_{-2}, \alpha_2]$ is a concatenation of $[\alpha_{-2}, \alpha]$ and $[\alpha, \alpha_2]$. Applying to it the iterates of f^2, we obtain, as in Case a), a chain of intervals containing infinitely many branch points.

Finally, since $\alpha \in \gamma^\circ$, it is dividing, so it is identified with the $\alpha-$fixed point introduced in §24.4.2. □

25.6.5. *Body and limbs of $\mathcal{T}$.* Recall from §9.1.5 the notion of branches and limbs of $\mathcal{T}$ at a point $z \in \mathcal{T}$ (note that the Hubbard tree is naturally centered at 0). Ler $\mathcal{L}_i$ be the *limbs* attached to the $\alpha-$fixed point, $i = 1, \dots, \mathfrak{q} - 1$, and let $\mathcal{L}'_i$ be the symmetric[1] *lateral limbs* attached to α' (if exist). Let us also consider the *body* $\mathcal{B}$ of $\mathcal{T}$, which is the closure of the component of $\mathcal{T} \smallsetminus \{\alpha, \alpha'\}$ containing 0.

[1]Meaning that $\mathcal{L}_i$ and $\mathcal{L}'_i$ have symmetric local branches at α and α' respectively.

EXERCISE 25.32. *The limbs of the Hubbard tree can be labeled so that*

- *For* $i \in \{1, \dots, \mathfrak{q}-2\}$, F *embeds* $\mathcal{L}_i$ *and* $\mathcal{L}'_i$ *to* $\mathcal{L}_{i+1}$*;*
- F *embeds* $\mathcal{L}_{\mathfrak{q}-1}$ *and* $\mathcal{L}'_{\mathfrak{q}-1}$ *to* $\mathcal{B} \cup \bigcup \mathcal{L}'_i$*;*
- F *maps* $\mathcal{B}$ *to* $\mathcal{L}_1 \equiv \mathcal{L}_v$*; moreover, this map is an embedding on each branch of* $\mathcal{T}$ *at* 0.

In case when there are two branches at 0 (i.e., 0 is not a tip), we say that $F : \mathcal{B} \to \mathcal{L}_v$ is a *folding map*. Note that this notion can be applied to general maps between trees. In case of an interval maps, it is equivalent to (topological) unimodality.

More generally, we can consider the Hubbard branches $\mathcal{L}_i \equiv \mathcal{L}_i(\alpha)$ at a repelling periodic point $\alpha \in \mathcal{T}$ of some period p (where $\mathcal{L}_0 \equiv \mathcal{B}$ the corresponding *body* while $\mathcal{L}_i$, $i \neq 0$, are the *limbs*) .

LEMMA 25.33. *Each sector* $S_i \equiv S_i(\alpha)$ *centered at* α *contains exactly one Hubbard branch* $\mathcal{L}_i \equiv \mathcal{L}_i(\alpha)$. *Thus the number of Julia branches at* α (*equal to the number of external rays landing at* α) *is equal to the valence of* $\mathcal{T}$ *at* α (*which is the number of Hubbard branches at* α).

PROOF. Otherwise there is a sector S_i that does not contain postrcritcal points. Then it can be univalently pulled back to itself by an appropriate branch of $f^{-p\mathfrak{q}}$. Then S_i is invariant under $f^{p\mathfrak{q}}$ contradicting the leo property of the Julia set. $\square$

25.6.6. *Extended Hubbard tree.* The *extended Hubbard tree* $\mathcal{I} \equiv \mathcal{T}^e_f$ is the legal hull of the cycle $\mathbf{0}$, the non-dividing fixed point β, and the co-fixed point $\beta' = -\beta$.

EXERCISE 25.34. (i) *Show that* $\mathcal{I} = \mathcal{T}_f \cup \sigma_f$ (*where* $\sigma \equiv \sigma_f$ *is the spine*) *and that* $\mathcal{T}^e_f$ *is* f*-invariant.*
(ii) *Describe how the spine* σ *is located with respect to the Hubbard tree* $\mathcal{T}$*, i.e., given a Hubbard tree* $\mathcal{T}$*, construct the extended Hubbard tree* $\mathcal{I}$.

EXERCISE 25.35. *Show that any point* $z \in \mathcal{I}$*, except for* β *and* β'*, eventually lands on* $\mathcal{T} \cap \sigma$*, i.e., there exists an* $n \in \mathbb{N}$ *such that* $f^n z \in \mathcal{T} \cap \sigma$.

It follows, in particular, that the skeleton of f can be obtained by taking preimages of the Hubbard tree:

$$\mathcal{S}k \equiv \bigcup_{n=0}^{\infty} f^{-n}(\sigma) = \bigcup_{n=0}^{\infty} f^{-n}(\mathcal{T}).$$

Notice that the branch points of $\mathcal{S}k$ are preimages of the branch points on $\mathcal{I}$ and the centers of the $\mathcal{D}_k$. As the former have the same valence as the corresponding points on $\mathcal{I}$, while the latter have infinite valence but lie outside the Julia set, we conclude (using Lemma 25.33) that the Hubbard tree fully captures branching of the Julia set:

COROLLARY 25.36. *For a superattracting polynomial* f*, any branch points* ζ *of* $\mathcal{J}(f)$ *is a preimage of a periodic branch point* z *of the extended Hubbard tree* $\mathcal{I}$. *The number of rays landing at* ζ *is equal to the valence of* z *in* $\mathcal{I}$.

EXERCISE 25.37. *The* β*−fixed point does not belong to* $\mathcal{T}$.

25.6.7. *$\alpha-$rotational type.* Let us say that $\mathcal{T}$ is of *$\alpha-$rotational type* if the Hubbard tree $\mathcal{T}$ does not contain the co-fixed point α'. In this case, the body $\mathcal{B}$ coincides with the central branch $\mathcal{L}_0 \ni 0$ of $\mathcal{T}$, implying that $f(\mathcal{L}_0) \subset \mathcal{L}_1$. It follows that the branches $\mathcal{L}_i$ at α are cyclically permuted by f (Exercise 25.32). Under these circumstances, the branches at α are also called (*extended satellite*) *little Hubbard trees* $\mathcal{T}_k^e \equiv \mathcal{I}_k$. We see that they touch at the $\alpha-$fixed point, but otherwise disjoint, cyclically permuted by the dynamics, and altogether cover the whole Hubbard tree $\mathcal{T}$.

The simplest examples of such a tree is a star centered at α whose branches are permuted with rotation number $\mathfrak{p}/\mathfrak{q}$. The corresponding quadratic polynomials are the centers of the hyperbolic components of the Mandelbrot set attached to the main cardioid (see §§33.5, 37.4 below).

Exercise 25.38. *Show that that each little satellite Hubbard tree $\mathcal{I}_k$ is the legal hull of the $\alpha-$fixed point and the postcritical points $0_{k+\mathfrak{q}n}$ that belong to the correspodning sector S_k (bounded by two rays landing at α).*

25.6.8. *Non-rotational type: puzzle.* Let us now consider an ($\alpha-$)non-rotational Hubbard tree $\mathcal{T} \ni \alpha'$. In this case, let us consider further preimages of α' in the extended Hubbard tree $\mathcal{I}$: namely, let $A^{(n)} := f^{-n}(\alpha)$, $n \in \mathbb{N}$ (where f stands for the action on $\mathcal{I}$). They tessellate $\mathcal{I}$ into subtrees $I_k^{(n)}$ called *puzzle trees of depth n*. For the moment, they are labeled arbitrary, ecccept that we let $I^{(n)} \equiv I_0^{(n)} \ni 0$ be the *central* puzzle tree, while $I_v^{(n)} \equiv I_1^{(n)} \ni v$ be the *valuable* one. [Since 0 and v are separated by the $\alpha-$fixed point (Lemma 25.31), these trees are well defined.] These tilings $\mathfrak{I}^{(n)}$ satisfy the following easy to check but crucial properties:

Exercise 25.39. (i) *The tilings $\mathfrak{I}^{(n)}$ are nested: each puzzle tree $I^{(n)}$ of depth n is tesselated by some puzzle trees $I_k^{(n+1)}$ of depth $n+1$.*

(ii) *This nest is invariant in the following sense: Any tile $(I_k^{(n)}, \partial I_k^{(n)})$ of depth $n>1$ with marked boundary is properly[2] mapped by f to some marked tile $(I_{j(k)}^{(n-1)}, \partial I_{j(k)}^{(n-1)})$ of depth $n-1$.*

(iii) *The above map is a proper embedding if $I_k^{(n)}$ is off-central $(k \neq 0)$ and is a proper folding map otherwise (for $k=0$).*

Let us define the *Principal Nest of Hubbard trees*, $I^0 \supset I^1 \supset \dots$ as follows: $I^0 := I^{(0)}$, and I^{n+1} is defined inductively as the pullback of I^n under the first return of the critical orbit to I^n.

Exercise 25.40. *Show that:*

(i) *It is indeed a nest;*

(ii) *The first return map $g_n : I^n \to I^{n-1}$ is a proper folding map.*

(iii) *$I^{n+1} \Subset I^n$ for n sufficiently big.*

See §31 for a detailed discussion of the Yoccoz puzzle and its Principal nest for the whole Julia set. The interval case, which is a special case of Hubbard tree, will be discussed in §31.11.

In the following two sections, we will proceed with analyzing the non-rotational type.

[2] In the adjusted sense of §50.3.2.

25.6.9. *Relative exactness.*

LEMMA 25.41. *In the non-rotational case, we have:*
(i) *There is a moment* $\mathbf{n} \in \mathbb{Z}_+$ *such that* $f^{\mathbf{n}}$ *stretches some interval* $[\alpha, \gamma] \subset [\alpha, 0)$ *onto* $[\alpha, \alpha'] \subset \mathcal{T}$.
(ii) *There is a sequence of iterated preimages of* α, $\gamma_k \in A^{(n_k)}$ *converging to* α *(with* $\gamma_0 \equiv \gamma$*), such that* $\gamma_{k+1} \in [\alpha, \gamma_k)$ *and* $f^{n_{k+1}-n_k}$ *homeomorphically stretches* $[\alpha, \gamma_{k+1}]$ *onto* $[\alpha, \gamma_k]$.

PROOF. (i) Since $\mathcal{T} \ni \alpha'$, there is an $\mathbf{n}$ such that $[\alpha, 0_{\mathbf{n}}) \ni \alpha'$, implying the assertion.

(ii) Let $\gamma_0 := \gamma$, $n_0 := \mathbf{n}$, and let us define $\gamma_1 \in [\alpha, \gamma)$ as the preimage of γ of smallest possible depth $n_1 > n_0$ and closest to α among such. Then $f^{n_1-n_0}$ maps $[\alpha, \gamma_1]$ onto $[\alpha, \gamma_0]$, while the intermediate iterates $f^m[\alpha, \gamma_1]$, $m = 0, 1 \ldots, n_1 - n_0 - 1$, do not contain γ_0. All the more, they do not contain 0, so the map $f^{n_1-n_0} : [\alpha, \gamma_1] \to [\alpha, \gamma_0]$ is homeomorphic.
Proceed inductively to construct further preimages γ_k. □

Let us now consider the closure of the grand backward orbit of the $\alpha-$fixed point:

$$A_- := \overline{\mathrm{Orb}_-(\alpha)} = \mathrm{cl} \bigcup_{n=0}^{\infty} A^{(n)}.$$

LEMMA 25.42. *In the non-rotational case, we have:*
(i) *For any local branch* T *of* $\mathcal{T}$ *at* α, *there is an* $n \in \mathbb{N}$ *such that* $f^n(T) = \mathcal{T}$.
(ii) *The set* A_- *is perfect.*
(iii) *For any interval* $V \subset \mathcal{I}$ *that intersects* A_- *there is an* $n \in \mathbb{N}$ *such that* $f^n(V) = \mathcal{T}$.

PROOF. (i) Since the local branches at α are cyclically permuted (Exercise 25.32), we can assume that $T \subset [\alpha, 0]$. Then Lemma 25.27 (density of the basin) together with Lemma 25.41(ii) imply that T contains some component of the basin $\mathcal{D}(\mathbf{0}) \cap \mathcal{T}$. Moreover, this component is homeomorphically mapped by some f^m onto the central component $\mathcal{D}_0 \cap \mathcal{T} \ni 0$. Hence $f^m(T) \ni 0$. Since the orbit of 0 spans the whole tree $\mathcal{T}$, we conculde that

$$\bigcup_{n=0}^{p-1} f^n(T) \supset \mathcal{T}.$$

Finally, let us consider a neighborhood $V := \bigcup_{n=0}^{\mathfrak{q}-1} f^n(T)$ of α in $\mathcal{T}$, where $\mathfrak{q}$ is the valence of α (we use Exercise 25.32) once again). Since α is repelling, $f^{\mathfrak{q}}(V) \supset V$, provided T is sufficiently small. It follows that $f^n(V) \supset \mathcal{T}$, as desired.

Assertions (ii) and (iii) readily follow. □

We refer to property (iii) above as *relative* (*topological*) *exactness* of f (rel the complement of A_-) or *relative leo property.*

EXERCISE 25.43. *What happens in the* $\alpha-$*rotational case?*

25.6.10. *Little Hubbard trees.* The set $\mathcal{I} \smallsetminus A_-$ is an open invariant subset of $\mathcal{I}$ whose components are (open) subtrees of $\mathcal{I}$. Let $\mathcal{I}^1_k$ be the closures of these components containing the postcritical points 0_k, respectively. They are permuted by the dynamics with some period p_1 dividing p. We call them (*extended*) *little Hubbard trees.* (According to our general convention, the little Hubbard tree $\mathcal{I}^1 \equiv \mathcal{I}^1_0 \ni 0$ is the *central* one, while $\mathcal{I}^1_1 \equiv \mathcal{I}^1_v \ni v$ is the *valuable* one.) Lemma 25.42(ii) implies:

Corollary 25.44. *In the non-rotational case, the little Julia trees $\mathcal{I}^1_k$ are pairwise disjoint.*

This allows us to define the *valence* of a little Hubbard tree $\mathcal{I}^1_k$ as the number of components $\mathcal{I} \smallsetminus \mathcal{I}^1_k$.

Lemma 25.45. *In the non-rotational case, we have:*

(i) *The valuable Hubbard tree $\mathcal{I}^1_v$ has valence one.*

(ii) *The central Hubbard tree, $\mathcal{I}^1_0$, has valence two. Moreover, the boundary points of $\mathcal{I}^1$ (rel $\mathcal{I}$) are $\beta^1 \equiv \beta^1_0$ and $(\beta^1)' \equiv (\beta^1)'$, where β^1 is a periodic point of period p_1, while $(\beta^1)'$ is the symmetric point.*

(iii) *Any other little Hubbard tree has valence one or two.*

Proof. Let us collapse all the components of $\mathcal{T} \smallsetminus A_-$ to points (a "devil staircase" construction). We obtain a tree with the induced action by f (an "abstract Hubbard tree": compare §25.6.12 below). We can now follow the proof of Proposition 25.25 to get (i) and (iii). Moreover, it leaves only options of valence one or two for the central Hubbard tree. Let us show that it is decided in favor of "two".

Since the valuable tree $\mathcal{I}^1_v$ has valence one, it has a single boundary point β^1_v and this point must be fixed by the return map f^{p_1}. Its preimages are the only boundary points for the central tree $\mathcal{I}^1$, so $\mathcal{I}^1$ has only one or two boundary points. If there was only one, β^1, then there would be only one branch T of $\mathcal{I}$ attached to β^1. But this branch could not contain both points α and α', contradicting the property $\mathcal{I} \ni \alpha'$.

Finally, as β^1 and $(\beta^1)'$ are 0-symmetric, one of them, β^1_0, is fixed for the corresponding return map f^{p_1}, while the other one, $(\beta^1_0)'$, is co-fixed. Moreover, since the little trees $\mathcal{I}^1_k$, $k = 0, \dots, p_1 - 1$, are disjoint, p_1 is the smallest period of β^1_0. $\square$

Corollary 25.46. *In the non-rotational case, the central Hubbard tree $\mathcal{I}^1$ is the convex hull of the postcritical points $0_{p_1 i}$, $i = 0, 1, \dots, p/p_1 - 1$, and the boundary points β^1, $(\beta^1)'$.*

Proof. If a tree $\mathcal{I}$ is spent by some points 0_k then any subtree $\mathcal{I}'$ is spent by the points 0_k contained in $\mathcal{I}'$ and the boundary points of $\mathcal{I}'$. $\square$

Taking the convex hull of the points 0_k contained in $\mathcal{I}^1$ we obtain the *little* (*unextended*) *Hubbard tree* $\mathcal{T}^1 \subset \mathcal{I}^1$. Its images $\mathcal{T}^1 := F^k(\mathcal{T}^1)$, $k = 0, 1, \dots, p_1$, form a *cycle* of little (unextended) Hubbard trees.

Exercise 25.47. *Show that the principal Hubbard trees shrink to the little Hubbard tree: $\bigcap I^n = \mathcal{I}^1$.*

25.6.11. *Nest of little trees.* If the Hubbar tree $\mathcal{T}^1$ is non-trivial (i.e., contains more than one point 0), we can apply the above construction to it, and obtain little Hubbard trees $\mathcal{I}^2 \supset \mathcal{T}^2$ of level two, of some period p_2, etc. In this way we obtain the nest of the central Hubbard trees,

$$\mathcal{T} \equiv \mathcal{T}^0 \supset \mathcal{I}^1 \supset \mathcal{T}^1 \cdots \supset \mathcal{I}^n \supset \mathcal{T}^n = \{0\}. \tag{25.5}$$

of some periods $1 \equiv p_0 \,|\, p_1 \,|\, \dots \,|\, p_n \equiv p$. A tree $\mathcal{T}^k$, $k > 0$, is classified as *satellite* or *primitive* depending on whether it is rotational in $\mathcal{T}^{k-1}$ or not.

A tree $\mathcal{T}$ is called *prime* if $n = 1$.

EXERCISE 25.48. *The following properties are equivalent:*

- *$\mathcal{T}$ is prime.*
- *$\mathcal{T}$ does not contain a strictly smaller periodic subtree $\mathcal{T}' \ni 0$ spanned by several (more than one) marked points 0_k which are not cut by the $\alpha-$fixed point.*
- *The iterated preimages of the $\alpha-$fixed point are dense in the complement of the basin $\mathcal{D}(\boldsymbol{\alpha})$.*

This nest of little Hubbard trees will be later linked to the nest of little Julia sets of quadratic-like renormalizations (compare §31.10). In the interval case, it will be discussed in §30.7.

25.6.12. *Abstract Hubbard tree.* An *abstract Hubbard tree* (with periodic critical point) is a topological tree $\mathcal{T}$ endowed with the following data:

(i) Marked points $0 \equiv 0_0, \dots, 0_{p-1}$ such that all the tips of $\mathcal{T}$ are marked. The point 0 is called *critical* or *central*, while $v \equiv 0_1$ is called *valuable*.

(ii) Cyclic order of local branches at any vertex. This is equivalent to the choice of an embedding $\mathcal{T} \to \mathbb{C}$ up to isotopy.

(iii) A continuous *Hubbard map* $F \equiv F_{\mathcal{T}} : (\mathcal{T}, (0_k)) \to (\mathcal{T}, (0_{k+1}))$ (where $k \in \mathbb{Z}/p\mathbb{Z}$) such that the restrictions of F to the branches $\mathcal{T}_\pm$ at 0 are injective. This map is determined uniquely up to *Thurston equivalence*: $F \sim \tilde{F}$ if there exists a homeomorphism[3] $h : (\mathcal{T}, (c_k)) \to (\tilde{\mathcal{T}}, (\tilde{c}_k))$ such that $h^{-1} \circ \tilde{F} \circ h$ is homotopic to F rel (c_k).

According to our general convention, a "topological tree" is considered up to a homeomorphism $h : \mathcal{T} \to \tilde{\mathcal{T}}$ respecting the above data, i.e., $(h(0_k)) = (\tilde{0}_k)$ and h preserves the cyclic order of local branches.

EXERCISE 25.49. *Show that such an h is a Thurston equivalence, i.e., it conjugates F to $\tilde{F}$ up to homotopy rel* $\mathbf{0}$.

Similarly to Proposition 25.25, we have:

PROPOSITION 25.50. *The valuable point v is a tip. The center* 0 *has valence at most two.*

Thus, 0 is is not a branch point for $\mathcal{T}$: It is either a tip or a a regular point. In the latter case, 0 is a *folding point* for F, i.e., F is two-to-one covering of a punctured neighborhood of 0 over a one-sided punctured neighborhood of v.

[3]Here we mark with tilde objects associated with $\tilde{F}$.

We can now carry out an *abstract* construction of the *extended Hubbard trees* $\mathcal{I}$, whose geometric counterpart was discussed in §25.6.6. It is obtained by attaching to $\mathcal{T}$ the *abstract spine* $\sigma = [\beta, \beta']$.

25.6.13. *Markov* (*revisited*) *and piecewice linear models.* Like for an actual quadratic map on its Hubbard tree, we can naturally associate to an abstract Hubbard map $F : \mathcal{T} \to \mathcal{T}$ a Markov matrix $A \equiv A_{\mathcal{T}}$ and the corresponding Markov chain σ_A. Obviously, A is independent of the particular map F in the Thurston class.

Exercise 25.51. *For a prime Hubbard tree we have:*
(i) *In the $\alpha-$rotational case, the Markov matrix A is a cyclic permutation.*
(ii) *In the non-rotational case, the Markov matrix A is primitive.*

Incorporating the last assertion into the Perron-Forbenius Theorem, we conclude that in the prime non-rotational case the Markov matrix A has a unique leading eignevalue $\lambda > 1$ to which corresponds a unique (up to scaling) positive eigenvector.

Exercise 25.52. *The matrix A has leading eigenvalue* 1 *iff all the little Hubbard trees* (25.5) *are satellite. In this case, the corresponding Markov graph* (see §19.19) *is a graph of a map.*

We say that under these circumstances, the Hubbard tree $\mathcal{T}$ (and the corresponding quadratic polynomial) is of *molecule* type.

Notice next that a Thurston class can be represented by a *piecewise linear* map. For instance, realize each Markov tile J_s (i.e., the closure of a component of $\mathcal{T} \setminus (\mathbf{b} \cup \mathbf{0})$) as the unit interval and make the map linear on each of these components.

Exercise 25.53. *Show that the Markov chain σ_A is nicely* (*in what sense*?) *semi-conjugate to this piecewise linear map.*

The Perron-Frobenius theory allows us to make a better choice of a piecewise linear model. Namley, let us realize the I_s as the intervals of length l_s that represent the leading eigenvector for the transpose to the Markov matrix $A = (a_{ts})$:

$$\sum l_t \cdot a_{ts} = \lambda \cdot l_s.$$

On this tree the Hubbard map admits a piecewise linear model *with the constant slope* $\lambda > 1$:

Exercise 25.54. *Any prime non-rotational superattracting Thurston class can be modeled by a unique piecewise linear leo map with constant slope $\lambda > 1$.*

25.6.14. *Real case.* In the case when the Hubbard tree is an interval (which corresponds to real polynomials), we obtain a *saw-like model* for a superattracting unimodal map:

Exercise 25.55. *Let f be a superattracting real quadratic map. Then:*
(i) *The Hubbard tree $\mathcal{T}$ is the interval $[v, v_1] \equiv [0_1, 0_2]$ containing all postcritical points.*
(ii) *The topological Markov chain σ_A is naturally semi-conjugate to the dynamics on $\mathcal{T} \setminus \mathcal{D}(\mathbf{0})$. Modify the construction so that the dynamics on $\mathcal{T} \setminus \mathcal{D}(\mathbf{0})$ is actually conjugate to a Markov chain σ_B.*

(iii) *In the prime $\alpha-$rotational case, A is a cyclic permutation matrix of order* 2.

(iv) *In the prime non-rotational case, A is primitive and f on $\mathcal{T}$ is naturally modeled by a leo saw-like map with constant slope $\lambda > 1$.*

25.7. Characteristic rays and Topological Model for a superattracting map.

25.7.1. *Characteristic rays, strips, and associated objects.* We keep assuming that f is a superattracting quadratic polynomial. As the Julia set $\mathcal{J}(f)$ is locally connected (Theorem 25.4), it can be modeled by a geodesic lamination $\mathcal{L} \equiv \mathcal{L}_f$ in the disk (see §9.4). Next, we will give an explicit description of this lamination.

Let us consider the immediate basin $\mathcal{D}_1 \equiv \mathcal{D}(v)$ of the critical value, and let β_1 be its *root*, i.e., the fixed point of the return map $f^p : \mathcal{D}_1 \to \mathcal{D}_1$. Proposition 25.25 (ii) implies that $\{\beta_1\} = \mathcal{T} \cap \mathcal{D}_1$ (since $\mathcal{T} \cap \mathcal{D}_1$ is a single point invariant under f^p). In particular, β_1 lies in the arc $(0, v)$ of the Hubbard tree, so it is a cut-point for the Julia set $\mathcal{J}$. It allows us to apply the theory of *ray configurations* developed in §24.5.

Let $\mathfrak{q} \geq 2$ be the number of external rays landing at β_1. These rays divide the complex plane into $\mathfrak{q}$ sectors. By Proposition 25.25 (ii), the sector S_{ch} that contains the basin $\mathcal{D}_1 \ni v$ is the *innermost valuable* sector, so it is *characteristic* in the sense of §24.5.1. According to our convention, all associated objects are also called "characteristic" or "valuable":

- the *characteristic rays* $\mathcal{R}^{\pm}_{\text{ch}} = \mathcal{R}^{\theta_\pm}$ that bound S_{ch}, as well as their angles,
$$0 < \theta_- < \theta_+ < 1;$$
- the *characteristic cut-line* $L_{\text{ch}} := \mathcal{R}^+_{\text{ch}} \cup \mathcal{R}^-_{\text{ch}} \cup \{\beta_1\}$; and
- the *characteristic geodesic* (or *leaf*)[4] $\gamma^{\text{ch}} \equiv \gamma^{\text{ch}}(f)$ in $\mathbb{D} \approx \mathbb{H}^2$ connecting the *characteristic angles* θ_- and θ_+ (which is also referred to as a *minor leaf*).

Lemma 25.56. *The critical value $v = c$ is the only point of the cycle $\mathbf{0}$ contained in the characteristic sector S_{ch}.*

Proof. Assume there exists another postcritical point 0_n, $1 < n < p$, in S_{ch}. Let $\gamma \subset \mathcal{K}$ be the legal path connecting 0_n to the root β_1. Since β_1 is the only intersection point of $\mathcal{K}$ and L_{ch}, γ is contained in $\overline{S}_{\text{ch}}$. Then it ends with the internal radius of $\mathcal{D}_1$ connecting v to β_1, for otherwise γ would be separated from $\mathcal{D}_1$ by a ray landing at β_1 (while by definition, S_{ch} does not contain such rays). It follows that $v \in \operatorname{int} \gamma$ contradicting the property that v is a tip of the Hubbard tree. □

The characteristic leaf L_{ch} lifts by f to two symmetric *critical leaves*, L_0 and L'_0, respectively passing through the root β_0 and the co-root β'_0 of $\mathcal{D}_0$. They bound an (open) topological strip Π_0 containing $\mathcal{D}_0$. Moreover, $f : \Pi_0 \to S_{\text{ch}}$ is a double branched covering, and 0 is the only point of $\mathbf{0}$ contained in Π_0 (by Lemma 25.56).

Pulling Π_0 further along $\mathbf{0}$, we obtain strips $\Pi_n \supset D_n$, $n = 1, \dots, p-1$ with the properties that 0_n is the only point of $\mathbf{0}$ contained in Π_n and f^{p-n} univalently maps Π_n onto Π_0. In particular, we obtain the *characteristic strip* $\Pi_{\text{ch}} \equiv \Pi_1 \supset \mathcal{D}_1$ contained in S_{ch}. Moreover, $f^p : \Pi_{\text{ch}} \to S_{\text{ch}}$ is a double branched covering.

The characteristic strip slices two arcs, I^{ch}_1 and I^{ch}_2, at the circle at infinity, each of which is homeomorphically mapped by T^p onto the characteristic arc J_{ch}

[4]We will sometimes refer to a cut-line as a "leaf" as well.

sliced by $S_{\rm ch}$. This puts us into the context of Lemma 19.60 yielding a T^p-invariant Cantor set $K \subset \bar{I}_1^{\rm ch} \cup \bar{I}_2^{\rm ch}$ and a Devil K-Staircase semiconjugacy $h : K \to \mathbb{T}$ between $T^p | K$ and the doubling map T. This Staircase has a clear dynamical meaning:

LEMMA 25.57. *Under the above circumstances, the Cantor set K consists of the angles $\theta \in \mathbb{R}/\mathbb{Z}$ whose rays $\mathcal{R}^\theta$ land on $\partial\mathcal{D}_1$. The natural projection $K \to \partial\mathcal{D}_1$ is a Devil K-Staircase semi-conjugating $T^p | K$ to $f^p | \partial\mathcal{D}_1$ (while the latter is conjugate to the doubling map by the Riemann uniformization).*

25.7.2. *Topological model.* Let us lift $\gamma_{\rm ch}$ to a pair of 0-symmetric geodesics in $\mathbb{D}$, γ_0 and γ_0', that separate $\gamma_{\rm ch}$ from $1 \in \mathbb{T}$. Let σ be the diameter of $\mathbb{D}$ connecting $e(\theta_+/2)$ to $-e(\theta_+/2)$. Now, let us pull the geodesics γ_0 and γ_0' further back under the iterated doubling map $T : \mathbb{T} \to \mathbb{T}$ (see Appendix, §32.5.1). We say that a geodesic $\gamma_{\rm ch}$ in $\mathbb{D}$ *generates* a geodesic lamination $\mathcal{L}$ if all its pullbacks under the iterated doubling map T are pairwise disjoint, and their closure is equal to $\mathcal{L}$.

THEOREM 25.58. *For a superattracting quadratic polynomial f, the characteristic geodesic $\gamma_{\rm ch}$ generates the lamination $\mathcal{L}_f$ that gives a model for the Julia set $\mathcal{J}$ of f as a quotient of $\mathbb{T}$. The dynamics on $\mathcal{J}$ is modeled by the quotient of the doubling map $T : \mathbb{T} \to \mathbb{T}$ (with the inverse Böttcher map $B^{-1} : \mathbb{T} \to \mathcal{J}$ semi-conjugating T to f). Accordinglty, the filled Julia set $\mathcal{K}$ admits the pinched disk model corresponding to $\mathcal{L}_f$.*

PROOF. Let us consider a leaf L of the lamination $\mathcal{L}_f$ comprising a pair of rays, $\mathcal{R}_1$ and $\mathcal{R}_2$, landing at the same cut point $a \in \mathcal{T}$ and bounding a minimal sector S (so there are no other rays in this sector landing at a). By Lemma 25.27, a can be approximated by components D_{-k} of the basin $\mathcal{D}(\mathbf{0})$, preimages of the immediate basin $\mathcal{D}_0$. There are two possibilities:

(i) $a \in \partial D_{-k}$ for some component $D_{-k} \subset S$. Iterating forward, we can assume that $a \in \partial\mathcal{D}_v$. But since the critical value v is not a branch point of $\mathcal{T}$ (Proposition 25.25), the immediate basin $\mathcal{D}_v$ intersects the Hubbard tree $\mathcal{T}$ at a single point, its root. But then L itself is the characteristic leaf $L_{\rm ch}$.

(ii) There is a sequence of components $D_{-k} \subset S$ of order $n_k \to \infty$ converging to a along the Hubbard tree. Let L_{-k} be the corresponding preimages of the characteristic leaf $L_{\rm ch}$ landing at the roots of D_{-k}. As $k \to \infty$, these curves converge to some curve $L_{-\infty}$ comprising two rays in $\overline{S}$ landing at a. But L is the only such a curve, so $L_{-\infty} = L$.

In either case, L is approximated by preimages of the leaf $L_{\rm ch}$. By Proposition 24.15, so is any leaf of the lamination $\mathcal{L}_f$. □

The geodesic γ_0[5] divides the circle $\mathbb{T}$ into two semi-circles $\mathbb{T}_\pm$ producing a coding of angles $\theta \in \mathbb{T}$ by binary sequences $\bar{\varepsilon}$.

EXERCISE 25.59. *Two angles are equivalent in the above lamination model, $\theta \underset{\mathcal{K}}{\sim} \theta'$, iff they have the same itinerary: $\bar{\varepsilon} = \bar{\varepsilon}'$.*

REMARK 25.60. A more general discussion of combinatorial models for Julia sets will be given in §32. In particular, the above lamination should be compared with $\mathcal{L}_\theta$ from §32.1.2.

[5] This choice is arbitrary: the geodesic γ_0', can be selected instead.

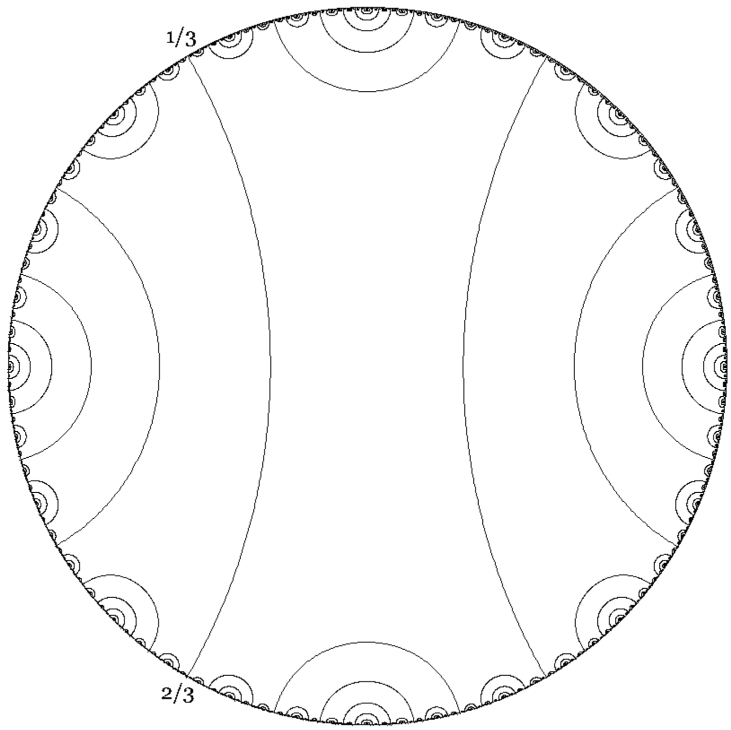

FIGURE 25.1. Geodesic lamination for the Basilica.

PROPOSITION 25.61. *For a hyperbolic quadratic polynomial f, the characteristic angles $\theta_\pm$ determine the Hubbard tree $\mathcal{T}$, and the other way around.*

PROOF. By Theorem 25.58, the characteristic angles determine the topological model for the filled Julia set, which in turn, determines the Hubbard tree $\mathcal{T}$.

Vice versa, from the Hubbard tree $\mathcal{T}$, we can recover the extended Hubbard tree $\mathcal{I}$ (Exercise 25.34). The orbit $f^n(\mathcal{R}^\pm_{\rm ch})$, $n = 0, 1, \dots, l-1$, of each characteristic ray can be encoded by a binary sequence $(\varepsilon_n)_{n=0}^{l-1}$ according to whether $f^n(\mathcal{R}^\pm_{\rm ch})$ lies above or below the spine σ. Extending this sequence periodically (to the right), we obtain the dyadic expansion of the corresponding angle $\theta_\pm$. $\square$

EXERCISE 25.62. *Describe the Basilica and Douady Rabbit laminations depicted on* Figures 25.1 *and* 25.2. *Relate then to the corresponding Julia sets depicted on* Figures20.4 *and* 20.5. *Show that each of these laminations comprises only pullbacks of the characteristic leaves* (*i.e., no non-trivial accummulation leaves exist*).

PROBLEM 25.63. (i) *Show that the last property of* Exercise 25.62 *holds if and only if the corresponding hyperbolic map is of prime satellite type.*
(ii) *Show that the laminatoin comprises countably many leaves iff the corresponding hyperbolic map is of molecule type.*

PROBLEM 25.64. *Let $\beta_k = f^{k-1}(\beta_{\rm ch})$ be the cycle of the characteristic repelling point $\beta_{\rm ch} \in \partial\mathcal{D}_v$, $k = 1, \dots, p$. Then there exists a Green puzzle piece P_k around*

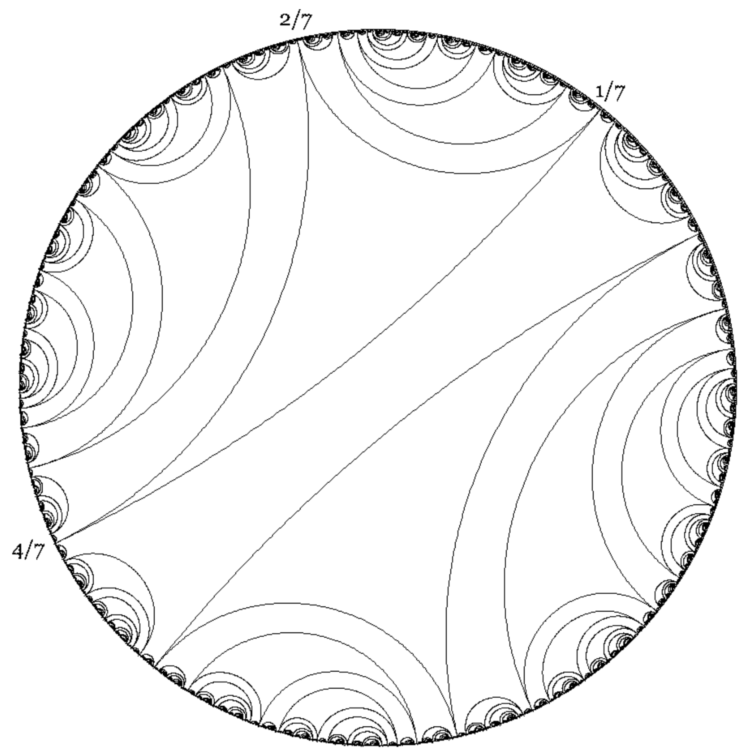

FIGURE 25.2. Geodesic lamination for the Douady Rabbit.

each point β_k whose external rays are preimages of the characteristic rays and such that f^p univalently maps P_k onto a strictly bigger disk $f^p(P_k) \Supset P_k$.

25.7.3. *Markov partition of the Julia set.* Let us consider a 0-symmetric dipole cut-line L passing through the periodic point α_0 and its symmetric $\alpha_0' \equiv -\alpha_0$ which is composed of two external rays $\mathcal{R}^{\theta_\mp/2}$ (landing at $\pm\alpha_0$ respectively) and the legal arc in $\overline{\mathcal{D}}_0$ connecting $\pm\alpha_0$. Let L_{-n}, $n = 0, \dots, p-1$, be the pullback of L under the branch of f^{-n} that maps $\mathcal{D}_0$ to $\mathcal{D}_{-n}$ (where p is the period of the characteristic rays).

EXERCISE 25.65. *The cut-lines L_{-n} do not cross each other*[6] *and cut the Julia set into pieces that form a Markov partition. Moreover, the corresponding Markov coding is one-to one, except over the iterated preimages of α_0. Over the latter, it is* q-*to*-1*, where* q *is the number of rays landing at α.*

25.8. Attracting-superattracting surgery. We will now describe a *qc surgery* that turns a hyperbolic polynomial $f \equiv f_c$ to a superattracting one, $f_\circ$. It will be our first encounter with this method; many more are to come. See §29.1 below for a general idea of the method.

[6]they can touch, though

25.8.1. *Surgery.* By Corollary 25.20, there is a homeomorphism

$$h_0 : (\overline{\mathcal{D}}_0, 0) \to (\overline{\mathbb{D}}, 0) \tag{25.6}$$

that admits a qc extension to the whole plane and conjugates the return map $f^p|\,\partial\mathcal{D}_0$ to the doubling map $f_0 : z \mapsto z^2$ on the unit circle $\mathbb{T}$. Any other component D of the basin $\mathcal{D}_f$ is univalently mapped onto $\mathcal{D}_0$ by some f^n with $n = n_D > 0$, so we can mark the "center" $0_D \in D$ as the preimage of 0 by f^n. We can also mark the "root" $\beta_D \in \partial D$, the preimage by f^n of the root β_0 of $\mathcal{D}_0$ (which is the fixed point of the return map $f^p : \partial\mathcal{D}_0 \to \partial\mathcal{D}_0$). For any $D \neq \mathcal{D}_0$, let us consider the Riemann mapping

$$h_D : (\overline{D}, 0_D, \beta_D) \to (\overline{\mathbb{D}}, 0, 1).$$

Then

$$h_D = h_{f(D)} \circ f \tag{25.7}$$

since both maps are the Riemann mappings $\overline{D} \to \overline{\mathbb{D}}$ normalized in the same way. For any component $\mathcal{D}_k$ of the immediate basin, $k = 0, \dots, p-1$, we let $0_k \equiv 0_{\mathcal{D}_k}$ and $\beta_k \equiv \beta_{\mathcal{D}_k}$.

Notice also that

$$f|\,\partial\mathcal{D}_0 = h_1^{-1} \circ (f_0|\,\mathbb{T}) \circ h_0. \tag{25.8}$$

Indeed, using notation $\prod^{\circ}$ for the composition read from the right to the left, we have:

$$f_0|\,\mathbb{T} = h_0 \circ (f^p|\,\partial\mathcal{D}_0) \circ h_0^{-1} = \prod_{0\le k\le p-1}^{\circ} h_{k+1} \circ (f|\,\partial\mathcal{D}_k) \circ h_k^{-1} = h_1 \circ (f|\,\partial\mathcal{D}_0) \circ h_0^{-1},$$

where the last equality follows form (25.7)

Let us now replace $f : \mathcal{D}_0 \to \mathcal{D}_1$ with $F \equiv F_f := h_1^{-1} \circ (f_0|\mathbb{D}) \circ h_0$. Formula (25.8) shows that F matches with f on $\partial\mathcal{D}_0$. Hence by letting $F = f$ on the complement of $\mathcal{D}_0$, we obtain a global double branched covering $\mathbb{C} \to \mathbb{C}$. Moreover,

$$F^p|\mathcal{D}_0 = h_0^{-1} \circ (f_0|\,\mathbb{D}) \circ h_0 \tag{25.9}$$

since $F^p|\mathcal{D}_0$ is equal to

$$f^{k-1} \circ (F|\,\mathcal{D}_0) = \left(\prod_{1\le k\le p-1}^{\circ} h_{k+1}^{-1} \circ h_k \right) \circ\, h_1^{-1} \circ (f_0|\,\mathbb{D}) \circ h_0 = h_0^{-1} \circ (f_0|\,\mathbb{D}) \circ h_0.$$

So, the return map to $\mathcal{D}_0$ is topologically superattracting, with the superattracting cycle $\mathbf{0} = (0_k)_{k=0}^{p-1}$.

Notice finally that the superattracting basin $\mathcal{D}_F(\mathbf{0})$ coincides with the attracting basin $\mathcal{D}_f(\boldsymbol{\alpha})$, and the set of the centers 0_D is completely invariant under F.

25.8.2. *Superattracting model.*

PROPOSITION 25.66. *For a hyperbolic quadratic polynomial $f \equiv f_c$, let $F \equiv F_f$ be the quasiregular map obtained from f by the above surgery. Then F is conjugate to a superattracting polynomial $f_\circ \equiv f_{c_\circ}$ by a qc map h coinciding with the Böttcher conjugacy on the basin of ∞.*

Proof. We will follow the notation from the above surgery construction. Let us consider a conformal structure $\nu_0 := h_0^*(\sigma)$ on $\mathcal{D}_0$. Since $h_0 : \mathcal{D}_0 \to \mathbb{D}$ conjugates the return map $F^p|\,\mathcal{D}_0$ to $f_0|\,\mathbb{D}$, where the latter is holomorphic, the structure ν_0 is F^p-invariant.

Let us spread ν_0 around the whole basin $\mathcal{D} \equiv \mathcal{D}_F(\mathbf{0})$ by the dynamics. Namely, any other component D of the basin $\mathcal{D}$, is univalently mapped onto $\mathcal{D}_0$ under some iterate of $F^n = f^n$, $n = n_D \in \mathbb{Z}_+$. Let $\nu_D = (F^n)^*(\nu_0) = (f^n)^*(\nu_0)$ be the corresponding conformal structure on D. Since f is holomorphic, ν_D has the same dilatation as ν_0.

Putting these structures together, we obtain a conformal structure ν with bounded dilatation on the whole basin $\mathcal{D}$. Moreover, it is invariant under F. Indeed, by construction, $\nu_D = F^*(\nu_{F(D)})$ for any component $D \neq \mathcal{D}_0$, while $\nu_0 = F^*(\nu_1)$ by the F^p-invariance of ν_0 (where $\nu_1 \equiv \nu_{\mathcal{D}_1}$).

Let us now extend ν to the whole plane $\mathbb{C}$ by letting $\nu = \sigma$ outside $\mathcal{D}$. Since $F = f$ is holomorphic outside $\mathcal{D}$, we obtain an invariant conformal structure with bounded dilatation on the whole plane $\mathbb{C}$.

By the Measurable Riemann Mapping Theorem, there is a qc map $h : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ such that $\nu = h^*(\sigma)$. Moreover, h is conformal on $\mathcal{D}_f(\infty)$. Normalized to be tangent to id at ∞, h conjugates F to a superattracting polynomial $f_\circ \equiv f_{c_\circ}$. $\square$

25.8.3. *Attracting Hubbard tree and characteristic rays.* With the set of centers $0_D \in \mathcal{D}_F(\mathbf{0}) = \mathcal{D}_f(\mathbf{0})$ in hands, we can define legal arcs, the spine and the Hubbard tree $\mathcal{T} \equiv \mathcal{T}_F \equiv \mathcal{T}_f$ in the same way as for a superattracting quadratic polynomial. We call $\mathcal{T}$ an *attracting Hubbard tree.* It is a subset of $\mathcal{K}(F) = \mathcal{K}(f)$ which is invariant under the quasiregular map F but not under the map f itself.

Since F is topologically conjugate to the superattracting model $f_\circ$, all the properties from §25.6.1 (which are purely topological) are valid for $\mathcal{T}_f$ as well. They can also be easily proved directly in the f-plane, and in particular, we suggest to the reader to check directly the analogue of Lemma 25.27:

Exercise 25.67. *Prove without using the superattracting model $f_\circ$ that the basin $\mathcal{D} \cap \mathcal{T}$ is dense in the attracting Hubbard tree $\mathcal{T}$.*

We also see that the root β_1 of $\mathcal{D}_1 \ni v$ is a cut point for $\mathcal{J}_c$, and we can define the *characteristic* (*valuable*) *rays* $\mathcal{R}^\pm \equiv \mathcal{R}^{\theta_\pm}$ landing at β_1 for f as in the superattracting case. Since the conjugacy between F and $f_\circ$ is Böttcher on the basin of ∞, the characteristic angles $\theta_\pm$ for the map f and its superattracting model $f_\circ$ are the same.

25.9. Real hyperbolic polynomials. The results of this section are an easy adaptation of the above complex results and arguments to the real symmetric setting, so we will present them as a series of exercises.

Let us consider a real-symmetric quadratic polynomial $f = f_c : x \mapsto x^2 + c$ with $c \in [-2, 1/4] \equiv \mathcal{M}_\mathbb{R}$. It has an invariant interval $\mathcal{I} \equiv \mathcal{I}_c = [-\beta, \beta]$, where $\beta \equiv \beta_c$ is the fixed point with positive multiplier. The map f is called (*real*) *hyperbolic* if it has a real attracting cycle $\boldsymbol{\alpha} = (\alpha_n)_{n=0}^{p-1}$. Our remark in §21.2.4 shows that this notion is consistent with the complex one: *A real quadratic polynomial $f_c : \mathcal{I} \to \mathcal{I}$ with $c \in \mathcal{M}_\mathbb{R}$ is real hyperbolic iff its complexification $f_c : \mathbb{C} \to \mathbb{C}$ is complex hyperbolic.*

There is one noteworthy difference between the real and complex situations. In the complex case all simply attracting germs are locally topologically equivalent, while in the real case, there is an obvious topological invariant: the *sign of the multiplier* (see §19.8.1).

Let $\mathcal{J}_{\mathbb{R}}(f) := \mathcal{I} \smallsetminus \mathcal{D}(\boldsymbol{\alpha}) = \mathcal{J}(f) \cap \mathbb{R}$ be the "real Julia set". Since $\mathcal{J}_{\mathbb{R}}(f) \subset \mathcal{J}(f)$, the dynamics on $\mathcal{J}_{\mathbb{R}}(f)$ is expanding.

Exercise 25.68. *For a real hyperbolic quadratic map f_c, $c \in [-2, 1/4]$, the real Julia set $\mathcal{J}_{\mathbb{R}}(f) \subset \mathcal{I}$ is nowhere dense, and in fact, uniformly porous in all scales. Thus, $\omega(x) = \boldsymbol{\alpha}$ for an open dense set of $x \in \mathcal{I}$ of full measure.*

Let us now consider the superattracting case:

Exercise 25.69. *Let f_c be a superattracting real polynomial, so $f^p(0) = 0$. Then:*

(i) *the Hubbard tree $\mathcal{T}_f$ is the interval $[v, v_1]$ containing all the points*

$$0_k, \quad k = 0, 1, \dots, p-1.$$

(ii) *The dynamics on $\mathcal{J}(f) \cap \mathcal{T}_f = \mathcal{J}_{\mathbb{R}}(f) \cap \mathcal{T}_f$ is conjugate to a Topological Markov Chain.*

(iii) *In the prime case, this Markov chain is primitive for $p > 2$ and a singleton for $p = 2$ (rotational case).*

In fact, the last statement is valid for general hyperbolic maps:

Exercise 25.70. *Let f be a hyperbolic real quadratic polynomial. Then the dynamics on $\mathcal{J}_{\mathbb{R}}(f) \cap \mathcal{T}_f$ is conjugate to a topological Markov chain* (*with the same specifications as in the previous section*).

Exercise 25.71. (i) *For real hyperbolic maps, describe nest of little Hubbard trees* (25.5) *in terms of the kneading theory;*

(ii) *Describe the corresponding real Julia set.*

Note finally that in the real hyperbolic case, the attracting-superattracting surgery can be done in the $\mathbb{R}$-symmetric way (beginning with an $\mathbb{R}$-symmetric qc homeomorphism h_0 (25.6)). Then the superattracting quasiregular map F is $\mathbb{R}$-symmetric as well, making the real Julia sets $\mathcal{J}_{\mathbb{R}}(f)$ and $\mathcal{J}_{\mathbb{R}}(f_\circ)$ (from Proposition 25.66) topologically equivalent. In particular, this remark reduces Exercise 25.70 to 25.69.

26. Parabolic maps

Parabolic maps are quite similar to hyperbolic ones, and can be well controlled. The reason is that the place where the map loses hyperbolicity is precisely localized: it is the parabolic cycle itself. In this section, we will summarize basic property of parabolic maps emphasizing differences with hyperbolic ones, while leaving various details to the reader.

Let us start with the simplest and most important example:

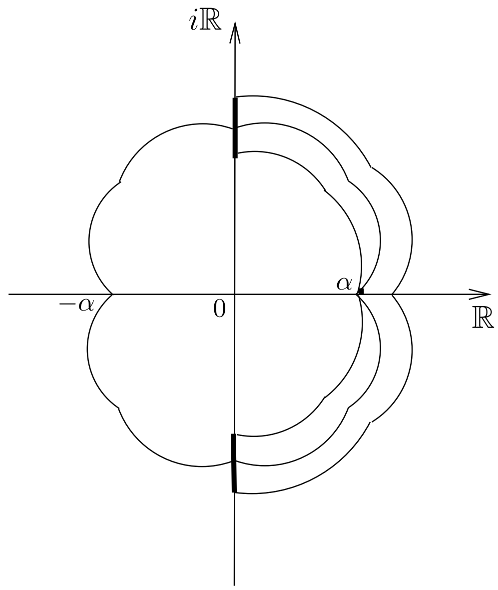

FIGURE 26.1. Domain P.

26.1. Cauliflower. Recall that this is the name for the quadratic polynomial $f \equiv f_{1/4} : z \mapsto z^2 + 1/4$ (and its Julia set depicted on Figure 20.3) corresponding to the cusp $c = 1/4$ of the Mandelbrot set (see §33.3 below). This is the only quadratic map that has a parabolic fixed point (at 1/2) where two fixed points, α and β, merge. We will use both notations for this fixed point. (Informally, when it is viewed from *inside* of the cauliflower, it is α, while viewed from *outside*, it becomes β.)

26.1.1. *Jordan property.*

THEOREM 26.1. *The cauliflower Julia set $\mathcal{J} \equiv \mathcal{J}_{1/4}$ is a Jordan curve.*

It happens that 1/4 is a real parameter, so the cauliflower map is $\mathbb{R}$-symmetric, which helps to analyze its dynamics. In particular, it follows from the dynamical description on $\mathbb{R}$ (Exercise 20.14) that $\mathcal{J} \cap \mathbb{R} = \{\pm\alpha\}$. Taking one more preimage, we obtain: $\mathcal{J} \cap i\mathbb{R} = \pm i\sqrt{3}/2$, where the latter are two preimages of $-\alpha$.

EXERCISE 26.2. *There exists an (open) $\mathbb{R}$-symmetric smooth rectangle P as depicted on* Figure 26.1 *whose boundary intersects $\mathcal{J}$ at three points, α and $\pm i\sqrt{3}/2$ and which is invariant under the inverse branch of f^{-1} fixing α.*

LEMMA 26.3. *We have: $\bigcap f^{-n}(\overline{P}) = \{\alpha\}$, where f^{-n} are the iterates of the inverse branch from* Exercise 26.2.

PROOF. The family of inverse branches f^{-n} is normal on P. Since $f^{-n}x \to \alpha$ for $x \in P \cap \mathbb{R}$, $f^{-n}z \to \alpha$ for all $z \in P$, locally uniformly on P. Since $f(\overline{P} \smallsetminus \{\alpha\}) \subset P$, $f^{-n}z \to \alpha$ locally uniformly in a neighborhood of $\overline{P} \smallsetminus \{\alpha\}$. The convergence is also locally uniform on $\overline{P}$ near α, since a relative neighborhood of α in $\overline{P}$ is contained in a uniformly repelling petal. Hence $f^{-n}z \to \alpha$ uniformly on $\overline{P}$, and the conclusion follows. $\square$

COROLLARY 26.4. *The cauliflower is weakly locally connected at* α.

PROOF. It follows from Exercise 9.16 that $\mathcal{J} \cap \overline{P}$ is connected. Hence so are all $\mathcal{J} \cap \overline{P}_n$ where $P_n := f^{-n}(P)$ (with the above inverse branches). Note P_n can be included into an open Jordan disk $Q_n \ni \alpha$ such that $\overline{Q}_n \cap \mathcal{J} = \overline{P}_n \cap \mathcal{J}$ and $\operatorname{diam} Q_n \to 0$ (just "thicken" P_n slightly near α), which implies the desired. $\square$

PROPOSITION 26.5. *The cauliflower* $\mathcal{J}$ *is locally connected.*

PROOF. It is enough to prove that $\mathcal{J}$ is weakly locally connected at any point. Since it is true at α, it is also true at all iterated preimages of α. For any other point $z \in \mathcal{J}$, there is a sequence $n_k \to \infty$ such that $f^{n_k} z \to \zeta \neq \alpha$. Let us consider the exterior dyadic grid $\Delta^n_{\bar{i}}$ in $\mathbb{C} \smallsetminus \overline{\mathbb{D}}$ from §23.6.2 and transfer it by the Böttcher map to $\mathbb{C} \smallsetminus \mathcal{K}$. Call the corresponding exterior dyadic tiles $Q^n_{\bar{i}}$. Lemma 26.3 implies that there is a tyle $\overline{Q} \equiv \overline{Q}^n_{\bar{i}} \ni \zeta$ disjoint from α. Then all the inverse branches $f^{-n}_{\bar{i}}$ are well defined in a neighborhood of $\overline{Q}$. By the Shrinking Lemma, $\operatorname{diam} f^{-n_k}(\overline{Q}) \to 0$ so the exterior dyadic tiles around z shrink. It follows that the inverse Böttcher function $B^{-1} : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus \mathcal{K}$ extedns continuously to the Julia set $\mathcal{J}$, and we are done. $\square$

Now, by Proposition 24.16, all cut-points of the cauliflower $\mathcal{J}$ belong to $\mathcal{S}k^\circ = \bigcup f^{-n}(-\beta, \beta)$. But this interval is contained in the Fatou set, so there are no cut points at all. As a hull without cut-points must be a Jordan disk, this completes the proof of Theorem 26.1.

Note, however, that *cauliflower is not a quasidisk* as it contains a petal at α with the cusp. This geometrically distinguishes the cauliflower from hyperbolic maps f_c with an attracting fixed point (i.e., for c *inside* the *main cardioid*, see §33.3 below).

26.1.2. *Dyadic dynamics on the Julia set.* Proposition 26.5, together with Theorem 23.25 and the Conformal Schönflies Theorem, imply that the Böttcher function extends to a homeomorphism $B : \mathcal{J} \to \mathbb{T}$ conjugating $f \equiv f_{1/4}$ to the doubling map $T : \mathbb{T} \to \mathbb{T}$. This produces a dyadic coding of $\mathcal{J}$, which can be also described directly as follows.

Let us consider the fixed and co-fixed points for f, $\beta = 1/2$ and $\beta' = -1/2$. They partiton the Jordan curve $\mathcal{J}$ into the upper and the lower arcs I_0 and I_1. Each of these arcs is almost homeomorohically mapped onto the whole curve $\mathcal{J}$ (i.e., its interior is mapped homeomorphically onto the punctured curve $\mathcal{J} \smallsetminus \{\beta\}$). Thus, they form a Bernoulli partition of $\mathcal{J}$ that leads to a coding $h : \mathcal{J} \to \mathbb{T}$ semi-conjugating f to T (compare §19.13.2). Moreover, the proof of Proposition 26.5 (or continuity of $B^{-1} : \mathbb{T} \to \mathcal{J}$) shows that the corresponding dyadic intervsls shrink, implying that h is a homeomorphism (obviously, coinciding with the extension of B).

We let $\beta^0 \equiv \beta$, $\beta^1 \equiv \beta'$, and for $n \geq 2$, we let $\beta^n_{i_1 \dots i_{n-1}}$, $i_k \in \{0, 1\}$, be the f^n-preimages of β dynamically labeled according to the partition of $\mathcal{J}$ into the upper and lower arcs, $\mathcal{J}_0 = \mathcal{J} \cap \overline{\mathbb{H}}_+$, $\mathcal{J}_1 = \mathcal{J} \cap \overline{\mathbb{H}}_-$. These points correspond to the dyadic points $m/2^n$ of order n on $\mathbb{T}$ (m is odd), so we will refer to them as *dyadic points of order* n *on* $\mathcal{J}$ (where β and β' are the dyadic points of order 0 and 1, respectively).

EXERCISE 26.6. *In the above discussion, the dynamical coding* $\beta^n_{i_1 \dots i_{n-1}}$ *corresponds to the binary expansion* $[i_1 \dots i_{n-1} 1]$ *of* $m/2^n$.

Let us summarize the above discussion:

PROPOSITION 26.7. *The cauliflower dynamics on the Julia set, $f : \mathcal{J} \to \mathcal{J}$, is topologically conjugate to the doubling map $T : \mathbb{T} \to \mathbb{T}$.*

COROLLARY 26.8. *The cauliflower dynamics on the Julia set is expansive.*

26.1.3. *Cauliflower checkerboard.* Let us now descrie a dynamical checkerboard tesselation of the cauliflower that helps to visualize its dynamics. Let us tessellate the interior of the filled Julia set, $\mathcal{K}^\circ \equiv \operatorname{int} \mathcal{K}_{1/4}$ into two symmetric tiles,

$$T_0^0 := \mathcal{K}^\circ \cap \overline{\mathbb{H}}_+ \quad \text{and} \quad T_1^0 := \mathcal{K}^\circ \cap \overline{\mathbb{H}}_-$$

cut off by the invariant horizontal interval $\gamma^0 := [\beta', \beta] = \mathcal{K} \cap \mathbb{R}$. They form an unbranched Bernoulli tiling of $\mathcal{K}^\circ$ (see §19.11.3): each $\operatorname{int} T_i^0$ is univalently mapped by f onto

$$\mathcal{K}^\circ \smallsetminus [v, \beta) \supset \operatorname{int} T_0^0 \cup \operatorname{int} T_1^0, \quad \text{where } v = 1/4, \beta = 1/2.$$

Taking the dynamical pullbacks of this tiling, we obtian a nest of tilings $\mathcal{T}^n$ of $\mathcal{K}^\circ$ by dynamical dyadic tiles $T^n_{i_0 \dots i_{n-1}}$, $i_k \in \{0, 1\}$. The limit $\mathcal{T}^\infty$ of these tilings form a tesselation of $\mathcal{K}^\circ$ by tiles $T^\infty_{\bar{i}}$, where $\bar{i} \in \Sigma_2$ ends with $\bar{0} := (000\dots)$ or $\bar{1} := (111\dots)$ (see Figure 23.5 and Exercise 26.10 below).

It is instructive to look closer at several first steps of this process. For instance, the tesselation $\mathcal{T}^1$ is obtained by cutting $\mathcal{K}^\circ$ into four symmetric tiles by the horizontal and vertical intervals, γ^0 and $\gamma^1 := \mathcal{K}^\circ \cap i\mathbb{R}$. The full preimage $\Gamma^2 := f^{-1}(\gamma^1)$ comprises two proper arcs $\gamma_i^2 \subset T_i^0$ in $\mathcal{K}^\circ$, $i \in \{0, 1\}$, crossing γ^1 at precritical points c_i^1, respectively. The arc γ_i^2 lands at the consecutive dyadic points $\beta^3_{ij} \in \mathcal{J}$, $j \in \{0, 1\}$, of order 3. By cutting the previous four tiles by these arcs, we obtain eight tiles of the tessellation $\mathcal{T}^2$.

In general, let $\Gamma^n = f^{-n}(\gamma_0)$. It is a finite lamination in $\mathcal{K}^\circ$ comprising 2^{n-1} proper arcs $\gamma^n_{i_1 \dots i_{n-1}}$, $i_k \in \{0, 1\}$, landing at pairs of consecutive dyadic points on $\mathcal{J}$ of order $n+1$. Altogether, they form a 4-valent tree Γ that tessellates $\mathcal{K}^\circ$ into limiting checkerboard tiles $T^\infty_{\bar{i}}$.[7]

Two most prominent tiles of this tessellation, $T_{\bar{0}} \equiv T_{00\dots}$ and $T_{\bar{1}} \equiv T_{11\dots}$, attached to the real interval $[0, \beta]$, are f-invariant. Color them black and white. All others are univalent pullbacks of these: they inherit the corresponding coloring from their parents, turning the tessellation into a checkerboard.

EXERCISE 26.9. *Justify inductively the above description, labeling the objects dyadically so that:*

(i) *For $n \geq 2$, the arc $\gamma^n_{i_1 \dots i_{n-1}}$, $i_k \in \{0, 1\}$, of the lamination Γ^n lands at the pair of two consecutive dyadic points $\beta^{n+1}_{i_1 \dots i_{n-1} i_n} \in \mathcal{J}$ of order $n+1$, $i_n \in \{0, 1\}$.*

(ii) *$\gamma^n_{i_1 \dots i_{n-1}}$, $n \geq 2$, contains two precritical points of order n, which can be labeled $c^n_{i_1 \dots i_{n-1} i_n}$, $i_n \in \{0, 1\}$.*

(iii) *$\gamma^{n+1}_{i_1 \dots i_{n-1} i_n}$ crosses $\gamma^n_{i_1 \dots i_{n-1}}$ at the precritcal point $c^n_{i_1 \dots i_{n-1} i_n}$, and it does not cross any other arcs of level $\leq n+1$. Thus, each arc $\gamma^n_{i_1 \dots i_{n-1}}$ of order $n \geq 1$ intersects only three arcs of the tree Γ.*

[7] In what follows, "$\gamma^n_{i_1 \dots i_{n-1}}$ for $n = 1$" will mean "γ^1".

EXERCISE 26.10. (i) *The invariant tile* $T_{\bar{0}}^{\infty}$ *has vertices* $\beta, c^0 \equiv 0, c_0^1, c_0^2, \dots$ (*countered clockwise*). *Its closure intersects the Julia set* $\mathcal{J}$ *at a single point,* β. *Similarly, for the invariant tile* $T_{\bar{1}}^{\infty}$.
(ii) *Two tiles,* $T_{i_1\dots i_{n-1}1\bar{0}}^{\infty}$ *and* $T_{i_1\dots i_{n-1}0\bar{1}}$ *are attached to the dyadic point* $\beta_{i_1\dots i_{n-1}}^n$ *in* $\mathcal{J}$.
(iii) *List the edges of any tile as above.*

There is a quick recepie for labeling our tree Γ. Let us call 0 the *root* of Γ and orient Γ from the root towards the Julia set. Let $e_0^1 \subset \gamma_0^1$ be the edge of Γ that starts at the root and gows up, while $e_1^1 \subset \gamma_1^1$ be the edge that goes down. They end at the vertices c_0^1 and c_1^1 respectively. If we turn right from a vertex c_i^1 we label the corresponding edge as e_{i0}^2, while we label it e_{i1}^2 if we turn left. Then the edge e_{ij}^2 ends at the vertex c_{ij}^2. By turning left or right at these vertices, we obtain edges e_{ijk}^3 ending at the vertices c_{ijk}^3. And so on. In this way we will label all the *compact edges*[8] of Γ as $e_{i_1\dots i_n}^n$, and will recover the labeling of all the vertices $c_{i_1\dots i_n}^n$.

It us instructive to look at the above picture in another coordinate system:

EXERCISE 26.11. (i) *Show that the cauliflower map is conformally conjugate to the map* $z \mapsto z + 1/z + 1$.
(ii) *Depict the cauliflower tessellation in this coordinate.*

26.2. Parabolic Blaschke model for the return map. Let now $f \equiv f_c$ be a parabolic quadratic polynomial, let $\boldsymbol{\alpha} = (\alpha_k \equiv f^k \alpha)_{k=0}^{p-1}$ be its parabolic cycle with rotation number $\theta = \mathfrak{p}/\mathfrak{q}$ and multiplier $\rho = e(\theta)$, and let $\mathcal{D} \equiv \mathcal{D}_0$ be the component of the immediate basin $\mathcal{D}(\boldsymbol{\alpha})$ containing 0 (see Theorem 21.25). We can label points α_k so that $\alpha \in \partial\mathcal{D}$. Note that the Leau-Fatou flower attached to α comprises $\mathfrak{q}$ petals (see Corollary 21.26) and the period of $\mathcal{D}$ is equal to $p\mathfrak{q}$.

26.2.1. *Parabolic Blaschke model.* In the parabolic case, it is more convenient to uniformize the basin $\mathcal{D} \equiv \mathcal{D}(\alpha)$ by the upper half-plane model sending α to ∞. Let
$$\phi : (\overline{\mathcal{D}}, 0, \alpha) \to (\overline{\mathbb{H}}_+, i, \infty)$$
be such an uniformization, and let
$$g = \phi \circ f^{p\mathfrak{q}} \circ \phi^{-1} : \overline{\mathbb{H}}_+ \to \overline{\mathbb{H}}_+$$
be the corresponding model for the first return map to $\mathcal{D}$. Exercise 3.3 implies:

EXERCISE 26.12. *The uniformization* ϕ *can be normalized so that*
$$g(z) = z - \frac{1}{z}. \tag{26.1}$$

We will refer to the map g as the *parabolic Blaschke map.* It is an example of a *parabolic rational map* (with the Julia set $\mathbb{T}$).

LEMMA 26.13. *The parabolic Blaschke map* g (26.1) *has the following properties:*
(i) $g^n z \to \infty$ *in* $\mathbb{H}_{\pm}$;
(ii) $g : \mathbb{T} \to \mathbb{T}$ *is topologically conjugate to the doubling map* T (*and hence is expansive*).

[8] Notices that we refrained from going straight, keeping non-compact edges unlabeled: the reader is welcome to create his/her own labeling system for those ones.

In case of the cauliflower map $f \equiv f_{1/4}$, the parabolic basin $\mathcal{D}(\alpha)$ is a Jordan domain on which f is conformally conjugate to the parabolic Blaschke map g, while the basin of infinity $\mathcal{D}(\infty) = \hat{\mathbb{C}} \smallsetminus \overline{\mathcal{D}}(\alpha)$ is a Jordan disk on which f is conformally conjugate to the square map $f_0 : \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}} \to \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$. The gluing between these two models is fulfilled by a circle homeomorphism $h : \mathbb{T} \to \mathbb{T}$ conjugating $g \,|\, \mathbb{T}$ to $f_0 \,|\, \mathbb{T}$. Thus, we can say that f is the *mating* between the parabolic Blaschke map on $\overline{\mathbb{D}}$ and the square map on $\hat{\mathbb{C}} \smallsetminus \mathbb{D}$ by means of the homeomorphism h. However, as the cauliflower is not a quasidisk, this mating is not quasiconformal (see also Exercise 21.19).

26.2.2. *Parabolic spine.* Notice that the imaginary axis $i\mathbb{R}_+$ is invariant under the parabolic Blaschke map g, contains the critical point (i), and lands at the parabolic point (∞) and its preimage (0). The corresponding arc Γ is f-invariant and contains the critical point 0. Moreover, it lands at the parabolic point $\alpha \in \partial\mathcal{D}$ and the co-parabolic point $\alpha' \equiv -\alpha$. Indeed, let γ be the arc of Γ bounded by 0 and v. Then $f^n(\gamma) \to \alpha$ by definition of the parabolic basin. So, the "positive end" of Γ lands at α. By symmetry, the "negative end" lands at α'.

Hence the closure of Γ if a closed invariant arc containing the critical point. It is called the *parabolic spine* for f:

$$\sigma^{\mathrm{par}} \equiv \sigma_f^{\mathrm{par}} := \mathrm{cl}(\phi^{-1}(i\mathbb{R}_+)) = \phi^{-1}(i\mathbb{R}_+) \cup \{\pm\alpha\}.$$

Notice that for the cauliflower map $f_{1/4}$, the parabolic spine coincides with

$$\mathcal{I} = [-1/2, 1/2] \equiv [-\beta, \beta].$$

26.2.3. *Blaschke maps of* 2nd *type.* We have seen two types of degree two Blaschke maps maps $g : \mathbb{D} \to \mathbb{D}$: with an attracting fixed point inside $\mathbb{D}$, and maps $g : \mathbb{H}_+ \to \mathbb{H}_+$ with a parabolic point on $\partial\mathbb{H}_+$. There are two more type of Blaschke maps:

Problem 26.14. *Show that the map*

$$g : (\hat{\mathbb{C}}, \mathbb{H}_+) \to (\hat{\mathbb{C}}, \mathbb{H}_+), \quad g : z \mapsto z + 1 - \frac{1}{z},$$

is a double branched covering of $\mathbb{H}_+$ over itself. It has a parabolic fixed point at ∞ such that $f^n z \to \infty$ without landing at ∞ for all $z \in \hat{\mathbb{C}}$ except a Cantor set $\mathcal{J} \subset \hat{\mathbb{R}}$. The restriction $g|\,\mathcal{J}$ is an expanding map topologically conjugate to the Bernoulli shift σ_2. (Of course, $\mathcal{J}$ is the Julia set of this rational map.)

Problem 26.15. *Work out a similar problem for Blaschke maps*

$$g : (\hat{\mathbb{C}}, \mathbb{H}_+) \to (\hat{\mathbb{C}}, \mathbb{H}_+), \quad g : z \mapsto \lambda z - \frac{1}{z}, \quad \text{where } \lambda > 1,$$

with an attracting fixed point at ∞.

The dynamical structure of Blaschke maps can be compared with the structure of Fuchsian groups described in §2.4.8.

26.3. Local connectivity of general parabolic Julia sets.

Exercise 26.16. *There is a shrinking nest of Green puzzle pieces $(P_k^n)_{n\in\mathbb{N}}$ around each parabolic point α_k, $k = 0, \dots, p-1$.*

Theorem 26.17. *The Julia set $\mathcal{J}(f)$ of a parabolic quadratic polynomial is locally connected.*

PROOF. Local connectivity at each parabolic point α_k follows from Exercise 26.16. The rest of the argument follows the lines of the Cauliflower case (Lemma 26.5). □

As in the hyperbolic case, we immediately conclude:

COROLLARY 26.18. *For a parabolic map, the inverse Böttcher map*

$$B^{-1} : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus \mathcal{K}$$

extends to a continuous map $\mathbb{T} \to \mathcal{J}$ (denoted in the same way) semi-conjugating the doubling map T to $f|\,\mathcal{J}$. Moreover, $B^{-1}(e(\theta)) \in \mathcal{J}$ is the landing point of the external ray $\mathcal{R}^\theta$.

COROLLARY 26.19. *Given a parabolic map f, let D_i be the components of* $\operatorname{int}\mathcal{K}$ *(arbitrary labeled). Then any component D_i of* $\operatorname{int}\mathcal{K}$ *is a Jordan disk, and* $\operatorname{diam} D_i \to 0$.

As in the cauliflower case, the return map $f^{\mathfrak{pq}} : \overline{\mathcal{D}}_0 \to \overline{\mathcal{D}}_0$ admits the parabolic Blaschke model:

COROLLARY 26.20. *The Riemann mapping $\phi : \overline{\mathcal{D}}_0 \to \overline{\mathbb{D}}$ can be normalized so that $g := \phi \circ f^{\mathfrak{pq}} \circ \phi^{-1} : \overline{\mathbb{H}}_+ \to \overline{\mathbb{H}}_+$ becomes the parabolic Blaschke map* (26.1).

COROLLARY 26.21. *The conformal map $\psi : (\overline{\mathcal{D}}_0, 0, \beta_0) \to (\mathcal{K}_{1/4}, 0, 1/2)$ conjugates the return map $f^{\mathfrak{pq}}|\,\mathcal{D}_0$ to the cauliflower map $f_{1/4}|\,\mathcal{K}_{1/4}$.*

26.4. Bounded distortion and quasi-self-similarity.

26.4.1. *Bounded distortion.* The following self-similarity property allows us to control the geometry of parabolic maps at arbitrary small scales near almost all points:

LEMMA 26.22. *For any parabolic polynomial f, there exists an $\varepsilon > 0$ such that for any point $z \in \mathcal{J}(f)$ whose orbit does not land in the parabolic cycle $\boldsymbol{\alpha}$, there exists an infinite sequence of moments n_k such that the inverse branches*

$$f^{-n_k} : (\Delta_k, z_{n_k}) \to (D_k, z)$$

are well defined in the disk $\Delta_k := \mathbb{D}(z_{n_k}, \varepsilon)$ (where $z_n \equiv f^n z$), have an absolutely bounded distortion:

$$\frac{|Df^{-n_k}(\zeta)|}{|Df^{-n_k}(z_{n_k})|} \asymp 1 \quad \forall\, \zeta \in \Delta_k,$$

and $D_k := f^{-n_k}(\Delta_k)$ are shrinking ovals of absolutely bounded shape around z.

PROOF. The local picture of the Leau-Fatou flower (see Proposition 21.27) implies existence of $\varepsilon_0 > 0$ with the property that for any point $z \in \mathcal{J}(f)$ whose orbit does not land in the parabolic cycle $\boldsymbol{\alpha}$, there exists an infinite sequence of moments n_k such that $\operatorname{dist}(z_{n_k}, \boldsymbol{\alpha}) \geq \varepsilon_0$. Since $f^n(0) \to \boldsymbol{\alpha}$, it follows that $\operatorname{dist}(z_{n_k}, \overline{\mathcal{P}}_f) \geq 2\varepsilon > 0$ for some $\varepsilon > 0$. Hence the desired inverse branches of f^{-n_k} are well defined in $\mathbb{D}(z_{n_k}, 2\varepsilon)$, and the Koebe Distortion Theorem completes the proof. □

LEMMA 26.23. (i) *For any $\varepsilon > 0$, the first transit map through the ε-neighborhood of the parabolic cycle $\boldsymbol{\alpha}$ is expanding on the Julia set.*

(ii) *The Julia set is uniformly porous at all points in all scales.*

Proof. (i) Combine the above lemma with the Shrinking Lemma.

(ii) Porousity at the parabolic points follows from the Leau-Fatou Flower picture. For any other point $z \in \mathcal{J}$ and any sufficiently small scale $\varepsilon > 0$, take the first moment n such that the outer radius of $\Delta_n := f^n(D(z,\varepsilon)$ centered at z_n is at least ε_0 or the first moment n when Δ_n contains a parabolic point (whichever happens first. If Δ_n does not contain parabolic points, then the disk $D(z_n, \varepsilon_0)$ (where the Julia set is uniformly porous) can be univalently and with a bounded distortion pulled back to z, implying the desired. If Δ_n contains a parabolic point, take one pullback of it, $\Delta_{n-1} \ni z_{n-1}$. The Julia set is still porous at this oval while it does not anymore contain parabolic points, so the previous argument can be applied to it. □

26.5. Area of parabolic Julia set. Applying Lemma 26.23 and the Lebesgue Density Points Theorem, or Theorem 22.2 and Proposition 21.27, we obtain:

Proposition 26.24. *For any parabolic quadratic polynomial f,* area $\mathcal{J}(f) = 0$.

26.6. Parabolic Hubbard tree. We will now attach a Hubbard tree $\mathcal{T}_f$ to any parabolic quadratic polynomial f. To this end we will turn it, by means of a surgery, to a topological superattracting map F. This surgery is similar to the attracting-superattracting surgery from §25.8, except that it is not quasiconformal this time.

26.6.1. *Parabolic-attracting surgery.* By Proposition 26.12 and Lemma 26.13, the return map $f^{p\mathfrak{q}}$ to the boundary of the immediate basin $\mathcal{D}_0$ is topologically conjugate to the doubling map $f_0 : z \mapsto z^2$ on the unit circle $\mathbb{T}$. Let us extend this conjugacy to a homeomorphism

$$h_0 : (\overline{\mathcal{D}}_0, 0) \to (\overline{\mathbb{D}}, 0). \tag{26.2}$$

Using this homeomorphism, we can do a surgery exactly as in §25.8 turning f to a topological double branched covering F coinciding with f on $\mathbb{C} \smallsetminus \mathcal{D}_0$ with a superattracting periodic point of period p at 0.

Exercise 26.25. *Work out details.*

We let $\mathbf{a} := (F^k 0)_{k=0}^{p-1}$ be the superattracting cycle of F. Notice that $\mathcal{D}_F(\mathbf{a}) = \mathcal{D}_f(\boldsymbol{\alpha})$. Below we will refer to this basin as $\mathcal{D}$.

We will see below that there is a hyperbolic quadratic polynomial $\tilde{f}$ which is topologically conjugate to F (see Exercise 35.28). However, this conjugacy cannot be quasiconformal (by Exercise 21.16(ii))). In the primitive case one can say even more:

Exercise 26.26. *If the parabolic cycle under consideration has multiplier* 1*, then there are no qc homeomorphisms $H : (\mathbb{C}, \mathcal{J}(f)) \to (\mathbb{C}, \mathcal{J}(\tilde{f}))$.*

26.6.2. *Parabolic Hubbard tree.* With the set of centers c_D in hands, we can define legal arcs, the spine and the Hubbard tree $\mathcal{T}_F$ for F in the usual way. We call $\mathcal{T}_f \equiv \mathcal{T}_F$ the *parabolic Hubbard tree.*

The first properties from §25.6.1 readily extend to this case. It is noteworthy that the *root* of the parabolic component $\mathcal{D}_0 \ni 0$ is our parabolic periodic point $\alpha \equiv \alpha_0$. Indeed, it is the only fixed point for the return map $F^{p\mathfrak{q}} : \overline{\mathcal{D}}_0 \to \overline{\mathcal{D}}_0$

(coinciding with $f^{p\mathfrak{q}}$ on the boundary $\partial\mathcal{D}_0$). Accordingly, the root of $\mathcal{D}_1 \ni v$ is $\alpha_1 \equiv f(\alpha)$.

Less straightforward is the parabolic version of Lemma 26.27:

LEMMA 26.27. *The basin $\mathcal{D} \cap \mathcal{T}$ is dense in the model Hubbard tree $\mathcal{T}$ of a parabolic map f.*

PROOF. Let us consider the postcritical set $\overline{\mathcal{P}} \equiv \overline{\mathcal{P}}_f = \operatorname{orb} 0 \cup \boldsymbol{\alpha}$ and the associated hyperbolic metric on $\mathbb{C} \smallsetminus \overline{\mathcal{P}}$. As in Lemma 26.27, it induces a hyperbolic distance on each component of $\mathcal{T} \smallsetminus \overline{\mathcal{D}}^{\bullet}$. The map f expands this metric, and the expansion is uniform away from $\boldsymbol{\alpha} \subset \overline{\mathcal{D}}^{\bullet}$.

For an arc $L \subset \mathcal{T} \smallsetminus \overline{\mathcal{D}^{\bullet}}$, we let $d_{\text{hyp}}(L)$ stand for the hyperbolic distance (25.3) between the endpoints of L. The above expanding property implies that if there an arc L as above such that $d_{\text{hyp}}(L) \geq \varepsilon > 0$ then

$$d_{\text{hyp}}(f(L)) \geq \lambda \cdot d_{\text{hyp}}(L), \quad \text{where } \lambda = \lambda(\varepsilon) > 1.$$

Hence $d_{\text{hyp}}(f^n(L) \to \infty$.

Moreoover , there is a subsequence of momenbts $n_k \to \infty$ such that the intervals $f^{n_k}(L)$ stay a definite distance away from the postcritical set $\overline{\mathcal{P}}_f$. Indeed, in the primitive case the boundary of the basin $\mathcal{D}^{\bullet}$ is repulsive, while in the satellite case $\overline{\mathcal{P}} \Subset \overline{\mathcal{D}^{\bullet}}$.

It follows that the Euclidean length of the intervals $f^n(L)$ goes to ∞ as well, which is of course impossible. □

By definition, the parabolic Hubbard tree $\mathcal{T}$ is invariant under the quasiregular superattracting map F. In fact, with a bit of extra care, it can be also made invariant under the original parabolic map f. Recall that the initial conjugacy h_0 (26.2) is the composition of the conformal isomorphism $\phi : (\overline{\mathcal{D}}_0, \sigma_0) \to (\mathcal{K}_{1/4}, \mathcal{I}_{1/4})$ and a homeomorphism $H : \mathcal{K}_{1/4} \to \overline{\mathbb{D}}$ conjugating $f_{1/4}|\,\mathcal{J}_{1/4}$ to $f_0|\,\mathbb{T}$. By selecting the latter $\mathbb{R}$-symmetric, we make the spine of F in $\mathcal{D}_0$ invariant under $f^{p\mathfrak{q}}$, implying the desired.

26.7. Characteristic rays and Topological Model for parabolic maps. We can now define all the *characteristic objects* for a parabolic map f: the characteristic rays $\mathcal{R}^{\pm}_{\text{ch}}$ landing at the parabolic root α_1 of the immediate component $\mathcal{D}_1 \ni v$, the characteristic sector $S_{\text{ch}} \supset \mathcal{D}_1$ bounded by these rays, and the characteristic geodesic γ_{ch} in $\mathbb{D}$. It leads to a topological model for f on $\mathcal{D}(\infty) \cup \mathcal{J}$ in the same way as in the hyperbolic case:

THEOREM 26.28. *For a parabolic quadratic polynomial f, the characteristic geodesic γ_{ch} generates the lamination $\mathcal{L}_f$ that gives a model for the Julia set $\mathcal{J}$ of f as a quotient of $\mathbb{T}$. The dynamics on $\mathcal{J}$ is modeled by the corresponding quotient of the doubling map $f_0 : \mathbb{T} \to \mathbb{T}$ (with the inverse Böttcher map $B^{-1} : \mathbb{T} \to \mathcal{J}$ semi-conjugating f_0 to f). Accordinglty, the filled Julia set $\mathcal{K}$ admits the pinched disk model corresponding to $\mathcal{L}_f$.*

In particular, we obtain:

COROLLARY 26.29. *Let f be the quadratic polynomial that has a parabolic fixed point with multiplier $e(\mathfrak{p}/\mathfrak{q})$. Then the Julia set $\mathcal{J}$ is modeled by the quotient of the doubling map $f_0 : \mathbb{T} \to \mathbb{T}$ by the lamination $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ generated by the characteristic geodesic of the rotation set $\Theta_{\mathfrak{p}/\mathfrak{q}} \subset \mathbb{T}$. The filled Julia set $\mathcal{K}$ is modeled by the corresponding pinched disk.*

26.8. Real parabolic maps. Let us now continue the discussion of real parabolic maps that began in §21.3.7. In this case, we define the *real Julia set* as the complement to the parabolic basin, $\mathcal{J}_{\mathbb{R}}(f) := \mathcal{I} \smallsetminus \mathcal{D}(\boldsymbol{\alpha})$. Exercise 21.29 implies that in the primitive parabolic case, $\mathcal{J}_{\mathbb{R}} = \mathcal{J} \cap \mathbb{R}$, while in the satellite parabolic case, $\mathcal{J}_{\mathbb{R}} = (\mathcal{J} \cap \mathbb{R}) \smallsetminus \mathrm{Orb}_{-}(\boldsymbol{\alpha})$.

EXERCISE 26.30. *Let $f : \mathcal{I} \to \mathcal{I}$ be a real parabolic quadratic polynomial. Then:*
(i) *The dynamics on $\mathcal{J}_{\mathbb{R}} \cap [v, v_1]$ is conjugate to a Topological Markov Chain.*
(ii) length $\mathcal{J}_{\mathbb{R}} = 0$*, so the orbits of almost all points $x \in \mathcal{I}$ converge to the parabolic cycle.*

Note in conclusion that the parabolic-attracting surgery can be done in the $\mathbb{R}$-symmetric way showing that the real parabolic Julia sets $\mathcal{J}_{\mathbb{R}}$ is topologically equivalent to the corresponding hyperbolic one.

27. Other special classes

27.1. Critically preperiodic (Misiurewicz) maps.

27.1.1. *First observations.* A quadratic polynomial f is called (*critically*) *preperiodic* or *Misiurewicz* if 0 itself is not periodic but $\mathrm{orb}\, 0$ lands in some cycle $\boldsymbol{\alpha}$ (i.e., $f^n(0) \in \boldsymbol{\alpha}$ for some $n \in \mathbb{Z}_+$). In this case, $\boldsymbol{\alpha}$ is called the *postcritical cycle*.

The simplest and most popular example is the Chebyshev map $z \mapsto z^2 - 2$. The next one is $z \mapsto z^2 + i$ (see Figure 20.6).

Postcritically preperiodic maps are the closest relatives of hyperbolic maps, and good part of the theory can be obtained by adapting the corresponding hyperbolic ideas.

EXERCISE 27.1. *For a preperiodic map:*
(i) *All cycles are repelling;*
(ii) *The filled Julia set has empty interior* (*and thus,* $\mathcal{J} = \mathcal{K} \ni 0$)*;*
(iii) *The Julia set has zero area.*

27.1.2. *Associated orbifold.* In fact, any preperiodic map f is expanding with respect to some metric that has finitely many cone singularities, so f can be viewed as an *orbifold expanding map.*[9]

As a model, let us take one more glance at the Chebyshev map $T : z \mapsto 2z^2 - 1$ (affinely equivalent to $f_{-2} : z \mapsto z^2 - 2$). It satisfies the functional equation

$$T(\mathrm{Cos}\, z) = \mathrm{Cos}\, 2z, \text{ where } \mathrm{Cos}\, z := \cos 2\pi z.$$

The function $\mathrm{Cos} : (\mathbb{C}, \mathbb{Z}) \to (\mathbb{C}, \pm 1)$ is the orbifold universal covering over the orbifold with the underlying space $\mathbb{C}$ and two cone singularities of order 2 at points ± 1. The Euclidean metric on $\mathbb{C}$ pushes down to the flat orbifold metric on $\mathbb{C}$ with cone singularities at ± 1 with angle π.

The map T infinitesimally expands this metric by 2 at all points except 0 and ± 1 (where the differential of T can be interpreted only in the orbifold sense).

In general, to any preperiodic map f we can associate a *Riemann orbifold* $\mathcal{O}_f$ by assigning weight 2 to all post-valuable points $z \in \mathrm{orb}\, v$, and weight 1 to the rest.

[9]We refrain from calling f an "orbifold hyperbolic map" as it can suggest that the corresponding orbifold, rather than the map, is "hyperbolic".

LEMMA 27.2. *The orbifold $\mathcal{O} \equiv \mathcal{O}_f$ is hyperbolic for all critically preperiodic quadratic polynomials $f \equiv f_c$ except the Chebyshev map (for which $\mathcal{O}$ is parabolic). In the hyperbolic case, the multivalued inverse map f^{-1} lifts to a single-valued holomorphic map $G \equiv \hat{f}^{-1}$ of the universal covering $\hat{\mathcal{O}} \approx \mathbb{D}$. Moreover, this lift is non-invertible.*

$$\begin{tikzcd} \mathbb{D} \arrow[d, "\pi"'] & \mathbb{D} \arrow[l, "G"'] \arrow[d, "\pi"] \\ \mathbb{C} \arrow[r, "f"'] & \mathbb{C} \end{tikzcd}$$

PROOF. Let $k = |\overline{\mathcal{P}}_f|$ be the size of the post-valuable set. Then $k > 2$, unless f is Chebyshev. Hence the Euler characteristic of $\mathcal{O}$ is negative, $\chi(\mathcal{O}) = 1 - k/2 < 0$, so $\mathcal{O}$ is hyperbolic. Moreover, $\mathcal{O}$ is good by Theorem 2.93. Hence its universal covering is isomorphic to $\mathbb{D}$.

Furthermore, f^{-1} admits local lifts to $\mathbb{D}$ everywhere (Exercise 1.112). By the Monodromy Theorem, it admits a global lift G to $\mathbb{D}$. As G has critcal points (at the fiber points over $v = c$), it is non-invertible. $\square$

By the Schwarz Lemma, for any preperiodic quadratic f (except for Chebyshev) the inverse map $\hat{f}^{-1} : \mathbb{D} \to \mathbb{D}$ is contracting with respect to the hyperbolic metric. Hence $\|Df(z)\|_{\text{hyp}} > 1$ for any $z \in \mathbb{C} \smallsetminus \operatorname{orb} 0$, where the norm is taken with respect to the orbifold hyperbolic metric).

EXERCISE 27.3. *Show that for any $R > 0$ there exists $\rho > 1$ such that*

$$\|Df(z)\|_{\text{hyp}} \geq \rho \quad \textit{for all } z \in \mathbb{D}_R \smallsetminus \operatorname{orb} 0.$$

27.1.3. *$\mathcal{J}$ is a dendrite.*

THEOREM 27.4. *The Julia set of a Misiurewicz map is locally connected; hence it is a dendrite.*

PROOF. It follows the lines of the proof of Theorem 25.4 using the orbifold expansivity (Exercise 27.3). Namely, show that the inverse Böttcher map $B^{-1} : \mathbb{C} \smallsetminus \bar{\mathbb{D}} \to \mathbb{C} \smallsetminus \mathcal{J}$ is Hölder continuous in any annulus $\mathbb{A}(1, R)$, $R > 1$, from the Euclidean metric on the domain to the orbifold metric in the target. $\square$

EXERCISE 27.5. *Any two points in a Misiurewicz dendrite can be separated by a cut-line through some preimage of a postcritical periodic point.*

27.1.4. *Hubbard tree.* The Hubbard tree $\mathcal{T} \equiv \mathcal{T}_f$ for a preperiodic map f is defined in the same way as for a superattracting one, as the legal hull of $\mathbf{0} \equiv \operatorname{orb}(0)$ (which is a finite subset of $\mathcal{J}$). Note that since the filled Julia has empty interior, any two points of $\mathcal{J}$ can be connected by a unique arc in $\mathcal{J}$, so the adjective "legal" is actually redundant in this context. Let $\boldsymbol{\beta}$ be the set of branch points of $\mathcal{T}$. We mark on $\mathcal{T}$ the points of $\mathbf{0}$ and $\boldsymbol{\beta}$.

Similarly to the superattracting case, we have:

EXERCISE 27.6. (i) *The marked Hubbard tree $(\mathcal{T}, \mathbf{0} \cup \boldsymbol{\beta})$ is invariant under f; hence all branch points of $\mathcal{T}$ are (pre-)periodic; moreover, $f : \mathcal{T} \to \mathcal{T}$ is surjective.*
(ii) *The critical value $v = c$ is a tip of $\mathcal{T}$; the critical point 0 is a regular point.*

As in the superattracting case (§25.6.11), we say that the Hubbard tree $\mathcal{T}$ for a Misiurewicz map is *prime* if it does not contain a non-trivial strictly smaller subtree $\mathcal{T}' \subset \mathcal{T}$ containing 0 and invariant under some iterate f^p, $p > 1$.

EXERCISE 27.7. *Show that the following properties are equivalent:*

(i) *The Hubbard tree is prime;*

(ii) *The preimages of the $\alpha-$fixed point are dense in $\mathcal{T}$;*

(iii) *The dynamics on $\mathcal{T}$ is topologically exact* (*leo*)*.*

PROJECT 27.8. *Develop a theory of little Hubbard trees leading to an analogue of nest* (25.5)*:*

$$\mathcal{T} \equiv \mathcal{T}^0 \supset \mathcal{I}^1 \supset \mathcal{T}^1 \cdots \supset \mathcal{I}^n \supset \mathcal{T}^n, \tag{27.1}$$

where the return map $f^p : \mathcal{T}^n \to \mathcal{T}^n$ is topologically exact (*and with relarive exactness for intermediate trees*)*.*

If we puncture out all the points $0_k \in \mathbf{0}$ and $b_j \in \boldsymbol{\beta}$ from the Hubbard tree $\mathcal{T}$, the rest will be disjoint union of topological intervals J_s°. Their closures J_s form a tiling $\mathfrak{J}$ of $\mathcal{T}$.

PROBLEM 27.9. (i) *The tiling $\mathfrak{J}$ is Markov. In the prime case, the Markov matrix A is primitive.*

(ii) *It generates a natural semi-conjugacy $h : \Sigma_A \to \mathcal{T}$. This semi-conjugacy is one-to-one over all points of $\mathcal{T}$ except the points of $\mathbf{0} \cup \mathbf{b}$ and their iterated preimages.*

(iii) *If $\mathcal{T}$ is prime then $f \,|\, \mathcal{T}$ is conjugate to a piecewise-linear model with constant slope.*

(iv) *Translate tree structure* (27.1) *into a combinatorial structure of the matrix A.*

27.1.5. *Topological model.* Since the critical value $v = c$ is a preperiodic point for a Misiurewicz map f, there are finitely many *characteristic* (or *valuable*) rays $\mathcal{R}^i_{\mathrm{ch}} \equiv \mathcal{R}^{\theta_i}_{\mathrm{ch}}$ landing there. As in the hyperbolic case (§25.7), they generate a nice topological model for the map:

PROBLEM 27.10. *Let f be a Misiurewicz quadratic polynomial, and let $\mathfrak{R}$ be the periodic rays configuration landing at the postcritical cycle. It generates a completely invariant lamination $\mathcal{L}$ whose pinched disk model is topologically conjugate to $f \,|\, \mathcal{J}$. This lamination is determined by any characteristic angle θ_i as well as by the abstarct Hubbard tree of f.*

Let us now lift any characteristic ray $\mathcal{R}^i_{\mathrm{ch}}$ to a cut-line L_i through the critical point. Select any of them, $L \equiv L_i$. The corresponding geodesic $\gamma \equiv \gamma_i \subset \bar{\mathbb{D}}$ divides the circle $\mathbb{T}$ into two semi-circles $\mathbb{T}_\pm$ producing a coding of angles $\theta \in \mathbb{T}$ by binary sequences $\bar{\varepsilon}$.

EXERCISE 27.11. *Two angles are equivalent in the above lamination model, $\theta \underset{\mathcal{J}}{\sim} \theta'$, iff they have the same itinerary: $\bar{\varepsilon} = \bar{\varepsilon}'$.*

EXERCISE 27.12. *Describe the topological model for $z \mapsto z^2 + i$.*

27.1.6. *Real Misiurewicz maps.* Specifying the above discussion to the interval situation, we obtain:

EXERCISE 27.13. *Let f be a postcritically preperiodic real quadratic map. Then:*

(i) *The Hubbard tree $\mathcal{T}$ is the interval $[v, v_1]$ containing all postcritical points.*

(ii) *The dynamics on $\mathcal{T}$ is naturally semi-conjugate (in fact, "almost conjugate") to a Topological Markov Chain. In the prime case, the corresponding Markov matrix A is primitive.*

(iii) *If $\mathcal{T}$ is prime then $f\,|\,\mathcal{T}$ is a topologically exact map conjugate to a saw-like model with constant slope.*

In the interval situation, the liittle trees of the nest (27.1) are periodic intervals. Moreover, the return map $f^p : \mathcal{T}^n \to \mathcal{T}^n$ to the deepest interval is topologically exact. Let $\mathcal{A} := \bigcup_{k=0}^{p-1} f^k(\mathcal{T}^n)$. It is the measure-theoretic attractor for our map:

PROBLEM 27.14. (i) *Show that for a.e. $x \in \mathcal{I}$, there is an $n \in \mathbb{N}$ such that $f^n x \in \mathcal{A}$.*

(ii) $\omega(x) = \mathcal{A}$ *for a.e. x.*

(iii) *The restriction $f\,|\,\mathcal{A}$ is ergodic with respect to the Lebesgue measure.*

A general theory of attractors for real quadratic maps will be developed in §46.

THEOREM 27.15. *Any real postcritically preperiodic quadratic map $f : \mathcal{I} \to \mathcal{I}$ is stochastic: it has a unique acim $d\mu = \rho\, dm$. Moreover,* $\operatorname{supp}\mu = \mathcal{A}$ *and the density ρ is real analytic outside post-valuable points, i.e., on $\mathcal{I} \smallsetminus \overline{\mathcal{P}}$. At the post-valuable points, ρ has $(1/\sqrt{\cdot})$-singularities.*

PROOF. The proof is modeled on the expanding circle case (Theorem 19.76). Namely, formula (19.2) for the densities ρ_n of the push-forward measures implies the following analogue of (19.11):

$$\rho_n(y) = \sum_{x \in g^{-n}y} \frac{\varepsilon_n(x)}{Dg^n(x)} = \sum_{i=1}^{2^n} \delta_{n,i}(y) Dg_i^{-n}(y), \tag{27.2}$$

where the ε_n and $\delta_{n,i}$ are the signs of the corresponding derivatives. Moreover, all the branches g_i^{-n} are well defined on each Markov interval J_s and extend analytically to the upper and lower half-planes. By the Koebe Distortion Theorem, we obtain oscillation bounds (19.12) on any smaller subinterval $J'_s \Subset J_s$, and in fact, on the whole disk $\mathbb{D}(J'_s)$ based on J'_s. Together with the normalization condition, it implies uniform bounds, and hence normality, for the ρ_n on $D_n(J'_s)$.

Thus, we can take limits for Cezaro averages to obtain densities of invariant measures. What we need to check is that these limits are not identically vanishing on the intervals $\operatorname{int} J_s$. But otherwise the corresponding limiting measure would concentrate on the postcritical cycle $\boldsymbol{\alpha}$. So, we have to estimate the mass of the $\rho_n\, dm$ near $\boldsymbol{\alpha} = (\alpha_n)_{n=0}^{p-1}$:

EXERCISE 27.16. *Show that $\rho_n(y) \asymp 1/\sqrt{|y - \alpha_k|}$ near each point α_k.*

The conclusion follows. $\square$

See Theorem 46.20 for a more general criterion, supplied with a different proof.

Now, the Ergodic Theorem implies that the above acim $d\mu = \rho\, dx$ governs the behavior of almost all points:

COROLLARY 27.17. *For a Misiurewicz map f, almost all orbits are equidistributed with respect the the absolutely continuous invarinat measure μ, i.e., for a.e. $x \in \mathcal{I}$ and for any continuous test function $\phi \in \mathbb{C}(\mathcal{I})$ we have:*

$$\frac{1}{N}\sum_{n=0}^{N-1} \phi(f^n x) \to \int \phi(x)\rho(x)\, dx.$$

COROLLARY 27.18. *For a Misiurewicz map f, almost points $x \in \mathcal{I}$ have a* positive Lyapunov exponent*:*

$$\frac{1}{N}\sum_{n=0}^{N-1} \log|Df^n x| \to \int \log|Df(x)|\, \rho(x)\, dx > 0 \quad \text{as } N \to \infty. \tag{27.3}$$

PROOF. The $(1/\sqrt{\cdot})$−asymptotics of the density ρ near singularities implies that $\log|x|$ is intergable with respect ot μ. Hence the Ergodic Theorem is applicable to this function: for a.e. $x \in \mathcal{I}$, we have

$$\frac{1}{N}\sum_{n=0}^{N-1} \log|f^n x| \to \int \log|x|\, \rho(x)\, dx > 0 \quad \text{as } N \to \infty.$$

Since $\log|Df(x)| = \log|x| + \log 2$ for a quadratic map $f \equiv f_c$, the last limit translates into the existence of the Lyapunov exponent (27.3). Its positivity follows from the expanding property for the orbifold metric (or from the Shrinking Lemma). □

The last two Corollaries motivate the name of *stochastic*. Indeed, typical orbits of our maps are moving in a random fashion over the phase space (albeit governed by a measure) and enjoy exponentilly fast instability. The Chebyshev map was our first stochastic example. Misiurewicz maps provide us with a countable supply of stochastic parameters c_i in the quadratic family.

PROJECT 27.19. *Fill in omitted details in this section* (§27.1)*. Develop a more complete theory along the lines of the hyperbolic theory.*

27.2. Subhyperbolic maps. A quadratic polynomial f which is neither hyperbolic nor parabolic is called *subhyperbolic postcritically non-recurrent* if its critical point is non-recurrent: $0 \notin \omega(0)$.

THEOREM 27.20. *Let f be a subhyperbolic quadratic polynomial. Then $0 \in \mathcal{J}$ and the critical orbit eventually lands in some hyperbolic set $K \subset \mathcal{J}$. Moreover, the filled Julia set $\mathcal{K}$ has empty interior (so $\mathcal{K} = \mathcal{J}$), all periodic points of f are repelling, and* $\operatorname{area} \mathcal{J} = 0$.

LEMMA 27.21. *Under assumptions of Theorem 27.20, there exist an $\varepsilon_0 > 0$ such that any pullback $D \equiv D_0, D_{-1}, D_{-2}, \dots$ of any disk $D = \mathbb{D}(z, \varepsilon_0)$ hits the critical point at most once.*

PROOF. Let $\Delta \equiv \Delta(t) \supset \mathcal{K}$ be the subpotential domain of a small level $t > 0$. It is contained in an ε-neighborhood of the filled Julia set $\mathcal{K}$ with $\varepsilon = \varepsilon(t) \to 0$ as $t \to 0$. Endow Δ with the hyperbolic metric.

Let $d = d(\varepsilon)$ be the hyperbolic distance from 0 to $\operatorname{orb}(v)$. Since 0 is non-recurrent, $d > 0$.

Claim 1. $d(\varepsilon) \to \infty$ as $\varepsilon \to 0$.

Indeed, if $0 \in \mathcal{J}$ then 0 is ε-close to $\partial\Delta$, and the Claim follows from the blow-up property of the hyperbolic metric near the boundary (Lemma 7.7).

Assume $0 \in \mathcal{F}$, and let D be the component of $\mathcal{F}$ containing 0. By assumption, D is neither attracting nor parabolic. It cannot be a Siegel disk either since the latter does not contain critical points. By Classification of periodic components (Theorem 21.47), D is not periodic.

Hence all components $f^n(D)$, $n \in \mathbb{Z}_+$, are different from D. It follows that any path connecting 0 to $\operatorname{orb}(v)$ must cross $\mathcal{J}$ and hence must pass ε-close to some point of $\partial\Delta$. The conclusion follows again from Lemma 7.7.

Let us say that a disk $\mathbb{D}(z,\delta) \subset \Omega$ is *protected* if the twice bigger disk $\mathbb{D}(z, 2\delta)$ is still contained in Ω. For any $n \in \mathbb{N}$ and $C > 0$, let us consider the following properties

Property $\operatorname{Diam}_\varepsilon[C,n]$. For any protected disk $D \subset \Omega_\varepsilon$ and any pullback

$$D \equiv D_0, D_{-1}, \dots, D_{-n}, \tag{27.4}$$

all the disks D_{-k}, $k = 0, 1, \dots, n$, have hyperbolic diameter in Ω_ε bounded by C.

Property $\operatorname{Deg}_\varepsilon[n]$. Any pullback (27.4) as above hits the critical point at most once.

We will prove inductively the following:

Claim 2. There exist an absolute $C > 0$ and $\varepsilon = \varepsilon(f) > 0$ such that Properties $\operatorname{Diam}_\varepsilon[n, C]$ and $\operatorname{Deg}_\varepsilon[n]$ are satisfied for all $n \in \mathbb{N}$.

We will do it in two shots:

a) First, let us show that for any $C > 0$ there exists an $\varepsilon = \varepsilon(C, f) > 0$ such that

$$\operatorname{Diam}_\varepsilon[C, n-1] \implies \operatorname{Deg}_\varepsilon[n].$$

By Claim 1, we can select an $\varepsilon > 0$ so that $d(\varepsilon) > 2C$, so 0 stays $2C$ away from $\operatorname{orb} v$. On the other hand, by assumption

$$\operatorname{diam}_{\text{hyp}} D_{-k} < C \quad \text{for } k = 0, 1, \dots n-1.$$

So, if one of these disks, say D_{-m}, contains 0, then

$$D_{-m} \cap \operatorname{orb} v = \emptyset \quad \text{and} \quad D_{-m} \cap D_{-k} = \emptyset \text{ for } k = 0, \dots, m-1. \tag{27.5}$$

It follows from the latter that only one of the disks D_{-k} can contain 0. If non of them do, then $\operatorname{Deg}_\varepsilon[n]$ is obvious. If $D_{-(n-1)} \not\ni v$, then $D_{-n} \not\ni 0$, and the conclusion follows again. Finally, if $D_{-(n-1)} \ni v$ then $D_{-m} \ni f^{n-1-m}(v)$ contradicting (27.5).

b) Let us now show that there exists an absolute $C > 0$ such that

$$\forall\, \varepsilon > 0 \quad \operatorname{Deg}_\varepsilon[n] \implies \operatorname{Diam}_\varepsilon[C, n].$$

For any protected disk $D \equiv \mathbb{D}(z, \delta)$, all disks $\mathbb{D}(\zeta, \delta/2)$ centered at points $\zeta \in D$ are also protected. By the assumption, any pullback of order n of any such disk hits the critical point at most once. Hence the pullbacks of order n of the twice smaller disk $\mathbb{D}(\zeta, \delta/4)$ have an absolutely bounded hyperbolic diameter (by some absolute constant C_0).

Furthermore, the disk D can be covered by an absolute number N of such disks $\mathbb{D}(\zeta, \delta/4)$. Therefore, any pullback of D of order n can be covered by at most $2N$ pullback disks (where 2 is the degree bound of the pullback from Assumption $\mathrm{Deg}_\varepsilon[n]$), so its hyperbolic diameter is bounded by $2NC_0$. $\square$

COROLLARY 27.22. *Let f be a subhyperbolic postcritically non-recurrent polynomial. Then for any $\delta > 0$ there exists an $\varepsilon > 0$ such that any f^n-pullback D_{-n} of any disk $\mathbb{D}(z, \varepsilon)$ centered at $z \in \mathcal{J}$ satisfies:* $\operatorname{diam} D_{-n} < \delta$.

COROLLARY 27.23. *Let f be a subhyperbolic polynomial. Then any compact invariant set $K \subset \mathcal{J}$ that does not contain 0 is hyperbolic.*

REMARK 27.24. See Lemma 45.7 for a related statement in the Puzzle context.

PROOF. Let $\delta := \operatorname{dist}(0, K) > 0$, and then find ε from the previous Corollary. Then for any $z \in K$ and any $n \in \mathbb{N}$, the f^n-pullback $\Delta_n(z)$ of the disk $\mathbb{D}(f^n z, \varepsilon)$ to z is univalent. By the Shrinking Lemma, $\operatorname{diam} \Delta_n(z) \to 0$ as $n \to \infty$ uniformly in $z \in K$, implying that $|Df^n(z)| \to \infty$ as $n \to \infty$ uniformly in $z \in K$. $\square$

Proof of Theorem 27.20. Since our map is not hyperbolic, its Julia set is connected, so $0 \in \mathcal{K}$.

The main assertion of the Theorem (that $\operatorname{orb} 0$ lands in a hyperbolic set) follows from Corollary 27.23 applied to $K := \operatorname{cl}(\operatorname{orb} v)$.

Our map does not have attracting and parabolic cycles by assumption. It does not have Cremer points either. Indeed, by Proposition 21.39, a Cremer point would belong to $\omega(0)$, making it repelling by the first assertion.

For a similar reason, there are no Siegel disks. Indeed, if D is a Siegel disk then ∂D would be contained in $\omega(0)$ by Proposition 21.39, which is hyperbolic by the first assertion. But if $z \in D$ is sufficiently close to ∂D then the whole $\operatorname{orb} z$ stays close to ∂D. By hyperbolicity of ∂D, it would imply that

$$|Df^n(z)| \to \infty \quad \text{as } n \to \infty,$$

which cannot happen inside a Siegel disk.

We conclude that all periodic points of f are repelling, and consequently (by Classification Theorem 21.47), f does not have periodic bounded Fatou components.

Similarly, we can deal with wandering components using Lemma 22.1. By that Lemma, $\omega(z) \subset \omega(0)$ for any $z \in D$. Since $\omega(0)$ is hyperbolic, $|Df^n(z)| \to \infty$, provided $\operatorname{orb} z$ does not pass through 0. On the other hand, by (22.1), $|Df^n(z)| \to 0$ at any point $z \in D$ – contradiction.

It follows that $\operatorname{int} \mathcal{K} = \emptyset$, and hence $0 \in \mathcal{J}$.

Finally, let us show that $\operatorname{area} \mathcal{J} = 0$. Let

$$K = \{z : \operatorname{dist}(f^n z, \omega(0)) \le \frac{1}{2} \operatorname{dist}(0, \omega(0)), \quad n = 0, 1, \dots\}.$$

It is a compact invariant set that does not contain 0. By Corollary 27.23, it is hyperbolic, and hence has zero area (see Exercise 25.24). It follows that the set $\bigcup_{n=0}^{\infty} f^{-n}(K)$ has zero area as well. On the other hand, by Theorem 22.2, this set has full area in $\mathcal{J}$. $\square$

27.3. Cremer maps: wild Julia sets. Cremer maps provide us with first *wild* examples of non-locally-connected Julia sets:

PROPOSITION 27.25. *The Julia set $\mathcal{J}(f)$ of any Cremer quadratic polynomial f is not locally connected.*

PROOF. Without loss of generality we can assume that the Cremer point α is fixed. Let $\phi : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathcal{D}(\infty)$ be the Riemann uniformization of the basin of ∞. If the Julia set is locally connected then ϕ extends continuously to a surjective map $\mathbb{C} \smallsetminus \mathbb{D} \to \overline{\mathcal{D}(\infty)}$ (where $\overline{\mathcal{D}(\infty)} = \mathcal{D}(\infty) \cup \mathcal{J}$).

Let $X := \phi^{-1}(\alpha) \subset \mathbb{T}$, and let $\delta > 0$. By continuity of ϕ, any point $\zeta \in \overline{\mathcal{D}(\infty)}$ sufficiently near α has a ϕ-preimage $z \in \mathbb{C} \smallsetminus \mathbb{D}$ which is δ-close to X. In turn, the latter has an f_0-preimage $z_{-1} \in \mathbb{C} \smallsetminus \mathbb{D}$ which is $(\delta/2)$-close to X. Let $\zeta_{-1} := \phi(z_{-1})$. Then $f(\zeta_{-1}) = \zeta$, and by continuity of ϕ, ζ_{-1} is close to α. Thus, $\zeta_{-1} = f^{-1}(\zeta)$, where f^{-1} is the local inverse branch of f near α.

Taking further iterated preimages of z_{-1}, we construct a backward orbit $(z_{-n}) \subset \mathbb{C} \smallsetminus \mathbb{D}$ for f_0 whose projection $(\zeta_{-n}) := (\phi(z_{-n}))$ to $\mathbb{C}$ is the orbit of ζ under the local branch f^{-1}. It follows that α is Lyapunov stable for $f^{-1} \,|\, \overline{\mathcal{D}(\infty)}$.

Let us show that it is Lyapunov stable on a full ε-neighborhood of α.[10] As $\mathcal{J}$ is lc, $\operatorname{diam} D_k \to 0$, where D_k are components of $\operatorname{int} \mathcal{K}$ (see Exercise 1.33). Hence for all but finitely many D_k, we have:

$$D_k \cap \mathbb{D}(\alpha, \varepsilon/2) \neq \emptyset \implies D_k \Subset \mathbb{D}(\alpha, \varepsilon).$$

For such a D_k, we have the Lyapunov stability by the Maximum Principle. Which leaves us only finitely many “big” D_k to deal with. By shrinking ε if needed, we can assume that all these big domain D_k contain α on the boundary. Take one of these domain, $D \equiv D_k$. Assume D is (pre-)periodic; then without loss of generality it can be assumed to be periodic. Then by Theorem 21.47 it is contained in an attracting or parabolic basin, or is a Siegel disk. But the former two cases are excluded by Theorem 25.2 and Lemma 26.23, while in the latter case, there are no fixed points on ∂D (since in the lc case, the dynamics on the boundary of a Siegel disk is conjugate to an irrational rotation). Finally, if D is wandering then it has a “small” iterate $f^{-n}(D)$, and the conclusion follows again.

So, α is Lyapunov stable under f^{-1}. Hence f^{-1} is locally linearizable near α, and then so is f. Contradiction. □

REMARK 27.26. For another class of a wild Julia sets, see Example 45.21.

Notes to §§25–27

Local connectivity of hyperbolic, parabolic, and postcritically finite Julia sets was proved in the Orsay Notes. The general postcritically non-recurrent case (Theorem 27.20) is due to Mañé [**Mane**] (with further refinements and simplifications by Carleson-Jones-Yoccoz [**CJY**] and Shishikura & Tan Lei [**ShTL**]). Non-local-connectivity of Cremer Julia sets (Prop. 27.25) was observed independently by Douady & Sullivan (see [**Bl**, Theorem 10.14]) and the author (see [**L1**, Example 1.9]).

---

[10] For the reader who is willing to accept at this point the complete description of the dynamics on the Fatou set (Theorem 29.11 below), let us mention that $\operatorname{int} \mathcal{K} = \emptyset$ in the Cremer case, so no further considerations are needed.

First remarks on the area of Julia sets (for sufficiently expanding Cantor sets) were made by Fatou [**F2**, p. 42]. Propositions 25.23, 26.24 and Exercise 27.1 (inspired by Fatou's remarks) follow from [**L6**].

Makov partitions for hyperbolic Julia sets were constructed by Guckenheimer and Jakobson around 1970 [**Gu3, Ja2**].[11] It was revisited in [**Ja3, PS**] with the idea of a *coding tree*. Coding trees became a basis for a sophisticated algebraic theory of *Iterated Monodromy Groups* launched by Bartholdi, Grigorchuk, and Nekrashevich in the 2000s (see [**BGN, N**]).

Hubbard trees appeared in the Orsay Notes [**DH2**] (announced in [**D1**]). See also Poirier [**Poir**]. The theory of pinched disk models and geodesic laminations for Julia sets was designed by Douady & Hubbard (see [**D3**]) and Thurston [**Th1**] (see Notes to §32 for a more detailed description).

The attracting-superattracting surgery was introduced by Douady and Hubbard [**DH2**] as a tool for the Multiplier Theorem. The parabolic-attracting surgery is based upon a similar idea but it cannot be performed quasiconformally. However, it can be done by means of *David surgery* (or rather, inverse to it): see Haïssinsky [**Has1, Has2**]. Quite a general theory of David attracting-parabolic surgery has been recently developed in [**LMMN**].

The idea of *mating* based on qc welding (as discussed in §25.4.3) goes back to Bers in the context of quasi-Fuchsian groups [**Bers1**]. It was adapted to Dynamics by Sullivan [**S2**]. A much more general and sophisticated construction was proposed by Douady and Hubbard [**D2**]. Matings of two quadratic polynomials were analyzed (making use of the Thurston Realization Theorem) by Tan Lei [**TL2**].

Bullet and Penrose discovered in the 1990s that algebraic correspondences can lead to matings between polynomials and Kleinian groups [**BP, BL**]. Recently, it was discovered that such kind of matings are produced in abundance by the dynamics of the *Schwarz reflections in quadrature domains* [**LLMM1, LLMM2**]. Moreover, they are often related to the David surgery [**LMMN**]. A simplest example of such a mating, produced by the Schwarz reflection in the deltoid, leads to the mating of z^2 with the modular group. It can also be described as the David welding by means of the Minkowski ? function from Exercise 19.106.

Stochasticity of postcritically preperiodic maps (Theorem 27.15) was proved by Ruelle [**R**]. The proof given here, implying analyticity of the density, is due to Ognev [**O**]. It was generalized to critically non-recurrent maps by Misiurewicz: see Theorem 46.20 and the associated Notes.

28. Quadratic-like maps and renormalization: first glance

28.1. The concept.

28.1.1. *Definition.* The notion of a quadratic-like map is a fruitful generalization of the notion of a quadratic polynomial.

Definition 28.1. A *quadratic-like map* $f : U \to U'$ (abbreviated as "ql map") is a holomorphic double branched covering between two conformal disks U and U' in $\mathbb{C}$ such that $U \Subset U'$.

By the Riemann-Hurwitz Theorem, any quadratic-like map has a single critical point, which is of course non-degenerate. We normalize f so that the critical point

[11] Recall that Markov partitions actually exist for general expanding maps [**Krz**].

sits at 0 (unless otherwise is explicitly stated). Note that any quadratic polynomial $f = f_c$ restricts to a quadratic-like map $f : f^{-1}(\mathbb{D}_R) \to \mathbb{D}_R$ whose range is a round disk of radius $R > |f(0)|$. More canonically, for any $r > |B_f(0)|$ (where B_f is the Böttcher function for f), the restriction of f to the subpotential domain $\Sigma_f(r)$ (see §23.5.4) provides us with a quadratic-like map $f : \Sigma_f(r) \to \Sigma_f(r^2)$.

From now on (unless otherwise is explicitly stated) we will make the following *Technical Conventions*: For any quadratic-like map $f : U \to U'$, we assume that the domains U and U' are 0-*symmetric*[12] and that f is *even* , i.e, $f(z) = f(-z)$ for all $z \in U$. Moreover, we assume that both domains are *quasidisks*.

Note that the last assumption can be secured by the following *elementary adjustment* of f:

Exercise 28.2. *Take any* 0*-symmetric topological disk* $V' \ni f(0)$ *such that* $U \subset V' \subset U'$*, and let* $V = f^{-1}(V')$*. Then the map* $f : V \to V'$ *is quadratic-like.* (*Of course,* V' *can be chosen so that its boundary is real analytic.*)

More generally, we say that f is an *adjustment* of $\tilde{f}$ if there is a sequence of ql maps $f = f_0, f_1, \dots f_n = \tilde{f}$ such that each f_{k+1} is an elementary adjustment of f_k or the other way around.

Sometimes we will refer to a ql map satisfying the above Technical Conventions as *conventional*. Such a map f extends continuously to $\overline{U}$, so we can assume this without loss of generality.

The annulus $A = \overline{U}' \smallsetminus U$ is called the *fundamental annulus* of f. (We will refer in the same way to the corresponding open and semi-open annuli as well.)

A *degenerate quadratic-like map* f is a holomorphic double branched covering between two conformal disks U and U' in $\mathbb{C}$ such that $U \subset U'$, but U is not compactly contained in U'. So, we do not have a "space" in between U and U' and the fundamental annulus degenerates. In this case, the domain of f may not be adjusted so that it becomes a nice Jordan disk.

28.1.2. *Space of ql maps.* Let $\mathfrak{Q}'$ stand for the space of ql maps $f : (U, 0) \to (U', v)$, perhaps degenerate, endowed with the Carathéodory topology (with the critical point 0 marked). In this topology, a sequence of maps $f_n : (U_n, 0) \to (U'_n, v_n)$ converges to a ql map $f : (U, 0) \to (U', v)$ if the pointed domains $(U_n, 0)$ Carathéodory converge to $(U, 0)$ and $f_n \to f$ uniformly on compact subsets of U (see §7.7).

We let $\mathfrak{Q}$ be the subspace of $\mathfrak{Q}'$ consisting of genuine ql maps.

Exercise 28.3. *If a sequence of maps* $f_n \in \mathfrak{Q}'$ *converges to a genuine ql map* $f \in \mathfrak{Q}$ *then the maps* f_n *are eventually genuine ql. Moreover,*

$$\lim \operatorname{mod} A_n = \operatorname{mod} A,$$

where the A_n *and* A *are the corresponding fundamental annuli.*

[12] Of course, if f is even then U is automatically 0-symmetric.

28.1.3. *Julia set.* The notion of quadratic-like map does not fit to the canonical dynamical framework, where the phase space is assumed to be invariant under the dynamics. In the quadratic- like case, some orbits *escape* through the fundamental annulus (i.e., $f^n z \in A$ for some $n \in \mathbb{N}$), and we cannot iterate them any further. However, there are still a plenty of non-escaping points, which form a dynamically significant object. The set of all non-escaping points is called the *filled Julia set* of f and is denoted in the same way as for polynomials:

$$\mathcal{K}(f) = \{z : f^n z \in U,\ n = 0, 1, \dots .\}$$

By definition, the *Julia set* of f is the boundary of the filled Julia set: $\mathcal{J}(f) = \partial\mathcal{K}(f)$. Dynamical features of quadratic-like maps are very similar to those of quadratic maps (in §40.2 we will see a good reason for it):

EXERCISE 28.4. *Check that all dynamical properties of quadratic polynomials established in in* §§20–21 *are still valid for quadratic-like maps. In particular,*

(i) *The filled Julia set* $\mathcal{K}(f)$ *is a completely invariant full compact subset of* U.

(ii) *Basic dichotomy:* $\mathcal{J}(f)$ *and* $\mathcal{K}(f)$ *are either connected or Cantor; the former holds if and only if the critical point is non-escaping:* $0 \in \mathcal{K}(f)$.

(iii) *Any periodic component of* $\operatorname{int}\mathcal{K}(f)$ *is either in the immediate basin of an attracting/parabolic cycle, or is a Siegel disk.*

(iv) f^n *has* 2^n *fixed points counted with multiplicity. In particular,* f *has two fixed points counted with multiplicity, and one cycle of period two* (*which can merge with one of the fixed points*).

(v) f *can have at most one attracting or parabolic cycle.*

(vi) *adjustments from* Exercise 28.2 *do not change the filled Julia set.*

For a degenerate ql map $f : U \to U'$, we can still define the filled Julia set $\mathcal{K}^{\max}(f)$ as the set of non-escaping points but this set may be non-compact. Moreover, there could be better candidates for this role, e.g., there could exist a completely invariant hull $\mathcal{K} \Subset U$ (non necessarily equal to $\mathcal{K}^{\max}$). In fact, existence of such a hull gives an intrinsic criterion for non-degeneracy (see Lemma 41.3 below).

Let us mention several important properties which are still valid in the degenerate case:

EXERCISE 28.5. *Let a* $f : U \to U'$ *be a ql map, perhaps degenerate.*

(i) *If* f *has an attracting cycle* $\boldsymbol{\alpha}$, *then* f *acts properly on the basin* $\mathcal{D}(\boldsymbol{\alpha})$ *and the latter contains the critical point* 0. *Thus,* f *may have at most one attracting cycle.*

(ii) *If* $Q \Subset U$ *is an invariant compact set which is disjoint from* $\omega(0)$, *then* f *is expanding on* Q.

(iii) f^n *has* 2^n *fixed points counted with multiplicity.*

28.1.4. *External rays and fixed points.* Let us define a *vertical arc* for a quadratic-like map $f : U \to U'$, as a proper arc in the annulus $U' \smallsetminus \mathcal{K}$ going from the inner to the outer end, i.e., an arc $\mathcal{R} : (0, 1] \to \overline{U}' \smallsetminus \mathcal{K}$ such that $\mathcal{R}(t) \in U \smallsetminus \mathcal{K}$ for $t \in (0, 1)$, $\mathcal{R}(1) \in \partial U'$, and $\mathcal{R}(t) \to \mathcal{J}$ as $t \to 0$. (Compare §6.3.1.)

A vertical arc $\mathcal{R}$ is called *periodic with period* p if $f^p(\mathcal{R} \cap U^n) = \mathcal{R}$, where $U^n := f^{-n}(U)$. Similarly to Theorem 24.3, we have:

EXERCISE 28.6. *Any periodic vertical arc* $\mathcal{R}$ *of period* p *lands at some periodic point* $\beta \in \mathcal{J}$ (*of periodic dividing* p). *This point is either repelling or parabolic.*

In particular, any invariant vertical arc $\mathcal{R}$ lands at some fixed point $\beta \in \mathcal{J}$.

EXERCISE 28.7. *The β-fixed point captured in this way is independent of the arc $\mathcal{R}$.*

More generally, us consider a foliation $\mathcal{F}_0$ by vertical arcs of the fundamental annulus $A = \overline{U}' \smallsetminus U$, and pull it back under the dynamics. We obtain vertical foliations $\mathcal{F}^n := (f^n)^*(\mathcal{F}_0)$ in the annuli $A^n := f^{-n}(A)$ that concatenate into a single vertical foliation $\mathcal{F}$ in $\overline{U}' \smallsetminus \mathcal{K}$. The leaves of $\mathcal{F}$ play a role of *external rays* for a ql map in question, and we will often refer to them as such (they depend on the choice of the initial foliation $\mathcal{F}_0$, but the results are usually robust with respect to this choice.)

EXERCISE 28.8. *Prove the analogue of* Theorem 24.5 *for ql maps.*

In particular, we conclude that, like in the polynomial case, *the second fixed point, called α, is either non-repelling, or dividing.*

REMARK 28.9. In §40.2 we will prove a *Straightening Theorem* that will reduce the whole theory of Topological Dynamics for ql maps (as partly outlined above) to that for polynomials.

28.1.5. *Real ql maps.* Let us consider a quadratic-like map $f : U \to U'$ with real symmetric domains U and U'. Since these domains are simply connected, their real slices

$$U_{\mathbb{R}} := U \cap \mathbb{R} \quad \text{and} \quad U'_{\mathbb{R}} := U' \cap \mathbb{R}$$

are open intervals; moreover, $U_{\mathbb{R}} \Subset U'_{\mathbb{R}}$. If additionally, the map f is real, i.e. $f(U_{\mathbb{R}}) \subset U'_{\mathbb{R}}$, then it is naturally called a *real-symmetric* (or just *real*) quadratic-like map. Note also that according to our Conventions, f extends continuously to ∂U, in particular to $\partial U_{\mathbb{R}}$, and we have $f(\partial U_{\mathbb{R}}) \subset \partial U'_{\mathbb{R}}$.

For a real-symmetric quadratic-like map f, we let $\mathcal{K}_{\mathbb{R}} \equiv \mathcal{K}_{\mathbb{R}}(f) := \mathcal{K}(f) \cap \mathbb{R}$ be the real slice of its filled Julia set. We also call it $\mathcal{I} \equiv \mathcal{I}(f)$.

EXERCISE 28.10. *For a real-symmetric quadratic-like map $f : U \to U'$ with connected Julia set $\mathcal{K}(f)$, the real slice $\mathcal{I} \equiv \mathcal{K}_{\mathbb{R}}(f)$ is a closed interval compactly contained in $U_{\mathbb{R}}$. Moreover, the restriction $f : \mathcal{I} \to \mathcal{I}$ is a proper unimodal map* (see §20.4.3), *and its boundary fixed point $\beta \in \partial \mathcal{I}$ is either repelling or parabolic with positive multiplier: $f'(\beta) \geq 1$. If $f'(\beta) > 1$, then there exists a second fixed point $\alpha \in \operatorname{int} \mathcal{I}$. If it is non-attracting then it has a negative multiplier.*

Following our conventions, we let $\mathfrak{Q}_{\mathbb{R}}$ be the space of real ql maps.

The above notions naturally extend to the case of *degenerate* ql maps. As in the non-degenerate case, we have:

EXERCISE 28.11. *Let $f : U \to U'$ be a real ql map, perhaps degenerate, that has an attracting cycle $\boldsymbol{\alpha}$. Then $\boldsymbol{\alpha} \subset \mathbb{R}$.*

Naturally, $\mathfrak{Q}'_{\mathbb{R}}$ will stand for the real slice of the space of perhaps degenerate ql maps.

28.2. Uniqueness of a non-repelling cycle. We will now give the first illustration of how useful the notion of a quadratic-like map is. It exploits the flexibility of this class of maps: small perturbations of a quadratic-like map are still quadratic-like (on a slightly adjusted domain):

EXERCISE 28.12 (compare Exercise 28.2). *Let $f : U \to U'$ be a quadratic-like map with the fundamental annulus A. Take a 0-symmetric smooth Jordan curve $\gamma' \subset A$ generating $H_1(A)$, and let V' be the domain bounded by γ'. Assume $f(0) \in V'$. Let ϕ be a bounded holomorphic function on U, and let $g = f + \phi$, $V = g^{-1}V'$. If $\|\phi\|_\infty$ is sufficiently small then $g : V \to V'$ is a quadratic-like map.*

THEOREM 28.13. *Any quadratic-like map* (*in particular, any quadratic polynomial*) *can have at most one non-repelling cycle.*

PROOF. Assume that a quadratic-like map $f : U \to U'$ has two non-repelling cycles $\boldsymbol{\alpha} = (\alpha_k)_{k=0}^{p-1}$ and $\boldsymbol{\beta} = (\beta_k)_{k=0}^{q-1}$. Let μ and ν be their multipliers. Take two numbers a and b to be specified below.

Using the Interpolation formulas, find a polynomial ϕ (of degree $2p + 2q - 1$) vanishing at points α_k and β_k, such that $\phi'(\alpha_0) = a$, $\phi'(\beta_0) = b$, while the derivatives of ϕ at all other points α_k and β_k ($k > 0$) vanish.

Let $f_\varepsilon = f + \varepsilon\phi$, where $\varepsilon > 0$. Then $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are periodic cycles for f_ε with multipliers

$$\lambda_\varepsilon = \lambda + a\varepsilon \prod_{k>0} f'(\alpha_k) \quad \text{and} \quad \mu_\varepsilon = \mu + b\varepsilon \prod_{k>0} f'(\beta_k)$$

respectively. Since $|\lambda| \le 1$ and $|\mu| \le 1$, parameters a and b can be obviously selected in such a way that $|\lambda_\varepsilon| < 1$ and $|\mu_\varepsilon| < 1$ for all sufficiently small $\varepsilon > 0$. Thus, the cycles $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ become attracting for f_ε. But for a sufficiently small ε, the map f_ε is quadratic-like on a slightly adjusted domain containing both cycles (see Exercise 28.12). As such, it is allowed to have at most one attracting cycle (Exercise 28.4 (v)) – contradiction. □

This result together with Exercise 28.4 (iii) immediately yields:

COROLLARY 28.14. *Any quadratic-like map* (*in particular, any quadratic polynomial*) *can have at most one cycle of components of* $\operatorname{int}\mathcal{K}(f)$.

28.3. Polynomial-like maps. A *polynomial-like map* of degree d is a holomorphic branched covering $f : U \to U'$ of degree d between two nested conformal discs $U \Subset U' \subset \mathbb{C}$. The basic theory of quadratic-like maps developed above extends to the higher degree case in the straightforward way (with the same differences as in the case of polynomials: e.g., the Basic Dichotomy is not valid any more).

EXERCISE 28.15. *Develop a basic theory of polynomial-like maps.*

As in the polynomial case, the primary reason why we need higher degree polynomial-like maps in this book is because they appear as iterates of quadratic-like maps:

EXERCISE 28.16. *Let $f : U' \to U$ be a quadratic-like map such that $f^n(0) \in U$ for some $n \in \mathbb{Z}_+$. Let $U^n := f^{-n}(U)$. Then the map $f^n : U^n \to U$ is a polynomial-like map of degree 2^n. In particular, if $\mathcal{K}(f)$ is connected then the iterates $f^n : U^n \to U$ are polynomial-like for all $n \in \mathbb{Z}_+$.*

28.4. Quadratic-like renormalization.

28.4.1. *Definition.* The primary motivation for introducing quadratic-like maps comes from the idea of renormalization, which is a central idea in contemporary theory of dynamical systems.

A quadratic-like map $f : U \to U'$ is called (*ql*) *renormalizable with period* p if

(R1) There is a topological disk $V \ni 0$ such that all the domains

$$f^i(V),\ i = 0, 1, \dots, p-1,$$

are contained in U, and the map $g := (f^p : V \to f^p(V))$ is *quadratic-like*;

(R2) The filled Julia set $\mathcal{K}(g)$ (or, equivalently the Julia set $\mathcal{J}(g)$) is *connected*;

(R3) *Non-Cutting Assumption* is satisfied: the images

$$K_i := f^i(\mathcal{K}(g)), \quad i = 1, \dots, p-1,$$

can touch $K_0 := \mathcal{K}(g)$ only at the β-fixed point of the latter.

The sets K_i, $i = 0, 1, \dots, p-1$, are called the *little filled Julia sets* (while their boundaries J_i are naturally called *little Julia sets*). If they are actually disjoint, then the renormalization is called *primitive*. Otherwise it is called *satellite*. See Figure 28.1 and Figure 28.2.

The quadratic-like map $g : V \to V'$ is called a *pre-renormalization* of f. The *renormalization* $R_p f$ will be defined later on the level of "quadratic-like germs" by allowing to adjust and to rescale the domains of g (see §40.5.2). We will often refer to g itself as "a renormalization of f", $g = R_p f$, unless a careful distinction between ql *maps* and *germs* is needed. Hopefully it will not lead to a confusion (though there are many choices of a pre-renormalization g with a given renormalization period p).

A quadratic polynomial $f = f_c$ is called *renormalizable* if it restricts to a renormalizable quadratic-like map. It is equivalent to saying that there is a quadratic-like map $g = (f^p : V \to V')$ with properties (R1)–(R3) listed above.

28.4.2. *Julia bouquets.* Let us consider the union of all little Julia sets:

$$\mathbf{K} := \bigcup_{i=0}^{p-1} K_i = \operatorname{orb} K, \quad \text{where } K \equiv \mathcal{K}(g),\ K_i := f^i(K).$$

EXERCISE 28.17. *Let $\beta \in K$ be the β-fixed point of g, and let $s = p/\mathfrak{q}$ be its period. Then $\mathbf{K}$ consists of s connected components*

$$\mathfrak{B}_m = \bigcup_{i=0}^{\mathfrak{q}-1} K_{m+is},\ m = 0, \dots, s-1,$$

where for each m, the little Julia sets K_{m+is} form a bouquet rooted at their common fixed point $\beta_m := f^m(\beta)$.

28.4.3. *Douady-Hubbard* (*DH*) *renormalization.* Assume that f has a repelling periodic cut-cycle $\boldsymbol{\alpha}$. In §24.5 we discussed the associated ray portrait $\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\alpha})$ and showed that it produces a double branched covering $f^p : \Pi_{\text{ch}} \to S_{\text{ch}}$ from the characteristic strip Π_{ch} to the characteristic sector $S_{\text{ch}} \ni v \equiv f(0)$ (where p is the period of the rays in $\mathfrak{R}(\boldsymbol{\alpha})$). We will now proceed to show that under some extra assumptions this makes f renormalizable. We will freely use notions and notation from §24.5.

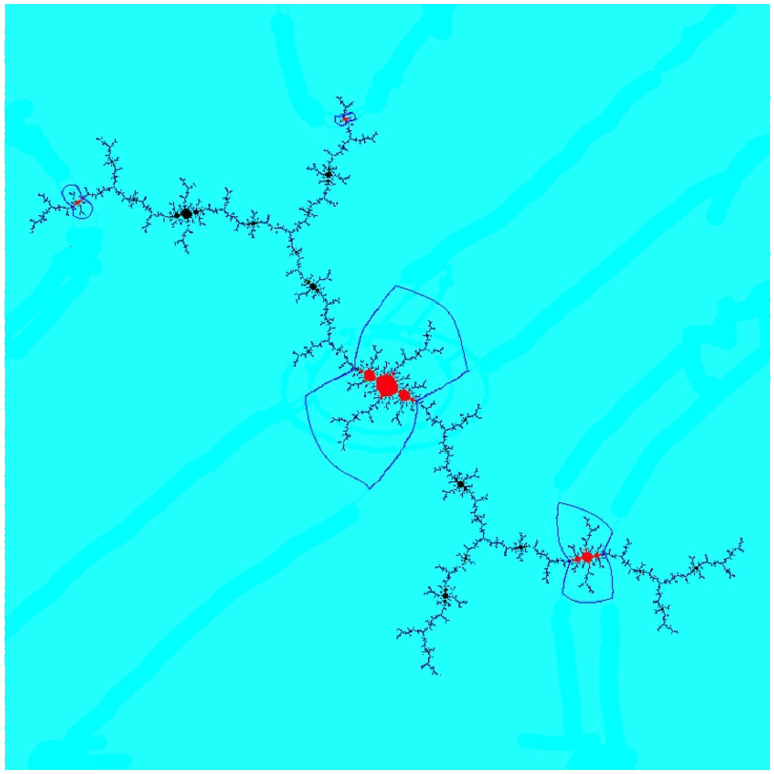

FIGURE 28.1. Domain for a degenerate primitive renormalization of period 4.

Select your favorite equipotential level $t > 0$, and let $t' = 2^p t$. Let W be the strip $\Pi_{\rm ch}$ truncated by the equipotential $\mathcal{E}^t$, and let W' be the sector $S_{\rm ch}$ truncated by the equipotential $\mathcal{E}^{t'}$. We obtain two nested topological disks, $W \subset W'$, such that $f^p : W \to W'$ is a double branched covering. So, it is a *degenerate* quadratic-like map: the only problem is that W and W' have a common piece of boundary (the truncated characteristic leaf $L_{\rm ch} = \mathcal{R}^+_{\rm ch} \cup \mathcal{R}^-_{\rm ch} \cup \{\alpha_{\rm ch}\}$).

EXERCISE 28.18. *Draw the degenerate renormalization picture for the rabbit, airplane, and other favorite hyperbolic maps of yours.*

To fix this problem, let us slightly *thicken* these domains, see Figure 28.3. Namely, one can replace the characteristic line $L_{\rm ch}$ with a nearby line Γ comprising pieces of two nearby rays bridged by a transverse arc δ selected so that $f^{\mathfrak{q}}(\delta) \cap \delta = \emptyset$ (see Exercise 23.13). Let $\Omega' \supset W'$ be the disk bounded by an arc of Γ and arc of the equipotential $\mathcal{E}^{t'}$. Pulling it back under f^p, we obtain a domain $\Omega \supset W$.

EXERCISE 28.19. *Show that* $\Omega \Subset \Omega'$, *and the map* $f^p : \Omega \to \Omega'$ *is quadratic-like. Moreover, its filled Julia set is contained in* W.

This quadratic-like map is a potential *valuable renormalization* of f. To bring it to the *critical point*, let us consider preimages $V := f^{-1}(\Omega)$ and $V' := f^{-1}(\Omega')$.

LEMMA 28.20. *Assume* $f(0) \in W \cup \{\alpha_{\rm ch}\}$. *Then:*

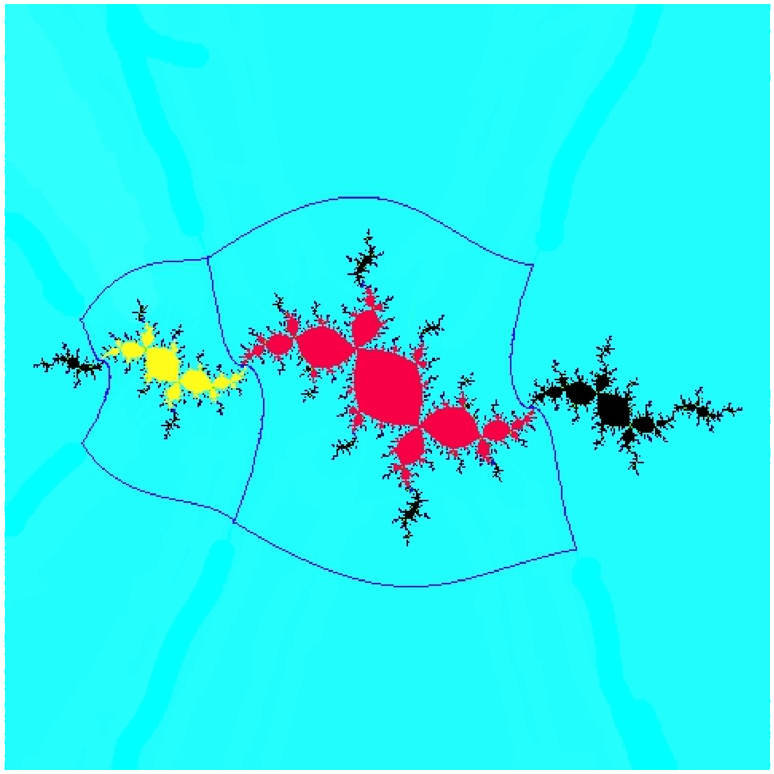

FIGURE 28.2. Domain for a degenerate satellite renormalization of period 2.

(i) *The map* $g = (f^p : V \to V')$ *is quadratic-like;*

(ii) *The (filled) Julia set* $\mathcal{K}(g)$ *is contained in the central ameba* $\overline{\Upsilon} = f^{-1}(\overline{\Pi}_{\mathrm{ch}})$*;*

(iii) *If* $\mathcal{K}(g)$ *is connected then* g *is a pre-renormalization of* f*.*

PROOF. The first two assertions follow from Exercise 28.19. For the last assertion, we only need to check the Non-Cutting Assumption for the little Julia sets. Two see this, notice that the strips $f^n(\Upsilon)$, $n = 0, 1, \dots, p-1$, can touch only at points of the cycle $\boldsymbol{\alpha}$, which are the β-fixed points for the little Julia sets. □

We refer to g as the *DH (pre-)renormalization* associated to the ray portrait $\mathfrak{R}$. Accordingly, the map f is called *DH renormalizable*.

PROBLEM 28.21. *The above renormalization is primitive if and only if* $\mathfrak{q} = 1$*. In this case, there are exactly two rays landing at each periodic point* $\alpha_n \in \boldsymbol{\alpha}$*. If the renormalization is satellite then the rays landing at any* α_n *are cyclically permuted under* f^p *(with rotation number* $\mathfrak{p}/\mathfrak{q}$*).*

EXERCISE 28.22. *Assume* f *is DH-renormalizable with the pre-renormalization* g*. Then its little Julia set* $\mathcal{K}(g)$ *is obtained by chopping off a family of disjoint sectors rooted at the* β*-fixed point* $\beta(g)$ *and all its preimages under the iterates of* g*. All these sectors are univalent pullbacks (under iterates of* f^p*) of a single sector rooted at* $\beta(g)$*.*

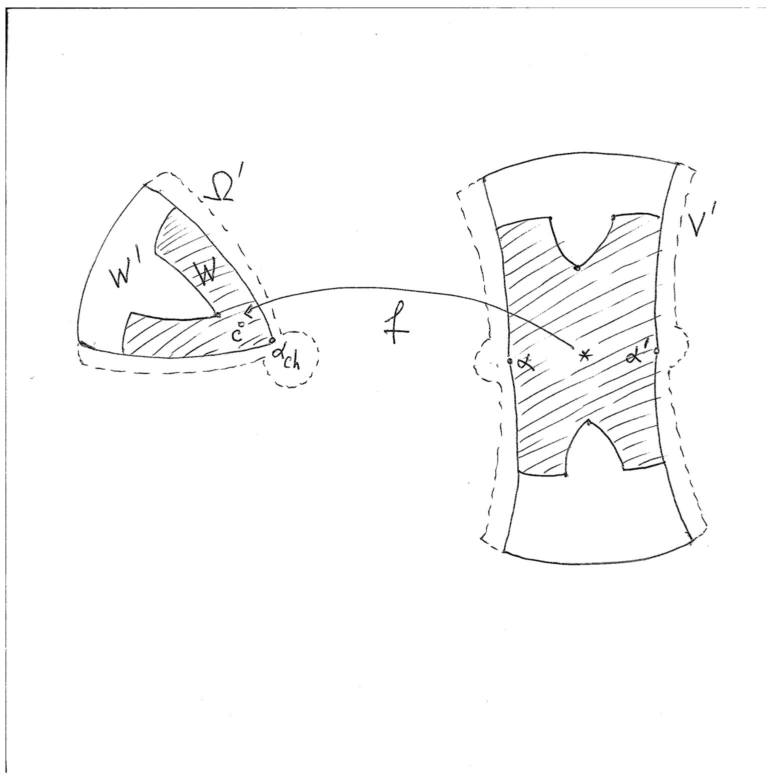

FIGURE 28.3. Thickening.

Applying the above discussion to the ray portrait $\mathfrak{R}(\boldsymbol{\beta})$ of the root cycle of a hyperbolic map (see §25.7.1), we obtain:

PROPOSITION 28.23. *Any hyperbolic ql map f with an attracting cycle of period $p>1$ is renormalizable with period p. Its renormalization $R_p f$ is a hyperbolic ql map with an attracting fixed point.*

28.4.4. *Parabolic case.* Good part of the above discussion is still valid in the case when the cut-cycle $\boldsymbol{\alpha}$ is parabolic. Then the corresponding ray portrait $\mathfrak{R}(\boldsymbol{\alpha})$ produces in the same way a valuable degenerate ql map $f^p : W \to W'$ called a *degenerate renormalization* of f. In the primitive case (when the immediate parabolic basins do not touch), this map can still be thickened to a ql map (since f^p is repelling in the exterior neighborhood of W near α_{ch}), producing its *DH renormalization.* However, in the satellite case, there are attracting petals in the complement of W, which makes thickening impossible. We leave to the reader filling in details of this discussion:

EXERCISE 28.24. *Let f be a parabolic quadratic polynomial, and let p be the period of external rays landing at the parabolic cycle. Then f is DH renormalizable with period p if and only if the cycle is primitive. In this case, the renormalization $R_p f$ is a ql map with a parabolic fixed point.*

In the satellite case, we say that f is *almost renormalizable* or *renormalizbale** with period p.

EXERCISE 28.25. *Let f be DH renormalizable with respect to a cut-cycle $\boldsymbol{\alpha} = (\alpha_k)_{k=0}^{p-1}$, and let K be the corresponding little Julia set. Show that α_0 is the only periodic point in K which is not a cut-point for K but a cut-point for the big Julia set $\mathcal{K}$.*

28.4.5. *Renormalization combinatorics: first encounter.* Each DH renormalization comes together with certain combinatorial data. It accounts for the renormalization period p and the "positions" of the little Julia sets K_i in the big one, $\mathcal{K} \equiv \mathcal{K}(f)$. The simplest way to record it at this stage is by prescribing the periodic ray portrait $\Theta \equiv \Theta(\boldsymbol{\alpha})$ of the cut-cycle $\boldsymbol{\alpha}$ defining the renormalization.

Later on we will introduce several other ways to record the renormalization combinatorics, by prescribing a superattracting parameter $c_\circ$, or a Hubbard tree $\mathcal{T}$, or a little Mandelbrot copy M (see §37.11.2).

28.4.6. *Immediately renormalizable maps.* Let us now apply the above discussion to the ray portrait $\mathfrak{R} \equiv \mathfrak{R}(\alpha)$ of the fixed cut-point α (assuming it is repelling or primitively parabolic). This configuration divides the plane into $\mathfrak{q}$ sectors S_i as described in §24.4.3. (Here we will use notation from that description.)

Let us also consider the symmetric configuration of rays, $\mathfrak{R}' \equiv \mathfrak{R}(\alpha')$ landing at α'. These two configurations together divide the plane into the central strip Π_0, the above $\mathfrak{q}-1$ sectors S_i, $i = 1, \dots, \mathfrak{q}-1$, and the symmetric sectors $S'_i \equiv S_{-i}$.

In this case, the recipe of §28.4.3, can be slightly refined. Select your favorite height $t > 0$ and let $t' = 2^{\mathfrak{q}} t$. Consider the subpotential domains $\Sigma \equiv \Sigma(t)$ and $\Sigma' \equiv \Sigma(t')$, and let

$$Y := \mathrm{cl}(\Pi_0 \cap \Sigma), \quad Y' := \mathrm{cl}(S_0 \cap \Sigma') \tag{28.1}$$

Note that these sets are puzzle pieces in the sense of §9.1.1. Moreover, $f^{\mathfrak{q}} : Y \to Y'$ is a degenerate quadratic-like map. It can be thickened to a quadratic-like map

$$g = (f^{\mathfrak{q}} : V \to V'), \tag{28.2}$$

with the little (filled) Julia set

$$K^\bullet \equiv \mathcal{K}(g) = \{z : \ f^{n\mathfrak{q}} z \in Y, \ n = 0, 1, 2, \dots\}. \tag{28.3}$$

This Julia set is connected if and only if the critical orbit does not escape from Y:

$$f^{n\mathfrak{q}}(0) \in Y, \quad n = 0, 1, \dots \tag{28.4}$$

In this case, the map f is called *immediately renormalizable* with the (pre-) renormalization g. Since the little Julia sets $K_i^\bullet := f^i(K^\bullet)$ touch at α, the immediate renormalization is of *satellite* type.

Moreover, the union of the little Julia sets $K_i^\bullet$ form *the Julia bouquet* centered at α. The Hubbard tree of this renormalization type is the star with $\mathfrak{q}$ edges cyclically

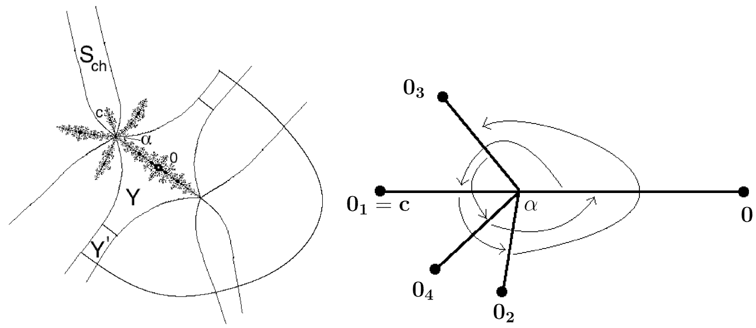

FIGURE 28.4. Bouquet of little Julia sets for an immediately renormalizable map with rotation number 2/5. On the right is the corresponding Hubbard tree.

permuted with rotation number $\mathfrak{p}/\mathfrak{q}$, the marked points being the tips of the star (see Figure 28.4).

This is the first renormalization possible at all:

LEMMA 28.26. *Assume f is renormalizable with some period p (not necessarily the smallest one), little Julia set $K \ni 0$, and associated little Julia sets $K_i = f^i(K)$, $i = 0, \dots, p-1$. Then*

(i) *Each little Julia set K_i is contained in the strip $\overline{\Pi}_0$ or one of the sectors $\overline{S}_{\pm j}$, $j = 1, \dots, \mathfrak{q}-1$. In particular, $K \subset \overline{\Pi}_0$ and $K_1 \subset \overline{S}_{\mathrm{ch}}$.*

(ii) *If f is immediately renormalizable then p is a multiple of $\mathfrak{q}$ and $K \subset K^\bullet$.*

(iii) *If $p = \mathfrak{q}$ then f is immediately renormalizable and $K = K^\bullet$.*

PROOF. (i) Assume some K_i is not contained in one of the sets on the list. Then it crosses the ray portrait $\mathfrak{R} \cup \mathfrak{R}'$ at one of the points α, α', so one of these two points is a cut-point for K_i. Applying $f^{\mathfrak{q}-i}$, we see that α is a cut-point for K. But then K_1 crosses K at α, contradicting the Non-Cutting Assumption (R3) of the definition of renormalization.

(ii) If f is immediately renormalizable, then for any $n \in \mathbb{N}$, we have $f^{n\mathfrak{q}}(0) \in \overline{\Pi}_0$, while $f^{n\mathfrak{q}+i}(0) \in \overline{S}_i$, $i = 1, \dots, \mathfrak{q}-1$. Let us show that the same inclusions hold for the whole little Julia sets:

$$f^{n\mathfrak{q}}(K) \subset \overline{\Pi}_0, \quad f^{n\mathfrak{q}+i}(K) \subset \overline{S}_i, \quad n \in \mathbb{N},\ i = 1, \dots, \mathfrak{q}-1. \tag{28.5}$$

Otherwise, some little Julia set $f^l(K)$ would be contained in some sector $\overline{S}'_i$, $i = 1, \dots, \mathfrak{q}-1$. But by (i), $f^l(0)$ does not belong to S'_i, so it must be equal to α'. Applying f, we obtain $f^{l+1}(0) = \alpha$.

But then all further iterates $f^m(K)$, $m \geq l+1$, contain α as well. By (i), they are contained in $\Pi_0 \cup \bigcup_{i=1}^{\mathfrak{q}-1} \overline{S}_i$ and hence cannot go back to $\overline{S}'_i$. This is a contradiction with periodicity of $f^l(K)$, which proves (28.5).

It immediately implies that p is a multiple of $\mathfrak{q}$, and the inclusion $K \subset \mathcal{K}^\bullet$ follows from (28.3).

(iii) If $p = \mathfrak{q}$ then by (i) we have:

$$f^{\mathfrak{q}n}(0) = f^{pn}(0) \in K \subset \Pi_0, \quad n = 0, 1, \dots$$

By definition (28.4), f is immediately renormalizable. By (ii), $K \subset \mathcal{K}^\bullet$.

Let us consider the thickened quadratic-like map $g := (f^{\mathfrak{q}} : V \to V')$ (28.2). Then $K^\bullet = \mathcal{K}(g)$, while K is a $g-$completely invariant subhull of $\mathcal{K}(g)$. The only such a subhull is $\mathcal{K}(g)$ itself. □

Exercise 28.27. *The little Julia sets K_i do not contain the β-fixed point of f.*

28.4.7. *Canonical Julia nest.* Lemma 28.26 concerning immediate renormalization can be generalized to an arbitrary DH renormalization:

Exercise 28.28. *Let $\mathfrak{R}(\boldsymbol{\alpha})$ be a periodic ray portrait associated with a cut-cycle $\boldsymbol{\alpha}$ of period r, and let q be the period of the rays in $\mathfrak{R}(\boldsymbol{\alpha})$. Assume f is renormalizable with some period $p > r$, with a pre-renormalization g and little Julia sets*

$$K_i, \ i = 0, \dots, p-1, \quad \text{where } K \equiv K_0 \ni 0.$$

Then:

(i) *Each little Julia set K_i is contained in the (closed) central ameba $\overline{\Upsilon}$ or in one of the strips $f^j(\overline{\Pi}_{\mathrm{ch}})$, $j = 0, \dots, q-2$ (in particular, $K \subset \overline{\Upsilon}$, $K_1 \subset \overline{\Pi}_{\mathrm{ch}}$). Moreover, $p \geq q$.*

(ii) *If f is also DH renormalizable with respect to $\mathfrak{R}(\boldsymbol{\alpha})$, with the pre-renormalization $g_{\boldsymbol{\alpha}}$ and the little Julia set $K_{\boldsymbol{\alpha}}$, then p is a multiple of q and*

$$K \subset K_{\boldsymbol{\alpha}}.$$

(iii) *If $p = q$ then f is DH renormalizable with respect to $\mathfrak{R}(\boldsymbol{\alpha})$, $K = K_{\boldsymbol{\alpha}}$, and g coincides with $g_{\boldsymbol{\alpha}}$ on K.*

Let $1 = p_0 < p_1 < \dots$ be the full sequence of DH renormalization periods for f (which can be finite or infinite). The ratios $q_n := p_n/p_{n-1}$ are called *relative renormalization periods* (letting $q_0 = 1$).

If the sequence (p_n) has "length" at least N (i.e., the period p_N is well defined) then f is *N times DH renormalizable*. In particular, if it has infinite length then f is *infinitely DH renormalizable*. If it has zero length (no non-trivial periods) then f is *DH non-renormalizable*.

Corollary 28.29. *Let (p_n) be the sequence of all DH renormalization periods of a map f. Then:*

(i) *For each n, the little (filled) Julia set $K^{[n]}$ of period p_n is canonically defined.*

(ii) *These sets are nested:*

$$\mathcal{K}(f) \equiv K^{[0]} \supset K^{[1]} \supset K^{[2]} \supset \dots \tag{28.6}$$

(iii) *Each p_n is a multiple of p_{n-1}.*

(iv) *For $n < N$, the map $f_n := R_{p_n} f$ is renormalizable with period $q_{n+1} = p_{n+1}/p_n$, and*

$$R_{q_{n+1}} f_n = f_{n+1}.$$

We will refer to $R_{p_1} f$ as the *first DH renormalization* of f, and we will usually reserve notation Rf for this one. Then $R_{p_n} f = R^n f$ is the *n-fold DH renormalization* of f.

If f is infinitely renormalizable with a bounded sequence (q_n) of relative periods then we say that f is of *bounded type*, or f has a *bounded combinatorics*.

REMARK 28.30. Later on, we will show that any ql renormalization is of DH type (see Theorem 31.22), so (p_n) is in fact the *full* sequence of renormalization periods and the nest (28.6) is the *full* nest of little Julia sets.

The (*full DH*) *renormalization combinatorics* of f is the sequence (finite or infinite) of the ray portraits $\Theta^{[n]}$ that describe the combinatorics on all renormalization levels $n = 0, 1 \dots$ (see §28.4.5). In the finitely renormalizable case, it is actually determined by the combinatorics $\Theta^{[N]}$ of the deepest level N. In the infinitely renormalizable case, it is a new combinatorial data:

$$(\Theta^{[0]}, \Theta^{[1]}, \dots) \tag{28.7}$$

Note that it is often more instructive to record the combinatorics of the *relative* (rather than *absolute*) renormalizations: see §43.4.

28.4.8. *Postcritical impression $\mathcal{O}_f$.* Assume f is infinitely DH renormalizable with periods p_n and little filled Julia sets $K^{[n]}$. Let

$$K_i^{[n]} := f^i(K^{[n]}),\ i = 0, 1, \dots, p_n - 1; \quad \mathcal{O} \equiv \mathcal{O}_f := \bigcap_{n=0}^{\infty} \bigcup_{i=0}^{p_n - 1} K_i^{[n]}. \tag{28.8}$$

(Compare with the real case (30.13) below.)

For the sequence $\mathbf{q} = (q_n)$ of relative renormalization periods, let us consider the adding machine $\tau_{\mathbf{q}}$ on the $\mathbf{q}$-adic ring $\mathbb{Z}_{\mathbf{q}} = \varprojlim \mathbb{Z}/p_n\mathbb{Z}$ (see §19.16.2).

EXERCISE 28.31. *There is a continuous map $\pi : \mathcal{O} \to \mathbb{Z}_{\mathbf{q}}$ semi-conjugating $f|\mathcal{O}$ to the adding machine $\tau_{\mathbf{q}}$. Each fiber of π is either a hull or a singleton. Moreover, $\mathcal{O}$ is a Cantor set if and only if π is a homeomorphism.*

EXERCISE 28.32. *Assume f is infinitely renormalizable. Then:*

(i) *Periodic points do not belong to $\mathcal{O}$;*

(ii) *All periodic points are repelling;*

(iii) $\operatorname{int} \mathcal{O} = \emptyset$.

28.4.9. *Area of little Julia sets.* We will now show that the little Julia sets dynamically capture the full area of the filled Julia set $\mathcal{K}$:

LEMMA 28.33. *Assume f is renormalizable with period p and the little* (*filled*) *Julia set $K \ni 0$. Then* $\operatorname{area}(\mathcal{K} \smallsetminus \operatorname{Orb} K) = 0$.

PROOF. Let $f^p : U \to U'$ be a quadratic-like renormalization of f with the Julia set K and the fundamental annulus $A = \overline{U}' \smallsetminus U$.

Since $\mathbf{K} \equiv \operatorname{orb} K \supset \overline{\mathcal{P}}$, Theorem 22.2 on the Global Measure-Theoretic Attractor implies that $\omega(z) \subset \mathbf{K}$ for almost all $z \in \mathcal{K}$. Since the little Julia sets form a cycle of period p, there is a sequence $n(j) \to \infty$ (with gaps at most p) such that $\operatorname{dist}(z_{n(j)}, K) \to 0$ (recall that $z_n \equiv f^n z$). It follows that $z_{n(j)} \in U$ for j sufficiently big (we can assume that this holds for all j).

Assume first that the renormalization is *primitive*. Then the renormalization domains can be selected so that $A \cap \mathbf{K} = \emptyset$, so $\operatorname{dist}(A, \mathbf{K}) \geq \varepsilon > 0$. Since $\operatorname{dist}(z_n, \mathbf{K}) < \varepsilon$ for n sufficiently big, we conclude that eventually $z_n \notin A$.

On the other hand, $z_{n(j)} \in U$. If $z_{n(j)} \notin K$ then $f^{pm}(z_{n(j)}) \in A$ for some m, and we arrive at a contradiction.

Assume now that the renormalization is *satellite*. We can also assume without loss of generality that it is the first renormalization, so f is immediately renormalizable. Then the little Julia sets K_i form a bouquet touching K at its β-fixed point, so the annulus A cannot be selected disjoint from $\mathbf{K}$. However, if $\omega(z) \not\ni \beta$ then it can be selected disjoint from $\omega(z)$, implying that eventually $z_n \notin A$. This is all needed to carry the above argument.

Assume finally that $\beta \in \omega(z) \subset \overline{\mathcal{P}}$. If $\operatorname{orb} z$ does not land in $\mathbf{K}$ then there is a sequence of moments $n(j) \to \infty$ such that $z_{n(j)+1} \to \beta$ and $z_{n(j)} \in S'_i$, where S'_i is a lateral sector attached to β'. (Note that $z_{n(j)+1} \in S_{i+1}$ where $i+1$ is taken mod the number $\mathfrak{q}$ of rays landing at β.) Then there is a sequence $m(j) \in \mathbb{N}$ such that $z_{n(j)+1+m(j)}$ belongs to the central sector S_0 and stays a definite distance away from β but within the range of its local linearizing coordinate. It follows that there is an $r > 0$ independent of j such that $\mathbb{D}_r(z_{n(j)+1+m(j)}) \subset S_0$ and this disk can be univalently pulled back with a bounded distortion to an oval $D_j \subset S_{i+1}$ of bounded shape around $z_{n(j)+1}$ whose size is comparable with its distance to β and which contains a gap in $\mathcal{J}$ of definite size. Taking its pullback by f, we obtain an oval $D'_j \subset S'_i$ of bounded shape around $z_{n(j)}$ whose size is comparable with its distance to β' and which also contains a gap in $\mathcal{J}$ of definite size. This oval is well separated from $\mathbf{K} \supset \overline{\mathcal{P}}$, so by the Koebe Theorem it can be pulled back to z with a bounded distortion. It follows that $\mathcal{J}$ is porous at z. By the Lebesgue Density Points Theorem, the set of such points has zero area. $\square$

28.4.10. *Appendix: Tuned rotation cycles.* Let us consider a ray configuration $\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\beta})$ landing on a cut-cycle cycle $\boldsymbol{\beta}$. In Lemma 24.17 we introduced a degenerate ql map $f^p : \Pi_{\mathrm{ch}} \to S_{\mathrm{ch}}$ from a strip Π_{ch} onto the characteristic sector S_{ch}. Let $\mathrm{Sh}_{\mathrm{ch}} = (\theta_-, \theta_+) \subset \mathbb{T}$ be the shadow of S_{ch} at infinity (see §9.1.1). Its preimage $T^{-p}(S_{\mathrm{ch}})$ under the iterated doubling map is the shadow $\mathrm{Sh}(\Pi_{\mathrm{ch}})$ of the strip Π_{ch}. It comprises two open intervals on $\mathbb{T}$; let I_0 and I_1 be their closures. Under T^p each of these intervlas is mapped injectively onto the closed characteristic shadow $\overline{\mathrm{Sh}}_{\mathrm{ch}} = [\theta, -\theta] \supset I_0 \cup I_1$. This brings us to the situation considered in Exercise 19.60 and Lemma 24.30. Consequently, we obtain a T^p-invariant Cantor set $K \subset I_0 \cup I_1$ on which T^p is monotonically semi-conjugate to T. Moreover, we obtain:

LEMMA 28.34. *Under the above circumstances, for any non-zero* $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, *there exists a unique tuned rotation cycle* $\Theta_{\mathfrak{p}/\mathfrak{q}}(\mathfrak{R}) \subset \mathrm{Sh}(\Pi_{\mathrm{ch}})$ *for* T^p *with rotation number* $\mathfrak{p}/\mathfrak{q}$. *Moreover, any finite* T^p*-invariant set* $\Theta \subset \mathrm{Sh}(\Pi_{\mathrm{ch}})$ *on which* T^p *preserves the cyclic order coincides with one of these tuned rotation cycles.*

We will refer to such a cycle as the *tuned rotation cycle with rotation number* $\mathfrak{p}/\mathfrak{q}$ *associated with the ray configuration* $\mathfrak{R}$.

Assume the multiplier of $\boldsymbol{\beta}$ is not equal to 1. Then by Exercise 28.5 the degenerate ql map $f^p : \Pi_{\mathrm{ch}} \to S_{\mathrm{ch}}$ has a unique fixed point $\alpha_{\mathrm{ch}} \in \Pi_{\mathrm{ch}}$ (generating a cycle

$\boldsymbol{\alpha}$ of period p for f).[13] Assuming α is either repelling or parabolic with non-zero rotation number, let $\mathfrak{R}(\alpha) = \{\mathcal{R}^{\theta_k}\}_{\theta_k \in \Theta}$, $\Theta \equiv \Theta(\alpha) \subset \Pi_{\text{ch}}$, be the ray configuration landing at α.

Lemma 28.35. *Under the above circumstances. the set $\Theta(\alpha)$ is the tuned rotation cycle with rotation number $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}^*_{\text{odd}}$, where $\mathfrak{p}/\mathfrak{q}$ is the combiantorial rotation number of α (and the corresponding cycle $\boldsymbol{\alpha}$).*

Proof. The map $f^p : \mathfrak{R}(\boldsymbol{\alpha}) \to \mathfrak{R}(\boldsymbol{\alpha})$ is a cyclic order preserving permutation of the rays. Hence the induced map $T^p : \Theta \to \Theta$ preserves the cyclic order on Θ. As $\Theta \subset \text{Sh}(\Pi_{\text{ch}})$, Lemma 28.34 tells us that Θ is a tuned rotation cycle with some rotation number $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}^*_{\text{odd}}$. By definition, this number is equal to the combinatorial rotation number of α (and of the corresponding cycle $\boldsymbol{\alpha}$). □

Assuming that our ray portrait makes f renormalizable, we obtain the following corollary from Lemma 9.5:

Lemma 28.36. *Let $g : \Pi_{\text{ch}} \to S_{\text{ch}}$ be a (degenerate) pre-renormalization of f (assoiciated with a ray portrait $\mathfrak{R}$) whose α-fixed point is either repelling or parabolic with multiplier different from 1. Then the number of accesses to α from $\mathbb{C} \smallsetminus \mathcal{K}(f)$ and from $\mathbb{C} \smallsetminus \mathcal{K}(g)$ is the same (and equal to $\mathfrak{q}$, the denominator of the combinatorial rotation number).*

Proof. Let $\mathcal{A}_f$ and $\mathcal{A}_g$ be the sets of accesses to α from $\mathbb{C} \smallsetminus \mathcal{K}(f)$ and from $\mathbb{C} \smallsetminus \mathcal{K}(g)$ repectively. By Lemma 9.5, there is a natural injection $i : \mathcal{A}_f \to \mathcal{A}_g$.

By Lemma 28.35, $\mathcal{A}_f$ is identified with the tuned rotation set $\Theta_{\mathfrak{p}/\mathfrak{q}}(\mathfrak{R})$. Moreover, the map $g|\,\mathfrak{R} = f^p|\,\mathfrak{R}$ acts on the image $i(\mathcal{A}_f)$ as a $\mathfrak{p}/\mathfrak{q}$-rotation.

By thickening g to a ql map, we can apply to it the discussion of §28.1.4. It implies (see Exercise 28.8) that the configuration $\mathcal{A}_g$ is cyclically permuted under g. Hence it does not contain smaller invariant subconfigurations, so $i(\mathcal{A}_f) = \mathcal{A}_g$. □

Notes. The notion of quadratic-like (and more generally: polynomial-like) map was introduced by Douady and Hubbard in their fundamental paper [**DH3**]. The application to the sharp bound on the number of finite non-repelling cycles for polynomials (by $d-1$, see Theorem 28.13) was given in [**D1**]. An analogous bound (by $2d-2$) for rational maps is much harder to prove; it was established by Shishikura [**Sh3**] by means of qc surgery. (See also Epstein [**Ep**] for an algebrageometric approach to this result.)

The idea of quadratic-like renormalization also appeared in [**DH3**] (without using this term). It became a basis for the Complex Renormalization Theory. The Non-Cutting Assumption was coined down by McMullen [**McM1**, §7.3].

29. Topological Dynamics on the Fatou set

29.1. Quasiconformal deformations.

[13] Note that the root β_{ch} of S_{ch} is the other fixed point of the thickened ql map $f^p : \Omega \to \Omega'$, but it is not contained in Π_{ch}.

29.1.1. *Pullbacks.* Consider a K-quasiregular branched covering $f : S \to S'$ between Riemann surfaces (see §14.9). Then any conformal structure μ on S' can be *pulled back* to a structure $\nu = f^*(\mu)$ on S. Indeed, as quasiregular maps are differentiable a.e. on S with non-degenerate derivative and are absolutely continuous, we can let $\nu(z) = (Df(z)^{-1})_*(\mu(fz))$ for a.e. $z \in S$ (compare with the last paragraph of §11.2). This structure has a bounded dilatation:

$$\frac{\|\nu\|_\infty + 1}{\|\nu\|_\infty - 1} \le K \frac{\|\mu\|_\infty + 1}{\|\mu\|_\infty - 1}.$$

If f is holomorphic then in any conformal local charts near z and $f(z)$ we have:

$$(f^*\mu)(z) = \frac{\overline{f'(z)}}{f'(z)}\,\mu(fz).$$

An obvious (either from this formula or geometrically) but crucial remark is that *holomorphic pullbacks preserve dilatation of conformal structures.* (Compare with the discussion in §§11.1–11.2.)

29.1.2. *QC surgeries and deformations.* Consider now a quasiregular (quasi-holomorphic) map $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ *preserving* some conformal structure μ on $\hat{\mathbb{C}}$. By the Measurable Riemann Mapping Theorem, there is a qc homeomorphism $h_\mu : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ such that $(h_\mu)_*(\mu) = \sigma$. Then $f_\mu = h_\mu \circ f \circ h_\mu^{-1}$ is a quasiregular map preserving the standard structure σ on $\hat{\mathbb{C}}$. By Weyl's Lemma, f_μ is holomorphic outside its critical points. Since the isolated singularities are removable, f_μ is holomorphic everywhere, so that it is a rational endomorphism of the Riemann sphere. Of course, $\deg f_\mu = \deg f$. Since h_μ is unique up to post-composition with a Möbius map, *f_μ is uniquely determined by μ up to a Möbius conjugacy.*

Thus, a qc-invariant view of a rational map of the Riemann sphere is a *quasiregular endomorphism $f : (S^2, \mu) \to (S^2, \mu)$ of a qc sphere S^2 that preserves some conformal structure μ.* This provides us with a powerful tool of holomorphic dynamics: the method of *qc surgery.* The recipe is to cook by hand a quasiregular endomorphism of a qc sphere with desired dynamical properties. If it admits an invariant conformal structure, then it can be realized as a rational endomorphism of the Riemann sphere.

It may happen that f itself is a rational map preserving a non-trivial conformal structure μ. Then f_μ is called a *qc deformation of f*. If f is polynomial, then let us normalize h_μ so that it fixes ∞. Then $f_\mu^{-1}(\infty) = \infty$ and hence the deformation f_μ is polynomial as well. If $f : z \mapsto z^2 + c$ is quadratic then let us additionally make h_μ fix 0. Then 0 is a critical point of f_μ, so that

$$f_\mu(z) = t(\mu) z^2 + b(\mu), \quad t(\mu) \in \mathbb{C}^*. \tag{29.1}$$

Composing h_μ with complex scaling $z \mapsto t(\mu)z$, we turn this quadratic polynomial to the normal form $z \mapsto z^2 + c(\mu)$.

29.1.3. *Holomorphic dependence.* Assume now that $\mu = \mu_\lambda$ depends holomorphically on parameter λ. By Theorem 14.6, the map $h_\lambda \equiv h_{\mu(\lambda)}$ is also holomorphic in λ. However, the inverse map h_λ^{-1} is not necessarily holomorphic in λ.

Exercise 29.1. *Give an example.*

It is a miracle that despite it, the deformation $f_\lambda \equiv f_{\mu(\lambda)}$ is still holomorphic in λ!

LEMMA 29.2. *Let $f_\lambda = h_\lambda \circ f \circ h_\lambda^{-1}$, where f and f_λ are holomorphic functions and h_λ is a holomorphic motion (of an appropriate domain). Then f_λ holomorphically depends on λ.*

PROOF. Taking $\partial_{\bar\lambda}$-derivative of the expression $h_\lambda \circ f_0 = f_\lambda \circ h_\lambda$, we obtain:

$$0 = \partial_{\bar\lambda} h_\lambda \circ f_0 = f'_\lambda \circ \partial_{\bar\lambda} h_\lambda + \partial_{\bar\lambda} f_\lambda \circ h_\lambda = \partial_{\bar\lambda} f_\lambda \circ h_\lambda.$$

To complete the proof, we need to verify (in the first place) that $f_\lambda(z)$ belongs to the Sobolev class $W^{1,1}_{\text{loc}}$ in the λ-variable as z gets frozen. Indeed, the map $\lambda \mapsto h_\lambda^{-1}(z)$ with a frozen z can be interpreted as the holonomy map from the cross-section $\{z = \text{const}\}$ to $\{\lambda = \lambda_\circ\}$ (along the lamination corresponding to the holomorphic motion h_λ). By Lemma 17.6, this map is quasiregular. Hence the map $\chi_z(t,\lambda) = h_t \circ f_\circ \circ h_\lambda^{-1}(z)$, for z and t fixed, is qr in λ as well (as a composition of a qr and qc maps). Thus, it belongs to the Sobolev class $W^{1,1}_{\text{loc}}$ in λ . On the other hand, it is holomorphic in t. It follows that is restriction to the line $\{t = \lambda\}$ belongs to $W^{1,1}_{\text{loc}}$ as well (which can be checked by taking smooth approximations to χ_z in $W^{1,1}_{\text{loc}}$). □

COROLLARY 29.3. *Consider a quadratic map $f : z \mapsto z^2 + c_0$. Let μ_λ be a holomorphic family of f-invariant Beltrami differentials on $\mathbb{C}$. Normalize the solution $h_\lambda : \mathbb{C} \to \mathbb{C}$ of the corresponding Beltrami equation so that the qc deformation $f_\lambda = h_\lambda \circ f \circ h_\lambda^{-1}$ assumes a form $f_\lambda : z \mapsto z^2 + c(\lambda)$. Then the parameter $c(\lambda)$ depends holomorphically on λ.*

PROOF. Consider first the solution $H_\lambda : \mathbb{C} \to \mathbb{C}$ of the Beltrami equation that fixes 0 and 1. It conjugates f to a quadratic polynomial of form (29.1). By Lemma 29.2, its coefficients $t(\lambda)$ and $b(\lambda)$ depend holomorphically on λ. The complex rescaling $T_\lambda : z \mapsto t(\lambda)z$ reduces this polynomial to the normal form with $c(\lambda) = t(\lambda)b(\lambda)$, and we see that $c(\lambda)$ depends holomorphically on λ as well. □

29.1.4. *Invariant extensions of conformal structures.* In applications, we usually start with an invariant conformal structure on a smaller Riemann surface and extend it to an invariant conformal structure on an ambient one. It can be done under very general circumstances.

Let S be a Riemann surfaces endowed with a holomorphic equivalence relation $\mathcal{R}$, and let U be an open subset of S. Let $\tilde U$ stand for the $\mathcal{R}-$saturation of U (see §29.4). For all practical purposes, the reader can think of the grand orbit equivalence relation for a holomorphic map f, so $\tilde U$ is just the grand orbit of U.

LEMMA 29.4. *Any $\mathcal{R}$-invariant conformal structures μ on $U \cup (S \smallsetminus \tilde U)$ admits a unique $\mathcal{R}$-invariant extension $\tilde\mu$ to S. Moreover* $\operatorname{Dil}\mu = \operatorname{Dil}\nu$.

In particular, if U is a fundamental domain for $\tilde U$, then any conformal structures μ on U admits a unique $\mathcal{R}$-invariant extension $\tilde\mu$ to S, and $\operatorname{Dil}\mu = \operatorname{Dil}\nu$.

PROOF. Since the set of critical points of $\mathcal{R}$ is at most countable, while the desired conformal structure has to be only measurable, we do not need to define it at the critical equivalence classes.

Let $\zeta_\circ \in \tilde U$ be a point in a regular equivalence class. By definition of the saturation $\tilde U$, it has an $\mathcal{R}$-equivalent point $z_\circ \in U$, hence there exists a local section ϕ of $\mathcal{R}$ such that $\phi(z_\circ, \zeta_\circ) = 0$. Since the class of $\zeta_\circ$ is regular, we can locally express z near $z_\circ$ as $\psi(\zeta)$ with a function ψ holomorphic near $\zeta_\circ$. Let

$\tilde{\mu} = \psi^*(\mu)$ near ζ. This definition is independent of the choice of the local section ϕ since $\mu|U$ is $\mathcal{R}$-invariant. $\square$

COROLLARY 29.5. *Any $\mathcal{R}$-invariant conformal structure μ on U admits a unique $\mathcal{R}$-invariant extension $\tilde{\mu}$ to S such that μ coincides with the standard structure σ on $S \smallsetminus \tilde{U}$. Moreover,* $\operatorname{Dil} \tilde{\mu} = \operatorname{Dil} \mu$.

In particular, if U is a fundamental domain for $\tilde{U}$, then any conformal structure μ on U admits a unique $\mathcal{R}$-invariant extension $\tilde{\mu}$ to S such that $\mu = \sigma$ on $S \smallsetminus \tilde{U}$, and $\operatorname{Dil} \tilde{\mu} = \operatorname{Dil} \mu$.

We will refer to the extension given in this Corollary as *canonical*.

COROLLARY 29.6. *Let $X \subset \hat{\mathbb{C}}$ be a wandering measurable set for a rational map $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ such that all the iterates $f^n | X$, $n \in \mathbb{N}$, are injective. Then any conformal structure μ on X admits the canonical f-invariant extension $\tilde{\mu}$ to the whole sphere, and* $\operatorname{Dil} \tilde{\mu} = \operatorname{Dil} \mu$.

29.2. No Wandering Domains Theorem. The definition of a wandering domain was given in §22.1.

THEOREM 29.7. *A quadratic polynomial f has no wandering domains.*

The rest of the section will be devoted to the proof of this theorem. The idea is to endow a wandering domain D with a 3-parameter family of conformal structures μ_λ, $\lambda \in \mathbb{R}^3$, then to promote it to a family of f-invariant conformal structures on the whole Riemann sphere $\hat{\mathbb{C}}$, and to consider the corresponding qc deformation f_λ of f. With some care this deformation can be made *efficient*, i.e., the map $\lambda \mapsto f_\lambda$ can be made injective. But this is certainly impossible since a 3D parameter domain cannot be embedded into $\mathbb{C}$.

Let us now supply the details. Since D is wandering, only one domain $D_n = f^n D$, $n \in \mathbb{N}$, can contain the critical point 0. By replacing D with $f^{n+1} D$, we can eliminate this possibility.

So, assume $\operatorname{orb} D$ does not contain 0. Then all the maps $f : D_n \to D_{n+1}$ are conformal isomorphisms (being unbranched coverings over simply connected domains, see Exercise 20.4). Hence D is a fundamental domain for its saturation $\operatorname{Orb} D$ by the grand orbit equivalence relation.

Let us now consider an arbitrary conformal structure μ_0 on D (as always, μ_0 is assumed to be measurable with bounded dilatation). By Corollary 29.6, μ_0 canonically extends to an invariant conformal structure μ on the whole sphere $\hat{\mathbb{C}}$, and moreover $\operatorname{Dil} \mu = \operatorname{Dil} \mu_0$.

EXERCISE 29.8. *Work out details of this canonical extension* (without making references to general statements of §29.4).

By the Measurable Riemann Mapping Theorem, there exists a qc map $h_\mu : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ such that $\mu = h_\mu^* \sigma$. Let

$$f_\mu = h_\mu \circ f \circ h_\mu^{-1} : \hat{\mathbb{C}} \to \hat{\mathbb{C}}.$$

By Corollary 29.3, h_μ can be normalized so that $f_\mu : z \mapsto z^2 + c_\mu$ is a quadratic polynomial holomorphically depending on μ.

We will now make a special choice of a 3-parameter family $\mu = \mu_\lambda$ of the initial conformal structures on D to ensure that the qc deformation f_λ is efficient.

Namely, we let $\mu_\lambda = (\psi_\lambda)_*\sigma$, where $\psi_\lambda : D \to D$ is a smooth 3-parameter family of diffeomorphisms that extend to the ideal boundary $\partial^i D$, and the family $\lambda \mapsto \psi_\lambda$ is efficient in the quotient $\mathrm{Aut}(D) \setminus \mathrm{Diff}_+(\partial^i D)$.[14]

EXERCISE 29.9. *Construct such a family of diffeomorphisms.*

Since the real dimension of the parameter space is bigger than 2, by the Implicit Function Theorem, there exists a one-parameter family of conformal structures μ_t (within our 3-parameter family) such that $c_t \equiv \mathrm{const}$. Let us take a base point τ in this family. Then $f_t = f_\tau$ for all t, and hence the homeomorphisms $H_t = h_t \circ h_\tau^{-1}$ commute with f_τ.

EXERCISE 29.10. *Let $H_t : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a one-parameter family of homeomorphisms commuting with a quadratic polynomial f such that $f_\tau = \mathrm{id}$ for some parameter τ. Then $H_t|\, J(f) = \mathrm{id}$ for all t.*

Since $\partial D \subset J(f_\tau)$ (here ∂D is the ordinary boundary, not the ideal one), we conclude that $H_t|\partial D = \mathrm{id}$ for all t. Hence $h_t|\,\partial D = h_\tau|\,\partial D$ and, in particular, $h_t(D) = h_\tau(D) =: \Delta$.

Note now that since $(\psi_t)_*\sigma = \mu_t|\, D = (h_t)^*\sigma$, the map $h_t \circ \psi_t : D \to \Delta$ is conformal. Hence the map

$$\psi_t^{-1} \circ h_t^{-1} \circ h_\tau \circ \psi_\tau : D \to D$$

is a conformal automorphism of D. But on ∂D it coincides with $\psi_t^{-1} \circ \psi_\tau$. By the Carathéodory Prime Ends theory, these two maps have the same extension to the ideal boundary $\partial^i D$, contradicting the efficiency of ψ_t.

The theorem is proved.

29.3. Complete picture of the dynamics on the Fatou set. Putting together No Wandering Domains Theorem and Theorem 21.47, we obtain:

THEOREM 29.11. *For any point $z \in \mathcal{F}(f)$, $\mathrm{orb}\, z$ either converges to an attracting or parabolic cycle, or else lands in a Siegel disk.*

29.4. Appendix: Holomorphic equivalence relations. Holomorphic equivalence relations provide an adequate general set-up for various situations we face. However, we do not exploit it in a serious way, but just use it occasionally as a convenient language.

Let S be a Riemann surface, and let $\mathcal{R}$ be an equivalence relation on S with countable classes. We say that $\mathcal{R}$ is *holomorphic* if there exists a countable family Φ of holomorphic functions $\phi_n(z,\zeta)$ in two variables such that two points $z, \zeta \in S$ are $\mathcal{R}$-equivalent if and only if $\phi_n(z,\zeta) = 0$ for some $\phi_n \in \Phi$. The functions ϕ_n are called *local charts* of $\mathcal{R}$. We assume that local charts ϕ_n are *primitive* in the sense that they are not powers of other holomorphic functions, $\phi_n \neq \psi^k$ for $k \geq 2$. If a local chart has a form $z = \mathbf{h}(\zeta)$ for some holomorphic function h then it is called a *local section* of $\mathcal{R}$.

For instance, orbits of a discrete subgroup $\Gamma \subset \mathrm{Aut}\, S$ (e.g., consider a Fuchsian group acting on $\mathbb{D}$) form a holomorphic equivalence relation. For the context of this book, the most important type of a holomorphic equivalence relation is the *grand orbit relation* generated by a holomorphic map f, e.g., by a rational endomorphism

[14] In other words, $\psi_\lambda = \phi \circ \tilde{\psi}_\lambda \circ \phi^{-1}$, where $\phi : \mathbb{D} \to D$ is the Riemann uniformization and $\tilde{\psi}_\lambda$ is a family of diffeomorphisms $\bar{\mathbb{D}} \to \bar{\mathbb{D}}$ such that $\psi_t|\,\mathbb{T} \neq M \circ \psi_\lambda|\,\mathbb{T}$ for any $t \neq \lambda$, $M \in \mathrm{PSL}^\#(2,\mathbb{R})$.

of the Riemann sphere $\hat{\mathbb{C}}$ (but we will also deal with partially defined maps). More generally, one can consider *algebraic equivalence relations* generated by composing various branches of a global algebraic function $f(z,\zeta) = 0$.

REMARK 29.12. Even more generally, one can consider relations generated by holomorphic pseudo-groups or pseudo-semigroups.

A *critical point* of $\mathcal{R}$ is a point $z_\circ$ such that $\partial_z \phi_n(z_\circ.\zeta_\circ) = 0$ for some local chart $\phi_n \in \Phi$. A *critical equivalence class* is a class containing a critical point. Since the local charts are primitive, the critical points of any section are isolated, and hence altogether there are at most countably many critical points. Non-critical points are called *regular*.

Any equivalence relation on S can be restricted to a subset $D \subset S$. An open subset D is called a *fundamental domain* for $\mathcal{R}$ if the restriction $\mathcal{R}$ to D is trivial (in other words, D contains at most one point of any equivalence class) while its restriction to the closure $\bar{D}$ is complete (i.e., any equivalence class crosses $\bar{D}$). Under these circumstances, the closure $\bar{D}$ will also be referred to as a "(closed) fundamental domain" for $\mathcal{R}$.

The $\mathcal{R}$-*saturation* $\tilde{D}$ of a set $D \subset S$ is the union of all equivalence classes that cross D.

Terms "fundamental domain for f", "f-saturation of D" etc. mean the corresponding objects for the grand orbit equivalence relation generated by f. For instance, f-saturation of a set X is its grand orbit $\operatorname{Orb} X := \bigcup_{n\in\mathbb{N}} f^{-n}(\operatorname{orb} X)$.

EXERCISE 29.13. *A domain $D \subset S$ is a fundamental domain for a map f restricted to the f-saturation of D if and only if D is wandering and the iterates $f^n | D$ are injective, $n \in \mathbb{N}$.*

EXERCISE 29.14. *Show that the saturation of an open subset by a holomorphic equivalence relation is open.*

A conformal structure on S (or a more general tensor field) is called $\mathcal{R}$-*invariant* if it is invariant under all local sections near regular points.

29.5. Notes. Sullivan's No Wandering Domains Theorem for rational functions appeared in [**S1**]. Since then, it appeared in every basic text book on the subject. Besides establishing this important fact, it introduced to Holomorphic Dynamics the powerful method of *quasiconformal deformations*.

The method was upgraded to *quasiconformal surgery* by Douady and Hubbard [**D1, D2, DH2**] who found numerous striking applications for it (some of which are described in this book). It was further developed by Shishikura [**Sh3**], followed by many other people. See the book by Branner and Fagella [**BF**] (containing contributions by other people) with a thorough introduction to the method and its various applications.

Our exposition of the No Wandering Domains Theorem extends without changes to the case of higher degree polynomials. For rational functions, the proof is exactly the same for simply connected components of $\mathcal{F}(f)$ but some extra analysis is needed to rule out multiply-connected domains (that can be actually done by a direct geometric argument, attributed to N. Baker, that avoids qc deformations).

The No Wandering Domains Theorem is analogous to the Ahlfors Finiteness Theorem for Kleinian groups [**Ah2**]. It manifested a deep connection between the

Iteration Theory of rational functions and the Theory of Kleinian groups, which became known as the *Sullivan Dictionary*. Remarkably, Fatou anticipated this connection. On p. 22 of [**F4**], he wrote: "L'analogie remarqueé entre les ensembes de points limites des groupes kleineens et ceux qui sont constitués par les frontières des régions de convergence des itérées d'une fonction rationnelle ne parait d'ailleurs pas fortuite et il serait probablement possible d'en faire la syntèse dans une théorie générale des groupes discontinus des substitutions algrébriques".

Note that wandering domains appear in *Transcendental Dynamics*. The first example of an entire function with a wandering domain was given by Baker in the 1970s [**Ba2**]. More examples were constructed in the early 1980s by Eremenko-Lyubich [**EL2, EL3**] and Herman [**He1**]. In particular, [**EL2, EL3**] contains an example of an *oscillating* wandering domain (accumulating on some finite point). Recently, Bishop has developed a qc surgery techniques that allowed him to construct new interesting oscillating examples [**Bi**].

As periodic components are concerned, entire functions may exhibit one more type, *Baker domains*, where the orbits converge to ∞ (see [**F5, Ba1, EL3**]). However, they do not appear in the class of functions with bounded singular set [**EL3**]. (This class has recently drawn quite a bit of attention, see [**Bi, Re2**].)

However, there is a nice class of transcendental functions *of finite type*, or *Speizer class*, (including λe^z and $\lambda \sin z$) that enjoy exactly the same description of the dynamics on the Fatou set as their polynomial counterparts [**EL2, EL4**] (see also [**BaR, GK**]).

Note in conclusion that the field of *Holomorphic Dynamics* can be extended beyond *Iteration* of holomorphic maps and actions of *Kleinian groups*, incorporating *Holomorphic Foliations* [**Ve, GM, IY**][15] & *Laminations* [**Can, S6, LMin**], *Algebraic Correspondences* [**BP, BL, LLMM2**], and *Schwarz reflections* in quadrature domains [**LLMM1, LLMM3**]. Fatou's dream is getting fulfilled!

30. Topological dynamics of real quadratic maps

In this section we will describe topological dynamics for a real quadratic polynomial $f \equiv f_c$, $c \in \mathcal{M}_{\mathbb{R}} \equiv [-2, 1/4]$, restricted to its maximal invariant interval $\mathcal{I} \equiv \mathcal{I}_c = [-\beta_c, \beta_c]$ (see Exercise 20.10). We will prove that f has a unique topological attractor $\mathcal{A}^{\mathrm{t}}$ that describes the ω-limit set of a generic point $x \in \mathcal{I}$. Moreover, $\mathcal{A}^{\mathrm{t}}$ is either a limit cycle, or a solenoid (a Feigenbaum attractor), or a cycle of transitive intervals. The geometric foundation for this description is provided by Real *a priori* Bounds and No Wandering Intervals Theorem. A notion of Real Renormalization will emerge naturally.

30.1. Classes of maps.

30.1.1. *Class $\mathfrak{G}$ and Epstein class $\mathfrak{E}$.* Though our main theme is the quadratic family, we will develop the theory for real, perhaps degenerate, quadratic-like maps. The reason is essentially the same as the one that motivated introduction of ql maps in the first place: iterates of quadratic polynomials are not quadratic any more, but their appropriate restrictions can be quadratic-like. Passing to limits, these maps may degenerate, leading to various classes of degenerate ql maps.

Recall from §28.1.5 that $\mathfrak{Q}'_{\mathbb{R}}$ stands for the space of ql maps $f : U \to U'$, perhaps degenerate, endowed with the Carathéodory convergence. We will also use

[15] In [**GM**], the method of qc deformations was applied to holomorphic foliations.

notation $\mathfrak{G}'$ for this space. A map $f : U \to U'$ of class $\mathfrak{G}'$ restricts to a unimodal map $f : U_{\mathbb{R}} \to U'_{\mathbb{R}}$, where $U_{\mathbb{R}} = U \cap \mathbb{R}$, $U'_{\mathbb{R}} = U' \cap \mathbb{R}$ are the real slices of the corresponding domains. As f extends continuously to the closure of $U_{\mathbb{R}}$, it has a well defined invariant interval

$$\mathcal{I} \equiv \mathcal{I}(f) := \{x : f^n x \in \overline{U}_{\mathbb{R}},\ n = 0, 1, \dots\},$$

its "*real filled Julia set*".

Let $\mathfrak{G}$ be the space of maps $f : U \to U'$ of class $\mathfrak{G}'$ such that $\mathcal{I} \Subset U'_{\mathbb{R}}$.

Exercise 30.1. *Let $f : U \to U'$ be a map of class $\mathfrak{G}'$. Then:*

(i) *The restriction $f\,|\,\mathcal{I}$ is a proper unimodal map;*[16]

(ii) *$f \in \mathfrak{G}$ iff $\mathcal{I} \Subset U_{\mathbb{R}}$;*

(iii) *If $f \in \mathfrak{G}$ then $f\,|\,\mathcal{I}$ is repulsive.*

Furthermore, let $\mathfrak{E}'$ be the subspace of $\mathfrak{G}'$ for which $U' = \mathbb{C}(U'_{\mathbb{R}}) \equiv \mathbb{C} \smallsetminus (\mathbb{R} \smallsetminus U'_{\mathbb{R}})$, and let $\mathfrak{E}$ stands for the similar subspace of $\mathfrak{G}$. These classes are called *Epstein*.

Note that any quadratic polynomial f_c with connected Julia set restricts to a map of Epstein class with $U_{\mathbb{R}}$ being an arbitrary open 0-symmetric interval containing $\mathcal{I}_c$. With this remark in mind, we view all f_c, , $c \in \mathcal{M}_{\mathbb{R}}$, as maps of Epstein class.

30.1.2. *Real modulus.* For a map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$, we let

$$\operatorname{mod}_{\mathbb{R}} f := \frac{\operatorname{dist}(\partial U', \mathcal{I})}{|\mathcal{I}|},$$

so for $f \in \mathfrak{E}$, we have: $\operatorname{mod}_{\mathbb{R}} f = \operatorname{mod}_{\mathbb{R}}(U'_{\mathbb{R}} : \mathcal{I})$ (see §6.3.5).

For $\mu > 0$, we let

$$\mathfrak{G}(\mu) := \{f \in \mathfrak{G} :\ \operatorname{mod}_{\mathbb{R}} f \geq \mu\},$$

and similarly we define the Epstein subclass $\mathfrak{E}(\mu)$.

Exercise 30.2. (i) *Any map $f \in \mathfrak{G}(\mu)$ is conformally equivalent to a map $\tilde{f}$ of Epstein class $\mathfrak{E}(\nu)$ with $\nu = \nu(\mu) > 0$.*

(ii) *The conjugacy has a bounded distortion on $\mathcal{I}$ depending only on μ.*

(iii) *The distortion goes to 0 in small scales with the rate depending on μ only.*

This remark reduces the theory of real ql maps (perhaps degenerate) to maps of Epstein class.

30.1.3. *Compactness.*

Lemma 30.3. *For any $\mu > 0$, the space of unimodal maps $f : \mathbb{I} \to \mathbb{I}$ of class $\mathfrak{G}(\mu)$ (or of $\mathfrak{E}(\mu)$) is compact in the C^∞-topology.*

Proof. Let us represent our maps $f : U \to U'$ as $h \circ f_0$, where $f_0(z) = z^2$ and h is a conformal diffeomorphism from $V = f_0(U) \supset [0, 1]$ onto U'. Note that the map f_0 is quasisymmetric in the sense that

$$\operatorname{dist}(z, \mathbb{I}) > \varepsilon > 0 \quad\Longrightarrow\quad \operatorname{dist}(f_0(z), [0, 1]) > \delta(\varepsilon) > 0.$$

[16] If $\partial\mathcal{I} \subset \partial U_{\mathbb{R}}$, our definitions do not require f to be real analytic in a real neighborhood of $\mathcal{I}$, but in all occurrences of such maps (as almost renormalizations of maps of class $\mathfrak{G}$) this will be the case.

It follows that $\mathrm{mod}(V \smallsetminus [0,1]) \geq \nu(\mu) > 0$. Therefore, the family of diffeomorphisms h is compact in the Carathéodory topology, implying that the family of their restrictions to $[0,1]$ is C^∞-compact. $\square$

30.1.4. *Inverse branches.* Given an $\mathbb{R}$-symmetric topological disk U and an open interval $I \subset U$, we let

$$U(I) := U \smallsetminus (\mathbb{R} \smallsetminus I)$$

be the corresponding slit domain (compare (2.8)).

LEMMA 30.4. *Let $f : U \to U'$ be a map of class $\mathfrak{G}$, and let $I \subset U$, $I' \subset U'$ be two open intervals . Assume for some $n \in \mathbb{Z}_+$, I is diffeomorphically mapped under f^n onto I'. Then there is an $\mathbb{R}$-symmetric domain $V \subset U(I)$ containing I which is conformally mapped by f^n onto $U'(I')$.*

PROOF. First, let $n = 1$. Since $U'(I')$ does not contain the critical value v, the inverse map $f^{-1} : I' \to I$ extends to the whole domain $U'(I')$. Its image is the desired domain V.

The general case follows from the Telescoping Lemma. $\square$

LEMMA 30.5. *Let $f : U \to U'$ be a map of class $\mathfrak{G}$, and let $I \subset U$, $I' \subset U'$ be two open intervals, where $I \ni 0$ is 0-symmetric. Assume for some $n \in \mathbb{Z}_+$, the image $f(I)$ is contained in an interval J which is mapped diffeomorphically under f^{n-1} onto I'. Then there is an $(\mathbb{R}, 0)$-symmetric domain $V \subset U(I)$ containing I such that $f^n : V \to U'(I')$ is a double branched covering.*

PROOF. By Lemma 30.4, there is a domain $W \subset U'(J)$ which is conformally mapped under f^{n-1} onto $U'(I')$. Then $V := f^{-1}(W)$ is the desired domain. $\square$

30.2. Regular dynamics. The theory that we have already developed for real hyperbolic, parabolic and critically preperiodic quadratic polynomials can be easily extended to class $\mathfrak{G}$.

30.2.1. *Hyperbolic case.* A map $f : U \to U'$ of class $\mathfrak{G}$ is called *hyperbolic* if it has an attracting cycle $\boldsymbol{\alpha}$. For such a map, we naturally define the *real basin*

$$\mathcal{D}_\mathbb{R}(\boldsymbol{\alpha}) := \{x \in U_\mathbb{R} : f^n x \to \boldsymbol{\alpha}\},$$

the *real immediate basin* $\mathcal{D}^\bullet_\mathbb{R}(\boldsymbol{\alpha})$, and the *real Julia set* $\mathcal{J}_\mathbb{R}(f) := \mathcal{I} \smallsetminus \mathcal{D}_\mathbb{R}(\boldsymbol{\alpha})$.

EXERCISE 30.6. *For a hyperbolic map $f \in \mathfrak{G}$, the real immediate basin of an attracting cycle contains the critical point 0. Hence f can have at most one attracting cycle, and this cycle is real.*

The following assertion generalizes Exercise 25.68 to class $\mathfrak{G}$:

EXERCISE 30.7. *Any real hyperbolic map $f \in \mathfrak{G}$ is expanding on its real Julia set $\mathcal{J}_\mathbb{R}(f)$. It follows that $\mathcal{J}_\mathbb{R}(f)$ is nowhere dense in $\mathcal{I}$, and in fact, uniformly porous in all scales. Thus, $\omega(x) = \boldsymbol{\alpha}$ for an open set of $x \in \mathcal{I}$ of full measure.*

30.2.2. *Parabolic case.* A map $f \in \mathfrak{G}$ is called *parabolic* if it has a parabolic cycle $\boldsymbol{\alpha}$. Again, we naturally define the *real parabolic basin*

$$\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha}) := \operatorname{int}\{x \in U_{\mathbb{R}} : f^n x \to \boldsymbol{\alpha}\},$$

the *real immediate parabolic basin* $\mathcal{D}^{\bullet}_{\mathbb{R}}(\boldsymbol{\alpha})$, and the *real Julia set* $\mathcal{J}_{\mathbb{R}}(f) := \mathcal{I} \smallsetminus \mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ (see §26.8).

Exercise 30.8. *For a parabolic map $f \in \mathfrak{G}$, the real immediate basin of a parabolic cycle contains the critical point* 0*. Hence f can have at most one parabolic cycle, and this cycle is real.*

Exercise 30.9. *For any real parabolic map $f \in \mathfrak{G}$ and any $\varepsilon > 0$, we have:*
(i) *The first transit map through the ε-neighborhood of $\boldsymbol{\alpha}$ is expanding on the real Julia set $\mathcal{J}_{\mathbb{R}}(f)$;*
(ii) *$\mathcal{J}_{\mathbb{R}}(f)$ is nowhere dense in $\mathcal{I}$, and in fact, uniformly porous in all scales.*
(iii) *$\omega(x) = \boldsymbol{\alpha}$ for an open set of $x \in \mathcal{I}$ of full measure.*

30.2.3. *Regular maps.* So, in both hyperbolic and parabolic cases, almost all orbits converge to a cycle. Such a dynamics (and the corresponding map $f : \mathcal{I} \to \mathcal{I}$) is called *regular*. Thus, we have:

Theorem 30.10. *A unimodal map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$ is regular if and only if it is either hyperbolic or parabolic.*

In conclusion, let us give a more general version of Exercises 30.7 and 30.9. Let us say that an invariant compact *set* $K \subset \mathcal{I}$ is *parabolic* under a unimodal map $f : \mathcal{I} \to \mathcal{I}$ if K does not contain the critical point but contains a parabolic cycle of f.

Exercise 30.11. *Let $f : \mathcal{I} \to \mathcal{I}$ be a unimodal map of class $\mathfrak{G}$, and let $K \subset \mathcal{I}$ be a hyperbolic or parabolic set for f. Then* $\operatorname{length} K = 0$.

30.2.4. *Misiurewicz case.*

Exercise 30.12. *Following the lines of* §27.1.6, *develop a measure-theoretic theory of dynamics for Misiurewicz maps of class $\mathfrak{G}$.*

The conclusion of the theory is that Misiurewicz maps of class $\mathfrak{G}$ are *stochastic*.

30.3. Fixed points. Given a map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$, let $\beta \equiv \beta_g$ be its boundary fixed point with multiplier $\rho(\beta) = f'(\beta)$. (Note that $\beta > 0$ under our standing convention that 0 is the minimum point.) Since $\mathcal{I}$ is repulsive, $\rho(\beta) \geq 1$ (compare §20.4.3).

Many elementary properties of real quadratic maps (see §20.4) extend to this bigger class, and in particular, we have:

Lemma 30.13. (i) *If $\rho = 1$ then* $\operatorname{int} \mathcal{I}$ *is the real immediate basin of β.*
(ii) *If $\rho > 1$ then* $\operatorname{int} \mathcal{I}$ *contains exactly one fixed point, $\alpha \equiv \alpha_f$.*
(iii) *If α is attracting then* $\operatorname{int} \mathcal{I}$ *is its real immediate basin.*
(iv) *If α is non-attracting then $\rho(\alpha) \leq -1$.*

Moreover, $\mathcal{I}$ is the only invariant interval on which f is proper and repulsive.

Proof. (i) By Exercise 28.5 (iii), f has at most two real fixed points counted with multiplicity. If $\rho = 1$ then β is a multiple point, so there are no other fixed points in $\mathcal{I}$. Since $f(\beta') = \beta > \beta'$, we conclude that $f(x) > x$ for all $x \in [-\beta, \beta)$. Hence $f^n x \to \beta$ for all $x \in \mathcal{I}$, and the conclusion follows.

(ii) In this case, $f(x) < x$ for $x \in \operatorname{int} \mathcal{I}$ near β. Since $f(\beta') > \beta'$, there should be a fixed point in between.

We leave the rest as an Exercise. □

We say that a map $f \in \mathfrak{G}$ is *Myrberg* if its α-fixed point is parabolic with multiplier -1.

30.4. Forward orbits of intervals.

30.4.1. *Wandering intervals and homtervals.* According to the general terminology, an interval $L \subset I$ is called *wandering* if $f^n(L) \cap f^m(L) = \emptyset$ for $m > n \geq 0$. A wandering interval is viewed to be *trivial* if its orbit converges to a cycle. We consider wandering intervals of all types: open, semi-open, and closed.

An interval S is called *periodic* of period $p \in \mathbb{Z}_+$ if $f^p(S) \subset S$, while the intervals $f^n(S)$, $n = 0, 1, \dots, p-1$ have pairwise disjoint interiors. In this case, the union $\bigcup f^n(S)$ is called a *cycle of intervals.*

Lemma 30.14. *Let L be a non-wandering interval. Then $\mathcal{L}^+ := \bigcup_{n=0}^{\infty} f^n(L)$ is a finite union of intervals. Hence $f^m(L)$ is eventually absorbed by a cycle of intervals contained in $\mathcal{L}^+$.*

Proof. Since L is non-wandering, there exist $n \in \mathbb{Z}_+$ and $p \in \mathbb{N}$ such that $f^n(L) \cap f^{n+p}(L) \neq \emptyset$. Applying the iterates of f^p, we see that

$$f^{n+kp}(L) \cap f^{n+(k+1)p}(L) \neq \emptyset \quad \text{for } k = 0, 1, 2, \dots.$$

Hence $S := \bigcup_{k=0}^{\infty} f^{n+kp}(L)$ is an interval. Obviously, it is invariant under f^p. Taking the component of $\bigcup_{n=0}^{p-1} f^n(S)$ containing S, we obtain a desired periodic interval of some period dividing p. □

An interval $L \subset I$ is called *homterval* if all iterates $f^n : L \to I$, $n \in \mathbb{N}$, are monotone. Again, we consider homtervals of all types.

A homterval is called *maximal* if it is not contained in the interior of a bigger one. Taking the closures of all open maximal homtervals, we obtain all closed maximal homtervals.

Exercise 30.15. (i) *Maximal open homtervals are components of*

$$\mathcal{I} \smallsetminus \overline{\operatorname{Orb}_-(0)}$$

(*So, maximal closed homtervals are the closures of those components.*)

(ii) *Any open homterval is contained in a unique maximal open homterval.*

(iii) *If L is a maximal homterval such that $0 \notin \partial L$ then $f(L)$ is also a maximal homterval.*

There is a close relation between homtervals and wandering intervals:

EXERCISE 30.16. (i) *If L is an open homterval, then either L is wandering or it is contained in the basin of a non-repelling cycle.*

(ii) *If L is a wandering interval then some iterate $f^n(L)$ is a homterval.*

(iii) *If L is a non-trivial wandering homterval then the maximal homterval (of either type) containing L is also a non-trivial wandering interval.*

30.4.2. *Orbits of intervals around α.* Recall that for $c \in \mathcal{M}_{\mathbb{R}} \equiv [-2, 1/4)$, f_c has a unique fixed point in $\mathcal{I}_c^\circ$, called α_c, and for $c < 0$ this point has negative multiplier (Exercise 20.15). Recall also that for $c \in [-2, -3/4)$, the map f_c has a periodic cycle of period 2, and for $c \in [-2, -1]$ it has negative multiplier (see Exercise 20.18). In the latter case, $\mathcal{T}_c = [c, f(c)] \equiv [v, f(v)]$ is the minimal (closed) invariant interval containing 0. More generally, we have:

EXERCISE 30.17. *For a map $f \in \mathfrak{G}$, assume the α-fixed point is repelling and there are no attracting cycles of period 2 with non-negative multiplier (so $c \in [-2, -1)$ in the quadratic case). Then the interval $[\alpha, 0]$ is monotonically mapped under f^2 onto a bigger interval, $[\alpha, f(v)]$. Moreover, $f^2(x) > x$ for all $x \in (\alpha, 0]$.*

LEMMA 30.18. *For a map $f \in \mathfrak{G}$, assume the α-fixed point is repelling and there are no attracting cycles of period 2 with non-negative multiplier. For any interval $L = [\alpha, \alpha + \varepsilon] \subset [\alpha, 0]$, there exists an $m \in \mathbb{N}$ such that f^{2m} monotonically maps L onto an interval containing $[\alpha, 0]$.*

PROOF. Exercise 30.17 implies that all points $x \in (\alpha, 0]$ eventually escape the interval $(\alpha, 0]$ under the iterates of f^2. Hence $f^{2m}(L) \supset [\alpha, 0]$ for some $m \in \mathbb{N}$. For the smallest such m, the map $f^{2m}| L$ is monotonic. □

COROLLARY 30.19. *Under the above circumstances, for any interval $L \subset \mathcal{I}$ containing α (maybe, on the boundary) we have $f^n(L) \cup f^{n+1}(L) \supset \mathcal{T} \equiv [v, f(v)]$ for some $n \in \mathbb{N}$. If $\operatorname{int} L \ni \alpha$ then $f^n(L) \supset \mathcal{T}$ for some $n \in \mathbb{N}$.*

PROOF. Let $\alpha \in \partial L$. Since α has a negative multiplier, we can assume without loss of generality that L lies on the right of α. By the above lemma, $f^{2m}(L) \supset [\alpha, 0]$ for some m. Applying two more iterates, we obtain the desired.

The last assertion follows by considering two halves of L,

$$L^+ = L \cap \{x \geq \alpha\} \text{ and } L^- = L \cap \{x \leq \alpha\}.$$

□

In conclusion, let us identify the smallest invariant interval for $f \in \mathfrak{G}$ (compare Exercise 20.18):

EXERCISE 30.20. *Let $f : \mathcal{I} \to \mathcal{I}$ be a map of class $\mathfrak{G}$ that has a cycle of period 2 with negative multiplier. Then $\mathcal{T} \equiv \mathcal{T}_f := [v, f(v)]$ is the smallest invariant interval of f containing the critical point 0. Moreover, for any $x \in \operatorname{int} \mathcal{I}$, there exist an n such that $f^n x \in \mathcal{T}$.*

30.5. Wandering intervals approach 0. If J is a wandering interval the $|f^n(J)| \to 0$ and hence the limit set $\omega(x)$ is independent of a point $x \in J$. Let us call it $\omega(J)$.

LEMMA 30.21. *If J is a wandering interval then $\omega(J) \ni 0$.*

PROOF. If $\omega(J) \not\ni 0$ then eventually the intervals $J_n = f^n(J)$ stay away from 0; without loss of generality, we can assume that it is so from the very beginning:

$$\operatorname{dist}(f^n(J), 0) > 2\varepsilon_0 > 0, \quad n \in \mathbb{N}. \tag{30.1}$$

We can also assume that:

• The endpoints of J are not precritical;

• $|J_n| < \varepsilon_0/K$ for all $n \in \mathbb{N}$ and $\sum |J_n| < 1/(K+1)$,
where K is a big absolute constant, to be specified below;

• J is a maximal homterval: see Exercise 30.16.

Let us consider an interval L attached to J of the same length as J. Since $L \cup J$ is not a homterval, some interval $\operatorname{int} f^N(L) = \operatorname{int} L_N$ contains 0. By (30.1), $|L_N| \geq \varepsilon_0 \geq K\,|J_N|$ at that moment. Let us take the first moment $n \leq N$ for which $|L_n| \geq K\,|J_n|$. Then

$$\frac{|L_n|}{|J_n|} \geq K = K\frac{|L|}{|J|} \tag{30.2}$$

On the other hand,

$$\sum_{k=0}^{n} |L_k \cup J_k| \leq (K+1) \sum_{k=0}^{n} |J_k| \leq 1,$$

so by the Denjoy Distortion Estimate (Exercise 19.69), the distortion of f^n on $L \cup J$ is bounded by some absolute constant K_0. So, if $K > K_0$ then we arrive at a contradiction with (30.2). □

30.6. Maximal periodic interval. Let us now identify the *maximal* (*closed*) *periodic interval* $\mathcal{I}'$ (of the smallest period $p > 1$) around the critical point. We let $I^0 = [\alpha, -\alpha]$.

PROPOSITION 30.22. *For a map $f \in \mathfrak{G}$, assume the fixed point α is either repelling or parabolic with multiplier -1 (so $c \in [-2, -3/4]$ in the quadratic case).*
(i) *If $f^2(0) \in I^0$ then $p = 2$ and $\mathcal{I}' = I^0$. Moreover, $f^2\,|\,\mathcal{I}'$ is a proper unimodal map.*

Otherwise, let us consider the set $A_- = \overline{\operatorname{Orb}_-(\alpha)}$.

(ii) *If $A_- \not\ni 0$ then $p > 2$ and $\mathcal{I}' = \overline{S}$, where S is the connected component of $\mathcal{I} \setminus A_-$ containing 0. Moreover, $f^p\,|\,\mathcal{I}'$ is a proper unimodal map.*

(iii) *If $A_- \ni 0$ then there are no periodic intervals around 0 of period $p > 1$.*

PROOF. Note first that $\mathcal{I}'$ is 0-symmetric since a periodic interval remains such after symmetrization. Next,

$$\alpha \notin \operatorname{int} f^n(\mathcal{I}'), \quad n = 0, 1, \ldots, p-1, \tag{30.3}$$

since otherwise $\operatorname{int} f^{n+1}(\mathcal{I}') \cap \operatorname{int} f^n(\mathcal{I}') \ni \alpha$, contradicting the definition of a periodic interval. Hence $\mathcal{I}' \subset I^0$.

(i) We have: $f(I^0) = [v, \alpha]$, so $f(I^0) \cap I^0 = \{\alpha\}$. Moreover, if $v_1 \equiv f(v) \in I^0$ then

$$f^2(I^0) = f[v, \alpha] = [\alpha, v_1] \subset I^0.$$

We conclude that in this case, I^0 is a periodic interval of period 2, and as we have noticed above, it is the maximal one. The last assertion, on the proper unimodality, is obvious.

(ii) Let S_n be the component of $\mathcal{I} \smallsetminus A_-$ containing $f^n(S)$. Since $\mathcal{I} \smallsetminus A_-$ is an open invariant set, the S_n are open intervals and $f(S_n) \subset S_{n+1}$.

If the intervals S_n are pairwise disjoint then S is a wandering interval with $\omega(S) \not\ni 0$, contradicting Lemma 30.21. Hence $S_n = S_{n+p}$ for some $n \geq 0$, $p > 0$. Let us select the smallest n and p with these properties. If $n > 0$, then S_n is a periodic homterval, i.e., $f^p : S_n \to S_n$ is a homeomorphism. Then it contains an attracting or neutral periodic point a of period p or $2p$ (see Exercise 19.30). Furthermore, the real immediate basin $D^\bullet_{\mathbb{R}}(a)$ is contained in S_n since $\partial S_n \subset A_-$, and the latter does not intersect any basin (note that α is repelling in the case under consideration). But then S_n must contain a critical point (see Exercises 30.6 and 30.8) – contradiction.

Hence $n = 0$, so S is a periodic interval itself, and so is $\mathcal{I}' = \overline{S}$. Finally,

$$f(\mathcal{I}') \subset [v, \alpha], \quad f^2(\mathcal{I}') \subset [-\alpha, v_1],$$

so both of these intervals are disjoint from $\operatorname{int} \mathcal{I}' \subset \operatorname{int} I^0$. It follows that $p \geq 3$.

Let us now check that $g := f^p \,|\, \mathcal{I}'$ is a proper unimodal map. Since

$$f^n(\mathcal{I}') \cap \operatorname{int} \mathcal{I}' = \emptyset,$$

all the maps $f : f^n(\mathcal{I}') \to f^{n+1}(\mathcal{I}')$, $n = 1, \dots p-1$, are diffeomorphisms, implying that g is unimodal. Moreover, since the set A_- is forward invariant, the boundary $\partial \mathcal{I}'$ is g-invariant, so g is proper.

(iii) In this case, the iterated preimages of α accumulate on 0. Hence for any interval $L \ni 0$, there is an $m \in \mathbb{N}$ such that $\operatorname{int} f^m(L) \ni \alpha$. By (30.3), L cannot be periodic of period $p > 1$. □

30.7. Real renormalizations.

30.7.1. *Definition.* A map $f \in \mathfrak{G}$ is called *really renormalizable* if f has a repulsive periodic interval $S \ni 0$ in $\mathcal{I}$ of period $p > 1$.

If we drop the repulsiveness condition then we say that f is *almost renormalizable* or, for brevity, *renormalizable**. In this case, ∂S must contain a parabolic periodic point with multiplier 1. (However, as we will discuss momentarily, the reverse is not true: even when ∂S contains a parabolic point, f may be *genuinely* renormalizable.)

Remark 30.23. These notions are designed to match the corresponding complex ones. In the purely real theory, the repulsiveness of $\mathcal{I}'$ is not usually required for the map to be renormalizable.

As in the complex situation, real renormalization comes with its *combinatorics*: the order in which the intervals $S_k \equiv f^k(S)$, $k = 0, 1, \dots, p-1$, appear on the real line. More formally speaking, it is a permutation $(k_m)_{m=0}^{p-1}$ of p symbols such that $k = k_m$ if the interval S_k appears m's from the left. (This can also be formulated naturally in terms of the kneading theory, compare §37.11.3 .)

Exercise 30.24. *Show that being renormalizable* with a given combinatorics is a closed condition in the space $\mathfrak{G}$.*

30.7.2. *First renormalization.* For a renormalizable map, Proposition 30.22 supplies us with the biggest periodic interval $\mathcal{I}'$ of some period $p > 1$. Let $g = f^p : \mathcal{I}' \to \mathcal{I}'$ be the corresponding return map. As it is proper, it has a boundary fixed point $\beta_g \in \partial\mathcal{I}'$ with some multiplier $\rho_g \geq 1$.

LEMMA 30.25. (i) *If $f : \mathcal{I} \to \mathcal{I}$ is a renormalizable* map of class $\mathfrak{G}$, then the return map $g : \mathcal{I}' \to \mathcal{I}'$ is its first renormalization* of class $\mathfrak{G}$.*

(ii) *The renormalization* is not genuine iff it is doubling with parabolic fixed point β_g. In this case, f is Myrberg: its $\alpha-$fixed point α_f is parabolic with multiplier -1 (so $c = -3/4$ in the quadratic case). Moreover, $\beta_g = \alpha_f$.*

(iii) *In the genuinely renormalizable case, if β_g is parabolic then it is non-degenerate as in* (20.2) *from* §20.4.3.

PROOF. (i) We know that $g : \mathcal{I}' \to \mathcal{I}'$ is a proper unimodal map. Let us show that it belongs to class $\mathfrak{G}$.

Let $\mathcal{I}'_n \supset f^n(\mathcal{I}')$ be the pullback of $\mathcal{I}'$ under f^{p-n}, $n = 1, \dots, p$. Then each map $f^{p-n} : \mathcal{I}'_n \to \mathcal{I}_0$ is a diffeomorphism, and the inverse branch $f^{-(p-n)} : \mathcal{I}'_0 \to \mathcal{I}_n$ admits an analytic extension to the whole domain U'. Let $U_n := f^{-(p-n)}(U')$ and let $V := f^{-1}(U_1)$. Then $f^p : V \to U'$ is a ql map, perhaps degenerate, that gives us an analytic extension of $g : \mathcal{I}' \to \mathcal{I}'$. Moreover, $\mathcal{I}' \Subset V \subset U'$, implying that g is a map of class $\mathfrak{G}$.

(ii)–(iii) If $p = 2$ then case (i) of Proposition 30.22 holds, so $\mathcal{I}' = I^0 = [\alpha, -\alpha]$, where $\alpha \equiv \alpha_f$, and $g = f^2 \,|\, I^0$. We see that $\beta_g = \alpha$ and $g'(\beta_g) = f'(\alpha)^2$.

If α is repelling then clearly g is repulsive on $\mathcal{I}'$, so our renormalization* is genuine. Otherwise, α is parabolic with multiplier -1, so f is Myrberg. In this case, it attracts under f the whole real neighborhood of itself, so $\mathcal{I}'$ is attractive under g, and our renormalization* is not genuine.

If $p > 2$, then we are in case (ii) of Proposition 30.22, and hence $\beta_g \neq \alpha_f$. Since $\operatorname{int}\mathcal{I}'$ is a component of $\mathcal{I} \smallsetminus A_-$, iterated preimages of α_f accumulate on β_g. It follows that g cannot be attracting on the exterior of $\mathcal{I}'$. Hence it must be repulsive (independently of whether β_g is repelling or parabolic).

In the case of parabolic β_g, each cycle of parabolic-attracting petals must contain the critical point. Hence there is only one such a cycle (of period 1 under g), so β_g is non-degenerate. □

For a really renormalizable map $f \in \mathfrak{G}$, the return map $g : \mathcal{I}' \to \mathcal{I}'$ is called the *first real pre-renormalization* of f. The *first renormalization* of f is obtained by rescaling $\mathcal{I}'$ to the unit size:

$$Rf : \mathbb{I} \to \mathbb{I}, \quad Rf(x) = \sigma^{-1} g(\sigma x), \quad \text{with } \sigma = \pm|\mathcal{I}'|, \tag{30.4}$$

where the sign of σ is selected so that 0 is the *minimum* point of Rf.

A similar terminology applies to the renormalizable* situation.

REMARK 30.26. This distinction between "pre-renormalization" and "renormalization" becomes important when R is considered as an operator in the space of unimodal maps. This viewpoint will not play a role until volume III, so meanwhile we will often skip the prefix "pre-" in our terminology.

30.7.3. *Further renormalizations.* If the renormalization Rf is itself renormalizable, we say that f is *twice really renormalizable*. Then we obtain the *second real pre-renormalization* $f_2 : \mathcal{I}^2 \to \mathcal{I}^2$ and the corresponding second real renormalization $R^2 f : \mathbb{I} \to \mathbb{I}$. Proceeding this way, we can define *n-times renormalizable maps*, their pre-renormalizations $f_n = f^{p_n} : \mathcal{I}^n \to \mathcal{I}^n$ and renormalizations $R^n f : \mathbb{I} \to \mathbb{I}$. If this happens for all $n \in \mathbb{Z}_+$, we say that f is *infinitely renormalizable*.

If some $R^n f$ is a Myrberg map, then it has a *doubling renormalization** $R^{n+1} f$, and the latter is not renormalizable* anymore. In this case, f is only *n times renormalizable*, but *$n+1$ times renormalizable**. We will still use notation $f_{n+1} : \mathcal{I}^{n+1} \to \mathcal{I}^{n+1}$ and $R^{n+1} f$ for the last (pre-)renormalization*.

The sequence of renormalizations* f_n captures return maps to all interesting periodic intervals:

THEOREM 30.27. *Let $f \in \mathfrak{G}$, and let $f_n : \mathcal{I}^n \to \mathcal{I}^n$ be its real renormalizations*.*
(i) *Then all of these renormalizations, except perhaps the last one, are genuine. If the last one, f_N is not, then the previous map f_{N-1} is a Myrberg* (*and in particular, the last renormalization, $f_N = Rf_{N-1}$, is doubling*).
(ii) *Any proper return map $g : T \to T$ to a periodic interval T coincides with some f_n.*

PROOF. (i) follows from Lemma 30.25.

(ii) Let p_n be the periods of the $\mathcal{I}^n$, and let p be the period of T. Since the sequence (p_n) is strictly increasing, there is the biggest n such that $p_n \le p$. The corresponding renormalization interval $\mathcal{I}^n$ is the smallest one containing T.

If $p > p_n$, then by Proposition 30.22, $T = \mathcal{I}^{n+1}$ and $g = f_{n+1}$, contradicting the choice of n. Hence $p_n = p$, and then $T = \mathcal{I}^n$ since $\mathcal{I}^n$ is the only f_n-invariant interval on which f_n is proper. □

30.7.4. *Real vs complex renormalizations.* The following statement shows that for real ql maps, notions of real and complex renormalizations match:

PROPOSITION 30.28. *Let $f : U \to U'$ be a real-symmetric quadratic-like map. Then f is DH renormalizable* with some period p if and only if its restriction to the real line is really renormalizable* with the same period p. Moreover, the real renormalization* of f is obtained by taking the real slice of the corresponding complex renormalization*.*

The complex renormalization* can be thickened to the genuine one iff the corresponding real renormalization* can. It is the case, unless f is Myrberg (*and, in particular, the renormalization R is doubling*).

PROOF. We let $\mathcal{K} \equiv \mathcal{K}(f)$, so $\mathcal{I} = \mathcal{K} \cap \mathbb{R}$.

Assume f is DH renormalizable* with period p, and let $f^p : \Omega \to \Omega'$ be its degenerate ql renormalization associated with a periodic cut-point α (as in §28.4.3).

Since f is real-symmetric, $g \equiv f^p : \overline{\Omega} \to \overline{\Omega}'$ is a degenerate ql map associated with a periodic cut-point $\bar{\alpha}$ (where "bar" stands for the complex conjugacy). Since the DH renormalization* with period p is canonically defined (Corollary 28.29), the point α is real and the domains Ω, Ω' are real-symmetric.

Hence the little filled Julia set $\mathcal{K}' \equiv \mathcal{K}(g)$ is also real-symmetric, with the real slice $\mathcal{I}' = [\alpha, -\alpha]_\#$. The interval $\mathcal{I}'$ is a periodic interval of period p such that

$f^p\,|\,\mathcal{I}'$ is a proper unimodal map with the fixed point $\alpha \in \partial\mathcal{I}$ being repelling or parabolic. Hence it is the real renormalization* of $f : \mathcal{I} \to \mathcal{I}$ with period p.

Vice versa, assume that the map $f : \mathcal{I} \to \mathcal{I}$ is really renormalizable* with period p, and let $g \equiv f^p : \mathcal{I}' \to \mathcal{I}'$ be the corresponding renormalization. Then $\mathcal{I}' = [\beta, -\beta]_\#$, where β is either repelling or parabolic periodic point. In either case, we obtain the associated $\mathbb{R}$-symmetric ray portrait $\mathfrak{R}(\beta)$ and the corresponding $\mathbb{R}$-symmetric puzzle piece $\Omega' \ni 0$.

Since $\mathcal{I} \subset \mathcal{K}$, the rays under consideration do no cross $\mathcal{I}$, implying that $\Omega'_\mathbb{R} = \mathcal{I}'$. Since f is really renormalizable* $f^{pn}(0) \in \mathcal{I}'$, $n = 0, 1, \dots$. All the more, $f^{pn}(0) \in \Omega'$, $n = 0, 1, \dots$, implying that f is DH renormalizable* with period p.

Finally, $g : \Omega \to \Omega'$ can be thickened to a genuine ql renormalization unless α is a satellite parabolic point. On the other hand, $g : \mathcal{I}' \to \mathcal{I}'$ belongs to class $\mathfrak{G}$ unless α is a flip parabolic point. But for real maps, α is satellite parabolic iff it is flip parabolic. Thus, the complex renormalization* is genuine (after the thickening) iff the real one is. □

c

30.8. Structure of non-renormalizable* interval maps.

30.8.1. *No wandering intervals.* Trivial wandering intervals obviously exist. However, we have:

Lemma 30.29. *Let $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$ be a non-renormalizable* interval map with both fixed points repelling. Then:*

(i) *All periodic points of f are repelling;*

(ii) *f does not have wandering intervals* (*equivalently: f does not have homtervals*).[17]

Proof. (i) By assumption, both fixed points of f are repelling. But so is any periodic point of period $p > 1$, for otherwise the immediate basin of this point would be a cycle of intervals of period $p > 1$ containing 0 (see Exercises 30.6 and 30.8), and the map would be renormalizable.

Thus, f does not have trivial wandering intervals.

(ii) Let L be a non-trivial wandering interval. Without; loss of generality, we can assume that L is a maximal wandering homterval (see Exercises 30.15 and 30.16).

By Lemma 30.21, the orbit of L accumulates on 0. Let us consider the moments $0 = m_0 < m_1 < m_2 < \dots$ of closest approached of the $\operatorname{orb} L$ to 0, i.e., for any $k \in \mathbb{Z}_+$, the interval $f^{m_k}(L)$ lies closer to 0 than $f^{m_{k-1}}(L)$, while all the intervals $f^n(L)$, $n = m_{k-1} + 1, \dots m_k - 1$, lie farther away. Then $f^{m_k}(L) \to 0$.

Since $|f^{m_k}(L)| \to 0$, we can take a further subsequence $Z = (m_{k(j)})$, satisfying

$$|f^{m_k}(L)| < |f^{m_{k-1}}(L)| \quad \text{for } m_k \in Z. \tag{30.5}$$

Let us use notation L_k and U_k respectively for the intervals $f^{m_k}(L)$ and $f^{m_{k-1}}(L)$ with $m_k \in Z$, i.e., satisfying (30.5). We let $U'_k = -U_k$ be the symmetric interval. Since the intervals $f(L_k)$ and $f(U'_k) = f(U_k)$ are disjoint, the intervals L_k and U'_k are disjoint as well. So, we obtain three disjoint intervals, U_k, L_k and U'_k, such that L lies in between U_k and U'_k

[17]In the presence of a non-repelling fixed point, f has still only trivial wandering intervals.

Let us consider the smallest 0-symmetric interval $S_k \supset L_k$ containing U_k. Let $H_k^{\pm}$ be the components of $S_k \smallsetminus L_k$. By (30.5),

$$|H_k^{\pm}| \geq |U_k| > |L_k|. \tag{30.6}$$

Let us now pull the interval S_k back along the orbit $(f^n(L))_{n=0}^{m_k}$. We obtain a pair of intervals (Q_k, L) mapped under f^{m_k} to (S_k, L_k).

Lemma 30.30. *The map $f^{m_k} : Q_k \to S_k$ is a homeomorphism.*

Proof. Otherwise, there exists a moment $n < m_k$ such that $f^n(Q_k) \ni 0$. Let V be the smallest 0-symmetric interval containing $f^n(L)$. Since $U_k \equiv f^{m_{k-1}}(L)$ is closer to 0 than any other interval $f^n(L)$ with $n < m_k$, V contains S_k, and hence $f^{n+1}(Q_k) \supset f(V) \supset f(S_k)$. It follows that

$$S_k = f^{m_k-n-1}(f^{n+1}(Q_k) \supset f^{m_k-n}(V) \supset f^{m_k-n}(S_k).$$

Thus, S_k is invariant under f^{m_k-n}. But non-renormalizable* maps with both fixed points repelling do not have periodic intervals around 0 of period > 1. □

Since L is a maximal homterval, the intervals Q_k shrink to L. Let $W_k^{\pm}$ be the preimages of the $H_k^{\pm}$ under $f^{m_k} : Q_k \to S_k$. Then $|W_k^{\pm}| \to 0$. Together with (30.6) this contradicts Corollary 20.29. □

30.8.2. *Topological exactness (leo property).*

Proposition 30.31. *Let $f : I \to I$ be a non-renormalizable map of class $\mathfrak{G}$ with both fixed points repelling. Then its restriction to the minimal invariant interval $\mathcal{T}$ is topologically exact (and hence topologically transitive).*

Proof. For any interval $L \subset \mathcal{T}$, we need to find an $n \in \mathbb{N}$ such that $f^n(L) = \mathcal{T}$. If $\operatorname{int} L \ni \alpha$, this is the content of Corollary 30.19. Since f is non-renormalizable, the preimages of α accumulate on 0, so the conclusion holds if $L \ni 0$.

By Lemma 30.29, f does not have homtervals. Hence for any interval L, we have $f^n(L) \ni 0$ for some $n \in \mathbb{N}$, and the conclusion follows. □

Corollary 30.32. *For a non-renormalizable* map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$ with both fixed points repelling, we have:*

(i) *For any $x \in \mathcal{T}$, the grand backward orbit $\operatorname{Orb}_-(x)$ is dense in $\mathcal{T}$;*

(ii) *The set of periodic points is dense in $\mathcal{T}$.*

Proof. (ii) We see that for any interval $L \subset \mathcal{T}$, $f^n(L) \supset L$ for some $n \in \mathbb{N}$. Then L contains a periodic point of period n (not necessarily minimal), by the Intermediate Value Theorem. □

30.9. Renormalizable case: real hyperbolic sets.

Proposition 30.33. *Let $f : \mathcal{I} \to \mathcal{I}$ be a renormalizable* map of class $\mathfrak{G}$, and let $g \equiv f^p : \mathcal{I}' \to \mathcal{I}'$ be its first renormalization*. Assume that the boundary fixed point $\beta' \in \partial\mathcal{I}'$ of g is repelling. Then:*

(i) *The set $Q := \{x \in \mathcal{T} : \ f^n x \notin \operatorname{int} \mathcal{I}', \quad n = 0, 1, 2, \dots\}$ is hyperbolic;*

(ii) *The restriction $f|Q$ is topologically conjugate to a Markov shift σ_A;*

(iii) *The restriction of f to $Q \cap \mathcal{T}$ is topologically exact (leo property);*

(iv) *All orbits eventually land either in the cycle $\operatorname{orb} \mathcal{T}'$ or in Q.*

Together with Proposition 25.23 (or rather its version for class $\mathfrak{G}$) this implies:

COROLLARY 30.34. *Under the above circumstances, almost all orbits eventually land in the cycle of* $\mathcal{I}'$.

PROOF. (i) Let $\mathcal{T}' \subset \mathcal{I}'$ be the minimal invariant interval of g, and let

$$\Upsilon' := \operatorname{Orb} \mathcal{T}' \equiv \bigcup_{k=0}^{p-1} f^k(\mathcal{T}') \subset \bigcup_{k=0}^{p-1} f^k(\mathcal{I}') \equiv \operatorname{orb} \mathcal{I}' =: O'. \tag{30.7}$$

This is a forward invariant compact set containing the postcritical set $\overline{\mathcal{P}}_f$. Moreover, $\Omega' := \mathbb{C} \smallsetminus \Upsilon'$ is connected and $f^{-1}(\Upsilon')$ is strictly bigger than Υ' (e.g., $f^{-1}(\Upsilon') \smallsetminus \Upsilon'$ contains the intervals that are 0-symmetric with $\operatorname{int}(f^k(\mathcal{T}'))$, $k \neq 0$). By Corollary 20.32, $Q \subset \Omega$ is hyperbolic.

(ii) The doubling case is elementary (see Exercise 30.35 below), so assume that $p > 2$. Let L_i be the components of $\mathcal{I} \smallsetminus \operatorname{int} O'$, and let $X_i = L_i \cap Q$. Since $\partial \mathcal{I} \cup \partial O'$ is forward invariant and since $0 \in \operatorname{int} O'$, the sets X_i form a Markov partition of Q: each X_i is mapped homeomorphically onto the union of some of them. The conclusion follows from Lemma 19.92.

(iii) For $p > 2$, this follows from the exactness of $f \,|\, \mathcal{T}$ rel $\mathcal{I}'$ (Lemma 25.42). The doubling case is trivial: see Exercise 30.35 below.

(iv) By definition, all orbits eventually land either in O' or in Q. By Exercise 30.20, the former orbits eventually land in Υ'. $\square$

EXERCISE 30.35. *Show that the above hyperbolic set Q is either the $\beta-$fixed point and its preimage β' (in the doubling case) or a Cantor set (otherwise). Moreover, in the latter case, the corresponding Markov matrix A has spectral radius $r(A) > 1$.*

EXERCISE 30.36. *Prove that the above results remain true when the g-fixed point β' is parabolic (except that Q is not hyperbolic anymore).*

30.10. Structure of at most finitely renormalizable maps. Summarizing the above results, we obtain a good topological description of maps that are at most finitely renormalizable:

THEOREM 30.37. *Let $f : \mathcal{I} \to \mathcal{I}$ be a map of class $\mathfrak{G}$ which is exactly n times renormalizable*, $n \in \mathbb{N}$. Let $g = f^p : \mathcal{I}' \to \mathcal{I}'$ be its last renormalization*, and $\mathcal{T}' \subset \mathcal{I}'$ be its smallest invariant interval. Then*

(i) *either g has a non-repelling fixed point, and $\operatorname{int} \mathcal{I}'$ is its immediate basin of attraction,*

(ii) *or else g is topologically exact (leo) on $\mathcal{T}'$, and hence f is topologically transitive on the corresponding cycle of intervals, $\mathcal{A}^{\mathrm{t}} \equiv \mathcal{A}^{\mathrm{t}}_f := \operatorname{orb} \mathcal{T}'$.*

Moreover, the set

$$Q := \{x : \ f^n x \notin \operatorname{int} \mathcal{I}', \ n = 0, 1, 2, \dots\}$$

is a Cantor set of zero length on which the dynamics is topologically conjugate to a Markov chain. If the $g(\beta') \neq 1$, where $\beta' \in \partial \mathcal{I}'$ is the boundary fixed point of g, then Q is hyperbolic.

EXERCISE 30.38. *Under the above circumstances, the point β' is f-periodic with period p or $p/2$. The letter happens iff $R^{n-1} f$ is renormalizable with period two.*

We will refer to $\mathcal{T}'$ as a *leo periodic interval*.

COROLLARY 30.39. *The cycle $\mathcal{A}^{\mathfrak{t}}$ of the leo periodic interval $\mathcal{T}'$ is the unique topological attractor of f, i.e., $\omega(x) = \mathcal{A}^{\mathfrak{t}}$ for a generic point $x \in \mathcal{I}$.*

30.11. Renormalization Filtration. Let us summarize the structural information accumulated so far.

Assume that a map $f \in \mathfrak{G}$ is exactly N times renormalizable, where $N \in [0, \infty]$. Then we have two intertwined nests of periodic intervals

$$\mathcal{I} \equiv \mathcal{I}^0 \supset \mathcal{T}^0 \supset \mathcal{I}^1 \supset \mathcal{T}^1 \cdots \supset \mathcal{I}^N \supset \mathcal{T}^N \tag{30.8}$$

of periods p_n such that $g_n := f^{p_n}|\mathcal{I}^n$ is the n-fold (pre-)renormalization of f. (In the case of $N = \infty$, the nests are infinite). The p_n are the *renormalization periods* of f. Moreover, the p_{n+1} are multiples of the p_n, and the ratios $q_{n+1} := p_{n+1}/p_n$ are the *relative renormalization periods* of f. Note that $R^n f$ is renormalizable with period q_{n+1} (compare §28.4.7).

The orbits of these intervals form two intertwined nests of cycles of intervals,

$$O^n := \operatorname{orb} \mathcal{I}^n \quad \text{and} \quad \Upsilon^n := \operatorname{orb} \mathcal{T}^n.$$

All orbits in $\operatorname{int} O^n$ eventually land in Υ^n. For $n < N$, the orbits of Υ^n that never land in $\operatorname{int} \mathcal{I}^{n+1}$ form an invariant compact set Q^{n+1} on which the dynamics is topologically conjugate to an irreducible Markov chain. It is called the *basic set of level n*. In the doubling case ($q_{n+1} = 2$) Q^{n+1} is a repelling periodic cycle; otherwise it is a Cantor set. Moreover, the dynamics on this Cantor set expanding, except when $n = N$ and f has a primitive parabolic point. (Notice that Q^0 is the $\beta-$fixed point.)

If $N < \infty$ then Υ^n is either contained in the immediate basin on an attracting or a parabolic cycle, or else, it is a cycle of a leo interval. Notice that in the former case, the parabolic cycle can be satellite (flip), in which case it belongs to $\operatorname{int} \Upsilon^n$ and $R^N f$ is renormalizable* with period 2.

This structure will be called the *Renormalization Filtration*.

30.12. Real *a priori* bounds.

30.12.1. *Extensions.* Let us consider a renormalizable* map $f \in \mathfrak{G}$ with a renormalization* $g : \mathcal{I}_0 \to \mathcal{I}_0$ of period p (not necessarily the first one). Let

$$\mathcal{I}_k := f^k(\mathcal{I}_0), \quad k = 1, \dots, p.$$

Note that $\mathcal{I}_1$ and $\mathcal{I}_2$ are the extreme intervals among the $\mathcal{I}_k$ (in the sense of the order on $\mathbb{R}$). For $k \neq 1, 2$, let $\hat{\mathcal{I}}_k$ be the maximal open interval containing $\mathcal{I}_k$ but not containing other intervals $\mathcal{I}_j$, $j \neq k$ (i.e., it is the convex hull of two intervals $\operatorname{int} \mathcal{I}_j$ that lie next to $\mathcal{I}_k$). For an integer $k \in [1, p]$, let $\Delta^k \ni v$ be the pullback of $\hat{\mathcal{I}}_k$ under f^{k-1}.

LEMMA 30.40. *Each map*

$$f^{k-1} : (\Delta^k, \mathcal{I}_1) \to (\hat{\mathcal{I}}_k, \mathcal{I}_k), \quad k = 1, \dots, p,$$

is a diffeomorphism.

PROOF. Otherwise, there exists a moment $m \in [0, k-2]$ such that $f^m(\Delta^k) \ni 0$. But then $f^m(\Delta^k)$ contains one-half, $\mathcal{I}_+$ or $\mathcal{I}_-$, of the interval $\mathcal{I}_0$ (say, $\mathcal{I}_+$). Then

$$f^{m+1}(\Delta^k) \supset f(\mathcal{I}_+) = \mathcal{I}_1. \tag{30.9}$$

Applying f^{k-m-2}, we conclude that $\hat{\mathcal{I}}_k \supset \mathcal{I}_{k-m-1}$, contradicting the definition of $\hat{\mathcal{I}}_k$. □

COROLLARY 30.41. *For any integer $j \in [1, p-1]$, the map*

$$f^{p-j} : \mathcal{I}_j \to \mathcal{I}_p$$

admits a diffeomorphic extension with range $\hat{\mathcal{I}}_p$.

PROOF. This extension is given by $f^{p-j} : f^{j-1}(\Delta^p) \to \hat{\mathcal{I}}^p$. □

30.12.2. *Map dependent bounds.*

REAL A PRIORI BOUNDS. *Let $f \in \mathfrak{G}(\mu)$ be a renormalizable map with period p (not necessarily the smallest one), and let $g = f^p : \mathcal{I}_0 \to \mathcal{I}_0$ be its renormalization. Then:*

(i) *$g \in \mathfrak{G}(\nu)$, with $\nu > 0$ depending only on μ.*

(ii)

$$\operatorname{mod}_{\mathbb{R}}(\hat{\mathcal{I}}_j : \mathcal{I}_j) \geq \delta(\mu) > 0, \quad j = 0, 1, \ldots, p-1, \tag{30.10}$$

where mod_R *is defined by* (6.2).

(iii) *The maps $f^{p-j} : \mathcal{I}_j \to \mathcal{I}_p$ are diffeomorphisms with bounded distortion (that depends on μ only).*

PROOF. Due to Exercise 30.2, we can assume without loss of generality that f belongs to Epstein class $\mathfrak{E}(\mu)$.

Let $\mathcal{I}_l$ be a minimal length interval in the cycle $(\mathcal{I}_k)_{k=0}^{p-1}$. Assume first that $l \neq 1, 2$. Then we have the well defined enlargement $\hat{\mathcal{I}}_l$, and by the minimality property,

$$\operatorname{mod}_{\mathbb{R}}(\hat{\mathcal{I}}_l : \mathcal{I}_l) \geq 1. \tag{30.11}$$

By Lemma 30.40, the map $f^{l-1} : \mathcal{I}_1 \to \mathcal{I}_l$ extends to a diffeomorphism $f^{l-1} : \Delta^l \to \hat{\mathcal{I}}_l$. By Corollary 6.18, the space of (30.11) around $\mathcal{I}_l$ can be pulled back to an absolute space around a valuable interval $\mathcal{I}_1$:

$$\operatorname{mod}_{\mathbb{R}}(\Delta^l : \mathcal{I}_1) \geq \varepsilon > 0.$$

Since each monotonicity branch of f is $\kappa(\mu)$-qs (and f is even),[18] this space can be pulled further back to a definite space around $\mathcal{I}_0$:

$$\operatorname{mod}_{\mathbb{R}}(\hat{\mathcal{I}}_0 : \mathcal{I}_0) \geq \delta_0(\mu) > 0. \tag{30.12}$$

If $l \in \{1, 2\}$ then let $\hat{\mathcal{I}}_l$ be the convex hull of $\mathcal{I}_l$ and $\operatorname{int} \mathcal{I}_j$, where $\mathcal{I}_j$ is the interval next to $\mathcal{I}_l$. Pulling it back by f^l, which is $\kappa(\mu)$-qs, we obtain (30.12) once again.

By Corollaries 6.18 and 30.41, the space around $\mathcal{I}_0$ can be pulled back to a space around any $\mathcal{I}_j$, yielding (30.10). This proves (i) and (ii).

Assertion (iii) follows from the Koebe Distortion Theorem. □

[18] Under these circumstances, we will somewhat loosely say that f itself is κ-qs.

Let $L : \bigcup_i \Delta_i \to \mathcal{I}_0$ be the first landing map to $\mathcal{I}_0$. Here the intervals Δ_i are the pullbacks of $\mathcal{I}_0$ corresponding to the first landings $f^m x \in \operatorname{int} \mathcal{I}_0$ of points $x \in \mathcal{I}$ (that ever land). The intervals Δ_i have disjoint interiors.

Exercise 30.42. *Under the above circumstances, we have:*

(i) *Each branch $L : \Delta_i \to \mathcal{I}_0$ admits an extension to a diffeomorphism $\hat{\Delta}_i \to \hat{\mathcal{I}}_0$.*

(ii) *The branches $L : \Delta_i \to \mathcal{I}_0$ have a bounded distortion (depending only on μ in the Real A Priori Bounds statement).*

See §§31.2 and 31.11 below for a related discussion of the first landing maps in the puzzle context.

30.12.3. *Beau bounds.* Let us say that *Real Beau Bounds are valid over class* $\mathfrak{G}$ if there exists an absolute $\nu > 0$ with the following property: For any $\mu > 0$ there exists $N = N(\mu)$ such that for any $(n \geq N)-$times renormalizable map $f \in \mathfrak{G}(\mu)$, the renormalization $R^n f$ belongs to $\mathfrak{G}(\nu)$.

Real Beau Bounds. *Real Beau Bounds are valid over class $\mathfrak{G}$. In particular, the bounds are absolute over the quadratic family.*

Proof. Let us go back to the proof of real *a priori* bounds in the previous section. First, due to Exercise 30.2(iii), we can still assume that f belongs to the Epstein class $\mathfrak{E}(\mu)$.

Next, notice that the bounds are absolute as long as $|\mathcal{I}_0| < \varepsilon |\mathcal{I}|$ with some $\varepsilon > 0$ depending only on μ. Indeed, dependence of the bounds on μ appears only through (30.12), which depends only on the qs-dilatation of $f^l \,|\, \mathcal{I}_0$, $l \in \{1, 2\}$. The latter is small as long as $\mathcal{I}_0$ is sufficiently small rel $\mathcal{I}$ (depending on μ only).

Finally, $|\mathcal{I}^n| \leq \sigma |\mathcal{I}^{n-1}|$ with some $\sigma = \sigma(\mu) \in (0, 1)$. Hence $|\mathcal{I}^n| \leq \varepsilon \cdot |\mathcal{I}|$ for $n \geq \log \varepsilon / \log \sigma$, and the conclusion follows. $\square$

In global terms, beau bounds can be formulated as follows:

Class $\mathfrak{G}$ contains an absorbing compact subset $\mathfrak{K}$, so, for any compact subset $Y \subset \mathfrak{E}$ there is an N such that for any $(n \geq N)-$renormalizable map $f \in Y$, we have $R^n f \in \mathfrak{K}$.

Remark 30.43. There is a version of *a priori* bounds for non-renormalizable maps: see §46.2 below.

30.13. Infinitely renormalizable maps.

30.13.1. *Intersection of nested cycles of intervals.* Assume that a map $f \in \mathfrak{G}$ is infinitely renormalizable.

Let $\mathcal{I}^n_k := f^k(\mathcal{I}^n)$, $k = 0, 1, \dots, p_n - 1$, so these intervals (for a given n) form a cycle O^n_f of period p_n containing the postcritical set $\overline{\mathcal{P}}_f$. The intersection of these cycles,

$$O_f := \bigcap_{n=0}^{\infty} O^n_f \tag{30.13}$$

is an invariant compact set containing the postcritical set $\overline{\mathcal{P}}_f$ as well. We will show that these two sets actually coincide, and have empty interior (so they are Cantor). By the following observation, this amounts to showing that f does not have wandering intervals.

LEMMA 30.44. *Any component of* $\operatorname{int} O_f$ *is a non-trivial wandering interval.*

PROOF. Let J be a component of $\operatorname{int} O_f$. Then for any $n \in \mathbb{N}$, J is contained in some periodic interval $\mathcal{I}^n_{k_n}$ of period p_n. Hence the intervals $f^m(J) \subset I^n_{k_n+m}$, $m = 0, 1, \dots, p_n - 1$, are pairwise disjoint. Since $p_n \to \infty$, J is wandering.

If J was trivial, it would be contained in the basin $\mathcal{D}(\boldsymbol{\alpha})$ of some non-repelling cycle of some period q. But then periodic intervals $\mathcal{I}^n_{k_n}$ would be eventually contained in $\mathcal{D}(\boldsymbol{\alpha})$ as well, which is impossible for $p_n > q$ (exercise). □

Note in conclusion that the set O_f is the real slice of the postcritical impression $\mathcal{O}_f$ (28.8), so according to our notational conventions, $O \equiv \mathcal{O}_{\mathbb{R}}$, and we can refer to O as the *real postcritical impression*.

30.13.2. *Solenoids.* We will derive the absence of wandering intervals from the real *a priori* bounds.

LEMMA 30.45. *A real infinitely renormalizable map* $f \in \mathfrak{G}$ *does not have wandering intervals.*

PROOF. Let J be a wandering interval. Without loss of generality, we can assume that J is a maximal (open) homterval (see Exercise 30.15). Let T^n be the intervals of monotonicity of the iterates f^n containing J. They form a nest shrinking to J (see Exercise 30.16).

For any $n \in \mathbb{N}$, let us consider the periodic interval $\mathcal{I}^n \ni 0$. Since $\partial\mathcal{I}^n$ consists of a periodic point and its f^{p_n}-preimage, the intervals $J_m \equiv f^m(J)$ do not intersect $\partial\mathcal{I}^n$. So, either $J_m \subset \operatorname{int} \mathcal{I}^n$ or $J_m \cap \mathcal{I}^n = \emptyset$ (for any m).

By Lemma 30.21, $\omega(J) \ni 0$, implying that for any $n \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ such that $J_m \subset \mathcal{I}^n$. Let $m = m_n$ be the first such landing moment.

By Exercise 30.42, the map $f^m : J \to J_m$ admits an extension to a diffeomorphism $f^m : \hat{\Delta}^n \to \hat{\mathcal{I}}^n$. Let Δ^n be the pullback of $\mathcal{I}^n$ under this diffeomorphism.

By the real *a priori* bounds (30.10) and Koebe, the maps $f^m : \Delta^n \to \mathcal{I}^n$ have a bounded distortion. Hence

$$\frac{|\Delta^n \smallsetminus J|}{|J|} \asymp \frac{|\mathcal{I}^n \smallsetminus J_m|}{|J_m|}.$$

However, the left-hand side goes to 0 as $n \to \infty$ (since the intervals Δ^n shrink to J), while the right hand side is bounded by 1 from below (since J_m is contained in one-half of the interval $\mathcal{I}^n$). □

For the relative renormalization periods $q_n = p_n/p_{n-1}$, let us consider the adding machine $\tau_{\mathbf{q}}$ on the adic ring $\mathbb{Z}_{\mathbf{q}}$ (see §19.16.2).

COROLLARY 30.46. (i) *The set* O_f (30.13) *is a Cantor set coinciding with the postcritical set* $\overline{\mathcal{P}} \equiv \overline{\mathcal{P}}_f$. *In particular,* $\overline{\mathcal{P}} \ni 0$.
(ii) *The dynamics on* $\overline{\mathcal{P}}$ *is topologically conjugate to the adding machine* $\tau_{\mathbf{q}}$ (*odometer*). *In particular,* $f | \overline{\mathcal{P}}$ *is minimal, so the orbit of any point* $x \in \overline{\mathcal{P}}$ *is recurrent and dense in* $\overline{\mathcal{P}}$.

The sets $O_f = \overline{\mathcal{P}}_f$ are called *solenoids*. At the same time, they could also be called the *adding machines* or *odometers*, as well as *Feigenbaum attractors* (see next section).

PROJECT 30.47. Lemma 30.45 *can be proved without using a priori bounds, along the lines of* Lemma 30.29. *Try to work it out.*

30.13.3. *Feigenbaum attractor.* We are ready to prove that $\overline{\mathcal{P}} \equiv \overline{\mathcal{P}}_f$ is a unique attractor for f, in both topological and measure-theoretic sense:

THEOREM 30.48. *Let $f \in \mathfrak{G}$ be infinitely renormalizable. Then $\omega(x) = \overline{\mathcal{P}} \ni 0$ for a.e. $x \in \mathcal{I}$ and for a generic $x \in \mathcal{I}$.*

PROOF. Let O^n be the cycle of intervals $\bigcup_{k=0}^{p_n-1} \mathcal{I}_k^n$ of level n (compare (30.7)). By Proposition 30.33 and Corollary 30.34, each set $X_n := \{x : \omega(x) \subset O^n\}$ is everywhere dense and has full Lebesgue measure. Hence the intersection

$$\bigcap_n X_n = \{x : \omega(x) \subset \overline{\mathcal{P}}\}$$

is a set of full Baire category and full Lebesgue measure. Thus, $\omega(x) \subset \overline{\mathcal{P}}$ for a generic $x \in \mathcal{I}$ and for a.e. $x \in \mathcal{I}$.

Since the dynamics on $\overline{\mathcal{P}}$ is minimal (by Corollary 30.46),

$$\omega(x) \subset \overline{\mathcal{P}} \implies \omega(x) = \overline{\mathcal{P}}.$$

The conclusion follows. □

For this reason, the sets $O_f = \overline{\mathcal{P}}_f$ are also called *Feigenbaum attractors*.

30.13.4. *Bounded geometry.* An infinitely renormalizable map with bounded combinatorics is called *Feigenbaum*.

THEOREM 30.49. *Let $f \in \mathfrak{G}$ be a real Feigenbaum map. Then the attractor $O_f = \overline{\mathcal{P}}_f$ is a Cantor set with bounded geometry.*

PROOF. We let $\mathcal{T}^n$ be the smallest periodic interval of level n and $\mathcal{T}_k^n := f^k(\mathcal{T}^n)$, $k = 0, 1, \dots, p_n - 1$.

By the Real *A Priori* Bounds (i), the sequence of the renormalizations $\{R^n f\}_{n=0}^{\infty}$ belongs to a class $\mathfrak{G}(\nu)$ with some $\nu > 0$, and is hence pre-compact. Since the condition of being renormalizable with a bounded combinatorics is closed (by Exercise 30.24), the configurations of intervals $\mathcal{T}_k^{n+1}$ inside the $\mathcal{T}^n$, $n \in \mathbb{N}$, form a pre-compact family as well. Since the intervals $\mathcal{T}_k^{n+1}$ do not touch each other (for a given n), these configurations have a bounded geometry.

By means of the Real *A Priori* Bounds (iii), this property spreads around to all intervals $\mathcal{T}_j^n$: inside each of them, the configuration of the intervals $\mathcal{T}_k^{n+1}$ has a bounded geometry (uniformly over (n, j)). □

Since Cantor sets with bounded geometry have zero measure (see Exercise 19.117), we conclude:

THEOREM 30.50. *For any real infinitely renormalizable map $f \in \mathfrak{G}$,*

$$\operatorname{length} \overline{\mathcal{P}}_f = 0.$$

30.14. Topological Structure of unimodal maps.

30.14.1. *No Wandering Intervals.* Putting together Exercises 30.7 (hyperbolic case) and 30.9 (parabolic case), Lemmas 30.29 (at most finitely renormalizable case) and 30.45 (infinitely renormalizable case), we obtain the following fundamental result:

No Wandering Intervals Theorem. *A map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$ does not have non-trivial wandering intervals. In particular, an irregular map f* (in the sense of §30.2.3) *of class $\mathfrak{G}$ does not have wandering intervals at all.*

Corollary 30.51. *For an irregular map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$, the set* Crit^∞ *of precritical points is dense in $\mathcal{I}$.*

30.14.2. *Topological attractors: classification.* Topological attractors were defined in §19.7. Putting together Exercises 30.7 (hyperbolic case) and 30.9 (parabolic case), Corollary 30.39 (at most finitely renormalizable case) and Proposition 30.48 (infinitely renormalizable case), we obtain a *compete description* of topological attractors for quadratic polynomials:

Theorem 30.52. *Any unimodal map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}$ has a unique topological attractor $\mathcal{A}^{\mathfrak{t}} \equiv \mathcal{A}^{\mathfrak{t}}_f$ attracting generic orbits: $\omega(x) = \mathcal{A}^{\mathfrak{t}}$ for all $x \in \mathcal{I}$ except for a set of first Baire category. Moreover, $\mathcal{A}^{\mathfrak{t}}$ is either an attracting or parabolic cycle, or a cycle of a leo periodic interval, or a Feigenbaum attractor.*

30.14.3. *Topological structure of $\overline{\mathrm{Per}_{\mathbb{R}}(f)}$.* Let $\mathrm{Per}_{\mathbb{R}}(f)$ stand for the set of real periodic points of a map $f \in \mathfrak{G}$, and let $\Omega(f)$ be the set of *strongly non-wandering points* (which is the negation of the "weakly wandering" property introduced in §19.1):

$$\Omega_{\mathbb{R}}(f) = \{x \in \mathcal{I} : \text{ for any nbd } V \ni x \ \exists\ y \in V, n \in \mathbb{Z}_+ \text{ s.t. } f^n y \in V\}$$

i.e., $\Omega_{\mathbb{R}}(f)$ is the maximal set with the property that any neighborhood intersecting it contains a returning point.

Exercise 30.53. *$\Omega_{\mathbb{R}}(f)$ is a closed invariant set containing all recurrent* (*and hence all periodic*) *points.*

Putting together the Renormalization Filtration structure (§30.11) and Theorem 30.52, we obtain:

Theorem 30.54. *Let $f \in \mathfrak{G}$ be N time renormalizable map, $n \in [0, \infty]$. Then:*

(i) $\Omega_{\mathbb{R}}(f) = \overline{\mathrm{Per}_{\mathbb{R}}(f)} = \mathcal{A}^{\mathfrak{t}} \cup \bigcup_{n=0}^{N-1} Q^n$, *where $\mathcal{A}^{\mathfrak{t}}$ it the global topological attractor for f described in* Theorem 30.52*, and each basic set Q^n is an invariant compact set on which the dynamics is topologically conjugate to an irreducible Markov chain.*

(ii) *Each basic set Q^n is expanding, except of Q^N in the case when f has a primitive parabolic point.*

(iii) *The closure of real repelling periodic points is equal to $\Omega_{\mathbb{R}}(f)$ with the attracting or parabolic flip cycle removed* (*if exists: the latter is topologically attracting*)*.*

In §30.12 we defined the real Julia set $\mathcal{J}_{\mathbb{R}}(f)$ for a regular map $f \in \mathfrak{G}$. A natural version for an irregular map $f \in \mathfrak{G}$ would be the set of points whose orbits land in $\Omega_{\mathbb{R}}(f)$. This definition looks quite satisfactory for finitely renormalizable maps, but is more problematic for infinitely renormalizable ones:

EXERCISE 30.55. *The orbit of a point $x \in \mathcal{I}$ is Lyapunov unstable iff it lands in one of the basic sets Q^n or in the leo periodic interval (if exists).*

Also, even in the "satisfactory" cases there are some issues. Namely, for any irregular real polynomial $f = f_c$, except one special case (which one?)

– the "real Julia set" $\mathcal{J}_{\mathbb{R}}(f)$ is not the real slice of the complex one, $\mathcal{J}(f)$;

– $\mathcal{J}_{\mathbb{R}}(f)$ does not coincide with the closure of repelling periodic points.

So, we will refrain from attempting to select an "official" definition for the "real Julia set".

EXERCISE 30.56. *Let $f : \mathcal{I} \to \mathcal{I}$ be an irregular map of class $\mathfrak{G}$. Show that for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any interval $J \subset \mathcal{I}$ of length $\geq \varepsilon$ we have:*

$$|f^n(J)| \geq \delta, \quad n = 0, 1, 2, \dots.$$

30.15. Appendix: Negative Schwarzian derivative. Condition of negative Schwarzian derivative plays an important role in the one-dimensional dynamics. The first observations is that the maps $x \mapsto ax^d + c$, $d \in (1, +\infty)$ (and in particular, real quadratic polynomials!) have negative Schwarzian derivative. So, any C^3-map has this property near a non-degenerate critical point. Moreover, by the Chain Rule, the composition of two functions with negative Schwarzian derivative inherits this property. Hence the condition it is dynamically natural: all the iterates of a map with negative Schwarzian derivative inherit this property. Moreover, it allows one to control, in a uniform way, distortion of the iterates (similarly to what the Schwarz Lemma and the Koebe Distortion Theorem provide for us in the holomorphic case): below we formulate several statements to this effect (when we assume that $Sg > 0$, think of the inverse of a map with $Sf < 0$).

30.15.1. *Distortion Control.*

MINIMUM PRINCIPLE. *Let $f : I \to I'$ be a C^3 interval diffeomorphism with $Sf < 0$. Then $|f'|$ does not have local minima in $\operatorname{int} I$. Thus, the infimum of $|f'|$ is attained on the boundary of I.*

PROOF. Assume $|f'|$ attains a local minimum minimum at $0 \in \operatorname{int} I$. As replacing f with $af + b$ does not affect the Schwarzian derivative, we can normalize f so that $f(0) = 0$, $f'(0) = 1$. Then the Taylor expansion of f at 0 begins as $f(x) \sim x - bx^3$ with $b > 0$ (to ensure $Sf(0) < 0$). But then $f'(x) \sim 1 - 3bx^2$, making 0 a local maximum point for f'. □

LEMMA 30.57. *For a diffeomorphism f as above, any fixed point $\alpha \in \operatorname{int} I$ is topologically repelling.*

PROOF. Without loss of generality we can assume that f is orientation preserving. If α is not topologically repelling then $0 < f'(\alpha) \leq 1$, and on each non-repelling side of α there is a non-attracting fixed point β, $f'(\beta) \geq 1$. If $f'(\alpha) = 1$ then by the Minimum Principle $f'(x) > 1$ for all $x \in (\alpha, \beta)$, making $|f(\beta) - f(\alpha)| > |\beta - \alpha|$. If $f'(\alpha) < 1$ then there are non-attracting fixed points β, β' on both sides of α, contradicting the Minimum Principle. □

COROLLARY 30.58. *A diffeomorphism f as above, can have at most one fixed point in $\operatorname{int} I$.*

SCHWARZ CONTRACTION PROPERTY. *Let $g : I' \to I$ be a C^3 interval diffeomorphism. Let us supply I and I' with the hyperbolic metrics. Then g is strictly contracting iff $Sg > 0$.*

PROOF. Let us first show the macroscopic contraction:

$$\text{for any interval } J' \Subset \operatorname{int} I', \quad l_{\text{hyp}}(g(J')) < l_{\text{hyp}}(J'). \tag{30.14}$$

Since affine changes of variable do not affect the sign of the Schwarzian derivative, we can assume without loss of generality that $I = I' = \mathbb{I} \equiv [-1, 1]$ and that f is orientation preserving. Since Möbius changes of variable do not effect it either, we can assume that the left endpoints of J' and $g(J')$ are placed at 0.

By Corollary 30.58, ± 1 and 0 are the only fixed points of g, and 0 is topologically attracting (as $Sg > 0$). Hence $g(x) < x$ on $(0, 1)$ implying (30.14).

Let us now show the infinitesimal contraction:

$$\|Df(x)\|_{\text{hyp}} < 1 \quad \text{for all } x \in I'.$$

Again, normalize g as above with x placed at 0. Then $g((1/2)\,\|) \subset (\rho/2)\,\mathbb{I}$ with some $\rho < 1$. Since the map $\rho^{-1}\, g : (1/2)\,\mathbb{I} \to (1/2)\,\mathbb{I}$ contracts the hypebolic metric of $(1/2)\,\mathbb{I}$, we conclude that $\|Dg(0)\|_{\text{hyp}} \le \rho$.

□

Notice that if $I = [a, d]$, $J = [b, c] \Subset \operatorname{int} I$ then

$$l_{\text{hyp}}(J) = \log \frac{(c-a)(d-b)}{(b-a)(d-c)} = \log \left(1 + \frac{|J|}{|L|}\right) + \log \left(1 + \frac{|J|}{|R|}\right),$$

where $L = (a, b)$, $R = (c, d)$. So, the Schwarz Contraction Property tells us that maps with positive Schwarzian derivative contract certain cross-ratios (which is natural to expect, as maps with vanishing Schwarzian derivative, being Möbius, preserve cross-ratios). Similarly, such maps contract or expand five other possible cross-ratios that can be formed from four points $\{a, b, c, d\}$.

KOEBE DISTORTION PROPERTY. *Let $f : I \to I'$ be a C^3 interval diffeomorphism with $Sf < 0$. Let $J' \subset I'$ be a subinterval such that both components of $I' \smallsetminus J'$ have length $\ge \varepsilon |I'|$. Then*

$$\frac{|f'(x)|}{|f'(y)|} \le C(\varepsilon) \quad \forall x, y \in J \equiv f^{-}(J').$$

PROOF. As in the proof of the Schwarz Contraction Property, let us normalize g so that it is orientation preserving, $I = I' = \mathbb{I}$, and $x = 0$. Let $y > 0$ for definiteness. By Lemma 30.57, $g(x) < x$ for $x > 0$ near 0 and $g'(0) \le 1$. Moreover, since the hyperbolic length of $[x, y]$ in $\mathbb{I}$ is bounded by $L(\varepsilon)$, we have $y \le l = l(\varepsilon) < 1$.

If $g'(z) < g'(y)$ for some $z \in (y, 1)$, then g' would have a maximum point in $(0, 1)$, contradicting the Minimum Principle. So, $g'(z) \ge g'(y)$ on $[y, 1]$, implying that $g'(y) \le 1/(1-l)$ (since the image of $[l, 1]$ is contained in $[0, 1]$). □

Finally, let us recall class $\mathfrak{U}_{\mathbb{R}}$ from §7.2.1:

LEMMA 30.59. *Any diffeomorphism $g : I' \to I$ of class $\mathfrak{U}_{\mathbb{R}}$ has positive Schwarzian derivative.*

PROOF. By Corollary 7.2, g contracts the hyperbolic metric, so $Sg > 0$ by the Schwarz Contraction Property. □

30.15.2. *S-unimodal maps.* A C^3 unimodal map $f : \mathcal{I} \to \mathcal{I}$ is called *S-unimodal* if $Sf(x) < 0$ outside the critical point. The above piece of analytical theory allows one to develop the dynamical theory of S-unimodal maps to the same extent as the the theory of real quadratic polynomials. It starts with the following result:

Singer Theorem. *Let f be a proper S-unimodal map that has an attracting or parabolic cycle $\boldsymbol{\alpha} = (f^n\alpha)_{n=0}^{p-1}$. Then the immediate basin $\mathcal{D}^\bullet(\alpha)$ is non-empty and contains the critical point.*

Corollary 30.60. *An S-unimodal map has at most one non-repelling cycle.*

Proof. Assume $\boldsymbol{\alpha}$ is attracting. Then the immediate basin $I := \mathcal{D}^\bullet(\alpha)$ is an open interval. If $\mathcal{D}^\bullet(\alpha)$ does not contain 0, then the return map f^p to I is a diffeomorphism with negative Schwarzian derivative and with an attracting fixed point inside, contradicting Lemma 30.57.

Assume now $\boldsymbol{\alpha}$ is parabolic. If it is topologically repelling then the inverse map $g = f^{-p}$ near α is strictly contracting (by the Schwarz Contraction Property), making the cycle $\boldsymbol{\alpha}$ repelling (rather than parabolic) for f. The contradiction implies that the immediate basin $I \equiv \mathcal{D}^\bullet(\alpha)$ is a non-empty open interval.

If $\mathcal{D}^\bullet(\boldsymbol{\alpha})$ does not contain the critical point, then the return map $F := f^p$ to I is a diffeomorphism. After replacing F with F^2 if needed, we can assume that $F \mid I$ is orientation preserving. Then each $\beta \in \partial I$ is a fixed point with $F'(\beta) \geq 1$, while $F'(x) < 1$ at some point $x \in \operatorname{int} I$, contradicting the Minimum Principle. □

Project 30.61. *Develop a dynamical theory of S-unimodal maps.*

One can go further and develop a theory for C^2 unimodal maps with a power-type critical point (i.e., of the form $\phi(x^d)$ with $d \in (1, +\infty)$, where ϕ is a local C^2 diffeomorphism). It is done by combining the negative Schwarzian property of the power map near the critical point with a Denjoy distortion control away from it. One subtlety to be aware of is that there is no Singer Theorem in this generality, and in fact, a C^2 unimodal map can have infinitely many attracting cycles.

Project 30.62. *Develop a dynamical theory of C^2 unimodal maps.*

Notes. The topological structure of non-invertible one-dimensional maps was first analyzed by Myrberg [**Myr1, Myr2**] and Sharkovsky [**Sha2**] in the 1960s. It was further refined in the 1970-80s by Guckenheimer [**Gu1**], Misiurewicz [**Mi1**], Jonker and Rand [**JR**], van Strien [**vS**], Blokh [**B1, B2**], and others. The Structural Theorem (30.54) summarizes these developments in the unimodal case, emphasizing the Renormalization viewpoint.

The No Wandering Intervals Theorem is due Guckenheimer [**Gu1**]. In [**L6**] it was generalized to multimodal maps with negative Schwarzian derivative and non-degenerate critical points. A key idea of [**L6**] was distortion bounds along pullbacks of intervals (that were introduced there under the name of "chains of intervals"). Further generalizations appeared in [**BL4**] and [**MMS**].

Real *a priori bounds*, with an application to the measure of the Feigenbaum attractor, were first proved in the doubling case by Guckenheimer [**Gu2**]. The general combinatorics was treated independently by Blokh-Lyubich [**BL3**] and Sullivan (see [**MvS**]). Their importance for the Renormalization Theory was articulated by Sullivan in the late 1980s.

In fact, Real One-Dimensional Dynamics was originated in Poincaré's thesis [**Poi2**] in the context of circle homeomorphisms. Further important step was made by Denjoy [**Den**] who proved absence of wandering intervals for C^2-circle diffeomorphisms (and gave an example of a C^1-diffeomorphism that has a wandering interval). As we have already mentioned, the Denjoy Distortion Estimate (Exercise 19.69) appeared in that work. Lemma 30.21 (asserting that $\omega(J) \ni 0$) is an adaptation of Denjoy's machinery to unimodal maps.

The condition of negative Schwarzian derivative was introduced to Dynamics by Singer [**Si**]. The Minimum Principle and the Singer Theorem appeared in the same paper. At about the same time, the Schwarzian derivative was used by Herman (without calling it by name) in the context of circle diffeomorphisms [**He3**]. The machinery was further developed by Yoccoz [**Y2**], Guckenheimer [**Gu2**] (the Koebe Distortion Principle appeared here), Swiatek [**Sw**], de Melo & van Strien [**MvS**], and became the standard analytic tool in the area. (Albeit, it is not needed for polynomial dynamics as the classical Schwarz Lemma and Koebe Distortion Theorem do the job.)

The smooth dynamical theory was essentially reduced to the S-unimodal theory in the work of Oleg Kozlovski [**Koz**] who showed that an appropriated first return map around the critical point has negative Schwarzian derivative.

31. Yoccoz puzzle and its Principal Nest

Kids know well the "puzzle game" of cutting a picture into small pieces and then trying to put them back together. Such a game can be played with dynamical pictures like Julia sets and the Mandelbrot set as well. It turned out to be a very efficient way to describe the combinatorics of the corresponding dynamical systems and to control their geometry.

Our standing assumption will be that *both fixed points of f are repelling.*

31.1. Description of the puzzle. Let us fix some parameter wake $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$ attached to the main cardioid, and let $c \in \mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$, $f \equiv f_c$. (Note that c is allowed to lie outside the Mandelbrot set.)

The puzzle game starts by cutting the complex plane with the α-rays $\mathcal{R}_i$, $i = 1, \dots \mathfrak{q}$, landing at the α-fixed point (see §24.4.3). They are cyclically permuted by the dynamics. As usually, we let

$$\mathfrak{R} \equiv \mathfrak{R}(\alpha) := \{\alpha\} \cup \bigcup_{i=1}^{\mathfrak{q}} \mathcal{R}_i.$$

This configuration divides the plane into $\mathfrak{q}$ (open) sectors S_i as described in Lemma 24.10, where $S_0 \ni 0$ is the critical sector, and $S_1 \equiv S_{\text{ch}} \ni v \equiv f(0)$ is the characteristic, or valuable, one.

As in §28.4.6, let us select some equipotential $\mathcal{E} \equiv \mathcal{E}^{(0)} = \mathcal{E}^t$ of height $t > 0$ surrounding the critical value v (which is not automatic as we do not assume that $\mathcal{J}$ is connected). Let $\Sigma^{(0)} \ni f(0)$ be the (closed) subpotential disk bounded by $\mathcal{E}$. It is tiled by $\mathfrak{q}$ (closed) topological triangles

$$Y_i^{(0)} = \Sigma^{(0)} \cap \overline{S}_i, \quad i = 0, 1, \dots, \mathfrak{q} - 1,$$

called *puzzle pieces of depth* 0 (see Figure 31.1).

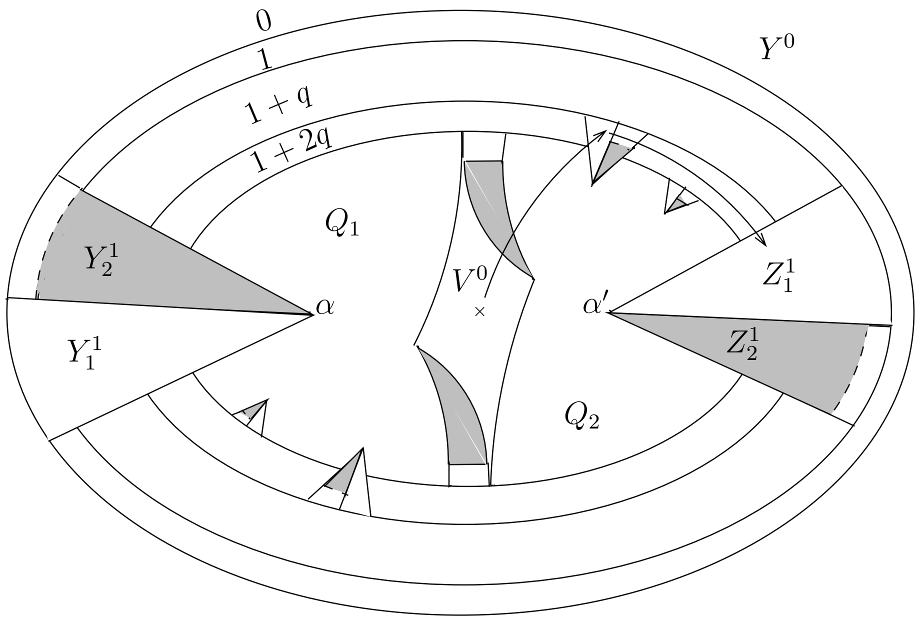

FIGURE 31.1. Initial puzzle.

We label them so that $Y^{(0)} \equiv Y_0^{(0)} \ni 0$ and $Y_{i+1}^{(0)} = f(Y_i^{(0)}) \cap \Sigma^{(0)}$, where i is considered mod $\mathfrak{q}$. (We will refer to $f_{\rm tr}(Y_i^{(0)}) := f(Y_i^{(0)}) \cap \Sigma^{(0)}$ as the *truncated image* of $Y_i^{(0)}$.) The puzzle piece $Y^{(0)}$ is naturally called *critical*, while its truncated image $Y_v^{(0)} \equiv Y_1^{(0)} \ni v$ is called *valuable*. We let $\mathfrak{R}^{(0)} := \mathfrak{R} \cap \Sigma^{(0)}$.

Take now the pullback $\mathcal{Y}^{(1)}$ of $\mathcal{Y}^{(0)}$ by f as follows. Let $\Sigma^{(1)} = f^{-1}(\Sigma^{(0)})$ be the subpotential disk bounded by the equipotential $\mathcal{E}^{(1)} := \mathcal{E}^{t/2}$. Cut it into pieces by the configuration $\mathfrak{R}^{(1)} := f^{-1}(\mathfrak{R}^{(0)})$ comprising $2\mathfrak{q}$ (arcs of) external rays landing at the points α and $\alpha' = -\alpha$. We obtain a tiling of $\Sigma^{(1)}$ by $2\mathfrak{q}-1$ closed topological disks $Y_i^{(1)}$ called *puzzle pieces of depth 1* ($2\mathfrak{q} - 2$ lateral triangles and one central 6-gone). We label them in such a way that

$$Y^{(1)} \equiv Y_0^{(1)} \ni 0, \quad Y_i^{(1)} \subset Y_i^{(0)}, \quad i = 0, 1, \dots, \mathfrak{q} - 1,$$

and we let

$$Z_i := -Y_i^{(1)}, \quad i = 1, \dots, \mathfrak{q} - 1. \tag{31.1}$$

Again, the puzzle piece $Y^{(1)}$ is called critical, while its truncated image

$$Y_v^{(1)} \equiv Y_1^{(1)} = f_{\rm tr}(Y^{(1)}) := f(Y^{(1)}) \cap \Sigma^{(1)}$$

is called valuable. (It contains the critical value v as long as the latter belongs to $\Sigma^{(1)}$.) See Figure 31.1.

Lemmas 24.10 and 24.11 imply that

$$f(Y^{(1)}) = Y_1^{(0)}; \quad f(Y_i^{(1)}) = f(Z_i) = Y_{i+1}^{(0)}, \quad i = 1, \dots, \mathfrak{q} - 1, \tag{31.2}$$

where $Y^{(0)}_{\mathfrak{q}} \equiv Y^{(0)}$. Moreover, the map $f : Y^{(1)} \to Y^{(0)}_1$ is a double branched covering, while all other maps, $f : Y^{(1)}_i \to Y^{(0)}_{i+1}$ and $f : Z_i \to Y^{(0)}_{i+1}$, $i = 1, \dots, \mathfrak{q} - 1$, are univalent.

If $f(0) \in \operatorname{int} \Sigma^{(1)} \smallsetminus \mathfrak{R}^{(1)}$, we can take the next pullback of $\mathcal{Y}^{(1)}$ by f to obtain puzzle $\mathcal{Y}^{(2)}$ of depth 2 supported on the disk $\Sigma^{(2)} := f^{-1}(\Sigma^{(1)})$ cut by the configuration $\mathfrak{R}^{(2)} := f^{-1}(\mathfrak{R}^{(1)})$ of (arcs of) the external rays landing at the points of $f^{-2}(\alpha)$, etc.

In general, if $f^n(0) \in \operatorname{int} \Sigma^{(0)} \smallsetminus \mathfrak{R}$ then we define the puzzle $\mathcal{Y}^{(n)}$ as the n-fold pullback of $\mathcal{Y}^{(0)}$. It is a tiling of the subpotential disk $\Sigma^{(n)} := f^{-n}(\Sigma^{(0)})$ bounded by the equipotential $\mathcal{E}^{(n)} := \mathcal{E}(t/2^n)$ obtained by cutting $\Sigma^{(n)}$ by the configuration $\mathfrak{R}^{(n)} := f^{-n}(\mathfrak{R})$ of the external rays landing at the points of $f^{-n}(\alpha)$. The tiles $Y^{(n)}_i$ of $\mathcal{Y}^{(n)}$ are called *puzzle pieces of depth n*. Among these puzzle pieces there is one, $Y^{(n)} \equiv Y^{(n)}_0$, whose interior contains the critical point 0. It is called *critical*. Its truncated image $Y^{(n)}_1 \equiv f_{\mathrm{tr}}(Y^{(0)}) := f(Y^{(0)}) \cap \Sigma^{(n)}$ is called *valuable*. (It contains the critical value $f(0)$ as long as $f(0) \in \Sigma^{(n)}$.)

Exercise 31.1. *Draw all possible puzzles in the wake $\mathcal{W}_{1/3}$ up to depth 4.*

Assume $n \equiv \operatorname{depth} P > 0$, and let $P' \supset P$ be the puzzle piece of depth $n-1$ containing P. We say that P is *protected* if $P \Subset P'$. For instance, on the first level, Z-pieces (31.1) are protected, $Z_i \Subset Y^{(0)}$, while the pieces $Y^{(1)}_i$, $i = 0, 1, \dots, \mathfrak{q} - 1$, are not (as they touch the corresponding pieces $Y^{(0)}_i \supset Y^{(1)}_i$ of zero depth along external rays of $\mathfrak{R}$).

A Jordan disk $D \Subset \mathbb{C}$ is called *nice* if $\operatorname{int} D \cap f^k(\partial D) = \emptyset,\ k = 0, 1, 2, \dots$ It is called *very nice* if $\overline{D} \cap f^k(\partial D) = \emptyset, \quad k = 1, 2, \dots$

The following lemma summarizes obvious but crucial properties of the puzzle pieces (that can be viewed as axioms of the puzzle):

Lemma 31.2. (i) *Puzzle pieces of depth n are closed Jordan disks with piecewise analytic boundary ("polygons") that meets $\mathcal{K}(f)$ at points of $f^{-n}\alpha$.*

(ii) *Under f, every puzzle piece $Y^{(n)}_i$ of depth $n > 0$ is mapped onto some puzzle piece $Y^{(n-1)}_j$ of depth $n - 1$. This map is univalent if $Y^{(n)}_i$ is off-critical, and is a double covering if $Y^{(n)}_i$ is critical (i.e., if $i = 0$).*

(iii) *Any two puzzle pieces P and Q are either nested or have disjoint interiors.*[19] *In the latter case, if $P \cap Q \neq \emptyset$ then $P \cap Q \cap \mathcal{K}$ is a singleton.*

(iv) *Markov Property: If $f(Y^{(n)}_i)$ intersects $\operatorname{int} Y^{(n)}_j$ then $f(Y^{(n)}_i) \supset Y^{(n)}_j$. Hence $f_{\mathrm{tr}}(Y^{(n)}_i)$ is tiled by several puzzle pieces $Y^{(n)}_j$.*

(v) *Any puzzle piece P is nice. If P is protected then it is very nice.*

Proof. (i) By definition, any $\operatorname{int} Y^{(n)}_i$ is the closure of a component of some $f^{-n}(\operatorname{int} Y^{(0)}_j)$. But for a polynomial map, the full preimage of an open Jordan disk is a disjoint union of Jordan disks. Since each $Y^{(0)}_j$ is a piecewise analytic triangle, each $Y^{(n)}_i$ is a piecewise analytic polygon. Its boundary meets $\mathcal{K}$ at points of $f^{-n}\alpha$ since the boundary of $Y^{(0)}_j$ meets $\mathcal{K}$ at α.

---

[19] Observe also that $P \cap \operatorname{int} Q \neq \emptyset \Longrightarrow \operatorname{int} P \cap \operatorname{int} Q \neq \emptyset$ since $P = \operatorname{cl}(\operatorname{int} P)$.

(ii) Since $\operatorname{int} Y_i^{(n)}$ is a component of some $f^{-1}(\operatorname{int} Y_j^{(n-1)})$, the map $f : Y_i^{(n)} \to Y_j^{(n-1)}$ is a branched covering. Since both pieces are simply connected, the conclusion follows from the Riemann-Hurwitz formula.

(iii) Since $f(\mathfrak{R}) = \mathfrak{R}$, we have: $f^{-n}(\mathfrak{R}) \supset f^{-(n-1)}(\mathfrak{R})$, $n = 1, \dots$. It follows that the tiling $\mathcal{Y}^{(n)}$ is a refinement of $\mathcal{Y}^{(n-1)}|\,\Sigma^{(n)}$. This implies the first assertion.

For the second assertion, observe that if $P \cap Q \neq \emptyset$ then $P \cap Q$ comprises two arcs of external rays meeting at a point of $\mathcal{J}$. (Compare Remark 9.3.)

(iv) It is obvious for $n = 0$, so let $n > 0$. Then by property (ii), $f(Y_i^{(n)}) = Y_k^{(n-1)}$ for some k. By property (iii), $Y_k^{(n-1)}$ contains $Y_j^{(n)}$.

(v) By definition, any puzzle piece $P = Y_i^{(n)}$ is a component of $\Sigma^{(n)} \smallsetminus f^{-n}(\mathfrak{R})$, so $\partial P \subset \partial \mathcal{E}^{(n)} \cup f^{-n}(\mathfrak{R})$. The iterated images $f^k(\mathcal{E}^{(n)})$, $k = 1, 2, \dots$, are disjoint from $\Sigma^{(n)}$, while the iterated images of $f^{-n}(\mathfrak{R})$ are contained in itself. The first assertion follows.

For the second assertion, notice that the above iterated images are in fact disjoint from the interior of any puzzle piece of depth $n - 1$. □

EXERCISE 31.3. *Two different symmetric puzzle pieces, Q and $Q' = -Q$, are disjoint, unless they touch at* 0 (*which is possible only when* orb 0 *lands at* α).

We say that two puzzle pieces P and Q are *essentially equal* if $P \cap \mathcal{K} = Q \cap \mathcal{K}$. In other words, they are different only by the equipotential level of their boundary arcs. Similarly, we say that P *is an essential pullback of* Q if P is essentially equal to a pullback of Q. For instance, puzzle pieces of depth one, $Y_i^{(1)}$ (in particular, the Z_j), are all essential pullbacks of $Y^{(0)}$. We also say that a family of puzzle pieces P_i form an *essential tiling* of a puzzle piece Q if the P_i have disjoint interiors and $\bigcup P_i \cap \mathcal{K} = Q \cap \mathcal{K}$. For instance, any puzzle piece $Y_j^{(n)}$ of depth n is essentially tiled by the puzzle pieces if $Y_i^{(n+1)}$ of depth $n+1$ contained in $Y_j^{(n)}$.

If the Julia set is connected and orb 0 does not land at α, then all puzzles $\mathcal{Y}^{(n)}$ are well defined, forming finer and finer tilings of nested neighborhoods $\Sigma^{(n)}$ of the filled Julia set $\mathcal{K}$ that nicely behave under the dynamics. In the rest of the section, we will describe how these puzzles capture the recurrence of the critical orbit.

If we consider below a puzzle $\mathcal{Y}^{(n)}$, we assume without mentioning that $f^n(0) \in \Sigma^{(0)} \smallsetminus \mathfrak{R}$, so that $\mathcal{Y}^{(n)}$ is well defined. (Not to be distracted by these details, we suggest the reader to assume in the first reading that the Julia set is connected and orb 0 does not land at α, so the above assumptions hold for all $n \in \mathbb{N}$.)

For $n \in \mathbb{N}$, let $\partial \mathcal{Y}^{(n)} := f^{-n}\alpha$ be the set of pre-α points of order n. These are exactly boundary points of puzzle pieces of depth n that belong to the Julia set $\mathcal{J}$. Let $\partial \mathcal{Y} := \bigcup \partial \mathcal{Y}^{(n)}$.

For $z \in \mathcal{J} \smallsetminus \partial \mathcal{Y}^{(n)}$, there is a unique puzzle piece of depth n containing z, and we will use notation $Y^{(n)}(z)$ for this piece. For $z \in \partial \mathcal{Y}^{(n)}$, we use notation $Y^{(n)}(z)$ for the union of all puzzle pieces containing z. It is a topological disk, and we will still refer to it as a "non-elementary" puzzle piece.

31.2. First Landing and Return maps.

31.2.1. *Pullbacks and the First Landing map.* Consider a puzzle piece P of depth $n \geq 1$ and a point z such that $f^m z \in \operatorname{int} P$ for some $m \geq 0$. The *pullback* of P along the orbit $\{f^k z\}_{k=0}^m$ (as defined in §19.1) is the puzzle piece Q of depth $n+m$ containing z. Notice that since the map $f^m : Q \to P$ is proper, $z \in \operatorname{int} Q$.

The map $f^m : Q \to P$ is a branched covering of degree 2^t, where t is the number of critical puzzle pieces among $f^k Q$, $k = 0, 1, \dots, m-1$. In particular, if there are no critical puzzle pieces among them, then $f^m : Q \to P$ is univalent.

Within this section, §31.2.1, *we keep notation* z, P, Q, n, *and* m *for the objects just described.*

Lemma 31.4. *If P is protected then so is Q. If additionally $m > 0$, then*

$$Q \cap \partial P = \emptyset.$$

Proof. Let $P' \Supset P$ be the puzzle piece of depth $n - 1 = \operatorname{depth} P - 1$, and let $Q' \supset Q$ be the puzzle piece of depth $n + m - 1 = \operatorname{depth} Q - 1$. Then

$$f^m : (Q', Q) \to (P', P).$$

Since the first map is proper, $Q \Subset Q'$.

Assume $Q \cap \partial P \neq \emptyset$. Applying f^m, we obtain that $P \cap f^m(\partial P) \neq \emptyset$, contradicting the very nice property of Lemma 31.2 (v). □

Lemma 31.5. *Let $f^m z$ be the first landing of* $\operatorname{orb} z$ *in* $\operatorname{int} P$. *Then:*

(i) *The puzzle pieces $f^k(Q)$, $k = 0, 1, \dots, m$, have pairwise disjoint interiors.*

(ii) *If P is protected then the puzzle pieces $f^k(Q)$, $k = 0, 1, \dots, m$, are pairwise disjoint.*

Proof. (i) Assume $\operatorname{int} f^k(Q) \cap \operatorname{int} f^l(Q) \neq \emptyset$ for some $0 \leq k < l \leq m$. By Lemma 31.2 (iii), $f^k(Q) \subset f^l(Q)$. Applying f^{m-l}, we obtain $f^s(Q) \subset P$, where $s = m - (l - k) < m$. In particular, $f^s z \in \operatorname{int} f^s(Q) \subset \operatorname{int} P$, contradicting the choice of m as the first landing time of $\operatorname{orb} z$ in $\operatorname{int} P$.

(ii) Assume now $f^k(Q) \cap f^l(Q) \neq \emptyset$ for some $0 \leq k < l \leq m$. Then by (i), $f^k(Q) \cap \partial(f^l(Q)) \neq \emptyset$. Applying f^{m-l}, we obtain $f^s(Q) \cap \partial P \neq \emptyset$, where $s = m - (l - k) < m$, contradicting Lemma 31.4. □

Under the above circumstances, the map $f^m : Q \to P$ is called a *branch of the first landing map to* P, while Q is called a *component of the first landing domain*. Notice that this terminology is slightly inconsistent since for a point $z \in \partial Q$, m does not have to be the moment of the *first* landing in P (it may land in ∂P earlier). Hopefully, this would not lead to a confusion. In case when P is protected, the terminology is unambiguous.

Corollary 31.6. (i) *If the puzzle piece P is critical then any branch $f^m : Q \to P$ of the first landing map is univalent.*

(ii) *If $f^m : Q \to P$ is a branch of the first landing map to P, where Q is critical* (*while P is arbitrary*), *then $f^{m-1} : f(Q) \to P$ is univalent.*

Lemma 31.7. *Let Q_1 and Q_2 be two different components of the first landing map to a piece P. Then* $\operatorname{int} Q_1 \cap \operatorname{int} Q_2 = \emptyset$. *Furthermore, if* $\operatorname{orb} 0$ *does not land at α and P is protected then $Q_1 \cap Q_2 = \emptyset$.*

Proof. It is easy to see that Q_1 and Q_2 cannot be nested, implying that $\operatorname{int} Q_1 \cap \operatorname{int} Q_2 = \emptyset$.

Furthermore, let $f^{m_1} : Q_1 \to P$ and $f^{m_2} : Q_2 \to P$ be the corresponding branches of the first landing map. If $m_1 \neq m_2$, say $m_1 < m_2$, then $f^{m_2 - m_1}$ is a domain of the first landing map to P intersecting ∂P, which contadicts Lemma 31.4. Otherwise, there is a moment $k \in [1, m_1 = m_2]$ such that $f^k(Q_1) = f^k(Q_2)$, so $W_1 := f^{k-1}(Q_1)$ and $W_2 := f^{k-1}(Q_2)$ are different symmetric puzzle pieces. The conclusion follows from Exercise 31.3. □

Putting all the above branches together, we obtain the *(full) landing map*

$$L \equiv L_P f : \bigsqcup Q_i \to P \tag{31.3}$$

(usually comprising infinitely many branches). One of the components Q_i is the original piece P on which $L = \mathrm{id}$. (In case when $P \ni 0$, it is naturally called critical or central and is labelled as Q_0.)

Moreover, $\operatorname{Dom} L = \bigsqcup \operatorname{int} Q_i$ is naturally embedded into $\mathbb{C}$, while the boundaries ∂Q_i may overlap creating a slight ambiguity for L to be defined on $\bigcup Q_i \subset \mathbb{C}$. In case of a protected P, the whole $\operatorname{Dom} L$ is embedded into $\mathbb{C}$.

When P is critical then any off-critical component $Q_i \neq P$ of $\operatorname{Dom} L$ is mapped by f univalently onto another component, Q_j. As for the critical component, the map $f : P \to f(P)$ is of course a double branched covering, and we have:

Exercise 31.8. *$f(P) \cap \operatorname{Dom} L$ is the union of some components Q_j, $j \neq 0$.*

31.2.2. *First Return map.*

Lemma 31.9. *Let $z \in P$, and let $f^m z$, $m > 0$, be the first return of* $\operatorname{orb} z$ *to* $\operatorname{int} P$. *Let $V \ni z$ be the pullback of P under f^m. Then:*

(i) *$V \subset P$ and the puzzle pieces $f^k(V)$, $k = 1, \dots, m$, have pairwise disjoint interiors.*

(ii) *If additionally P is protected, then $V \Subset P$ and the puzzle pieces $f^k(V)$, $k = 1, \dots, m$, are pairwise disjoint.*

Proof. (i) Note that due to the nice property of Lemma 31.2 (v), $z \in \operatorname{int} P$. Hence $V \cap \operatorname{int} P \neq \emptyset$, implying that the puzzle pieces V and P are nested. Since V is a deeper puzzle piece than P, we conclude that $V \subset P$.

Since the image $Q := f(V)$ is a component of the first landing domain to $\operatorname{int} P$, Lemma 31.5 (i) yields disjointness of the interiors of the puzzle pieces in question.

(ii) If P is protected then Lemma 31.5 (ii) implies in the similar way disjointness of the above puzzle pieces.

Moreover, $V \Subset P$ by Lemma 31.4. □

Under the above circumstances, the map $f^m : V \to P$ is called a *branch of the first return map to P*, while V is called a *component of the first return domain.* (Once again, for an unprotected P this terminology may be slightly inconsistent.)

It is particularly important to consider returns to critical puzzle pieces:

Corollary 31.10. *If the puzzle piece P is critical then a branch $f^m : V \to P$ of the first return map is a double branched covering or univalent depending on whether V is critical or not.*

Putting all the above branches together, we obtain the (*full*) *first return map*

$$T \equiv T_P f : \bigsqcup V_i \to P \tag{31.4}$$

(usually comprising infinitely many branches). As in the case of L, the boundaries ∂V_i may overlap creating a slight ambiguity for T, but if P is protected then there is no ambiguity.

If 0 returns to $\operatorname{int} P$ then it is contained in one of the domains V_i. We label it V_0 and call *critical* or *central*. We will also call it the *first kid* of P. The corresponding branch $T : V_0 \to P$ is a double branched covering. It is also called *critical* or *central*. All off-central branches $T : V_i \to P$ are univalent.

REMARK 31.11. If 0 does not return to $\operatorname{int} P$ then f is a Misiurewicz map which is already well understood (see §27.1). So, in what follows we usually assume that 0 returns to all puzzle pieces ("combinatorial recurrence").

Note that by definitions,

$$T = L \circ f \mid \operatorname{Dom} T, \tag{31.5}$$

and $\operatorname{Dom} T$ coincides with the preimage of $\operatorname{Dom} L$ under the double branched covering $f : P \to f(P)$. Together with Exercise 31.7 and Lemma 31.7, this implies:

COROLLARY 31.12. *The components $V_i \subset P$ of the first return domain have disjoint interiors. If P is protected then the components themselves are pairwise disjoint.*

PROOF. One remark is due: Two symmetric pullbacks V_i andV_i' of some non-valuable component $Q_j \subset f(P)$ are disjoint by Exercise 31.3. ☐

31.3. Generalized renormalization.

31.3.1. *Generalized quadratic-like maps.* Let $\{Q_i\}$ be a finite or countable family of disjoint (closed) Jordan disks compactly contained in a (closed) Jordan disk P.[20] We call a map $g : \bigcup Q_i \to P$ a *(generalized) polynomial-like map* if

- $g : Q_i \to P$ is a branched covering of finite degree which is univalent on all but finitely many Q_i;
- All the domains Q_i are visited by the critical orbits.

The DH polynomial-like maps correspond to the case of a single disk Q_0.

We define the *filled Julia set* $\mathcal{K}(g)$ as the set of all non-escaping points.

Let us say that a polynomial-like map g is *of finite type* if its domain consists of finitely many disks Q_i.

A generalized polynomial-like map is called *generalized quadratic-like* if it has a single (and non-degenerate) critical point. In such a case we will assume, unless otherwise is stated, that 0 is the critical point, and label the discs Q_i in such a way that $Q_0 \ni 0$.

The main example of a generalized ql map is the first return map $g : \bigcup Q_i \to P$ to a protected critical puzzle piece P restricted to the union of those disks Q_i that intersect $\operatorname{orb} 0$ (assuming that 0 returns to P). If 0 is not escaping under g, then g is called the *generalized renormalization* of f on P.

We will see that in many interesting cases, g is a map of finite type. For the moment, let us make just a simple observation:

[20] Recall from §50.3.1 that this means that $Q_i \subset \operatorname{int} P$.

Exercise 31.13. *The generalized renormalization of f on P is of finite type if and only if one of the following properties holds:*

(i) *any point $z \in \operatorname{orb} 0$ returns back to P in bounded time;*

(ii) *any point $z \in \overline{\mathcal{P}}_f$ returns back to P in bounded time.*

31.3.2. *Itineraries for the first landing map.* Let us consider the first return map $T : \bigsqcup V_j \to P$ to some critical puzzle piece P of positive depth. Let $z \in \operatorname{Dom} T \smallsetminus V_0$ be a point whose orbit lands in $\operatorname{int} V_0$, and let $T^l z \in \operatorname{int} V_0$ be the first landing point. Then

$$T^k z \in V_{i(k)}, \ i(k) \neq 0, \quad k = 0, 1, \dots, l-1, \tag{31.6}$$

and

$$\bar{i} \equiv \bar{i}_l(z) := (i(0), i(1), \dots, i(l-1)) \tag{31.7}$$

is called the *landing itinerary* for z. By the Telescoping Lemma, there is a puzzle piece $D^l(z) \ni z$ univalently mapped by T^l onto P. Hence $D^l(z) \ni z$ contains a puzzle piece $Q \equiv Q^l(z) \equiv Q(\bar{i})$, $\operatorname{int} Q \ni z$, univalently mapped onto V_0. This is a component of the first landing domain $L_{V_0} T$. On V_0 itself, we naturally let $L_{V_0} T := \mathrm{id}$ with the empty itinerary $(\emptyset)$. Altogether, these components form the domain for the first landing map:

$$\operatorname{Dom} L_{V_0} T = \bigsqcup Q(\bar{i}), \tag{31.8}$$

where the union is taken over all possible itineraries $\bar{i} = (i(0), \dots, i(l-1))$, $l \in \mathbb{Z}_+$, $i(k) \neq 0$, together with $(\emptyset)$.

Since T is the first return map to P, we have $(L_{V_0} T)(z) = (L_{V_0} f)(z)$ for any $z \in P$ whose orbit lands in V_0. In particular, $\operatorname{Dom} L_{V_0} T = P \cap \operatorname{Dom} L_{V_0} f$.[21] By Exercise 31.7, components $Q_{\bar{i}}$ have disjoint interiors. If V_0 is protected then the whole components are pairwise disjoint (Lemma 31.7).

31.3.3. *Renormalization of a generalized ql map.* Given the first return map $T : \bigcup V_i \to P$ to a critical puzzle piece P, let us take a closer look at the formation of the first return map to the central puzzle piece V_0.

Assuming the critical point returns to $\operatorname{int} P$, we have a well defined critical component $P' \equiv V_0 \ni 0$. Let us consider the first return map to P',

$$T' \equiv T_{P'} f : \bigcup V_i' \to P'.$$

Similarly to (31.5), we have:

$$T' = L' \circ T \,|\, V_0, \text{ where } L' \equiv L_{P'} T.$$

Hence $\operatorname{Dom} T'$ is equal to the full preimage of $\operatorname{Dom} L'$ by the double branched covering $T : P' \to P$. By (31.8), $\operatorname{Dom} T'$ is decomposed into pullbacks of the components $Q(\bar{i})$. One of these components, $Q(\bar{i}_v)$, is *valuable*: it contains the critical value $T(0)$. It has only one pullback, the central component V_0'. Any other component $Q(\bar{i})$ has two pullbacks, $V'_{\bar{i},\varepsilon}$, $\varepsilon = \pm 1$. Thus,

$$\operatorname{Dom} T' = V_0' \sqcup \bigsqcup_{\bar{i} \neq \bar{i}(v)} \bigsqcup_{\varepsilon = \pm 1} V'_{\bar{i},\varepsilon}.$$

[21] The latter should be understood as the union of components of $\operatorname{Dom} L_{V_0} f$ contained in P. Components that touch P along the boundary but do not overlap with $\operatorname{int} P$ should be disregarded.

Moreover, is V_0 is protected, then all these components are disjoint.

If the critical orbit is not escaping under T', then the restriction of T' to the union of puzzle pieces V_i' that meet $\operatorname{orb} 0$ is the *generalized renormalization* of T on P'. (This will be our default choice of "the generalized renormalization" of T unless P' is explicitly selected in a different way.) Note also that for the above discussion it is sufficient to know the restriction of T to the the union of puzzle pieces V_i that meet $\operatorname{orb} 0$, so the generalized renormalization acts on generalized ql maps (with non-escaping critical point).

31.3.4. *Koebe space.* To control distortion of the branches of landings and returns, we need an extension of these branches to bigger domains:

LEMMA 31.14. *Let P be a critical puzzle piece and let $P' \subset P$ be its first kid. Let $L' : Q' \to V'$ be a branch of the first landing map to P'. Then there is a disk $Q^K \supset Q'$ such that L' extends to a univalent map $L^K : Q^K \to P$.*

PROOF. Let us consider the first return map T to P (31.4), where $P' = V_0$. Let $L : Q \to P$ be the branch of the first landing map to P with $Q \supset Q'$, and let $W := L(Q')$. Then $L' = (T^l| W) \circ (L| Q')$ for some $l \in \mathbb{N}$.

Let $(i(0), \dots, i(l-1))$, $i(k) \neq 0$, be the itinerary of W through $\operatorname{Dom} T$, i.e.,

$$T^k(W) \subset V_{i(k)}, \ k = 0, \dots, l-1, \quad \text{while } T^l(W) = V_0.$$

Then $L(Q) \supset V_{i(0)}$ and $T(V_{i(k)}) \supset V_{i(k+1)}$, $k = 0, 1, \dots, l-1$, and the conclusion follows from the Telescoping Lemma. ☐

31.4. First escaping moment n. By (31.2), $f^{\mathfrak{q}}(0) \in Y^{(0)}$. So, if $f^{\mathfrak{q}}(0) \in U^1$, there are two options: either $f^{\mathfrak{q}}(0) \in Y^{(1)}$ (*central return*) or $f^{\mathfrak{q}}(0) \in Z_i$ for some $i \in \{1, \dots, \mathfrak{q} - 1\}$ (*non-central return*). In the former case, if $f^{2\mathfrak{q}}(0) \in \Sigma^{(1)}$, we obtain the same options: either $f^{2\mathfrak{q}}(0) \in Y^{(1)}$ or $f^{2\mathfrak{q}}(0) \in Z_i$ for some i, etc. So, either the critical point always returns to $Y^{(1)}$,

$$f^{n\mathfrak{q}}(0) \in Y^{(1)}, \quad n = 0, 1, \dots, \tag{31.9}$$

or else there exists the escaping moment $\mathbf{n} \in \mathbb{Z}_+$ such that

$$f^{\mathbf{n}\mathfrak{q}}(0) \in Z_i \quad \text{for some } i \in \{1, \dots, \mathfrak{q}\} \tag{31.10}$$

(provided $f^{\mathbf{n}\mathfrak{q}}(0) \in \Sigma^{(1)}$).

If option (31.9) takes place , the map f is *immediately renormalizable* in the sense of §28.4.6. Moreover, its little Julia set K (28.3) consists of all points that never escape from $Y^{(1)}$, and the family of all little Julia sets,

$$K_m = f^m(K), \quad m = 0, \dots, \mathfrak{q} - 1,$$

form a bouquet centered at α.

As we will see (Theorem 43.1) the set of parameters $c \in \mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$ for which f_c is immediately renormalizable assemble a little copy $\mathcal{M}_{\mathfrak{p}/\mathfrak{q}}$ of the Mandelbrot set attached to the main cardioid.

31.5. Principal Nest. We are now ready to introduce the *Principal Nest* of critical puzzle pieces,

$$V^0 \supset V^1 \supset V^2 \supset \cdots \ni 0, \tag{31.11}$$

and associated double coverings $g_n : V^n \to V^{n-1}$.

Assume f is not immediately renormalizable, and let $\mathbf{n}$ be the first escaping moment (31.10). We let $V_0 \ni 0$ be the pullback of $Z_{\mathbf{i}}$ along the orbit $\{f^k\}_{k=0}^{\mathbf{qn}}$. Since Z_i is protected, so is V^0

EXERCISE 31.15. *Show that V^0 is the most shallow (i.e., of the smallest depth) protected critical puzzle piece.*

Assume inductively that we have defined the nest up to V^{n-1}. If the orb(0) never returns to $\operatorname{int} V^{n-1}$ then the construction stops here. Otherwise consider the first return $f^{l_n}0$ of the critical point back to $\operatorname{int} V^{n-1}$. Let V^n be the pullback of V^{n-1} along this orbit and let $g_n = f^{l_n} : V^n \to V^{n-1}$. By Corollary 31.10, this map is a double covering. Being pullbacks of a protected piece, all the V^n are protected. By Lemma 31.9 (ii), $V^{n+1} \Subset V^n$ for all $n \in \mathbb{N}$. This completes the construction.

We call V^n the principal puzzle piece of *level* n (pay attention to the difference between the "level" and the "depth").

Whenever the principal piece V^{n-1} is well defined, we introduce the *principal first landing map* to V^{n-1}:

$$L_{n-1} : \bigcup_i Q_j^n \to V^{n-1}, \tag{31.12}$$

and the *principal first return* map g_n to V^{n-1}:

$$g_n : \bigcup_i V_i^n \to V^{n-1}, \tag{31.13}$$

where the components V_i^n are labeled so that $V_0^n = V^n$ (if exists). By Lemma 31.9(ii), the components V_i^n are pairwise disjoint and $V_i^n \Subset V^{n-1}$. By (31.5),

$$g_n \,|\, V_i^n = L_{n-1} \circ f \,|\, V_i^n. \tag{31.14}$$

The Principal Nest is infinite if and only if orb 0 visits all critical puzzle pieces. In this case, restricting the maps g_n to the union of puzzle pieces V_i^n that meet orb 0, we obtain the sequence of the *principal generalized renormalizations*. (We will use the same notation $g_n : \bigcup V_i^n \to V^{n-1}$ for them).

Topological annuli $A^n := V^{n-1} \smallsetminus \operatorname{int} V^n$ (which could be degenerate) are called *principal annuli*. (Their open or semi-closed counterparts are also called so.)

Note finally that one can consider similar nests beginning with any other critical piece V^0, as long as

$$V^0 \Supset V^1. \tag{31.15}$$

Such a nest will be referred to as the *Principal nest beginning with* V^0. Here is a natural choice:

EXERCISE 31.16. *Show that* (31.15) *implies $V^{n-1} \Supset V^n$ for any $n \in \mathbb{Z}_+$ (as long as the Principal Nest is defined). Show that* (31.15) *is satisfied for $V^0 := Y^{(1+(\mathbf{n}-1)\mathfrak{q})}$. (Recall that the above choice was $V^0 = Y^{(1+n\mathfrak{q})}$.)*

31.6. Central cascades and primitive renormalization.

31.6.1. *Central cascades.* There are two different combinatorial possibilities on every level which are important to distinguish. The return of the critical point to level $n-1$ (and the level itself) is called *central* if $g_n(0) \in \operatorname{int} V^n$ (see Figure 31.2). In this case, the critical orbit returns to level $n-1$ at the same time as to level n, so that $l_n = l_{n+1}$ and $g_{n+1} : V^{n+1} \to V^n$ is just the restriction of g_n to V^{n+1}. Central returns indicate the *fast* recurrence of the critical orbit.

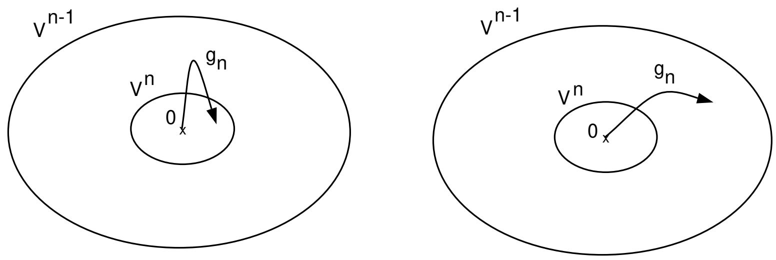

FIGURE 31.2. Central and non-central returns.

If $N+1$ consecutive levels, $m-1$, m,..., $m+N-1$, are central then the nest

$$V^{m-1} \supset V^m \supset \cdots \supset V^{m+N-1} \tag{31.16}$$

is called a *central cascade* of length $N+1$. In this case, $g^{l_m}0 \in V^{m+N}$ and the maps

$$g_{m+k} : V^{m+k} \to V^{m+k-1},\ k = 1, \ldots, N+1,$$

are just the restrictions of g_m to the corresponding puzzle pieces.

If this cascade is maximal then the levels $m-2$ and $m+N$ are non-central. In this case, $N+2$ is equal to the *escaping time* it takes for the critical orbit to escape V^m under the iterates of g_m:

$$g_m^k(0) \in V^{m+N-k+1} \smallsetminus V^{m+N-k+2},\ k = 1, \ldots, N+2.$$

If the return to level $m-1$ is non-central ($N=-1$), we will formally consider $\{V^{m-1}\}$ to be a "central cascade" of length 0. With this convention, the whole principal nest is decomposed into consecutive maximal central cascades. In fact, one of these cascades, the last one, can have infinite length,

$$V^{m-1} \Supset V^m \Supset \ldots, \tag{31.17}$$

in which case the critical point *never escapes* V^m under the iterates of g_m. As we will see momentarily, this happens exactly in the primitively renormalizable case.

31.6.2. *Primitive renormalization.*

THEOREM 31.17. *Let $f \equiv f_c$ be a quadratic polynomial with connected Julia set and both fixed points repelling. Assume f is not immediately renormalizable. Then f is either non-renormalizable or primitively renormalizable. Moreover,*

(i) *f is primitively renormalizable if and only if its Principal Nest is infinite and ends up with an infinite central cascade* (31.17). *In this case the map*

$$g \equiv g_m = (f^{l_m} : V^m \to V^{m-1})$$

is a pre-renormalization of f with period $p = l_m$.

(ii) *This renormalization is of DH type.*

(iii) *This is the first renormalization of f, i.e., if $h : V \to V'$ is another pre-renormalization with some period q, then q is a multiple of p and $\mathcal{K}(h) \subset \mathcal{K}(g)$.*

(iv) *If $q = p$ then $\mathcal{K}(h) = \mathcal{K}(g)$.*

PROOF. (i) Assume f is renormalizable with period q, a pre-renormalization $h : V \to V'$, and the little (filled) Julia set $K \equiv \mathcal{K}(h)$. The Non-Cutting Assumption (R3) from §28.4 implies that $K \not\ni \alpha$. Hence $K \subset Y^0$. It follows that K is contained in any critical pullback of Y^0. But Z_i is a pullback of Y^0, while all the pieces V^m, $m = 0, 1, \dots$, are pullbacks of Z_i, Hence

$$K \subset \bigcap V^m. \tag{31.18}$$

Consequently, for any $m \in \mathbb{Z}_+$, we have: $f^q(0) \in K \subset V^{m-1}$, so $l_m \le q$, where l_m is the first return time of the critical point to V^{m-1}. Thus, the sequence (l_m) eventually stabilizes, so all the returns are eventually central.

Vice versa, if the Principal Nest is well defined and ends up with an infinite central cascade (31.17), then $g : V^m \to V^{m-1}$ is a quadratic-like map with non-escaping critical point. Moreover, by Lemma 31.9, all the pieces

$$f^k(V^m), \quad k = 1, \dots, l_m,$$

are pairwise disjoint, so g is a primitive pre-renormalization of f.

(ii) Let $\mathcal{V}$ be the set of vertices of V^{m-1}. Let us define a map $\sigma : \mathcal{V} \to \mathcal{V}$ as follows. For a vertex $v \in \mathcal{V}$, consider two edges e_1 and e_2 of V^{m-1} attached to v, and the corresponding external rays $\mathcal{R}_1$ and $\mathcal{R}_2$, and let $L \equiv L(v) := \mathcal{R}_1 \cup \mathcal{R}_2 \cup \{v\}$ be the leaf comprising these two rays. There exists exactly one vertex $v' \in V^m$ such that the similar leaf $L' := L(v')$ separates L from 0. We let $\sigma(v) = g(v')$.

Any map of a finite set has a periodic point, so let us consider a periodic vertex v, $\sigma^s(v) = v$. Then there exists a vertex $v_1 \in V^{m-1+s}$ such that L_1 separates L from 0 and $g^s(L_1) = L$, where $L_1 = L(v_1)$.

Let Π be the strip bounded by L and L_1. It can be univalently lifted by g^s to a strip Π_1 attached to Π along L_1. Moreover, this strip is bounded by the leaf L_1 and a leaf $L_2 := L(v_2)$, where v_2 is a vertex of V^{m-1+2s} such that $g^s(v_2) = v_1$. Similarly, Π_1 can be univalently lifted to a strip Π_2 attached to Π_1 and bounded by the leaf L_2 and a leaf $L_3 := L(v_3)$, where $g^s(v_3) = v_2$, etc. In this way, we obtain a chain of strips Π_k attached one to another, bounded by leaves $L_k = L(v_k)$ and $L_{k+1} = L(v_{k+1})$, where v_k is a vertex of V^{m-1+pk}, and such that g^s univalently maps Π_{k+1} onto Π_k.

Let γ be an arc in Π connecting v to v_1. It is lifted by g^s to an arc γ_1 in Π_1 connecting v_1 to v_2 In turn, γ_1 is lifted by g^s to an arc γ_2 in Π_2 connecting v_2 to v_3, etc. In this way, we obtain a chain of arcs γ_k in Π_k attached one to another and such that g^s homeomorphically maps γ_{k+1} onto γ_k. They concatenate a curve Γ invariant under the relevant branch of g^{-s}. The standard argument using the hyperbolic metric in $V^{m-1} \smallsetminus \mathcal{K}(g)$ shows that this curve lands at some periodic point β of $\mathcal{K}(g)$. Moreover, the leaves L_k converge to a leaf $L(\beta)$ through β. So, β is a cut-point for the big Julia set $\mathcal{K}(f)$.

In fact, β is the g-fixed point, for otherwise the ray configuration of its cycle $\boldsymbol{\beta}$ would not have the structure described in §24.5.1. (The central component of $\mathbb{C} \smallsetminus \mathfrak{R}(\boldsymbol{\beta}) \cup \mathfrak{R}(\boldsymbol{\beta}')$ would not be a strip, compare Exercise 28.25.)

Thus, the period of β under f is equal to p. By Exercise 28.28 (iii), f is DH renormalizable with respect to $\mathfrak{R}(\boldsymbol{\alpha})$, and the corresponding DH renormalization is equal to g.

(iii) If f is renormalizable with period q and a pre-renormalization $h : V \to V'$, then by (31.18) we have:

$$\mathcal{K}(h) \subset \bigcap V^m = \mathcal{K}(g),$$

so g is the first renormalization of f.

(iv) follows from Exercise 28.28 (iii). □

In case of a non-renormalizable map, we say that the critical point is *combinatorially recurrent* if it visits the interiors of all critical puzzle pieces (or equivalently, if the Principal Nest is infinite).

Let us define the *height* of f as the number of the maximal central cascades in the principal nest. We see that *f is renormalizable if and only if it has finite height*.

Thus, the principal nest provides an algorithm to decide whether the map in question is renormalizable, whether this renormalization is satellite or primitive, and to capture this renormalization.

EXERCISE 31.18. *If f is a DH renormalizable quadratic-like map, then all its generalized renormalizations g_n are of finite type.*

On the negative side, the puzzle provides us with dynamical information only up to the first renormalization level. If we wish to penetrate deeper, we need to cut the little Julia sets into pieces and to go through the corresponding principal nest. This will be discussed in §31.9.

31.7. Initial Markov tiling. In this section we will construct a Markov tiling of $Y^{(0)}$ that captures combinatorics of the return of the critical orbit to the first piece V^0 of the Principal Nest.

31.7.1. *Returns to Y^0.* Let P_i be a finite or countable family of topological discs with disjoint interiors, and $g : \cup P_i \to \mathbb{C}$ be a map such that the restrictions $g|P_i$ are branched coverings onto their images. This map is called *Markov* if $g(P_i) \supset P_j$ whenever $\operatorname{int} g(P_i) \cap \operatorname{int} P_j \neq \emptyset$. A Markov map is called *unbranched* if all the restrictions $g|\, P_i$ are one-to-one onto their images. (Compare §19.14.2.)

A Markov map is called *Bernoulli* if there is a topological disc D such that $g(P_i) \supset D \supset \bigcup P_j$ for all i. Any such a D will be called a *range* of g. Similarly we can define an *unbranched* Bernoulli map. (Compare §19.11.3.)

Let us consider the *initial puzzle pieces of level one*, the critical (central) one, $Y^{(1)}$, and the lateral ones, $Z_i^{(1)} = -Y_i^{(1)}$, attached to α', $i = 1, \dots, \mathfrak{q}-1$ (see §31.1). Recall that they are labeled dynamically:

$$f(Y^{(1)}) \underset{\text{ess}}{=} Y_1^{(1)}, \quad f(Z_i^{(1)}) \underset{\text{ess}}{=} Y_{(i+1)}^{(1)},\ i = 1, \dots, \mathfrak{q}-2, \quad f(Z_{\mathfrak{q}-1}^{(1)}) = Y^{(0)}.$$

Let us now truncated these puzzle pieces by appropriate equipotentials and *accelerate* this dynamics as in §24.4.3:

$$Y^1 := Y^{(1)} \cap \Sigma^{(\mathfrak{q})}, \quad Z_i^1 := Z_i^{(1)} \cap \Sigma_i^{(\mathfrak{q}-i)},$$

$$F|\, Y^1 = f^{\mathfrak{q}}, \quad F|\, Z_i^1 = f^{\mathfrak{q}-i},\ i = 1, \dots, \mathfrak{q}-1. \tag{31.19}$$

This map is a double branched covering of the critical piece Y^1 over Y^0 and is a conformal isomorphism of each lateral piece Z_i^1 onto Y^0. So, it is a *Bernoulli map with range* Y^0. Moreover, $F(Y^1)$ double covers Y^1 and all the puzzle pieces Z_i^1 (recall Remark 1.107 for the technical meaning of the term "covers"). If $F(0) \in Y^1$ (central return) then the pullback of Y^1 by this map is the *initial critical* (*central*) *piece* $Y^2 \equiv Y^{(2\mathfrak{q})}$ *of level two*, while each Z_i^1 has two univalent pullbacks, *initial off-critical pieces* Z_{ij}^2 *of level two*, $j \in \{0, 1\}$ (see Figure 31.3 and Remark 31.19 below)

FIGURE 31.3. Initial tiling ($\mathfrak{p}/\mathfrak{q} = 1/3$, $\mathbf{n} = 2$).

Now, $F(Y^2)$ double covers all these puzzle pieces. If we again have a central return, (i.e., $F(0) \in Y^1$ and hence $F(0) \in Y^2$) then the degree two pullback of Y^2 by F gives us the *initial critical piece of level three*, Y^3, while the univalent pullbacks of the Z_{ij} produce four *off-critical initial pieces* Z^3_{ijk} *of level three*, $k \in \{0, 1\}$.

Repeating this procedure $\mathbf{n} - 1$ times (where $\mathbf{n}$ is the first escaping moment from §31.4) we obtain the *initial critical* (*central*) *nest*

$$Y^1 \supset Y^2 \supset \ldots \supset Y^{\mathbf{n}} \tag{31.20}$$

and a family of off-critical puzzle pieces

$$Z^m_{ij}, \quad m \in \{1, \ldots, \mathbf{n}\}, \quad i \in \{1, \ldots, \mathfrak{q} - 1\}, \quad j \in \{0, 1\}^{m-1}. \tag{31.21}$$

As an outcome, we obtain the tiling

$$Y^0 \underset{\text{ess}}{=} Y^{\mathbf{n}} \cup \bigcup_{m=1}^{\mathbf{n}} \bigcup_{ij} Z^m_{ij}. \tag{31.22}$$

Moreover, $F^{\mathbf{n}}(0) \in Z^1_{\mathbf{i}}$ for some $\mathbf{i} \in \{1, \ldots, \mathfrak{q} - 1\}$, and hence

$$F(0) \in Z^{\mathbf{n}}_{\mathbf{ij}} \text{ for some } \mathbf{j} \in \{0, 1\}^{\mathbf{n}-1}. \tag{31.23}$$

Let us *accelerate* this dynamics further on the pieces Z^m_{ij} by setting

$$G|\, Z^m_{ij} = F^m. \tag{31.24}$$

This map carries univalently each puzzle piece Z^m_{ij} onto Y^0.

REMARK 31.19. Let us finish with a comment on the *labeling* of the Z^m_{ij} by dyadic sequences $j \in \{0, 1\}^{m-1}$, which was not specified so far. Here is a natural way. Let us consider the central strip Π from §24.4.3. Since $\mathcal{K} \cap \Pi$ is a hull containing $\{\alpha, \alpha'\} = \partial\Pi \cap \mathcal{K}$, the complement $\Pi \smallsetminus \mathcal{K}$ comprises two domains, Π^0 and Π^1, representing the upper and lower ends of Π. This allows us to encode the orbit $F^n(Z^m_{ij})$, $n = 0, \ldots, m-2$ dynamically by dyadic sequences $j \in \{0, 1\}^{m-1}$. Interpreting these sequences as binary representations of numbers $j \in \{0, 1, \ldots, 2^{m-1} - 1\}$, we come up with labeling the pieces in question by dyadic angles $\theta_j = (2j + 1)/2^j \in \mathbb{R}/\mathbb{Z}$. Notice that the cyclic order of these angles corresponds to the cyclic order of the pieces Z^m_{ij} induced by the cyclic order of the external rays that bound these pieces.

31.7.2. *Escape from* Y^1. Let us look closer at the central piece $Y^{\mathbf{n}}$. Its image under $F = f^{\mathfrak{q}}$ two-to-one covers all the initial pieces of level $\mathbf{n}$. The pullback of $Z^{\mathbf{n}}_{\mathbf{ij}}$ from (31.23) gives us exactly $V^0 \ni 0$, the first puzzle piece in the Principal Nest (31.11). The pullbacks of the other pieces $Z^{\mathbf{n}}_{ij}$ provide some off-critical pieces

$$Z^{\mathbf{n}+1}_{ij}, \quad i \in [1, \mathfrak{q} - 1],\ j \in \{0, 1\}^{\mathbf{n}},\ (ij) \neq (\mathbf{ij}).$$

Finally, we have two univalent pullbacks Q_1 and Q_2 of $Y^{\mathbf{n}}$. Altogether, these pieces essentially tessellate the piece $Y^{\mathbf{n}}$:

$$Y^{\mathbf{n}} \underset{\text{ess}}{=} V^0 \cup (Q_1 \cup Q_2) \cup \bigcup_{ij} Z^{\mathbf{n}+1}_{ij}. \tag{31.25}$$

Let us extend map (31.24) to an *accelerated* (*branched*) *Markov map*

$$G : V^0 \cup (Q_1 \cup Q_2) \cup \bigcup_{m=1}^{\mathbf{n}+1} \bigcup_{ij} Z^m_{ij} \to Y^0, \tag{31.26}$$

by letting $G|\, V^0$ and $G|\, Z^{\mathbf{n}+1}_{ij}$ equal to $F^{\mathbf{n}+1}$, while $G|\, Q_k = F$. Under this map, each piece Z^m_{ij} is mapped univalently onto Y^0, each piece Q_1, Q_2 is mapped univalently onto $Y^{\mathbf{n}}$, while the piece V^0 is two-to-one mapped onto Y^0 (with branching at 0).

31.8. Solar system. We will now describe a Bernoulli scheme which performs a transit from the bottom to the top of a central cascade. Such transits will be treated as single steps of the generalized renormalization procedure (*cascade renormalization*).

Consider a central cascade (31.16) and let $g = g_m : V^m \to V^{m-1}$. Then the restrictions

$$g : V^k \setminus V^{k+1} \to V^{k-1} \setminus V^k, \quad k = m, \dots, m+N-1,$$

are double branched coverings. Pull the non-central puzzle pieces $V^m_i \subset V^{m-1} \setminus V^m$ from the top annulus to the consecutive annuli $V^{k-1} \setminus V^k$. We obtain a family of puzzle pieces $W^k_j \subset V^{k-1} \setminus V^k$ such that g^{k-m} univalently maps W^k_j onto some puzzle piece $V^m_i \equiv W^m_i$ (see Figure 31.4).

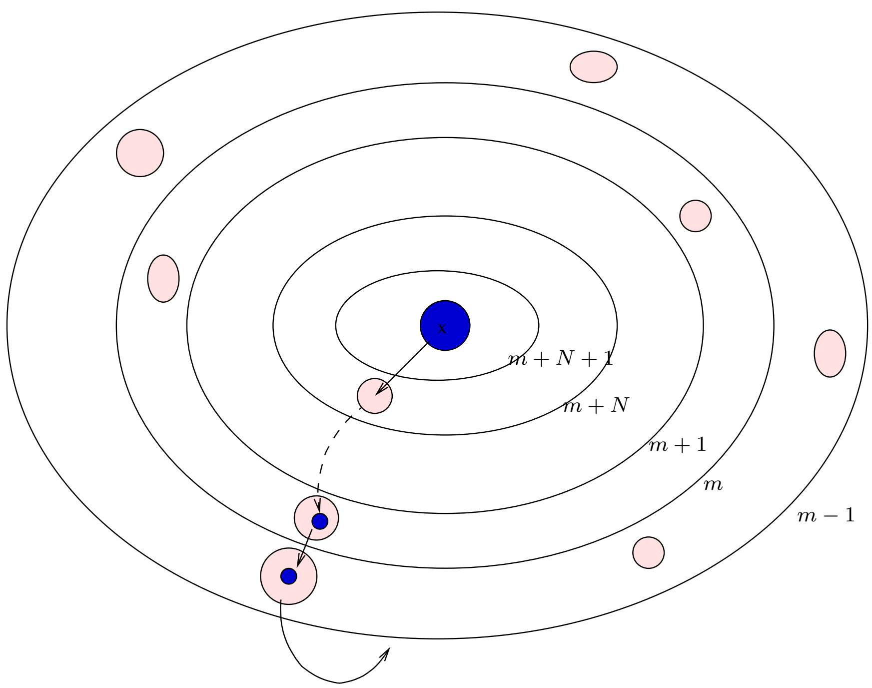

FIGURE 31.4. Solar system

Let us consider the following map:

$$G_m : \bigcup W^k_j \to V^{m-1}, \quad G\,|\, W^k_j = g_m \circ g^{k-m}. \tag{31.27}$$

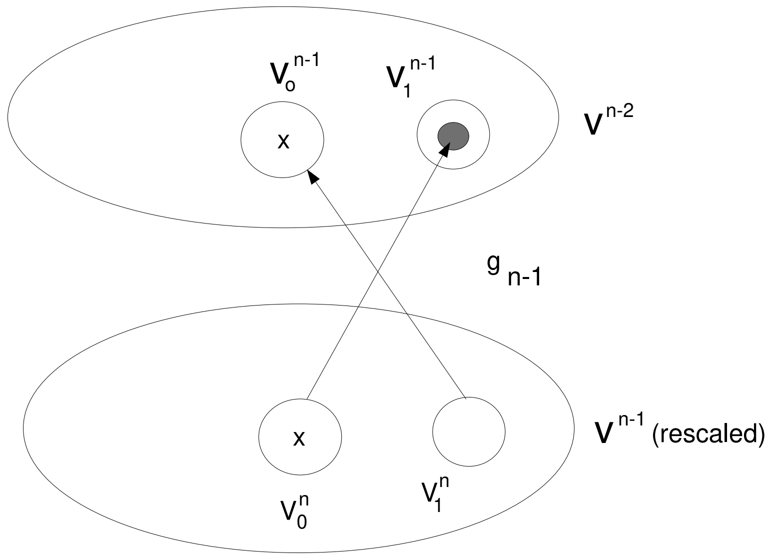

FIGURE 31.5. Fibonacci combinatorics

This map is (*unbranched*) *Bernoulli* in the sense that it univalently maps each domain W_j^k onto V^{m-1} (see §19.11.3). Together with the Telescoping Lemma, this picture implies:

LEMMA 31.20. (i) *Each branch of any first landing map*

$$L_k : \bigcup Q_i^k \to V^k, \quad k = m, \dots, m+N-1,$$

admits a univalent extension onto V^{m-1}.

(ii) *Each double covering*

$$g_n : V^n \to V^{n-1}, \quad n = m, \dots, m+N,$$

can be represented as $h_n \circ f = \phi_n \circ f_0$, *where* h_n *and* ϕ_n *are univalent maps with range* V^{m-1}.

(iii) *A similar representation* $h_{n,i} \circ f = \phi_{n,i} \circ f_0$ *holds for each univalent branch* $g_{n,i} : V_i^n \to V^{n-1}$, $i \neq 0$.

Note that in the last statement, both maps $f \,|\, V_i^n$ and $f_0 \,|\, V_i^n$ are univalent. However, they can still have big nonlinearity [if $\operatorname{diam} V_i^n$ is big compared to $\operatorname{dist}(V_i^n, 0)$].

31.8.1. *Fibonacci combinatorics.* Let us say that the *nth level* is *Fibonacci* if it is non-central (so, $g_n(0) \notin V^n$) but $g_n^2(0) \in V^n$. So, under g_n the critical point leaves the central piece V^n but then immediately returns back to it: this is the fastest combinatorial recurrence in the non-central case.

The following property justifies the name:

EXERCISE 31.21. *Let* t_n *be the return time of* 0 *to* V^n *under the original map* f. *Assume we have two Fibonacci levels in a row,* n *and* $n+1$. *Then* $t_{n+1} = t_n + t_{n-1}$.

31.9. Puzzle associated with a more general ray configuration. The key property we need for initiating the puzzle is to have an invariant ray configuration cutting the Julia set into several pieces. Above we used the first available one, $\mathfrak{R}(\alpha)$. Instead, we could start with any periodic ray configuration $\mathfrak{R} \equiv \bigcup \mathfrak{R}(\boldsymbol{\alpha}_i)$ associated with some repelling cut-cycles $\boldsymbol{\alpha}_i$.[22] Let $\mathcal{Y}[\mathfrak{R}]$ be the corresponding puzzle. This puzzle captures the biggest among little Julia sets K that are not cut by $\mathfrak{R}$ (if exists). If such a K does not exist, we say that f is *non-renormalizable with respect to* $\mathfrak{R}$. The whole theory comes through without essential changes except the specific construction of the first protected puzzle piece V^0 (see §31.5).

Assume now that f is n times DH renormalizable with little (filled) Julia sets $K^{[m]} = \mathcal{K}(R^m f)$, $m = 0, 1, \dots, n$. Let $\alpha_m, \beta_m \in K_m$ be their $\alpha-,\beta-$fixed points, and let $\boldsymbol{\alpha}_m, \boldsymbol{\beta}_m \in \mathcal{K}(f)$ be the corresponding f-cycles. Assume both fixed points α_n and β_n of $R^n f$ are repelling. Then we can start with a ray configuration

$$\mathfrak{R} := \bigcup_{m=0}^{n} \mathfrak{R}(\boldsymbol{\alpha}_m) \cup \mathfrak{R}(\boldsymbol{\beta}_m)$$

The corresponding puzzle tilings of the big Julia set $\mathcal{K}(f)$ induce tilings of the little Julia set $K^{[n]}$ that are exactly the same as the standard puzzle tilings of $K^{[n]}$ viewed as the Julia set of the renormalization $R^n f$. So, *if $R^n f$ is not immediately renormalizable*, then it has a protected puzzle piece as constructed in §31.5, which provides us with a protected puzzle piece $V^0[n]$ for the big puzzle $\mathcal{Y}[n] := \mathcal{Y}[\mathfrak{R}]$. It originates the principal nest

$$V^0[n] \supset V^1[n] \supset \dots \tag{31.28}$$

which is

- either finite (if f is combinatorially non-recurrent with respect to $\mathcal{Y}^{[n]}$),
- or captures the next little Julia set $K^{[n+1]} \subset K^{[n]}$ (if f is $n+1$ times renormalizable, with the $(n+1)$st renormalization being priitive),
- or has infinitely many non-central levels (if f is exactly n times renormalizable).

We say that the puzzle $\mathcal{Y}[n]$ is *associated with the nth renormalization level.* (Note that it is the puzzle for the original map rather than its renormalization!)

In the infinitely renormalizable case, we let $\mathcal{Y}[\infty] := \bigcup \mathcal{Y}[n]$.

31.10. Canonical Julia nest (revisited). Finally, we can identify all the renormalizations as DH, so that the canonical Julia nest described in §28.4.7 will become the full nest of little Julia set associated with various renormalizations. It will also provide us with the complex counterpart of Theorem 30.27.

THEOREM 31.22. *Any renormalization of a quadratic polynomial f is of DH type.*

PROOF. Let p be a renormalization period with little filled Julia set K. Let $1 < p_1 < p_2 < \dots$ be the DH renormalization periods with the nest of little Julia sets

$$\mathcal{K}(f) \equiv K^{[0]} \supset K^{[1]} \supset K^{[2]} \supset \dots \tag{31.29}$$

given by Corollary 28.29.

[22] We can also use non-cutting cycles as well if convenient for some reason, though they would not contribute to the refinement procedure.

Take the biggest n such that $p_n \le p$. Let $f_n : W_n \to W'_n$ be the degenerate ql map that produces $K^{[n]}$. By Exercise 28.28 (ii), p is a multiple of p_n and $K \subset K^n$.

If $p > p_n$, then by Theorem 31.17 (iii) (or rather, by its ql version applied to $R^n f$, see §31.9), $K = K^{[n+1]}$ and $p = p_{n+1}$, contradicting the choice of n.

So, $p = p_n$, and then $K = K^{[n]}$ by Exercise 28.28 (iii). $\square$

Thus, in Corollary 28.29 we can skip "DH" to obtain the canonical nest of little Julia sets, one for each renormalization period p_n:

$$\mathcal{K}(f) \equiv K^{[0]} \supset K^{[1]} \supset K^{[2]} \supset \dots . \tag{31.30}$$

Moreover, if f is n times renormalizable, and the α-fixed point of the last renormalization $R^n f$ is satellite parabolic, then we can complete this sequence with the *defenerate renormalization* $R^{n+1} f$ (which is a cauliflower ql map).

31.11. Real Puzzle and its Principal Nest. Assume now that f is a real ql map.[23] Recall that $\mathcal{I} = [-\beta, \beta]$, where β is the fixed point of f with multiplier $f'(\beta) \ge 1$. The other fixed point is called α. Assume both of them are repelling. Then α has a negative multiplier $f'(\alpha) < -1$ (see Exercises 28.4 (iv) and 28.10).

The real puzzle $\mathcal{Y}_\mathbb{R}$ is just the slice of the complex puzzle. It starts with a tessellation $\mathcal{Y}_\mathbb{R}^{(0)}$ of $\mathcal{I}$ into three intervals $Y_{\mathbb{R},i}^{(0)}$, the central interval $Y_\mathbb{R}^{(0)} \equiv Y_{\mathbb{R},0}^{(0)} = [-\alpha, \alpha]$ and two components of $\mathcal{I} \smallsetminus \operatorname{int} Y_R^{(0)}$. The further tessellations $\mathcal{Y}_\mathbb{R}^{(n)} = \{Y_{\mathbb{R},i}^{(n)}\}$ are defined as the pullbacks of $\mathcal{Y}_\mathbb{R}^{(0)}$ under the iterates f^n.

As in the complex situation, an interval J is called *nice* if

$$f^n(\partial J) \cap \operatorname{int} J = \emptyset, \quad n = 0, 1, \dots,$$

and *very nice* if $f^n(\partial J) \cap \bar{J} = \emptyset, \quad n = 0, 1, \dots$. One can also define *protected* real puzzle pieces in the same way as in the complex situation.

The real puzzle has the same basic properties as the complex one:

Exercise 31.23. *For any $f \in \mathfrak{Q}_\mathbb{R}$, we have:*

(i) *Real puzzle pieces of depth n are closed intervals whose boundary points belong to $f^{-n}\alpha$.*

(ii) *Under f, every real puzzle piece $Y_{\mathbb{R},i}^{(n)}$ of depth $n > 0$ is mapped to some puzzle piece $Y_{\mathbb{R},j}^{(n-1)}$ of depth $n-1$. This map is a diffeomorphism if $Y_{\mathbb{R},i}^{(n)}$ is off-critical, and is properly univalent if $Y_{\mathbb{R},i}^{(n)}$ is critical (i.e., if $i = 0$).*

(iii) *Any two real puzzle pieces P and Q are either nested or have disjoint interiors.*

(iv) *Markov Property. If $f(Y_{\mathbb{R},i}^{(n)})$ intersects $\operatorname{int} Y_{\mathbb{R},j}^{(n)}$, let $Y_{R,ij}^n := Y_{\mathbb{R},i}^{(n)} \cap f^{-1}(Y_{\mathbb{R},j}^{(n)})$. Then the map $f : Y_{R,ij}^{(n)} \to Y_{\mathbb{R},j}^{(n)}$ is either properly unimodal or a homeomorphism depending on whether $\operatorname{int} Y_{\mathbb{R},i}^{(n)} \ni 0$ or otherwise.*

(v) *Any real puzzle piece P is nice. If P is protected then it is very nice.*

Recall from §§28.4.6 and 31.4 the notion of *immediately renormalizable* maps. In the real case, there is only one immediately renormalizable combinatorics, the simplest one:

---

[23]In fact, the discussion can be easily adapted to more general classes of real unimodal maps, incuding class $\mathfrak{G}$.

Exercise 31.24. *The following properties of a real ql map $f : \mathcal{I} \to \mathcal{I}$ are equivalent:*

(i) *f is immediately renormalizable;*

(ii) *f is doubling renormalizable;*

(iii) $f^2(0) \in Y^{(1)}_{\mathbb{R}} \equiv [\alpha, \alpha']$.

Moreover, if f is not immediately renormalizable then:

a) $f^2(0) \in (\alpha', \beta]$, i.e., the first escaping moment $\mathbf{n}$ *is equal to* 1 *(see §31.4);*

b) If $f^2(0) \neq \beta$ (i.e., f is not Chebyshev) then orb 0 *eventually returns to $Y^{(0)}_{\mathbb{R}}$.*

The *Real Principal Nest* is the real slice of the complex one. It is the nest of intervals

$$I^0 \supset I^1 \supset \dots \tag{31.31}$$

such that I^{n+1} is inductively defined as the pullback of I^n under the first return map of the critical point to $\operatorname{int} I^n$. As in the complex situation, the choice of the initial interval I^0 is flexible. Our default choice will be $I^0 := [\alpha, -\alpha] \equiv Y^{(1)}_{\mathbb{R}}$ (compare Exercise 31.16).

As in the complex situation, the Principal Nest is infinite if and only if orb 0 visits all critical puzzle pieces (*combinatorial recurrence*).

Slicing *complex first landing maps* (31.3), we obtain *(principal) real first landing maps*

$$L_n : \bigsqcup J^n_i \to I^n \tag{31.32}$$

to the intervals I^n, where $\{J^n_i\}$ is a family of intervals with disjoint interiors mapped diffeomorphically onto I^n. One of these intervals coincides with I^n itself, call it J^n_0. Another one contains the critical value $f(0)$; call it J^n_1. Moreover, for $n \geq 1$, the intervals J^n_i are pairwise disjoint.

We can also consider *real first return maps*[24]

$$g_n : \bigsqcup I^n_i \to I^{n-1}, \tag{31.33}$$

to the intervals I^{n-1}. Then $g_n = L_n \circ f$, and in particular

$$g_n | \, I^n = h_n \circ f | \, I^n, \quad \text{where } h_n \equiv L_n | \, J^n_1. \tag{31.34}$$

Theorem 31.25. *Let a map $f : \mathcal{I} \to \mathcal{I}$ be non-renormalizable and let the critical point be combinatorially recurrent. Then the Principle Nest shrinks: $|I^n| \to 0$, and so do all the domains of the first landing maps:*

$$\max_i |J^n_i| \to 0.$$

Proof. The first assertion follows from Corollary 30.32 (which is a consequence of the topological exactness of the restriction $f | \, \mathcal{T}$). Exactness of $f | \, \mathcal{T}$ also implies shrinking of the intervals J^n_i contained in $\mathcal{T}$. To see this in general, notice that the intersection of a non-shrinking nest of intervals J^n_i would be a homterval. □

[24]It would be consistent to call real puzzle pieces $V^n_{\mathbb{R},i}$, but we prefer a simpler notation.

Finally, we can define a *real central cascade* of the Principal Nest,

$$I^{m-1} \supset I^m \supset \cdots \supset I^{m+N-1} \tag{31.35}$$

as the slice of a complex one. It is instructive to distinguish two types of cascades, *Chebyshev* (or *Ulam-Neumann*) and *parbolic* or (*saddle-node*), depending on whether the return is *high* (i.e., $\operatorname{int}(g_m(I^m)) \ni 0$) or *low*. In the former case, the return maps $g_m : I^m \to I^{m-1}$ with long central cascades are perturbations of Chebyshev maps, while in the latter case, they are perturbations of parabolic maps with multiplier 1.

Exercise 31.26. *Transfer the discussion of* §31.6 *to the real case.*

Exercise 31.27. *Review* §§25.6.7–25.6.12 *on Hubbard trees from the viewpoint of the interval theory just developed.*

Notes

On a brief history of the puzzle see Notes to §45.

The canonical nest of ql renormalizations and little Julia sets appeared in [**McM1**]. However, it was not identified with the DH renormalizations or with the renormalizations corresponding to the puzzle. It seems that a complete account of this folklore story provided by Theorem 31.17 has never been recorded before.

32. General combinatorial theory

In the first half of this section, we will discuss several models for the Julia set: lc, combinatorial, critical, and puzzle models. It could look confusing at first glance, but luckily, they all coincide under mild assumptions.

In the second half, we will discuss the Kneading Theory.

32.1. Various models.

32.1.1. *LC model.* In §9.4.2 we described a general locally connected model $K_{\rm lc}$ for a hull $K \subset \mathbb{C}$. It is obtained by taking the quotient of $(\mathbb{C}, \mathbb{D})$ by the clean geodesic lamination $\mathcal{L}_{\rm lc}$ generated by all cut-lines of K.

Applying this construction to the dynamical setting, we obtain a lamination $\mathcal{L}_{\rm lc}(f)$ and a lc model $\mathcal{K}_{\rm lc} \equiv \mathcal{K}_{\rm lc}(f)$ for the filled Julia set (and in particular, for the Julia set $\mathcal{J}_{\rm lc} \equiv \mathcal{J}_{\rm lc}(f)$). All these objects are endowed with dynamics:

Exercise 32.1. *The lamination $\mathcal{L}_{\rm lc}$ is completely invariant under the doubling map T. The map $f : (\mathbb{C}, \mathcal{K}, \mathcal{J}) \to (\mathbb{C}, \mathcal{K}, \mathcal{J})$ induces a natural topological double branched covering $F_{\rm lc} : (\mathbb{R}^2, \mathcal{K}_{\rm lc}, \mathcal{J}_{\rm lc}) \to (\mathbb{R}^2, \mathcal{K}_{\rm lc}, \mathcal{J}_{\rm lc})$. The map f is naturally semi-conjugate to $F_{\rm lc}$.*

32.1.2. *Critical model.* Let us say that a ray $\mathcal{R}^\theta$ (and its angle θ) is *valuable* if it converges to the puzzle end $E(v)$ of the critical value. (In particular, any ray landing at v is valuable.) As we know from Exercise 9.10 or from Lemma 32.37 (from Appendix), there is at least one valuable ray $\mathcal{R}^\theta$. Let $\gamma^\theta = (e(\theta/2), -e(\theta/2))$ be the corresponding diameter in $\mathbb{D}$. Taking the closure of its pullbacks and cleaning it up, we obtain a lamination $\mathcal{L}_{\rm crit}(f) := \mathcal{L}_\theta$ (see Appendix), to which we refer as the *critical lamination* of f. We let $F_{\rm crit} \equiv F_\theta$ be the corresponding model for f.

32.1.3. *Combinatorial model.* Let $f \equiv f_c$, $c \in \mathcal{M}$, be a quadratic polynomial with connected Julia set. Let us associate to f a geodesic lamination in $\mathbb{D}$ in the following way. For any periodic or pre-periodic cut-point $a \in \mathcal{J}$, let us consider all the rays $\mathcal{R}^{\theta_i}$ landing at a. By definition of a cut-point, there exist at least two such rays, and by Theorem 24.5, there exist only finitely many of them.

Mark points $e(\theta_i) \in \mathbb{T}$ on the unit circle and consider their hyperbolic convex hull in $\mathbb{D}$. We obtain a hyperbolic ideal polygon $P_a \subset \mathbb{D}$. The boundaries of all these polygons form the *rational* (*geodesic*) *lamination* $\mathcal{L}_{\mathbb{Q}}(f)$ of f. Its closure $\mathcal{L}_{\rm com}(f)$ is the *combinatorial* (*geodesic*) *lamination*. Obviously, it is completely invariant, so the doubling map projects to a topological double branched covering

$$F_{\rm com} : \mathbb{C}/\mathcal{L}_{\rm com} \to \mathbb{C}/\mathcal{L}_{\rm com}.$$

The quotient $\mathcal{K}_{\rm com} := \overline{\mathbb{D}}/\mathcal{L}_{\rm com}$ is called the *combinatorial filled Julia set of f*. Respectively, the quotient $\mathcal{J}_{\rm com} := \mathbb{T}/\mathcal{L}_{\rm com}$ is the *combinatorial Julia set of f*. The doubling map acts on each of them as a topological double branched covering.

Obviously, $\mathcal{L}_{\rm lc} \succ \mathcal{L}_{\rm com}$, so $F_{\rm lc}$ is naturally semi-conjugate to $F_{\rm com}$.

Two maps f and $\tilde{f}$ (and the corresponding parameters c and $\tilde{c}$) are called *combinatorially equivalent* if they have the same rational geodesic laminations: $\mathcal{L}_{\mathbb{Q}}(f) = \mathcal{L}_{\mathbb{Q}}(\tilde{f})$ (or equivalently, $\mathcal{L}_{\rm com}(f) = \mathcal{L}_{\rm com}(\tilde{f})$). In other words, they have the same landing pattern for the rational external rays.

Note that if the Böttcher conjugacy $h : \mathcal{D}(\infty) \to \tilde{\mathcal{D}}(\infty)$ extends continuously to the Julia sets, then the maps f and $\tilde{f}$ are combinatorially equivalent. In particular, by the $\mathcal{J}$-Stability Theorem 36.2, *for any component Δ of* $\operatorname{int}\mathcal{M}$*, any two parameters $c \in \Delta$ and $\tilde{c} \in \Delta$ are combinatorially equivalent.*

EXERCISE 32.2. *If two quadratic polynomials f and $\tilde{f}$ are topologically conjugate, then they are combinatorially equivalent.*

32.1.4. *Combinatorial laminations $\mathfrak{R}[n]$.* Let us consider a periodic ray configuration $\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\alpha})$ of some cut-cycle $\boldsymbol{\alpha}$. It generates a completely invariant lamination $\mathcal{L}(\mathfrak{R})$ as described in §25.7.2 for hyperbolic maps (see also §32.5.1 from the Appendix below). Namely, take the characteristic leaf $\gamma_{\rm ch}$ of this configuration and pull it back under the doubling map exactly as in the hyperbolic case. Taking the closure of all these leaves, we obtain the desired lamination.

LEMMA 32.3. *Assume that a quadratic polynomial f is DH renormalizable with respect to a periodic ray configuration $\mathfrak{R} \equiv \mathfrak{R}(\boldsymbol{\alpha})$ (as described in* §28.4.3*). Then the lamination $\mathcal{L}(\mathfrak{R})$ corresponds to the actual lamination of cut-lines for f. Moreover, none of these cut-lines cut through the little Julia sets.*

PROOF. Let $\mathcal{R}^{\theta_\pm}$ be the characteristic rays for $\mathfrak{R}$ (landing at the characteristic periodic point $\alpha_{\rm ch} \in \boldsymbol{\alpha}$). Let us consider a continuum concatenated of a ray $\mathcal{R}^{\theta_-/2}$ landing at the central periodic point $\alpha_0 \in \boldsymbol{\alpha}$, the little Julia set K, and the symmetric ray $\mathcal{R}^{\theta_-/2+1/2}$ landing at α_0'. It corresponds to the diameter σ in $\mathbb{D}$ connecting $\pm e(\theta_-/2)$.

Since the preimages of the characteristic cut-line $L_{\rm ch} = \mathcal{R}^{\theta_-} \cup \mathcal{R}^{\theta_+} \cup \alpha_{\rm ch}$ do not cut through K, the corresponding leaves in $\mathbb{D}$ do not cross the diameter σ. This is exactly the condition that determines the pullbacks of $\gamma_{\rm ch}$ (see §25.7.2). □

Under the circumstances of Lemma 32.3, we can also consider the lamination $\mathcal{L}_{\rm com}(f; \mathfrak{R})$ defined as the closure of all rational cut-lines through (pre-)periodic points that do not belong to little Julia sets corresponding to the ray configuration $\mathfrak{R}$. We call it the *combinatorial lamination* associated with the configuration $\mathfrak{R}$.

REMARK 32.4. As we will see later (Proposition 32.9), the above two laminations actually coincide: $\mathcal{L}_{\text{com}}(f;\mathfrak{R}) = \mathcal{L}(\mathfrak{R})$. We will also see in §37.3 that any periodic ray portrait $\mathfrak{R}$ corresponds to some hyperbolic map $f_\circ$, so the quotient of $\overline{\mathbb{D}}$ by the corresponding lamination $\mathcal{L}(\mathfrak{R})$ gives a topological model for $\mathcal{K}(f_\circ)$. The original map f will be interpreted as the "tuning" of $f_\circ$ by some quadratic map f_d (related to the renormalization of f), see §43.4.

If f is $n+1$ times renormalizable with $\mathfrak{R}$ generating the $(n+1)$st renormalization (where $n \in \mathbb{N}$), we will use notation $\mathcal{L}_{\text{com}}^{[n]} \equiv \mathcal{L}_{\text{com}}^{[n]}(f)$ for $\mathcal{L}_{\text{com}}(f;\mathfrak{R})$, and will refer to it as the *combinatorial lamination of* (*renormalization*) *level* n.

32.1.5. *Puzzle model.* Finally, let us consider the puzzle $\mathcal{Y}^{[n]}$ of some renormalization level n. Vertical sides of the corresponding puzzle pieces belong to a completely invariant family of cut-lines. The corresponding hyperbolic geodesics form a completely invariant geodesic lamination $\mathcal{L}_{\text{puz}}^{[n]}$ called the *puzzle lamination* of level n. The corresponding double branched covering

$$F_n : (\mathbb{R}^2, \mathcal{K}_{\text{puz}}^{[n]}) \to (\mathbb{R}^2, \mathcal{K}_{\text{puz}}^{[n]}).$$

is called the *puzzle model* for f of (renormalization) level n.

In this way, we obtain an increasing family of laminations. Their limit

$$\mathcal{L}_{\text{puz}} := \operatorname{cl} \bigcup \mathcal{L}_{\text{puz}}^{[n]}$$

is called the *puzzle lamination* for f. The corresponding quotient

$$F_{\text{puz}} : (\mathbb{R}^2, \mathcal{K}_{\text{puz}}) \to (\mathbb{R}^2, \mathcal{K}_{\text{puz}}), \quad \text{where } \mathcal{K}_{\text{puz}} := \overline{\mathbb{D}}/\mathcal{L}_{\text{puz}},$$

is called it the *puzzle model* for f. It is a topological double branched covering of the plane.

Obviously, $\mathcal{L}_{\text{com}} \succ \mathcal{L}_{\text{puz}}$.

LEMMA 32.5. *If f is periodically repelling then the lamination $\mathcal{L}_{\text{puz}}$ is polygonal.*

PROOF. By No Wandering Gaps Theorem (see Appendix), a non-polygonal gap eventually lands in a cycle $\mathcal{O}$ of gaps of some period p. By Proposition 32.41, $\mathcal{O}$ contains a central gap Q such that $0 \in \overline{Q}$. But then $f^p(Y) \supset Y$ for any critical puzzle piece Y, which is false for periodically repelling maps.

□

32.2. Comparison of different models. In the previous section we have associated to any quadratic polynomial f several completely invariant laminations:

$$\mathcal{L}_{\text{crit}} \text{ and } \mathcal{L}_{\text{lc}} \succ \mathcal{L}_{\text{com}} \succ \mathcal{L}_{\text{puz}}.$$

Here we will show that under fairly general conditions, all these laminations coincide.

32.2.1. $\mathcal{L}_{\rm com}$ *vs* $\mathcal{L}_{\rm puz}$. Let us start by showing that the puzzle *generates* the whole lamination $\mathcal{L}_{\rm com}$, meaning that any leaf of $\mathcal{L}_{\rm com}$ can be approximated by leaves corresponding to the vertical puzzle boundary.

LEMMA 32.6. (i) *Assume both fixed points of f are repelling. Let $a \in \mathcal{J}$ be a periodic point of f that does not belong to any little Julia set. Then*

$$\operatorname{diam} Y^{(m)}(a) \to 0.$$

(ii) *Assume f is n times renormalizable ($n \geq 1$). Let $a \in \mathcal{J}$ be a periodic point of f that does not belong to any little Julia set $K_i^{[n]}$ of renormalization level n. Then a is repelling and*

$$\operatorname{diam} Y^{[n-1](m)}(a) \to 0 \quad \text{as } m \to \infty.$$

(iii) *Assume f is exactly n times renormalizable ($n \geq 1$), with both fixed points of $R^n f$ repelling. Then any periodic point a is repelling and*

$$\operatorname{diam} Y^{[n](m)}(a) \to 0 \quad \text{as } m \to \infty. \tag{32.1}$$

(iv) *If f is infinitely renormalizable, then for any periodic point a there exists a renormalization level $n \in \mathbb{N}$ such that* (32.1) holds.

PROOF. (i) Let p be the period of a. Let us first show that *there exists a non-critical piece $Y \equiv Y^m(a)$ such that f^{-p} univalently maps Y into itself,* where f^{-p} is the branch of the inverse map fixing p.

Assume f is renormalizable. Let K be its little Julia set of top level, of period q, and let $\mathbf{K} := \bigcup_{i=0}^{q-1} f^i K)$ be its orbit . Then for any neighborhood $U \supset \mathbf{K}$, there exists a puzzle piece $Y^{(m)} \supset K$ such that the set $\mathbf{Y}^{(m)} := \bigcup_{i=0}^{q} f^i(Y^{(m)})$ is contained in U. (Note that this is true in both satellite and primitive cases.) Since $a \notin \mathbf{K}$, the set $\mathbf{Y}^{(m)}$ is disjoint from $\operatorname{orb} a$ for m big enough. It follows that the piece $Y^{(m)}(a)$, and its pullbacks along $\operatorname{orb} a$, are disjoint from $Y^{(m)}$, so $Y^{(m)}(a)$ has the desired property.

If f is non-renormalizable, let is consider the principal return maps $g_n = f^{l_n} : V^n \to V^{n-1}$. Since the pieces $f^k(V^n)$, $k = 0, 1, \dots, l_n - 1$, are pairwise disjoint, these domains do not contain periodic points of period $< l_n$. But $l_n \to \infty$ in the non-renormalizable case, implying that for n big enough,

$$\operatorname{orb} a \cap \bigcup_{k=0}^{l_n} f^k(V^n) = \emptyset.$$

Moreover, $V^{n-1} = Y^{(m)}$ for some m. It follows that the piece $Y^{(m)}(a)$, and its pullbacks along $\operatorname{orb} a$, are disjoint from V^n, implying the desired.

We can now slightly thicken the puzzle piece $Y \equiv Y^{(m)}$ to an open disk $D \supset Y$ such that $f^{-p}(D) \Subset D$ (compare §28.4.3). Applying the Schwarz Lemma, we complete the proof.

The proof of (ii) follows the same lines as (i) in the renormalizable case, using that the puzzle pieces $Y^{[n-1](m)}$ shrink (as $m \to \infty$) to the little Julia set $K^{[n]}$ of renormalization level n.

The proof of (iii) follows the same lines as (i) in the non-renormalizable case, using the puzzle $\mathcal{Y}[n]$ instead of $\mathcal{Y}[0]$.

In case (iv), a does not belong to the little Julia sets of some renormalization level n (see Exercise 28.32), so the conclusion follows from (ii). $\square$

COROLLARY 32.7. (i) *Repelling periodic points are perfectly rigid and well branched.*

(ii) *Yoccoz puzzle pieces associated with any periodic ray portrait* $\mathfrak{R}(\boldsymbol{\alpha})$ *are perfect.*

PROOF. By the above lemma, for any repelling periodic point β, there is a shrinking nest of puzzle pieces $P^m(\beta)$ whose interior contains β (these pieces could be finite unions of Yoccoz puzzle pieces that have β as a vertex). Hence β is rigid. It has only finitely many accesses by Theorem 24.5. Hence it is well branched by Corollary 9.13.

The initial puzzle pieces $Y_i^{(0)}$ are obtained by cutting a subpotential domain with the rays of $\mathfrak{R}(\boldsymbol{\alpha})$. The corresponding pieces of $\mathcal{K}$ are connected because the points of $\boldsymbol{\alpha}$ are well branched. Hence the puzzle pieces $Y_i^{(0)}$ are perfect. The deeper puzzle pieces are perfect as pullbacks of the $Y_i^{(0)}$.

Returning back to the repelling point β, we see that it is perfectly rigid since the puzzle pieces $P^m(\beta)$ are perfect. $\square$

LEMMA 32.8. *Assume f is periodically repelling. Then:*

(i) *If a puzzle impression $P_\infty \equiv P_\infty(E)$ (in the general sense of* §9.1.3) *is periodic, then P_∞ is a periodic point.*

(ii) *The lamination $\mathcal{L}_{\text{com}}$ is polygonal (and hence maximal).*

PROOF. (i) It is similar to the proof of the previous lemma. Let $p \in \mathbb{Z}_+$ be the period of our impression, so $f^p(P_\infty) \subset P_\infty$. It follows (as in Exercise 28.32) that P_∞ is not contained in the little Julia sets $K_i^{[n]}$ of sufficiently big renormalization level n. As in the proof of Lemma 32.6 (i), it implies existence of a Yoccoz puzzle piece $P \supset P_\infty$ which is univalently mapped under f^p onto itself. By thickening it and applying the Schwarz Lemma, we conclude that P_∞ contains a repelling periodic point a and the corresponding pullbacks $f^{-pn}(P) \ni a$ shrink to a.

(ii) Assume the lamination $\mathcal{L}_{\text{com}}$ has a non-polygonal gap Q. It corresponds to some puzzle end E. By the No Wandering Gaps Theorem (32.40), Q is preperiodic, so we can assume without loss of generality that it is periodic (of some period p). By (ii), Q is a periodic point a. By (i), the corresponding puzzle end is polygonal (maybe degenerate) – contradiction. $\square$

PROPOSITION 32.9. (i) *If both fixed points of f are repelling and f is non-renormalizable then the whole lamination $\mathcal{L}_{\text{com}}$ is generated by the puzzle $\mathcal{Y}$.*

(ii) *If f is n times renormalizable ($n \geq 1$) then the tuned lamination $\mathcal{L}_{\text{com}}^{[n-1]}$ is generated by the puzzle $\mathcal{Y}[n-1]$:*

$$\mathcal{L}_{\text{com}}^{[n-1]} = \mathcal{L}_{\text{puz}}^{[n-1]}.$$

(iii) *If f is exactly n times renormalizable, with both fixed points of $R^n f$ repelling, then the whole lamination $\mathcal{L}_{\text{com}}$ is generated by the puzzle $\mathcal{Y}[n]$:*

$$\mathcal{L}_{\text{com}} = \mathcal{L}_{\text{puz}}^{[n]} \equiv \mathcal{L}_{\text{puz}}.$$

(iv) *If f is infinitely renormalizable then the lamination $\mathcal{L}_{\text{com}}$ is generated by $\mathcal{Y}[\infty]$:*

$$\mathcal{L}_{\text{com}} = \mathcal{L}_{\text{puz}}.$$

Proof. (i) Let $a \in \mathcal{J}$ be any periodic point. Lemma 32.6 (i) implies that a can be separated by a puzzle cut-line from any other point $z \in \mathcal{J}$. Hence any ray landing at a can be approximated by puzzle cut-lines. Taking iterated pullbacks of all these rays, we obtain the family of all rational Green cut-lines. By definition, the corresponding hyperbolic geodesics generate $\mathcal{L}_{\text{com}}$.

(ii) Take any periodic point $a \in \mathcal{J}$ that does not belong to the little Julia sets $K_i^{[n]}$ of renormalization level n. Lemma 32.6 (ii) implies, in the same way as above, that any ray landing at a can be approximated by cut-lines of the puzzle $\mathcal{Y}[n-1]$. Iterated pullbacks of all these rays form the family of all rational Green cut-lines that do not intersect the $K_i^{[n]}$. By definition, the corresponding hyperbolic geodesics generate $\mathcal{L}_{\text{com}}^{[n-1]}$.

(iii) In this case, the rays landing at any periodic point $a \in \mathcal{J}$ are generated by the puzzle $\mathcal{Y}[n]$.

(iv) In this case, any periodic cycle $\mathbf{a}$ is repelling and lies outside some little Julia set $K^{[n]}$ (see Exercise 28.32). Hence the rays landing at a are generated by the puzzle $\mathcal{Y}[n]$. □

32.2.2. *All the laminations are the same.*

Theorem 32.10. *Let $f \equiv f_c$ be a periodically repelling quadratic polynomial with connected Julia set (i.e., $c \in \mathcal{M}$). Then:*

(i) *There exist at least one and at most finitely angles $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q}_{\text{odd}})/\mathbb{Z}$ such that $\mathcal{L}_{\text{com}}(f) = \mathcal{L}_\theta$. These angles are the external angles of the valuable dynamical rays $\mathcal{R}^\theta$. Moreover, if f is not critically preperiodic then there exists at most two such angles.*

(ii) *All of the above lamination coincide: $\mathcal{L}(f) := \mathcal{L}_{\text{crit}} = \mathcal{L}_{\text{lc}} = \mathcal{L}_{\text{com}} = \mathcal{L}_{\text{puz}}$.*

(iii) *The lamination $\mathcal{L}(f)$ is polygonal.*

Remark 32.11. We will see later (Theorem 47.15) that any polygonal lamination $\mathcal{L}_\theta$ can be realized as $\mathcal{L}_{\text{com}}$ for some quadratic polynomial.

Proof. By Proposition 32.9, $\mathcal{L}_{\text{com}} = \mathcal{L}_{\text{puz}}$. By Proposition 32.5, the latter is polygonal. Hence it is maximal among clean laminations. Since $\mathcal{L}_{\text{lc}} \succ \mathcal{L}_{\text{puz}}$, these two laminations coincide as well.

Let us show that $\mathcal{L}_{\text{crit}} = \mathcal{L}_{\text{com}}$. Let θ be a valuable angle, and let I_0 be the puzzle impression of the critical puzzle end. Since both rays $\mathcal{R}^{\theta/2}$ and $\mathcal{R}^{\theta/2+1/2}$ accumulate into I_0, the union

$$\Gamma^\theta := \mathcal{R}^{\theta/2} \cup I_0 \cup \mathcal{R}^{\theta/2+1/2}$$

is connected.

Let $L = \mathcal{R}^\phi \cup \{a\} \cup \mathcal{R}^\psi$ be a cut-line through any periodic point $a \in \mathcal{J}$. Since a is rigid, it is not contained in I_0. It follows that L is disjoint from Γ^θ, and hence the pairs $\{e(\phi), e(\psi)\}$ and $\{e(\theta/2), -e(\theta/2)\}$ are unlinked in $\mathbb{T}$. Hence the geodesics γ^θ and $(e(\phi), e(\psi))$ are disjoint in $\mathbb{D}$.

Thus, γ^θ is disjoint from all periodic leaves of $\mathcal{L}_{\rm com}$, and hence from their iterated pullbacks, which form $\mathcal{L}_{\mathbb{Q}}$. As the closure of $\mathcal{L}_{\mathbb{Q}}$ is $\mathcal{L}_{\rm com}$, we conclude that γ^θ is unlinked with $\mathcal{L}_{\rm com}$. By Proposition 32.5, the latter is polygonal, so there are two possibilities:

a) $\mathcal{L}_{\rm com}$ has a leaf passing though 0;

b) $\mathcal{L}_{\rm com}$ has a polygonal central gap Q_0;

In case a), the central leaf must coincide with γ^θ. Then by Lemma 32.38, $\mathcal{L}_\theta \succ \mathcal{L}_{\mathbb{Q}}$ and hence $\mathcal{L}_\theta \succ \mathcal{L}_{\rm com}$.

In case b), γ^θ is a diagonal of Q_0. By No Wandering Gaps Theorem, any gap Q in $\mathcal{L}_{\rm com}$ is an iterated preimage of Q_0. Since the gaps are dense in $\mathcal{L}$, so are the iterated preimages of γ^θ.

Remark 32.12. In case when Q_0 has > 4 sides, f is critically preperiodic, and the whole desired structure follows from the results of §27.1.

In either case, $\mathcal{L}_{\rm crit} \equiv \mathcal{L}_\theta = \mathcal{L}_{\rm com}$. $\square$

Corollary 32.13. *Let f be a periodically repelling quadratic polynomial with locally connected Julia set. Then it is topologically conjugate to its combinatorial model $F_{\rm com}$. This combinatorial model coincides with F_θ where θ is any valuable angle. Moreover, if f is not critically preperiodic then there exists at most two valuable angles.*

32.3. Cantor case. Let us now consider a quadratic polynomial $f \equiv f_c$ with Cantor Julia set $\mathcal{J}$, so $c \in \mathbb{C} \smallsetminus \mathcal{M}$. As was discussed in §23.5.3 and §24.3.3, it has a well defined valuable ray $\mathcal{R}_\theta$ landing at $v = c$ whose preimages, two symmetric critical rays $\pm\mathcal{R}^{\theta/2}$ crashing at 0, form a proper topological line L. Moreover, the further pullbacks of L form lines through the points of Crit^∞ comprising pairs of crashing rays. All other rays safely land at some points of $\mathcal{J}$.

Note that if $\theta \notin \mathbb{Q}_{\rm odd}/\mathbb{Z}$ (so the critical rays are not periodic), then the separatrices landing at 0 do not crash anymore, so they are safely land at some points of $\mathcal{J}$. Further pullbacks of these separatrices (landing at the precritical points) have the same property.

Let us partition the unit circle $\mathbb{T} \subset \mathbb{C}$ into two (open) half-circles $\mathbb{T}_0$ and $\mathbb{T}_1$ by the diameter connecting $\pm e(\theta/2)$, where $\mathbb{T}_0 \ni 1$ (if $\theta = 0$ or $1/2$ then let $\mathbb{T}_0$ be the upper half-circle).

For an angle $\gamma \in \mathbb{R}/\mathbb{Z}$ that does not land at θ under the iterates of the doubling map T, we can define the itinerary of γ rel this partition:

$$\bar{\varepsilon}(\gamma) = (\varepsilon_0, \varepsilon_1, \dots), \quad \text{where } \varepsilon_n \in \{0,1\},\ T^n\gamma \in \mathbb{T}_{\varepsilon_n}.$$

Let

$$\mathcal{P}^n := \mathbb{C} \smallsetminus \bigcup_{m=0}^{n} f^{-m}(\mathcal{R}^\theta).$$

Exercise 32.14. *Let P_0^1 and P_1^1 be components of $\mathcal{P}^1 \equiv \mathbb{C} \smallsetminus L$. The set $\mathcal{P}^n$ consists of 2^n components $P^n_{\varepsilon_0\dots\varepsilon_{n-1}}$, $\varepsilon_m \in \{0,1\}$, naturally labeled (for $n \geq 1$) by dyadic sequences according to their itineraries through P_0^1 and P_1^0.*

The following assertion describes the landing pattern of regular rays:

Proposition 32.15. *Under the above circumstances, two regular rays $\mathcal{R}^\gamma$ and $\mathcal{R}^{\gamma'}$ land at the same point of $\mathcal{J}$ iff $\bar{\varepsilon}(\gamma) = \bar{\varepsilon}(\gamma')$.*

Proof. Since the rays $\mathcal{R}^\gamma$ and $\mathcal{R}^{\gamma'}$ are regular, they are disjoint from the preimages of L. Hence each of them is contained in some component of $\mathcal{P}^n$, P_n and P_n' respectively. Since they land on $\mathcal{J}$, while L is disjoint from $\mathcal{J}$, the closure of our rays are contained in the same components. So, if they land at the same point of $\mathcal{J}$, then $P_n = P_n'$ for any $n \in \mathbb{N}$, which is equivalent for them to to having the same itineraries: $\bar\varepsilon(\gamma) = \bar\varepsilon(\gamma')$.

Let $D \equiv D^0$ be the (open) subpotential disk through the critical value $v = c$. Its preimage is the union of two disks D_1^0 and D_1^1, bounded by the figure-eight loops through 0. They are separated by the cut-line L, so $D_\varepsilon^0 \subset P_\varepsilon^1$, $\varepsilon \in \{0, 1\}$ (where P_0^1 and P_1^1 are the components of $\mathbb{C} \smallsetminus L$). Further preimages of these disks are organized into disks $D^n_{\varepsilon_0,\dots\varepsilon_{n-1}}$ according to their itineraries through $D_0^1 \subset P_0^1$ and $D_1^1 \subset P_1^1$, so $D^n_{\varepsilon_0\dots\varepsilon_{n-1}} \subset P^n_{\varepsilon_0\dots\varepsilon_{n-1}}$. (See the proof of Theorem 20.5 and Exercise 32.14.)

If the rays $\mathcal{R}^\gamma$ and $\mathcal{R}^{\gamma'}$ have the same itineraries $\bar\varepsilon = (\varepsilon_0\varepsilon_1\dots)$ then they belong to the same component $P^n_{\varepsilon_0\dots\varepsilon_{n-1}}$ and hence land in the same disk $D^n_{\varepsilon_0\dots\varepsilon_{n-1}}$. It follows that they land at the same point

$$a_{\bar\varepsilon} := \bigcap D^n_{\varepsilon_0\dots\varepsilon_{n-1}} \in \mathcal{J}.$$

□

32.4. Kneading Theory.

32.4.1. *Kneading data and conjugacy.* Let us now consider a real map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{G}'$. Assume for definiteness that its critical point is the *minimum*. The *kneading data* for f is the order of postcritical points 0_n, $n = 0, 1, \dots$, on the real line.

Two periodically repelling maps f and $\tilde f$ of class $\mathfrak{G}'$ are called $\mathbb{R}$-*combinatorially equivalent* if they have the same kneading data, i.e., the map $h : 0_n \mapsto \tilde 0_n$ is monotonic (with respect to the order induced from the real line).

Obviously, if two such maps $f, \tilde f \in \mathfrak{G}'$ are topologically conjugate on the real line then they are $\mathbb{R}$-combinatorially equivalent. The inverse statement is also true:

Proposition 32.16. *Two periodically repelling real maps of class $\mathfrak{G}'$ are $\mathbb{R}$-combinatorially equivalent if and only if they are topologically conjugate on the real line.*

Proof. Let f and $\tilde f$ be two periodically repelling maps with the same kneading data.

Let us consider the tiling $\mathcal{P}^n$ of $\mathcal{I}$ by the monotonicity intervals of f^n (obtained by dividing $\mathcal{I}$ by the critical points of f^n). Assume inductively that there is a homeomorphism $h_n : (\mathcal{I}, \mathcal{P}^n) \to (\tilde{\mathcal{I}}, \tilde{\mathcal{P}}^n)$ respecting these tilings, and let us construct h_{n+1}.

Let us take a monotonicity interval $J \in \mathcal{P}^n$, and as usually let $J_n \equiv f^n(J)$. If $\operatorname{int} J_n \not\ni 0$ then we let $h_{n+1} \,|\, J = h_n \,|\, J$. Otherwise 0 divides J_n into two intervals $J_n^\pm$ whose pullbacks $J^\pm$ to J under f^n are monotonicity intervals for f^{n+1}.

The interval J_n is bounded by post-valuable points v_i and v_j with $i, j \in [0, \dots, n-1]$, or by the fixed boundary point $\beta \in \partial\mathcal{I}$. To simplify notation, assume that the latter is not involved. Since $\tilde f$ has the same kneading data, the corresponding interval $\tilde J_n = [\tilde v_i, \tilde v_j]$ also contains 0 in its interior, so we obtain similar intervals $\tilde J_n^\pm$ and $\tilde J^\pm$. Moreover, the kneading data determines the order of the intervals $J_n^\pm$

on the real line and the orientation of the landing maps $J^\pm \to J^\pm_n$, and hence the order of the intervals $J^\pm$ as well. It follows that $\tilde{J}^\pm$ have the same order in $\mathbb{R}$ as $J^\pm$. Hence there exists a homeomorphism $h_{n+1} : J \to \tilde{J}$ coinsiding with h_n on ∂J and such that $h_{n+1}(J^\pm) = \tilde{J}^\pm$. Doing this on every interval J, we construct the desired h_{n+1}.

By the No Wandering Intervals Theorem, the monotonicity intervals $J \in \mathcal{P}^n$ and $\tilde{J} \in \tilde{\mathcal{P}}^n$ shrink as $n \to \infty$. Hence the homeomorphism h_n uniformly converge to a homeomorphism h (see Exercise 19.50). Since the h_n are equivariant on the respective boundaries $\partial\mathcal{P}^n$, h is a conjugacy. □

Let us now consider a superattracting map $f \in \mathfrak{G}'$ with the superattracting cycle $(0_n)_{n=0}^{p-1}$. Its kneading data is the interval $\mathcal{T}_f = [v, v_1]$ with the marked superattracting cycle (up to a natural equivalence relation), which coincides with the (abstract) Hubbard tree of f. As above, two such maps with the same kneading data are called $\mathbb{R}-$*combinatorially equivalent*.

Exercise 32.17. *Two superattracting maps $f, \tilde{f} \in \mathfrak{G}'$ are $\mathbb{R}$-combinatorially equivalent iff they are topologically conjugate on the real line.*

We refrain from extending the notion of $\mathbb{R}-$combinatorial equivalence to more general real hyperbolic and parabolic maps as there are nuances (indicated in the Exercises below) that would make it somewhat cumbersome and inclompatible with its complex counterpart.

Exercise 32.18. *Let f and $\tilde{f}$ be two real hyperbolic (but not superattracting) maps of class $\mathfrak{G}'$ with the same kneading data, whose attracting cycles have periods $p, \tilde{p}$ and multipliers $\rho, \tilde{\rho}'$, respectively.*

(i) *If $\operatorname{sign}\rho = \operatorname{sign}\tilde{\rho}$ then $\tilde{p} = p$, and the maps f, $\tilde{f}$ are topologically conjugacte on the real line.*

(ii) *If $\rho > 0 > \tilde{\rho}$ then $\tilde{p} = 2p$.*

Exercise 32.19. *Let f be a real parabolic map of class $\mathfrak{G}'$ whose parabolic cycle has period p and multiplier $\rho \in \{\pm 1\}$, and let $\tilde{f} \in \mathfrak{G}'$ be another map with the same kneading data that has a non-repelling cycle of period $\tilde{p}$ and multiplier $\tilde{\rho}$.*

(i) *If $\rho = 1$ then $\tilde{f}$ is parabolic or hyperbolic with $\tilde{p} = p$ and $\tilde{\rho} > 0$. Moreover, f & $\tilde{f}$ are topologically conjugate iff $\tilde{f}$ is parabolic.*

(ii) *If $\rho = -1$ then $\tilde{f}$ is parabolic or hyperbolic such that*
either $\tilde{p} = p$, $\tilde{\rho} < 0$, and f & $\tilde{f}$ are topologically conjugate on the real line,
or else $\tilde{p} = 2p$ and $\tilde{\rho} > 0$.

Remark 32.20. Note that Case (i) corresponds ot the saddle-node bifurcation, while Case (ii) corresponds to the doubling bifurcation.

32.4.2. *Itineraries.* Let $f \in \mathfrak{G}'$. Let us partition the interval $\mathcal{I}$ as

$$\mathcal{I} = \mathcal{I}_- \sqcup \mathcal{I}_+ \sqcup \{0\},$$

where $\mathcal{I}_- < \mathcal{I}_+$ are the components of $\mathcal{I} \smallsetminus \{0\}$ (Intrinsically, $\mathcal{I}_-$ is specified as the component whose orientation is reversed by f; it is "valuable": $\mathcal{I}_- \ni v$.) This allows us to encode any orbit $(x_n \equiv f^n x)_{n=0}^\infty$ by a sequence of three symbols, $\pm$ and 0.

$$\bar{\varepsilon}(x) := (\varepsilon_0(x), \varepsilon_1(x), \dots),$$

where $\varepsilon_n(x) = \pm$ or 0 depending on whether $x_n \in \mathcal{I}_\pm$ or $x_n = 0$. It is called the *itinerary* of x.

The itinerary depends *equivariantly* on the point: $\bar{\varepsilon}(fx) = \sigma(\bar{\varepsilon}(x))$, where $\sigma : (\varepsilon_0, \varepsilon_1, \dots) \mapsto (\varepsilon_1, \varepsilon_2, \dots)$ is the symbolic shift.

We let

$$\bar{\varepsilon}_N(x) := (\varepsilon_0(x), \varepsilon_1(x), \dots, \varepsilon_{N-1}(x)) \tag{32.2}$$

be the initial itinerary of length N.

EXERCISE 32.21. (i) *Points whose N-itineraries $\bar{\varepsilon}_N(x)$ contain* 0 *are pre-critical points of order $\le N-1$, i.e., the critical points of f^N.*

(ii) *If $\bar{\varepsilon}_N(x) = \bar{\varepsilon}_N(y)$, where $\bar{\varepsilon}_N(x)$ does not contain* 0*, then x and y belong to the same* (*open*) *monotonicity interval of f^N.*

(iii) *For two different points x and y, $\bar{\varepsilon}(x) = \bar{\varepsilon}(y)$ if and if x and y belong to the same* (*open*) *homterval.*

Let Σ be the total space of sequences of three symbols, $\pm$ and 0, as above. Let us endow Σ with the *twisted lexicographic order* as follows. Take two different sequences $\bar{\varepsilon} = (\varepsilon_n)$, $\bar{\delta} = (\delta_n)$. Let N be the first place where $\varepsilon_N \ne \delta_N$, and let $k < N$ be the number of "$-$" in the sequence $\bar{\varepsilon}_N = \bar{\delta}_N$. Then $\bar{\varepsilon} > \bar{\delta}$ if $(-1)^k \varepsilon_N > (-1)^k \delta_N$. (Thus, the rule is lexicographic in the case of even k, and reverse in the odd one.) This order is designed so that the itinerary depends *monotonically* on the point:

EXERCISE 32.22. *For two points x, y that do not belong to the same homterval, $x < y$ if and only if $\bar{\varepsilon}(x) < \bar{\varepsilon}(y)$.*

For a finite sequence (32.2), we let

$$\bar{\varepsilon}_N^{\mathrm{per}} = (\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{N-1})^{\mathrm{per}} \tag{32.3}$$

be its periodic extension.

EXERCISE 32.23. *If the itinerary of some point x is a periodic sequence $\bar{\varepsilon}_p^{\mathrm{per}}$ with the smallest period p, then x converges to a cycle of minimal period p or $2p$* (*where x can be periodic itself*)*.*

32.4.3. *Kneading sequence.* The *kneading sequence* Kn_f is defined as the *valuable code*

$$\mathrm{Kn}_f := (\mathrm{kn}_1, \mathrm{kn}_2, \dots), \quad \text{where } \mathrm{kn}_m := \varepsilon_m(0) = \varepsilon_{m-1}(v),$$

with the convention that it stops with the first appearance of 0. Thus, the kneading sequence keeps track of the itinerary of v through the intervals $\mathcal{I}_\pm$ until the possible landing at 0. It is finite if and only if f is superattracting, in which case

$$\mathrm{Kn}_f = (\mathrm{kn}_1, \dots, \mathrm{kn}_{p-2}, 0)$$

where p is the period of 0.

LEMMA 32.24. *Kneading data determines the kneading sequence, and the other way around.*

PROOF. The first statement is obvious. To see the reverse, notice that by the equivariance of the coding, the kneading sequence determines the itineraries of all the postcritical points 0_n, $n \in \mathbb{N}$. By Exercise 32.22, this determines their order on the real line. $\square$

32.4.4. *Kneading model.* For a given map $f : \mathcal{I} \to \mathcal{I}$, let us say that a sequence $\bar{\varepsilon} \in \Sigma$ is *admissible* if

$$\sigma^n(\bar{\varepsilon}) \le \sigma(\mathrm{Kn}_f), \quad n = 0, 1, \dots$$

REMARK 32.25. Notice that $\sigma^{n+1}(\bar{\varepsilon}) \le \sigma(\mathrm{Kn}_f) \implies \sigma^n(\bar{\varepsilon}) \ge \mathrm{Kn}_f$, $n \in \mathbb{N}$.

We let Σ_f be the space of admissible sequences for f.

PROPOSITION 32.26. *A sequence $\bar{\varepsilon} \in \Sigma$ is admissible iff it is the itinerary $\bar{\varepsilon}(x)$ of some point $x \in \mathcal{T}$.*

PROOF. Obviously, any sequence $\bar{\varepsilon}(x)$, $x \in \mathcal{T} = [v, v_1]$, is admissible.

Vice versa, let us show inductively that any finite admissible sequence $(\varepsilon_0, \varepsilon_1 \dots \varepsilon_{N-1})$ with $\varepsilon_n \ne 0$[25] is represented by some monotonicity interval of f^N. Take some admissible sequence $\bar{\varepsilon} = (\varepsilon_0, \varepsilon_1 \dots \varepsilon_N)$ of order $N + 1$. Since $\sigma(\bar{\varepsilon})$ is an admissible sequence of order N, by induction it is represented by some monotonicity interval $J \subset \mathcal{T}$.

Assume $\varepsilon_0 = -1$. Since the image of $I_- = (v, 0)$ monitonically covers the whole interval $\mathcal{T} \supset J$, I_- contains a pullback of J representing the desired monotonicity interval.

Assume $\varepsilon_0 = 1$. Let $\bar{\delta} = (1, \delta_1, \dots \delta_N)$ be the initial code for v_1. Then w $\bar{\varepsilon} \le \bar{\delta}$ implies that $(\varepsilon_1 \dots \varepsilon_N) \le (\delta_1 \dots \delta_N)$ It follows that the left end-point of J lies to the left of $v_2 = f(v_1)$. But then the monitonic image $f(I_+) = (v, v_2)$ overlaps with J, and the conclusion follows. $\square$

In particular, the kneading sequence itself must be admissible:

$$\sigma^n(\mathrm{Kn}_f) \le \sigma(\mathrm{Kn}_f), \quad n = 0, 1, \dots$$

We will see later (Theorem 33.9) that vice versa, any admissible kneading sequence Kn is realizable for some quadratic polynomial.

We let $\mathfrak{Kn}^{\mathrm{adm}}$ be the space of admissible kneading sequences.

The twisted lexicographic order induces an order on kneading sequences. We say that a kneading sequence Kn_1 as *stronger* than Kn_2, $\mathrm{Kn}_1 \succ \mathrm{Kn}_2$, if $\mathrm{Kn}_1 < \mathrm{Kn}_2$ (or, equivalently $\sigma(\mathrm{Kn}_1) > \sigma(\mathrm{Kn}_2)$). By Proposition 32.26, *the stronger a kneading sequence is, the larger its kneading model.*

EXERCISE 32.27. *Show that the ordered set of admissible kneading sequences has two extreme points, the strongest one* $(- + + + + \dots)$ (*corresponding to the Chebyshev combinatorics*) *and the weakest one* $(+ + + + \dots)$ (*corresponding to the cauliflower combinatorics*).

EXERCISE 32.28. *Let* $\mathrm{Kn} \in \Sigma$, *and let* $\mathrm{orb}^+(\mathrm{Kn})$ *be the set of its positive orbit points* $\sigma^m(\mathrm{Kn}) > 0$, *i.e., such that* $\mathrm{Kn}_{m+1} = +$ (*and similarly,* $\mathrm{orb}^-(\mathrm{Kn})$ *is the set of negative orbit points*)*. Show that admissibility of* Kn *is equivalent to unimodality of the shift* σ *on* $\mathrm{orb}(\mathrm{Kn})$, *i.e.,* σ *is monotonically increasing on the set* $\mathrm{orb}^+(\mathrm{Kn})$ *and monotonically decreasing on* $\mathrm{orb}^-(\mathrm{Kn})$.

Endow the space of three symbols $\pm, 0$ with (non-Hausdorff) topology whose non-trivial open subsets are $\{+\}$ and $\{-\}$. Then endow Σ, and hence each Σ_f, with the corresponding weak topology.

[25] We leave to the reader to adjust the argument to the case when $\varepsilon_n = 0$ for some n.

EXERCISE 32.29. *The itinerary map* $\bar{\varepsilon} : \mathcal{I} \to \Sigma_f$ *monotonically semi-conjugates* f *to the shift* $\sigma : \Sigma_f \to \Sigma_f$. *Its fibers are either singletons or maximal homtervals.*

32.4.5. *Real combinatorial equivalence vs complex one.* For real quadratic polynomials we have introduced two notions of combinatorial equivalence. Let us now show that they match:

PROPOSITION 32.30. *Two periodically repelling or superattracting real quadratic polynomials* $f \equiv f_c$ *and* $\tilde{f} \equiv f_{\tilde{c}}$ $(c, \tilde{c} \in \mathcal{M}_{\mathbb{R}})$ *are* $\mathbb{R}-$*combinatorially equivalent if and only if they are combinatorially equivalent.*

PROOF. In the superattracting case, this folllows from the corresponding statement for general Hubbard trees (see Theorem 25.58 and Proposition 25.61).

In the periodically repelling case, we know from §32.1.2 and Theorem 32.2.2 that the combinatorial lamination is generated by the pullbacks of the critical diameter γ^θ. The corresponding point on the model Julia set $\mathcal{J}_{\rm com}$ is the critical point for the model map F. It lies on the spine $\sigma_{\rm com}$ of $\mathcal{J}_{\rm com}$ corresponmding to the dialmeter $\mathbb{I} \subset \overline{\mathbb{D}}$. Moreover, the dynamics of F on $\sigma_{\rm com}$ is unimodal.

Every point $z \in \mathcal{J}_{\rm com}$ belongs to some leaf. The endpoints of the landing leaves give us *combinatorial external angles* for z.

EXERCISE 32.31. *The external angles of a point* $x \in \sigma_{\rm com}$ *determine its itinerary on the spine* $\sigma_{\rm com}$, *and the other way around.*

In particular, the valuable angle θ determines the kneading sequence, and the other way around.

□

From now on, we will sometimes allow ourselves to skip "$\mathbb{R}$" in the notion of combinatorial equivalence.

EXERCISE 32.32. *Develop the kneading theory for continuous unimodal maps. An interesting particular case are saw-like maps.*

32.5. Appendix: Invariant geodesic laminations.

Under construction

32.5.1. *Invariance.* For two points $a, b \in \mathbb{T}$, we let $(a, b)_\#$ be the (non-oriented) geodesic in $\mathbb{D}$ with endpoints a and b.

Here we will use notation T for the "doubling map" $z \mapsto z^2$ of $\mathbb{T}$. Given a non-diameter geodesic $\gamma = (a, b)_\#$, let $\hat{T}(\gamma) := (T(a), T(b))_\#$. In other words, this action is induced by the diagonal action of the doubling map on the torus $\mathbb{T}^2$, $(a, b) \mapsto (T(a), T(b))$. (If $\gamma = (-a, a)_\#$ is a diameter then we let $\hat{T}(\gamma) = \{T(a)\} \subset \mathbb{T}$ be a "degenerate geodesic".)

More generally, we can consider any convex subset Q of $\mathbb{D}$ with totally geodesic boundary, and define $\hat{T}(Q)$ as the convex hull of $T(\partial^I Q)$ in $\mathbb{D}$ (or a singleton in $\mathbb{T}$, in case when Q is a single diameter).

A geodesic lamination $\mathcal{L}$ in $\mathbb{D}$ is called T- *invariant* if for any non-diameter leaf $\gamma = (a, b)_\#$, the geodesic $\hat{T}(\gamma)$ is also a leaf of $\mathcal{L}$.

EXAMPLE 32.33. (i) Any single diameter $\gamma_\theta = (e(\theta/2), -e(\theta/2), \theta \in \mathbb{R}/\mathbb{Z}$, is an invariant lamination on its own right.

(ii) The next simplest example is provided by a *periodic ray portrait* $\mathfrak{R}$, which is a finite lamination whose leaves are cyclically permuted under T. Namely, let us consider a finite sets $\Theta \subset \mathbb{T}$ of period q decomposed into the union of p *unlinked* subsets Θ_i of cardinality $r \geq 2$ cyclically permuted under T. (So, $q = pr$.) Consider convex hulls Q_i of the Θ_i. *Assume the* Θ_i *are rotation sets under* T^p. Then the boundary geodesics of the Q_i's form a finite invariant lamination $\mathfrak{R}$ (compare §24.5).

Exercise 32.34. *Why is it needed for the* Θ_i *in the above Example to be* rotation sets*?*

Proposition 32.35. *Gaps are dense in any invariant geodesic lamination except* $\mathcal{L}_{Ч}$.

Proof. Otherwise, $\operatorname{supp}\mathcal{L}$ contains a saturated rectangle $\Pi = (I, J)$, where I and J are disjoint intervals on $\mathbb{T}$ (see Exercise 2.70). If none of the leaves in Π is a diameter, then the intervals $I_1 := T(I)$ and $J_1 := T(J)$ are disjoint, and there is a rectangle $\Pi_1 \subset \mathbb{D}$ foliated by leaves with endpoints in I_1 and J_1. The length of the intervals doubles, so this operation can be applied only finitely many times. Eventually we obtain a rectangle as above containing a diameter as a leaf. Cut this rectangle into two sub-rectangles by the diameter and apply $\hat{T}$ to any of them. We obtain a foliated topological sector in the support of $\mathcal{L}$. Iterating it further, we will eventually obtain a domain containing a foliated half-disk. One more iterate turns into the full foliated disk. This foliation has two ideal singular points, $a, b \in \mathbb{T}$, whose set is invariant under T.

Let us show that these points cannot form a cycle of period two. Indeed, one of them, say a, is the image of the diameter leaf γ of $\mathcal{L}$. This leaf should end at b, for otherwise a would have three preimages on $\mathbb{T}$ under T. But then γ crosses the leaves of $\mathcal{L}$ near b.

Hence one of the singular points a, b is fixed and the other is its preimage. This is the picture of the Chebyshev foliation. ☐

A lamination is called *symmetric* if it is invariant under the central reflection $z \mapsto -z$ of $\mathbb{D}$. A lamination is called *completely* T*-invariant* if:

- It is invariant and symmetric;
- For any leaf γ, there exists a leaf λ such that $\hat{T}(\lambda) = \gamma$;
- Any gap in $\mathcal{L}$ is mapped by $\hat{T}$ onto a gap or a leaf.

The simplest example of a completely invariant lamination is the Chebyshev one, $\mathcal{L}_{Ч}$ (see §2.5) (called so because T acts on the quotient of $\overline{\mathbb{D}}/\mathcal{L}_{Ч}$ as the Chebyshev map).

A systematic way of producing completely invariant laminations is by *dynamical saturations* of invariant laminations:

Lemma 32.36. *Let* $\mathcal{L}_0$ *be an invariant lamination containing a diameter* γ_θ, $\theta \in \mathbb{R}/\mathbb{Z}$. *Then* $\mathcal{L}_0$ *can be pulled back to laminations* $\mathcal{L}_n := (\hat{T}^*)^n(\mathcal{L}_0)$, *and*

$$\mathcal{L} := \operatorname{cl} \bigcup_{n=0}^{\infty} \mathcal{L}_n$$

is a completely invariant lamination.

PROOF. Consider any geodesic $\delta = e(\phi), e(\psi))_{\#}$ of $\mathcal{L}$, where $\phi \neq \psi \mod 1$. Then one of these angles is different from θ, so assume $\phi \neq \theta \mod 1$. Then each preimage of ϕ under T, $\phi' = \phi/2$ and $\phi'' = \phi' + 1/2$, is different mod 1 from both $\theta' \equiv \theta/2$ and $\theta'' \equiv \theta' + 1/2$. Moreover, the points $e(\phi')$ and $e(\phi'')$, being 0-symmetric, lie on the opposite sides of the diameter $\gamma \equiv \gamma_\theta$.

If $\psi \neq \theta$ as well, then for the same reason, the preimages $e(\psi')$ and $e(\psi'')$ of $e(\psi)$ lie on the opposite sides of γ. Hence there is a unique way to connect $e(\phi')$ to either $e(\psi')$ or $e(\psi'')$ in $\mathbb{D}$ so that the corresponding geodesic is disjoint from γ. This gives us one lift of δ. The other lift is 0-symmetric to the first one.

In case $\psi = \theta$, we connect each $e(\theta')$ and $e(\theta'')$ to both $e(\phi')$ and $e(\phi'')$, to obtain *four lifts* of δ.

If we have to geodesics δ_1 and δ_2 of $\mathcal{L}$, then their lifts are unlinked. Indeed, let λ_k be a lift of δ_k, $k = 1.2$. They are obviously unlinked If they lies in different components $\mathbb{D} \smallsetminus \gamma$. If they lie in the same component and are linked, then their images $\hat{T}(\lambda_k) = \delta_k$ would be linked as well.

Thus, the lifts of all the leaves of $\mathcal{L}$ form a lamination, which we call $\hat{T}^*(\mathcal{L})$. Since $\mathcal{L}$ is invariant, $\hat{T}^*(\mathcal{L}) \succ \mathcal{L}$.

Iterating this procedure, we obtain an increasing sequence of pullback laminations $(\hat{T}^*)^n(\mathcal{L})$, so their union is a pre-lamination. By Exercise 2.75, its closure is a genuine lamination. □

In particular, we can produce a completely invariant lamination from a single diameter $\gamma^\theta = (e(\theta/2, -e(\theta/2))_{\#})$ viewed as an invariant lamination. We call this completely invariant lamination $\mathcal{L}_\theta$.

We say that a complete geodesic γ is *unlinked* with $\mathcal{L}$ if it is disjoint from $\operatorname{supp} \mathcal{L}$.

LEMMA 32.37. *For any symmetric lamination $\mathcal{L}$, there exists a diameter γ_θ which is either a leaf of $\mathcal{L}$ or is unlinked with $\mathcal{L}$.*

PROOF. The origin $0 \in \mathbb{D}$ either belongs to $\operatorname{supp} \mathcal{L}$ or to a gap Q. In the former case, 0 belongs to some leaf γ of $\mathcal{L}$, which is necessarily a diameter. In the latter case, the ideal boundary $\partial^I Q$ is symmetric since $\mathcal{L}$ is such. Then any diameter $(a, -a)$ with $a \in \partial^I Q$ is contained in Q and hence is unlinked with $\mathcal{L}$. □

LEMMA 32.38. *Let us consider a diameter $\gamma \equiv \gamma^\theta$, $\theta \in \mathbb{R}/\mathbb{Z}$. Assume that a periodic ray portrait $\mathfrak{R}$ is disjoint from γ. Then for any leaf δ of $\mathfrak{R}$, there exists a sequence of pullbacks $(\hat{T}^*)^{n_k}(\gamma)$ converging to δ.*

LEMMA 32.39. *For any completely invariant lamination $\mathcal{L}$, there exists a continuous map $\hat{T}_{\mathcal{L}} : (\mathbb{C}, \overline{\mathbb{D}}) \to (\mathbb{C}, \overline{\mathbb{D}})$ with the following properties:*

(i) *$\hat{T}_{\mathcal{L}}$ is an extension of the doubling map $T : \mathbb{T} \to \mathbb{T}$;*

(ii) *$\hat{T}_{\mathcal{L}}$ collapses some diameter γ_θ from* Lemma 32.37 *to a singleton $e(\theta) \in \mathbb{T}$, and it homeomorphically maps each semi-disk $\overline{\mathbb{D}} \smallsetminus \gamma_\theta$ onto $\overline{\mathbb{D}} \smallsetminus \{e(\theta)\}$.*

(iii) *$\mathcal{L}$ is invariant under $\hat{T}_{\mathcal{L}}$.*

(iv) *The map $\hat{T}_{\mathcal{L}}$ descends to a topological double branched covering $F : (\mathbb{R}^2, K) \to (\mathbb{R}^2, K)$, where $(\mathbb{R}^2, K) \approx (\mathbb{C}/\mathcal{L}, \overline{\mathbb{D}}/\mathcal{L})$.*

32.5.2. *No Wandering Gaps.* A gap Q of a completely invariant lamination $\mathcal{L}$ is called *wandering* if all its forward iterates $\hat{T}^n(Q)$ are pairwise disjoint gaps.

THEOREM 32.40. *A completely invariant geodesic lamination does not have wandering gaps.*

Proof to be added

Thus, *some iterate $\hat{T}^n(Q)$ of a gap Q is either a periodic gap or a leaf.* In the latter case, there is a smaller iterate $\hat{T}^m(Q)$, $m < n$, which is either a 0-symmetric rectangle or a triangle based on a diameter of $\mathbb{D}$. To understand the former case, we need to classify periodic gaps.

32.5.3. *Periodic gaps.*

PROPOSITION 32.41. *Let $\mathcal{L}$ be a completely invariant geodesic lamination, and let Q be a periodic gap in $\mathcal{L}$ of period p that does not contain 0. Then the ideal boundary $\partial^I Q$ is a rotation set for T^p. In particular, some iterate $\hat{T}^n(Q)$, $0 \le n < p$, is a "central" gap containing 0 in its boundary.*

32.5.4. *Rotational laminations.* Any rotation set with irrational rotation number (see §24.7) gives rise to a *rotational lamination $\mathcal{L}_\theta$* corresponding to the diameter $(e(\theta/2), -e(\theta/2))$.

PROBLEM 32.42. *Let Θ_θ be the rotation set from* Problem 24.28. *Let $G \subset \mathbb{D}$ be the hyperbolic converx hull of Θ_θ, and let G_i be its pullbacs under T.*

(i) *Show that the G_i are attached one to another along geodesics forming a tree of domains.*

(ii) *Show that the G_i are the gaps of the lamination $\mathcal{L}_\theta$, and the latter consists of the geodesic boundaries of the G_i,.*

(iii) *Formulate the corresponding assertions for the tuned rotational laminations.*

Notes. The Milnor-Thurston Kneading Theory was developed in the mid 1970s but was not published until one decade later [**MT**]. Its preliminary version had been developed by Metropolis, Stein & Stein [**MeStSt**] (1973). In fact, the kneading sequences for hyperbolic parameters had been already introduced by Myrberd one decade earlier [**Myr2**].

Combinatorial theory of Julia sets was initiated by Douady and Hubbard (probably, inspired by the Kneading Theory); its foundations appeared in the fundamental Orsay Notes [**DH2**]. It was further detailed and refined by their scientific school (Lavaurs, Tan Lei, Schleicher, and others), and by Milnor with his school (Poirier, Kiwi, and others).

Geodesic laminations were introduced by Thurston in the mid-1980s. (For almost three decades this work had existed only as a preprint, until it finally appeared in [**Th1**].) In particular, it contains the analysis of quadratic invariant laminations. Good part of the theory was generalized to the higher degree case by Kiwi [**Ki**] (based on the 1997 Stony Brook thesis). However, Thurston's No Wandering Gaps Theorem turned out to be special for the quadratic case, as Blokh and Oversteegen demonstrated [**BO1**].

CHAPTER 5

Parameter plane

33. Definition and first properties

33.1. Notational convention. Recall that we usually label the objects corresponding to a map f_c by c, e.g., $\mathcal{J}_c = \mathcal{J}(f_c)$, $\mathrm{Per}(f_c) = \mathrm{Per}_c$. We often use notation $c_\circ \equiv \circ$ for a base parameter, so that $f_\circ = f_{c_\circ}$, $\mathcal{J}_\circ = \mathcal{J}_{c_\circ}$, etc.

33.2. Connectedness locus and polynomials $c \mapsto f_c^n(0)$. The Mandelbrot set presents at one glance the whole dynamical diversity of the complex quadratic family $f_c : z \mapsto z^2 + c$. Figures 0.1 and 0.2 from Preface show this set and its blow-ups in several places.[1] It is remarkable that all this intricate structure is hidden behind the following one-line definition.

Recall the Basic Dichotomy for the quadratic maps: the Julia set $\mathcal{J}_c$ is either connected or Cantor (Theorem 20.5). By definition, the *Mandelbrot set* $\mathcal{M}$ consists of those parameter values $c \in \mathbb{C}$ for which the Julia set $\mathcal{J}_c$ is connected. It is equivalent to saying that the orbit of the critical point

$$0 \mapsto c \mapsto c^2 + c \mapsto (c^2 + c)^2 + c \mapsto \dots \tag{33.1}$$

[1] The reader can also find cosmic-style animations of $\mathcal{M}$ on the YouTube.

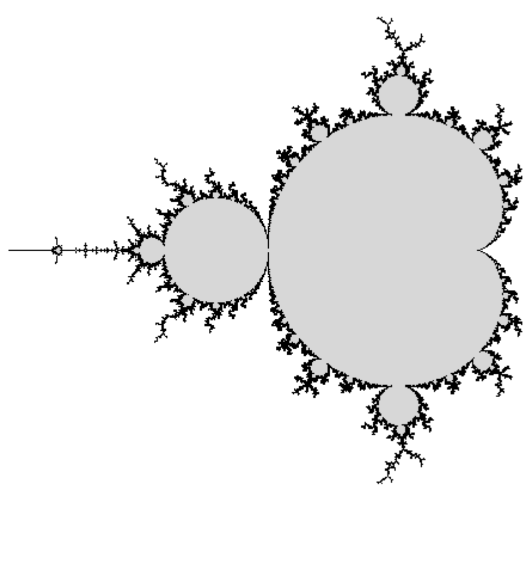

FIGURE 33.1. Mandelbrot set.

is not escaping to ∞. Let us denote the nth polynomial in (33.1) by $\upsilon_n(c)$, so that $\upsilon_0(c) \equiv 0$, $\upsilon_1(c) \equiv c$, and recursively

$$\upsilon_{n+1}(c) = \upsilon_n(c)^2 + c. \tag{33.2}$$

Note that $\deg \upsilon_n = 2^{n-1}$.

Though the polynomials υ_n are not iterates of a single polynomial, they behave in many respects similarly to the iterated polynomials:

Exercise 33.1 (Simplest properties of $\mathcal{M}$). *Prove the following properties:*

(i) *If $|\upsilon_n(c)| > 2$ for some $n \in \mathbb{N}$ then $\upsilon_n(c) \to \infty$ as $n \to \infty$. Hence $\mathcal{M} \subset \overline{\mathbb{D}}_2$.*

(ii) *$\upsilon_n(c) \to \infty$ locally uniformly on $\mathbb{C} \smallsetminus \mathcal{M}$. Hence $\mathcal{M}$ is compact.*

(iii) *$\mathbb{C} \smallsetminus \mathcal{M}$ is connected. Hence $\mathcal{M}$ is full and all components of* $\operatorname{int} \mathcal{M}$ *are simply connected.*

(iv) *The set of normality of the sequence (υ_n) coincides with $\mathbb{C} \smallsetminus \partial\mathcal{M}$.*

One says that the critical point 0 is *active* for a parameter $c_\circ$ if the sequence of polynomials $(\upsilon_n(c))$ is not normal near $c_\circ$, and is *passive* otherwise. We see that the set of active parameters coincides with $\partial\mathcal{M}$. For an active parameter, the behavior of the critical orbit highly sensitive to perturbations.

In the above discussion, one can already see a similarity between the Mandelbrot set (representing the whole quadratic family) and a filled Julia set of a particular quadratic map. It is just the first indication of a deep relation between dynamical and parameter objects.

33.3. Dependence of periodic points on c. What immediately catches the eye in the Mandelbrot set is the *main cardioid* with the *cusp* at $c = 1/4$. The cardioid bounds a domain Δ_0 of parameter values c such that f_c has an attracting fixed point (the *main hyperbolic component*).

Exercise 33.2. *Show that the main cardioid $\partial\Delta_0$ is given by the equation*

$$c = \frac{1}{2}e(\theta) - \frac{1}{4}e(2\theta) = \frac{1}{2}\rho - \frac{1}{4}\rho^2, \quad 0 \le \theta < 1,$$

where $\rho = e(\theta)$ is the multiplier of the neutral fixed point of f_c. The main cardioid has $3/2$-cusp at $1/4$.

Let us now take a look at how periodic points move with parameter:

Lemma 33.3. *Let $f_\circ$ has a cycle $(\alpha_k)_{k=0}^{p-1}$ of period p with multiplier $\rho_\circ \neq 1$. Then for nearby c, the maps f_c have a cycle $(\alpha_k(c))_{k=0}^{p-1}$ holomorphically depending on c. Its multiplier $\rho(c)$ holomorphically depends on c as well.*

Proof. Consider an algebraic equation $f_c^p(z) = z$. For $c = c_\circ$ it has roots $z = \alpha_k$, $k = 0, \dots, p-1$ (and maybe others). Since

$$\left.\frac{d(f_c^p(z) - z)}{dz}\right|_{c=c_\circ,\, z=\alpha_k} = \rho_\circ - 1 \neq 0,$$

the Implicit Function Theorem yields the first assertion. The second assertion follows from the formula for the multiplier:

$$\rho(c) = 2^p \prod_{k=0}^{p-1} \alpha_k(c).$$

$\square$

Thus, periodic points of f_c as functions of the parameter are algebraic functions branched at parabolic points only.

33.4. Hyperbolic components: first observations. A parameter value $c \in \mathbb{C}$ is called *hyperbolic/parabolic/Siegel* etc if the corresponding quadratic polynomial f_c is such.

PROPOSITION 33.4 (Hyperbolic components). *The set $\mathcal{H}$ of hyperbolic parameter values is the union of $\mathbb{C} \smallsetminus \mathcal{M}$ and some set of components of* $\operatorname{int} \mathcal{M}$.

PROOF. By definition, $\mathbb{C} \smallsetminus \mathcal{M} \subset \mathcal{H}$. Also, the property to have an attracting cycle is stable (see Lemma 33.3), hence $\mathcal{H} \cap \mathcal{M} \subset \operatorname{int} \mathcal{M}$.

Take now some hyperbolic parameter $c_\circ \in \mathcal{M}$ and let $\Delta_\circ$ be the component of $\operatorname{int} \mathcal{M}$ containing $c_\circ$. Let us show that $\Delta_\circ \subset \mathcal{H}$. The map $f_\circ$ has an attracting cycle of some period p. By Theorem 21.4, this cycle contains a point α_0 such that

$$\upsilon_{pn}(c_\circ) \equiv f_\circ^{pn}(0) \to \alpha_0 \text{ as } n \to \infty.$$

It is easy to see (Exercise!) that for nearby $c \in \Delta$ we have:

$$\upsilon_{pn}(c) \equiv f_\circ^{pn}(0) \to \alpha_0(c) \text{ as } n \to \infty,$$

where $\alpha_0(c)$ is the holomorphically moving attracting periodic point of f_c (Lemma 33.3). But the sequence of polynomials $\upsilon_{pn}(c)$, $n = 0, 1, \dots$, is normal in Δ (Exercise 33.1, (iv)). Hence it must converge in the *whole domain* Δ to some holomorphic function $\tilde{\alpha}(c)$ coinciding with $\alpha_0(c)$ near c_0. By analytic continuation, $\tilde{\alpha}(c)$ is a periodic point of f_c with period dividing p.

Moreover, the cycle of this point attracts the critical orbit persistently in Δ. It is impossible if this cycle is repelling somewhere. Indeed, a repelling cycles can only attract an orbit which eventually lands on it. This property is not locally persistent since otherwise it would hold for *all* $c \in \mathbb{C}$ (while it is violated, say, for $c = 1$).

If $\tilde{\alpha}(c)$ was neutral for some $c \in \Delta$, then it could be made repelling for a nearby parameter value. Thus, $\tilde{\alpha}(c)$ is attracting for all $c \in \Delta$, so that $\Delta \subset \mathcal{H}$. ☐

COROLLARY 33.5. *Neutral parameters lie on the boundary of $\mathcal{M}$.*

PROOF. Let $c_\circ$ be a neutral parameter, i.e., the map $f_\circ$ has a neutral cycle. This parameter can be perturbed to make the cycle attracting. If $c_\circ$ belonged to $\operatorname{int} \mathcal{M}$ then by Proposition 33.4 it would be hyperbolic itself – contradiction. ☐

EXERCISE 33.6. (i) *Any parameter $c \in \partial \mathcal{M}$ can be approximated by superattracting parameters.*
(ii) *Misiurewicz parameters form a countable dense subset of $\partial \mathcal{M}$.*

A component Λ of $\operatorname{int} \mathcal{M}$ is called *hyperbolic* if it consists of hyperbolic parameter values. Otherwise Λ is called *queer*. The reason for the last term is that it is generally believed that *there are no queer components*. In fact, it is a central conjecture in contemporary Holomorphic Dynamics:

CONJECTURE 33.7 (Density of Hyperbolicity). *There are no queer components. Hyperbolic parameters are dense in $\mathbb{C}$.*

Because of Exercise 33.6 (i), the second part of the conjecture would follow from the first one. It is also referred to as the *Fatou Conjecture*. See §38 for a further discussion.)

33.5. Primitive and satellite hyperbolic components.

PROPOSITION 33.8. *Let Δ be a hyperbolic component of period n of $\mathcal{M}$, let $\mathfrak{p}/\mathfrak{q} \neq 0 \mod 1$, and let $\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}} \in \partial\Delta$ be a parabolic parameter with rotation number $\mathfrak{p}/\mathfrak{q}$. Then there is a hyperbolic component Δ' of period $n\mathfrak{q}$ attached to Δ at $\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}}$.*

PROOF. We let $c_\circ \equiv r_{\mathfrak{p}/\mathfrak{q}}$, $f_\circ \equiv f_{c_\circ}$, and $g_c = f_c^n$. Let $\alpha_\circ$ be a parabolic fixed point for $g_\circ$. Since $g_\circ'(\alpha_\circ) \neq 1$, nearby maps g_c have a fixed point α_c depending holomorphically on c. Making a change of variable $z \mapsto z - \alpha_c$, we obtain a holomorphic family of quadratic polynomials that fix 0; we keep the same notation f_c for this family and its n-fold iterate g_c.

By Corollary 21.26, $g_\circ$ has $\mathfrak{q}$ parabolic petals attached to 0 that are cyclically permuted by $g_\circ$. Hence near the origin we have:

$$g_\circ^{\mathfrak{q}}(z) - z = b_{\mathfrak{q}+1} z^{\mathfrak{q}+1} + \dots, \quad b_{\mathfrak{q}+1} \neq 0.$$

So 0 is a fixed point of multiplicity $\mathfrak{q}+1$ for $g_\circ$, and hence nearby maps g_c have $\mathfrak{q}+1$ simple fixed points. One of them is 0 which is also fixed by f_c. Others are permuted by f_c. In fact, they form a single cycle of order $\mathfrak{q}$ since $f_c'(0) \approx e(\mathfrak{p}/\mathfrak{q})$ and hence f_c cannot have small cycles of order less than $\mathfrak{q}$.

The multiplier ρ_c of this cycle is a non-constant algebraic function of c equal to 1 at $c_\circ$. Hence there is a parameter domain attached to $c_\circ$ in which our cycle is attracting. It is contained in the desired hyperbolic component Δ'. $\square$

A hyperbolic component Δ' that was born from another hyperbolic component by the *period n-tupling bifurcation* described in Proposition 33.8 is called *satellite*. All other hyperbolic components of $\mathcal{M}$ are called *primitive*. They appear as a result of a *saddle-node* bifurcation. (See §§35.9.1 and 35.9.2 for a detailed discussion.)

Parabolic points on $\partial\Delta$ with multiplier 1 are called the *roots* of Δ. (In fact, we will see below (Theorem 35.3) that any hyperbolic component has a single root.) In particular, the bifurcation point $r_{\mathfrak{p}/\mathfrak{q}}$ is the root of the satellite component Δ'.

As we will see later (see §35.9), the type of a component can be easily recognized geometrically: satellite components are bounded by smooth curves, while primitive components have cusps at their roots.

33.6. Real quadratic family. Exercise 20.10 and its Corollary describe the *real slice of the Mandelbrot set*:

$$\mathcal{M}_{\mathbb{R}} := \mathcal{M} \cap \mathbb{R} = [-2, 1/4].$$

Moreover, for these parameters, the quadratic map f_c restricts to the maximal invariant interval $\mathcal{I}_c$,

$$f_c : \mathcal{I}_c \to \mathcal{I}_c, \quad x \mapsto x^2 + c, \quad c \in [-2, 1/4]. \tag{33.3}$$

We refer to this family as the *real quadratic family*.

33.6.1. *Real hyperbolic windows.* A *real hyperbolic window* $\Delta_{\mathbb{R}} \subset [-2, 1/4]$ is a component of the set of real hyperbolic parameters c. It is the real slice of some hyperbolic component Δ of $\operatorname{int} \mathcal{M}$. (Since this component Δ is $\mathbb{R}$-symmetric and simply connected, $\Delta \cap \mathbb{R}$ is an (open) interval.)

The *real version* of the *Fatou Conjecture* asserted that the hyperbolic windows are dense in $[-2, 1/4]$. It was confirmed in the 1990s (see §38.3 for a further discussion).

33.6.2. *Real quadratic family is full.* The following result partly explains the universal role played by the real quadratic family in the unimodal dynamics:

THEOREM 33.9. *Any admissible kneading sequence* Kn *is realizable for some real quadratic polynomial:* $\mathrm{Kn} = \mathrm{Kn}_c$ *for some* $c \in [-2, 1/4]$.

Let us do it in a more general framework of real analytic families. A family $(f_\lambda)_{\lambda \in L}$ of real unimodal maps $f_\lambda : \mathcal{I}_\lambda \to \mathcal{I}_\lambda$ over a parameter interval L is called *real analytic* if

• The interval $\mathcal{I}_\lambda$ depends continuously on λ (i.e, its endpoints move continuously with λ);

• The map $(\lambda, x) \mapsto f_\lambda(x)$ admits a real analytic extension to a neighborhood of $\bigcup_{\lambda \in L} \mathcal{I}_\lambda$ in $\mathbb{R}^2$.

We call it *non-trivial* if the kneading sequence $\mathrm{Kn}_\lambda \equiv \mathrm{Kn}(f_\lambda)$ is not identically constant for $\lambda \in L$.

EXERCISE 33.10. (i) *Superattracting parameters are isolated in any non-trivial real analytic family of unimodal maps.*

(ii) *If* $\lambda_\circ$ *is superattracting, then* $\mathrm{Kn}_\circ = (\varepsilon_1 \dots \varepsilon_{p-1} 0)^{\mathrm{per}}$, *and for* λ *near* $\lambda_\circ$,

$$\mathrm{Kn}_\lambda = (\varepsilon_1 \dots \varepsilon_{p-1}, \delta, \delta)^{\mathrm{per}} \quad \text{for some } \delta \in \{\pm\}$$

(*where* δ *stays constant on each side of* $\lambda_\circ$).

(iii) $\mathrm{Kn}_\circ$ *is the only admissible kneading sequence squeezed in between*

$$(\varepsilon_1 \dots \varepsilon_{p-1} -)^{\mathrm{per}} \quad \textit{and} \quad (\varepsilon_1 \dots \varepsilon_{p-1} +)^{\mathrm{per}}.$$

(iv) *If* $\lambda_\circ$ *is not superattracting, then the kneading function* $\mathrm{Kn} : \lambda \mapsto \mathrm{Kn}_\lambda$ *is continuous at* $\lambda_\circ$.

INTERMEDIATE VALUE THEOREM. *Let* $(f_\lambda)_{\lambda \in L}$ *be a real analytic family of unimodal maps* $f_\lambda \in \mathfrak{G}'$. *Let* Kn_0 *and* Kn_1 *be the kneading sequences for maps* f_0 *and* f_1 *respectively; assume* $\mathrm{Kn}_0 \neq \mathrm{Kn}_1$. *Then any intermediate admissible sequence* $\bar\varepsilon \in (\mathrm{Kn}_0, \mathrm{Kn}_1)_\#$ *is realizable as* Kn_λ *for some* $\lambda \in L$.

PROOF. Let us assume the contrary and apply, as in the case of the classical IVT, the dyadic subdivision method. It provides us with a shrinking nest of closed dyadic intervals I^n with the property that $\bar\varepsilon \in (\mathrm{Kn}_0^n, \mathrm{Kn}_1^n)_\#$, where the Kn_i^n are the kneading sequences at the endpoints of the I^n. (Let us assume for definiteness that $\mathrm{Kn}_0 < \mathrm{Kn}_1$; then the procedure can be designed so that $\mathrm{Kn}_0^n < \mathrm{Kn}_1^n$.)

Let $\bigcap I^n = \{\lambda_\circ\}$ and let $\mathrm{Kn}_\circ$ be the corresponding kneading sequence. Let

$$\mathrm{Kn}^+ := \liminf_{\lambda \searrow \lambda_\circ} \mathrm{Kn}_\lambda, \quad \mathrm{Kn}^- := \limsup_{\lambda \nearrow \lambda_\circ} \mathrm{Kn}_\lambda.$$

Then $\mathrm{Kn}^- \le \mathrm{Kn}_\circ \le \mathrm{Kn}^+$ and $\mathrm{Kn}^- \le \bar\varepsilon \le \mathrm{Kn}^+$. Let us now consider two cases:

a) Assume $f_\circ$ is not superattracting. Then the kneading function $\lambda \mapsto \mathrm{Kn}_\lambda$ is continuous at $\lambda_\circ$. Hence both sequences (Kn_0^n) and (Kn_1^n) converge to $\mathrm{Kn}_\circ$ implying that $\mathrm{Kn}^- = \mathrm{Kn}_\circ = \mathrm{Kn}^+$. As $\bar\varepsilon$ is squeezed in between Kn^- and $\mathrm{Kn}^=$, it must be equal to $\mathrm{Kn}_\circ$ as well.

b) Assume $f_\circ$ is superattracting. Then $\mathrm{Kn}_\circ = (\varepsilon_1 \dots \varepsilon_{p-1} 0)^{\mathrm{per}}$, and

$$\mathrm{Kn}_\delta = (\varepsilon_1 \dots \varepsilon_{p-1}, \delta)^{\mathrm{per}}, \quad \mathrm{Kn}_{-\delta} = (\varepsilon_1 \dots \varepsilon_{p-1}, -\delta)^{\mathrm{per}} \quad \text{for some } \delta \in \{\pm\}.$$

Since $\mathrm{Kn}_\circ$ is the only admissible sequence squeezed in between such two sequences, we again conclude that $\bar{\varepsilon} = \mathrm{Kn}_\circ$. □

A family $(f_\lambda)_{\lambda \in L}$ of unimodal maps is called *full* if any admissible kneading sequence $\mathrm{Kn} \in \mathfrak{Kn}$ is realizable in this family: there exists $\lambda \in L$ such that $\mathrm{Kn} = \mathrm{Kn}_\lambda$.

Corollary 33.11. *Let $(f_\lambda)_{\lambda \in L}$ be real analytic family of unimodal maps $f_\lambda \in \mathfrak{G}'$. If it contains the extremal kneading sequences $\mathrm{Kn}^{\max} = (- + + + +)$ and $\mathrm{Kn}^{\min} = (- - - -)$, then it is full. In particular, the real quadratic family is full.*

33.6.3. *Real structural stability.* Let us consider a real analytic family $(f_\lambda)_{\lambda \in L}$ of unimodal maps $f_\lambda : \mathcal{I}_\lambda \to \mathcal{I}_\lambda$ over a parameter interval L. A map $f_\circ \equiv f_{\lambda_\circ}$ (and the corresponding parameter $\lambda_\circ \in L$) is called (*really*) *structurally stable* in this family if for any $\lambda \in L$ sufficiently close to $\lambda_\circ$, the map f_λ is topologically conjugate to $f_\circ$, and moreover, the conjugacy $h_\lambda : \mathcal{I}_\circ \to \mathcal{I}_\lambda$ can be selected continuously in λ (in the uniform topology). By definition, the set of structurally stable parameters is open in Λ. The complementary closed set is called the (*real*) *bifurcation locus.*

Theorem 33.12. *For any real analytic family $(f_\lambda)_{\lambda \in L}$ of unimodal maps, the set of structurally stable parameters is dense in L. Any non-parabolic bifurcation parameter can be approximated by a superattracting one.*

Proof. Non-superattracitng hyperbolic parameters are structurally stable, so assume that $\lambda_\circ$ is neither hyperbolic nor parabolic. Then the corresponding kneading sequence $\mathrm{Kn}_\circ$ is infinite and aperiodic. If the kneading sequence Kn_λ is locally constant near $\lambda_\circ$ then the nearby maps f_λ admit the same model as $f_\circ$ (see §32.4.4), and hence are topologically conjugate to the latter. On the other hand, for the kneading sequence to change, one of the postcritical points $f^p_\lambda(0)$ must cross 0, creating a superattracting parameter nearby. □

A *queer interval* is a connected component $L_0 \subset L$ of the set of structurally stable parameters on which the maps f_λ are not hyperbolic. Equivalently, L_0 is the maximal open parameter interval on which the kneading function Kn_λ is an aperiodic constant.

Corollary 33.13. *The Real Fatou Conjecture is equivalent to the absence of queer intervals in the real quadratic family.*

33.7. Kneading model for the real quadratic family. Let us now describe the kneading model for the *whole* real quadratic family. Consider the ordered space $\mathfrak{Kn}^{\mathrm{adm}}$ of admissible kneading sequences and blow-up periodic non-superattracting ones to intervals. We obtain the *kneading parameter interval* $\mathfrak{Kn}$. The *kneading model* for the real quadratic family is the family of the combinatorial models for *all* unimodal maps (see §32.4.4) parametrized by $\mathfrak{Kn}$. The following statement summarizes our knowledge so far:

Proposition 33.14. *The quadratic family naturally projects onto its combinatorial model.*

A much deeper result asserts that this projection is monotonic (Theorem 37.33) and even a deeper one asserts that it is one-to-one (which is equivalent to the Real Fatou Conjecture).

Exercise 33.15. *Build up a "topological model" for the real quadratic family by "blowing up" periodic kneading sequences to intervals.*

34. Connectivity of $\mathcal{M}$

34.1. Uniformization of $\mathbb{C} \setminus \mathcal{M}$. In this section we will prove the first non-trivial result about the Mandelbrot set. The strategy of the proof is quite remarkable: it is based on the explicit uniformization of the complement $\mathbb{C} \setminus \mathcal{M}$ by $\mathbb{C} \setminus \overline{\mathbb{D}}$. Recall from Theorem 23.29 that for $c \in \mathbb{C} \setminus \mathcal{M}$, we have a well-defined function

$$\Psi_{\mathcal{M}}(c) := B_c(c), \tag{34.1}$$

where B_c is the Böttcher function for f_c extended to the domain Ω_c bounded by the critical figure-eight equipotential.

Theorem 34.1. *The Mandelbrot set $\mathcal{M}$ is connected. The function $\Psi_{\mathcal{M}}$ conformally maps $\mathbb{C} \setminus \mathcal{M}$ onto $\mathbb{C} \setminus \overline{\mathbb{D}}$. Moreover, it is tangent to the identity at ∞:*

$$\Psi_{\mathcal{M}}(c) \sim c \text{ as } c \to \infty.$$

We immediately obtain the parameter analogue of Corollary 23.28:

Corollary 34.2. *The Mandelbrot set has capacity 1.*

34.2. Basic Phase-Parameter Relation. Before passing to a proof of Theorem 34.1, let us make a couple of comments on its significance. Formula (34.1) reveals a remarkable relation between the dynamical and parameter planes of the quadratic family: *The Riemann position $\Psi_{\mathcal{M}}(c)$ of a parameter $c \in \mathbb{C} \setminus \mathcal{M}$ coincides with the Böttcher position $B_c(c)$ of the corresponding critical value $c \in \mathbb{C} \setminus \mathcal{J}(f_c)$.*

Recall from §23.5.3 that the polar coordinates of $B_c(z)$ are called the (dynamical) external coordinates of a point $z \in \Omega_c$. Similarly, the (*parameter*) *external coordinates* of a point $c \in \mathbb{C} \setminus \mathcal{M}$ are defined as the polar coordinates of $\Psi_{\mathcal{M}}(c)$.

We see that *the parameter external coordinates of a point $c \in \mathbb{C} \setminus \mathcal{M}$ coincide with its dynamical external coordinates* (in the f_c−dynamical plane).

Similarly to the dynamical situation (see §23.5.4), we can now introduce *parameter equipotentials* $\mathcal{E}^r_{\text{par}} \equiv \mathcal{E}^t_{par}$ (where $t = \log r$) and *parameter external rays* $\mathcal{R}^\theta_{\text{par}}$ by pulling back round circles (of radius r) and radial rays (of angle θ) by means of $\Psi_{\mathcal{M}}$. We obtain two (non-singular) foliations in $\mathbb{C} \setminus \mathcal{M}$. We conclude that

- For $c \in \mathcal{R}^\theta_{\text{par}}$ we have $c \in \mathcal{R}^\theta_c$;
- For $c \in \mathcal{E}^r_{\text{par}}$ we have $c \in \mathcal{E}^r_c$.

We use notation $\theta(c)$ and $r(c) = e^{t(c)}$ for the external angle and radius/height of a parameter $c \in \mathbb{C} \setminus \mathcal{M}$.

Also, as in §23.5.4, we define *parameter subpotential disks* of radius r (or, of height $t = \log r$) as

$$\Sigma_{\text{par}}(r) \equiv \Sigma_{\text{par}}(t) := \{c : \ |\Psi_{\mathcal{M}}(c)| \le r\},$$

where we let $|\Psi_{\mathcal{M}}| \equiv 1$ on $\mathcal{M}$.

Let us finish with a crucial observation:

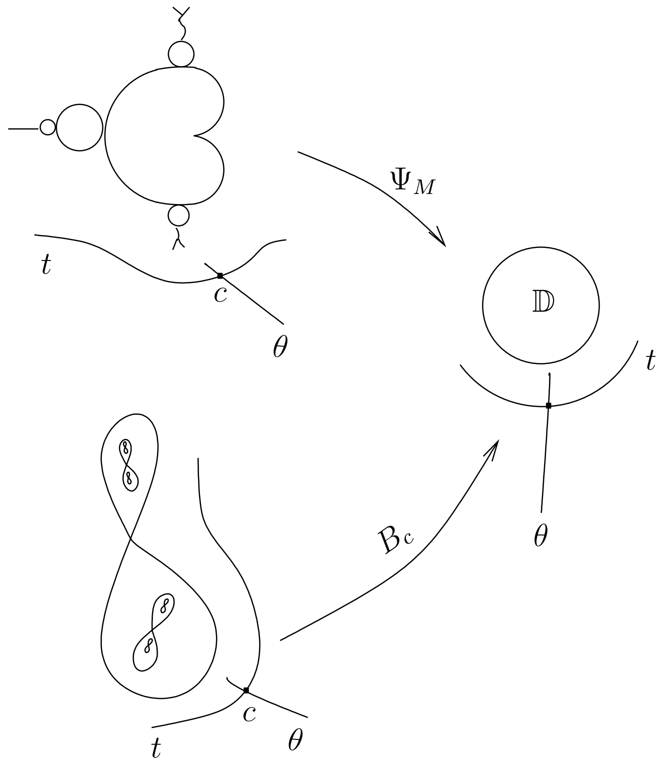

FIGURE 34.1. Uniformization of the Mandelbrot set.

LEMMA 34.3 (Criterion of ray crashing). *A dynamical ray $\mathcal{R}_c^\theta$ crashes at some precritical point if and only if $c \in \mathbb{C} \setminus \mathcal{M}$ and $\theta(c) = T^n\theta$ for some $n \geq 1$ (where T is the doubling map).*

PROOF. A dynamical ray $\mathcal{R}_c^\theta$ crashes at a precritical point of order $n \in \mathbb{N}$ iff $T^n\theta$ is a critical angle, or equivalently, $T^{n+1}\theta$ is the valuable angle. By the Phase-Parameter Relation, the latter is equal to $\theta(c)$. $\square$

34.3. An elementary proof of connectivity. We will give two proofs of Theorem 34.1. The first proof is short and elementary. The second proof, though longer and more demanding, illuminates the deeper meaning of formula (34.1) and the idea of qc deformations (see §34.5). Yet another proof will be provided in a more general context of quadratic-like families (see Corollary 42.4 below).

The first proof is based upon the explicit formula (23.11) for the Böttcher coordinate. It will imply that the function $\Psi_\mathcal{M}$ (34.1) is a holomorphic branched covering of degree 1.

Step 1: analyticity. By Corollary 23.37, the Böttcher function $B_c(z)$ is holomorphic on $\mathbf{\Omega} = \{(c, z) : \ z \in \Omega_c\}$. Hence its restriction to the diagonal $\{z = c\}$ is holomorphic on $\mathbb{C} \setminus \mathcal{M}$. But this is our function $\Psi_\mathcal{M}$.

Step 2: behavior at ∞. Let $v_n = f_c^n(c)$. Then $v_{n+1} = v_n^2(1+O(1/v_n))$, so there is a function $\delta(v) = O(1/v)$ such that

$$((1-\delta(v_n))v_n)^2 \le (1-\delta(v_n))v_{n+1} < (1+\delta(v_n))v_{n+1} \le ((1+\delta(v_n))v_n)^2,$$

Iterating these estimates backwards, we see that

$$\sqrt[2^n]{v_n} = c(1+O(1/c)) \quad \text{as } c \to \infty.$$

It follows that $\Psi_{\mathcal{M}}(c) = c(1+O(1/c)) \sim c$ as $c \to \infty$, so $\Psi_{\mathcal{M}}$ extends holomorphically to ∞, and is tangent to id at ∞.

Step 3: properness. Let us show that the map $\Psi_{\mathcal{M}} : \mathbb{C} \smallsetminus \mathcal{M} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ is proper:

$$|\Psi_{\mathcal{M}}(c)| \to 1 \text{ as } c \to \partial\mathcal{M}.$$

Let us define $n(c) \in \mathbb{N} \cup \{\infty\}$ as the last moment n such that $v_n(c) \in \overline{\mathbb{D}}_3$. By Exercise 33.1(i), $n(c) = \infty$ iff $c \in \mathcal{M}$. Moreover, $n(c) \to \infty$ as $c \to \mathcal{M}$. Otherwise there would exist $N \in \mathbb{N}$ and a sequence $c_k \to c \in \mathcal{M}$ such that $v_N(c_k) \in \mathbb{C} \smallsetminus \overline{\mathbb{D}}_3$, implying that $v_N(c) \in \mathbb{C} \smallsetminus \overline{\mathbb{D}}_3$ – contradiction.

Let us take a small neighborhood U of $\mathcal{M}$ such that $K_c \subset \overline{\mathbb{D}}_3$ for $c \in \overline{U}$ (equivalently, $n(c) > 0$ for $c \in \overline{U}$). Since the Green function is continuous on $\mathbf{\Omega}$,

$$L := \sup\{G_c(z) : \ (c,z) \in \overline{U} \times \mathbb{T}_3\} < \infty.$$

Since $z \mapsto G_c(z)$ is subharmonic on the whole plane $\mathbb{C}$ for any c, by the Maximal Principle we have $G_c(z) \le L$ for $(c,z) \in \overline{U} \times \overline{\mathbb{D}}_3$. Hence

$$G_c(c) = \frac{G_c(v_{n(c)}(c))}{2^{n(c)}} \le \frac{L}{2^{n(c)}} \to 0 \quad \text{as } c \to M.$$

It follows that $|B_c(c)| = e^{G_c(c)} \to 1$ as $c \to M$ $(c \in \mathbb{C} \smallsetminus \mathcal{M})$ as was asserted.

Conclusion. Thus, the map $\Psi_{\mathcal{M}} : \mathbb{C} \smallsetminus M \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ is a branched covering, so that, it has a well-defined degree. But ${\Psi_{\mathcal{M}}}^{-1}(\infty) = \{\infty\}$, and by Step 2, $\Psi_{\mathcal{M}}$ has local degree 1 at ∞. Hence $\deg \Psi_{\mathcal{M}} = 1$, and we are done.

34.4. Böttcher fibration, motion and foliation. In this section we will describe an interesting geometric structure hidden behind the above argument.

34.4.1. *Equivariance of a holomorphic motion.* Let us first introduce some general notions.

A holomorphic motion $h_c : X_\circ \to X_c$ of a set $X \subset \mathbb{C}$ over a parameter domain Δ is called *equivariant* if

$$h_c(f_\circ(z)) = f_c(h_c(z)) \tag{34.2}$$

whenever both points z and $f_\circ(z)$ belong to $X_\circ$. If the X_c are f_c-invariant, this just means that the maps h_c conjugate $f_\circ | X_\circ$ to $f_c | X_c$. (Of course, we can apply this terminology not only to the quadratic family.)

Notice that the equivariance property (34.2) means that the associated lamination (see §17.1) is invariant under the fibered dynamics

$$\mathbf{f} : (c,z) \mapsto (c, f_c(z)).$$

34.4.2. *Böttcher fibration of* $\mathbf{\Omega}$. Recall from §23.6.3 that $\mathbf{D} \equiv \mathbf{D}(\infty) \subset \mathbb{C}^2$ is the union of the basins $\mathcal{D}_c(\infty)$ over all $c \in \mathbb{C}$, and $\mathbf{\Omega} = \bigcup \Omega_c \subset \mathbf{D}$ is the subdomain for which $G_c(z) > G_c(0)$. Let us introduce the *fibered Böttcher function*

$$\mathbf{B} : \mathbf{\Omega} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}, \quad (c, z) \mapsto B_c(z). \tag{34.3}$$

Since $B'_c(z) \neq 0$ for $z \in \Omega_c$, $\mathbf{B}$ is a holomorphic submersion on $\mathbf{\Omega}$ whose fibers

$$L_b = \{(c, z) \in \mathbb{C}^2 : \ z \in \Omega_c \text{ and } B_c(z) = b\} \tag{34.4}$$

are local holomorphic graphs over the parameter plane. Hence they form a holomorphic foliation $\mathfrak{B}$ of $\mathbf{\Omega}$. We will call it the *Böttcher fibration*.

Notice that due to the Böttcher equation, the Böttcher fibration is invariant under the fibered dynamics $\mathbf{f}$.

We also consider the diagonal line in $\mathbf{\Omega}$:

$$\Gamma := \{(c, c) : \ c \in \mathbb{C} \smallsetminus \mathcal{M}\}, \tag{34.5}$$

associating to a *parameter* c the corresponding *critical value* $v = c$.

Lemma 34.4. (i) *For any* $b \in \mathbb{C} \smallsetminus \overline{\mathbb{D}}$*, the Böttcher fiber* L_b *is the graph of a holomorphic function* $\phi_b : \Sigma^{\circ}_{\text{par}}(r) \to \mathbb{C}$ *with* $r = |b|^2$ *(where* $\Sigma^{\circ}_{\text{par}}(r) \equiv \operatorname{int} \Sigma_{\text{par}}(r)$*).*
(ii) *For any* $r > 1$*, the fibers* L_b *with* $|b| > \sqrt{r}$ *restrict to an equivariant biholomorphic motion of the superpotential domains* $\Omega_c(\sqrt{r})$ *over* $\Sigma^{\circ}_{\text{par}}(r)$*.*
(iii) *The diagonal* Γ *is a global transversal to* $\mathfrak{B}$ *that intersects every fiber once.*

Proof. (i) As we have already observed, each Böttcher fiber L_b is a local holomorphic graph over the parameter plane. Let us show that it extends over the whole subpotential disk $\Sigma^{\circ}_{\text{par}}(r)$ with $r = |b|^2$. Indeed, for any $c \in \Sigma^{\circ}_{\text{par}}(r)$, we have:

$$|B_c(0)| = \sqrt{|B_c(c)|} = \sqrt{|\Psi_{\mathcal{M}}(c)|} < \sqrt{r} = |b|,$$

so the image $B_c(\Omega_c) = \{\zeta : |\zeta| > |B_c(0)|\}$ contains $\mathbb{C} \smallsetminus \overline{\mathbb{D}}_{|b|}$. Hence there exists a unique $z \in \Omega_c$ such that $B_c(z) = b$. Thus, L_b crosses once every vertical fiber $\{c\} \times \mathbb{C}$ over $c \in \Sigma^{\circ}_{\text{par}}(r)$, implying that it is a graph over the whole disk $\Sigma^{\circ}_{\text{par}}(r)$.

Remark 34.5. At the same time, L_b cannot be extended any further. Indeed, for any $c \in \partial\Sigma_{\text{par}}(r)$, we have

$$|B_c(0)| = \sqrt{|\Psi_{\mathcal{M}}(c)|} = \sqrt{r} = |b|.$$

Then $|B_{c'}(0)| > |b|$ for some c' near c, so the Böttcher function $B_{c'}$ does not assume value b in $\Omega_{c'}$. By definition, L_b is not defined over c'.

(ii) By the above result, the fibers L_b with $|b| > \sqrt{r}$ restrict to holomorphic graphs over $\Sigma^{\circ}_{\text{par}}(r)$. Thus, they form leaves of a holomorphic motion of the domains $\Omega_c(\sqrt{r})$ over $\Sigma^{\circ}_{\text{par}}(r)$. This motion is biholomorphic (i.e., transversally holomorphic) since its leaves form a holomorphic fibration over $\mathbb{D}_{\sqrt{r}}$ given by the Böttcher function $\mathbf{B}$.

(iii) Intersections of a Böttcher fiber L_b (34.4) with the diagonal Γ (34.5) are solutions of equation $B_c(c) = b$ (since condition $c \in \Omega_c(0)$ is automatically satisfied). By Theorem 34.1, this equation has a unique solution for every $b \in \mathbb{C} \smallsetminus \overline{\mathbb{D}}$. Thus, the diagonal Γ crosses once every leaf L_b of the Böttcher fibration of $\mathbf{\Omega}$, implying that it is a global transversal to $\mathfrak{B}$. $\square$

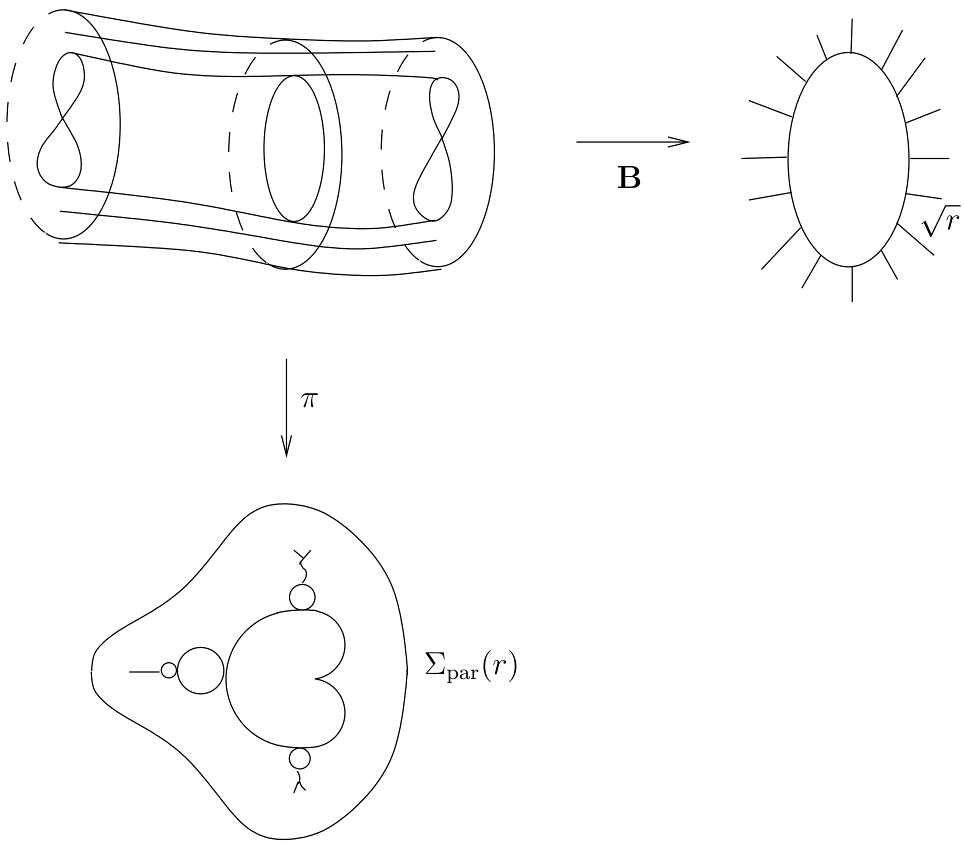

FIGURE 34.2. Böttcher fibration.

For any $c_\circ \in \mathcal{M}$, let us consider the *Böttcher holonomy*

$$h_\circ : \mathbf{\Omega} \to \{c\} \times (\mathbb{C} \smallsetminus \mathcal{K}_\circ\}$$

by sliding any point $(c, z) \in \mathbf{\Omega}$ along the fiber $L \ni (c, z)$ of $\mathfrak{B}$ to its unique intersection point with the vertical fiber $\{c_\circ\} \times \mathbb{C}$. It is given explicitly as $B_\circ^{-1} \circ B_c$, so for $c \in \mathcal{M}$, it coincides with the Böttcher conjugacy $\mathcal{D}_c \to \mathcal{D}_\circ$ defined above (see (23.10) in §23.5.2). In particular, for $c_\circ = 0$, the Böttcher holonomy $h_0 : \mathbf{\Omega} \to \mathbb{C} \smallsetminus \mathcal{K}_0$ is identified with the Böttcher map $\mathbf{B} : \mathbf{\Omega} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ itself.

From this point of view, the Riemann map $\Psi_\mathcal{M} : \mathbb{C} \smallsetminus \mathcal{M} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ is interpreted (up to identification $\{0\} \times \mathcal{D}_0 \approx \mathbb{C} \smallsetminus \overline{\mathbb{D}}$) as a composition of two maps:

- the *lift* of $\mathbb{C} \smallsetminus \mathcal{M}$ to the transversal Γ, $c \mapsto (c, c)$,
- and the Böttcher *holonomy* h_0 restricted to Γ, i.e.,

$$h_0 : (c, c) \mapsto (0, B_c(c).$$

34.4.3. *Böttcher foliation of* $\mathbf{D}$.

PROPOSITION 34.6. *The Böttcher fibration* $\mathfrak{B}$ *of* $\mathbf{\Omega}$ *extends to a non-singular holomorphic foliation (denoted in the same way) of the whole domain* $\mathbf{D}$. *Its leaves* *are branched coverings over the parameter plane with simple branched points at the precritical locus*

$$\mathbf{Crit}^\infty := \{(c,z) : \ z \in \mathrm{Crit}^\infty(f_c)\}. \tag{34.6}$$

PROOF. Since the initial Böttcher fibration $\mathfrak{B}$ of $\mathbf{\Omega}$ is $\mathbf{f}$-invariant, its pullback $\mathbf{f}^*(\mathfrak{B})$ extends itself to a holomorphic foliation of $\mathbf{f}^{-1}(\mathbf{\Omega})$, possibly singular. Since $\mathbf{f}$ is a local biholomorphism outside the zero section $\mathbf{0} := \{(c,0) : c \in \mathbb{C}\}$, the pullback $\mathbf{f}^*(\mathfrak{B})$ can be singular only at $\mathbf{0}$.

Let us show that $\mathfrak{B}$ is non-singular at these points either. Let $L = \{z = \phi(c)\}$ be a local leaf of $\mathfrak{B}$ through a point (c_0, c_0). Its pullback $\mathbf{f}^*(L)$ is given by equation

$$z^2 + c = \phi(c) \quad \text{near } (c_0, 0). \tag{34.7}$$

Since the diagonal $\Gamma = \{(c,c)\}$ is transverse to L, we have $\phi'(c_0) \neq 1$, so the holomorphic curve (34.7) is a graph of a holomorphic function $c = \psi(z^2)$, where $\psi'(0) \neq 0$. Thus, it is non-singular holomorphic curve double branched over the parameter plane.

Moreover, leaves near L form a holomorphic family $L_b = \{z = \phi_b(c)\}$, so by the IFT their pullbacks form a local holomorphic foliation.

Applying further $\mathbf{f}^*$-pullbacks (which are local biholomorphisms at the points of interest), we extend $\mathbf{f}^*(\mathfrak{B})$ to the whole domain $\mathbf{D}$.

To see that the leaves of the extended foliation are branched coverings over the first coordinate, it is sufficient to check the path lifting property. Let $\gamma : [0,1] \to \mathbb{C}$ be a c-path such that $\gamma[0,1)$ lifts to a path $\hat\gamma$ in a leaf L. We need to check that $\hat\gamma$ is compactly contained in $\mathbf{D}$. But L is contained in an equipotential set of the fibered Green function $\mathbf{G} : (c,z) \mapsto G_c(z)$ whose intersection with any tube $\{|c| < R\}$ is compactly contained in $\mathbf{D}$. □

34.4.4. *Böttcher motion of rays.* Given a parameter domain $\Lambda \subset \mathbb{C}$, let us consider all graphs of holomorphic functions $\phi : \Lambda \to \mathbb{C}$ whose images belong to leaves of the Böttcher foliation $\mathfrak{B}$. These graphs cannot cross since $\mathfrak{B}$ is non-singular. Hence they form a holomorphic motion, called the *Böttcher motion over* Λ. In case when Λ is simply connected, the Böttcher motion over Λ comprises all the leaves of $\mathfrak{B}$ that are unbranched over Λ.

LEMMA 34.7. *Assume for some $\theta \in \mathbb{R}/\mathbb{Z}$ and some parameter domain $\Lambda \subset \mathbb{C}$, we have*

$$\theta(c) \neq T^n\theta \text{ for any } n \geq 1 \text{ and } c \in \Lambda \smallsetminus \mathcal{M} \tag{34.8}$$

(where $\theta(c)$ stands for the external angle of c). Then the external ray $\mathcal{R}^\theta_c$ never crashes for $c \in \Lambda$ and moves holomorphically under the Böttcher motion over Λ,

$$h_c : \mathcal{R}^\theta_\circ(t) \mapsto \mathcal{R}^\theta_c(t), \quad t \in \mathbb{R}_+.$$

PROOF. By Lemma 34.3, conditions (34.8) is equivalent to saying that the external ray $\mathcal{R}^\theta_c$ never crashes for $c \in \Lambda$. Hence the maps h_c, $c \in \Lambda$, are well defined. Since the fibered Böttcher function $\mathbf{B}$ remains constant ($\equiv e^{t+2\pi i\theta}$) on the sets

$$\mathcal{O}_t := \{h_c(\mathcal{R}^\theta_\circ(t)) : \ c \in \Lambda\}, \quad t \in \mathbb{R}_+,$$

each $\mathcal{O}_t$ belongs to a leaf L_t of the Böttcher foliation. Moreover, $h_c = B_c^{-1} \circ B_\circ$, where B_c is the extension of the Böttcher function to the domain $\hat\Omega_c$ (see Problem

23.30), implying that h_c is holomorphic in c. So, $\mathcal{O}_t$ is a graph of a holomorphic function $z = \phi_t(c)$ contained in L_t. The conclusion follows. $\square$

Let

$$\hat{\mathbf{\Omega}} = \bigcup_{c\in\mathbb{C}} \hat{\Omega}_c, \tag{34.9}$$

where $\hat{\Omega}_c \subset \mathcal{D}_c(\infty)$ are the domains mapped by B_c onto the complements of the Levin-Sodin hedgehogs (see Problem 23.30).

Problem 34.8. *Show that*

(i) $\mathbf{\Omega}$ *is a domain in* $\mathbb{C}^2$*;*

(ii) *The fibered Böttcher function* $\mathbf{B}$ *extends to* $\hat{\mathbf{\Omega}}$ *as a single-valued holomorphic submersion;*

(iii) *The relative boundary of* $\hat{\mathbf{\Omega}}$ *in* $\mathbf{D}(\infty) \smallsetminus \mathbf{Crit}^\infty$ *is a 3D real analytic manifold.*

34.4.5. *Extended Phase-Parameter Relation.* Let a set $X_c \subset \mathcal{D}_c(\infty)$ moves holomorphically under the Böttcher motion over a pointed parameter domain $(\Lambda, c_\circ)$, where $c_0 \in X_\circ$. If every leaf L_z, $z \in X_\circ$, of this motion crosses the diagonal $\Gamma = \{(c,c) : \ c \in \Lambda\}$ then we have a phase parameter map $\psi : X_\circ \to \Lambda$ that associates to a point $z \in X_\circ$ the parameter $c \in \Lambda$ with the same external coordinates (i.e., $(c,c) \in L_z$).

Lemma 34.9. *Under the above circumstances, assume the set* $\{c : (c,c) \in X_c\}$ *is compactly contained in* Λ*. Then the phase-parameter map* ψ *extends to a qs homeomorphism of* $\overline{X}_\circ$ *onto the image.*

Proof. Lemma 34.4 implies that the diagonal Γ is transverse to the given Böttcher motion intersecting once every leaf. It follows that it is transversal to the extended motion of $\overline{X}_c$ as well, for otherwise it would have multiple intersections with nearby Böttcher leaves. The conclusion follow from Lemma 17.14. $\square$

34.5. Second proof of connectivity.

34.5.1. *Step 1: QC deformation.* The idea is to deform the map by moving around the Böttcher position of its critical value. To this end let us consider a two parameter family of diffeomorphisms $\psi_{\omega,q} : \mathbb{C} \smallsetminus \mathbb{D} \to \mathbb{C} \smallsetminus \mathbb{D}$ written in the polar coordinates as follows:

$$\psi = \psi_{\omega,q}(r,\theta) = (r^\omega, \theta + q\log r), \quad \omega > 0,\ q \in \mathbb{R}.$$

In terms of complex variable $a = re^{i\theta} \in \mathbb{C} \smallsetminus \mathbb{D}$ and complex parameter $\lambda = \omega + iq$, $\operatorname{Re}\lambda > 0$, this family can be expressed in the following concise form:

$$\psi_\lambda(a) = |a|^{\lambda-1}a. \tag{34.10}$$

This family commutes with $f_0 : a \mapsto a^2$: $\psi(a^2) = \psi(a)^2$, and acts transitively on $\mathbb{C}\smallsetminus\mathbb{D}$, i.e., for any $a_\circ$ and a in $\mathbb{C}\smallsetminus\mathbb{D}$, there exists a λ, such that $\psi_\lambda(a_\circ) = a$. (Note also that ψ_λ are automorphisms of $\mathbb{C} \smallsetminus \mathbb{D}$ viewed as a multiplicative semigroup.)

Take now a quadratic polynomial $f_\circ \equiv f_{c_\circ}$ with $c_\circ \in \mathbb{C} \smallsetminus \mathcal{M}$. Let us consider its Böttcher function $\phi_\circ : \Omega_\circ \to \mathbb{C} \smallsetminus \mathbb{D}_\circ$, where $\Omega_\circ \equiv \Omega_{c_\circ}$ is the complement of the figure eight equipotential (see §23.5.3) and $\mathbb{D}_\circ \equiv \mathbb{D}_{R_\circ}$ is the corresponding round disk, $R_\circ > 1$. Take the standard conformal structure σ on $\mathbb{C} \smallsetminus \mathbb{D}$ and pull it back by the composition $\psi_\lambda \circ \phi_\circ$. We obtain a conformal structure $\mu = \mu_\lambda$ in $\Omega_\circ$. Since

ψ_λ commute with f_0 while the Böttcher function conjugates $f_\circ$ to f_0, the structure μ is invariant under $f_\circ$.

Let us pull this structure back to the preimages of $\Omega_\circ$:

$$\mu^n \,|\, \Omega^n = (f_\circ^n)^*(\mu),$$

where $\Omega_\circ^n = f_\circ^{-n}\Omega_\circ$. Since μ is invariant on $\Omega_\circ$, the structures μ^{n+1} and μ^n coincide on $\Omega_\circ^n$, so that they are organized in a single conformal structure on $\bigcup \Omega_\circ^n = \mathbb{C} \smallsetminus J(f_\circ)$. Extend it to the Julia set $J(f_\circ)$ as the standard conformal structure.

We will keep notation $\mu \equiv \mu_\lambda$ for the conformal structure on $\mathbb{C}$ we have just constructed. By construction, it is invariant under $f_\circ$. Moreover, it has a bounded dilatation since holomorphic pullbacks preserve dilatation: $\|\mu_\lambda\|_\infty = \|(\psi_\lambda)^*(\sigma)\|_\infty < 1$.

By the Measurable Riemann Mapping Theorem, there is a qc map $h_\lambda : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ such that $(h_\lambda)_\circ(\mu_\lambda) = \sigma$. By Corollary 29.3, h_λ can be normalized so that it conjugates f_λ to a quadratic map $f_c \equiv f_{c(\lambda)} : z \mapsto z^2 + c(\lambda)$. Of course, the Julia set f_c is also Cantor, so that $c \in \mathbb{C} \smallsetminus \mathcal{M}$.

This family of quadratic polynomials is the desired qc deformation of $f_\circ$.

34.5.2. *Step 2: Analyticity.* We have to check three properties of the map $\Psi_\mathcal{M} : \mathbb{C} \smallsetminus M \to \mathbb{C} \smallsetminus \mathbb{D}$: analyticity, surjectivity, and injectivity. Let us take them one by one.

It is obvious from formula (34.10) that the Beltrami differential

$$\nu_\lambda = (\psi_\lambda)^*(\sigma) = \bar\partial\psi_\lambda / \partial\psi_\lambda$$

depends holomorphically on λ. Hence the Beltrami differential $(f_\circ)^*(\nu_\lambda)$ on $\Omega_\circ$ also depends holomorphically on λ (see Exercise 14.24). Pulling it back by the iterates of $f_\circ$ and extending it in the standard way to $J(f)$, we obtain by Lemma 14.23 a holomorphic family of Beltrami differentials μ_λ on $\mathbb{C}$. By Corollary 29.2, $c(\lambda)$ is holomorphic on λ as well.

34.5.3. *Step 3: Surjectivity.* Note that the map $\psi_\lambda \circ \phi_\circ \circ h_\lambda^{-1}$ conformally conjugates the polynomial $f_c \equiv f_{c(\lambda)}$ near ∞ to $f_0 : z \mapsto z^2$. By Theorem 23.23, these properties determine uniquely the Böttcher map ϕ_c of f_c, so that $\phi_c = \psi_\lambda \circ \phi_\circ \circ h_\lambda^{-1}$ with $c = c(\lambda)$. Since h_λ conjugates $f_\circ$ to f_c, we have: $h_\lambda(c_*) = c$ and hence

$$\Psi_\mathcal{M}(c) = \phi_c(c) = \psi_\lambda \circ \phi_\circ(c_\circ) = \psi_\lambda(a_\circ),$$

where $a_\circ$ is the Böttcher position of the critical value of $f_\circ$. Since the family $\{\psi_\lambda\}$ acts transitively on $\mathbb{C} \smallsetminus \mathbb{D}$, any point $a \in \mathbb{C} \smallsetminus \mathbb{D}$ can be realized as $\Psi_\mathcal{M}(c)$ for some $c = c(\lambda)$.

34.5.4. *Step 4: Injectivity.* We have to check that if

$$\phi_c(c) = a = \phi_{\tilde c}(\tilde c) \tag{34.11}$$

for two parameter values c and $\tilde c$ in $\mathbb{C} \smallsetminus \mathcal{M}$, then $c = \tilde c$. We let $f \equiv f_c$, $\phi \equiv \phi_c$, $\tilde f \equiv f_{\tilde c}$, $\tilde\phi \equiv \phi_{\tilde c}$. Similarly, we will mark with "tilde" the dynamical objects associated with $\tilde f$ that naturally correspond to dynamical objects associated with f.

Let $R = \sqrt{|a|}$. Then the maps ϕ^{-1} and $\tilde\phi^{-1}$ map $\mathbb{C} \smallsetminus \bar{\mathbb{D}}_R$ onto the domains $\Omega \equiv \Omega_c$ and $\tilde\Omega \equiv \Omega_{\tilde c}$ respectively. Moreover, they extend continuously to the boundary circle mapping it onto the boundary figures eight $\Gamma = \partial\Omega$ and $\tilde\Gamma = \partial\tilde\Omega$, and this extension if one-to-one except that

$$\phi^{-1}(\pm\sqrt{a}) = 0 = \tilde\phi^{-1}(\pm\sqrt{a}).$$

Hence the conformal map $h = \tilde{\phi}^{-1} \circ \phi : \Omega \to \tilde{\Omega}$ admits a homeomorphic extension to the closure of its domain:

$$h : (\mathrm{cl}(\Omega), 0) \to (\mathrm{cl}(\tilde{\Omega}), 0).$$

Consider a domain $\Omega^0 = f(\Omega)$ (exterior of the equipotential passing through c) and the complementary Jordan disk $\Delta^0 = \mathbb{C} \smallsetminus \Omega^0$. We will describe a hierarchical decomposition of Δ^0 into topological annuli A_i^n, $n = 1, \dots,$ $i = 1, 2, \dots, 2^n$. Let $\Omega^n = f^{-n}\Omega^0$ (so that $\Omega \equiv \Omega^1$). The boundary $\partial\Omega^n$ consists of 2^{n-1} disjoint figures eight. The loops of these figures eight bound 2^n (closed) Jordan disks Δ_i^n. The map f conformally maps Δ_i^n onto some Δ_j^{n-1}, $n \geq 1$. Let $A_i^n = \Delta_i^n \cap \mathrm{cl}(\Omega^{n+1})$. These are closed topological annuli each of which is bounded by a Jordan curve and a figure eight. They tile $\Delta^0 \smallsetminus J(f)$. The map f conformally maps A_i^n onto some A_j^{n-1}, $n \geq 1$.

Let us lift $h \equiv h_1$ to conformal maps $H_i : A_i^1 \to \tilde{A}_i^1$:

$$H_i \,|\, A_i^1 = (\tilde{f}|\tilde{A}_i^1)^{-1} \circ h \circ (f|A_i^1). \tag{34.12}$$

Since h is equivariant on the boundary of $\Omega^1 \smallsetminus \Omega^0$, it matches with the H_i on $\partial\Delta_i^1$. Putting these maps together, we obtain an equivariant homeomorphism $h_2 : \mathrm{cl}(\Omega^2) \to \mathrm{cl}(\tilde{\Omega}^2)$ conformal in the complement of the figure eight Γ:

$$h_2(z) = \begin{cases} h(z), & z \in \Omega^1, \\ H_i(z), & z \in A_i^1. \end{cases}$$

Since smooth curves are removable (recall §16), h_2 is conformal in $\Omega^2 \smallsetminus \{0\}$. Since isolated points are removable, h_2 is conformal in Ω^2. Thus h admits an equivariant conformal extension to Ω^2.

In the same way, h_2 can be lifted to four annuli A_i^2. This gives an equivariant conformal extension of h to Ω^3. Proceeding in this way, we will consecutively obtain an equivariant conformal extension of h to all the domains Ω^n and hence to their union $\cup\Omega^n = \mathbb{C} \smallsetminus J(f)$.

Since the Julia set $J(f)$ is removable (Theorem 16.11), this map admits a conformal extension through $J(f)$. Thus, f and $\tilde{f}$ are conformally equivalent, and hence $c = \tilde{c}$.

This completes the second proof of Theorem 34.1.

35. Hyperbolic components of $\mathcal{M}$

In this section we will prove that a hyperbolic quadratic polynomial is uniquely determined by its Hubbard tree and the multiplier of its attracting cycle. In other words, hyperbolic components of $\operatorname{int} \mathcal{M}$ can be labeled by Hubbard trees (or equivalently, by characteristic angles θ_-), while each of them can be conformally parametrized by the attracting multiplier.

35.1. Combinatorial Rigidity for superattracting maps. Here is our first rigidity result:

THEOREM 35.1. *If two superattracting quadratic polynomials f_c and $f_{\tilde{c}}$ have the same abstract Hubbard tree $\mathcal{T}$, or equivalently, if they have the same characteristic angles $\theta_\pm$,*[2] *then $c = \tilde{c}$.*

[2]See Proposition 25.61.

We let $f \equiv f_c$, $\tilde{f} \equiv f_{\tilde{c}}$, $\mathcal{D}(\infty) \equiv \mathcal{D}_f(\infty)$, $\mathcal{K} \equiv \mathcal{K}(f)$, etc, and we label the corresponding objects for $\tilde{f}$ with "tilde". By our assumption, $\mathcal{T} = \tilde{\mathcal{T}}$.

Let us split the proof in several steps.

Step 1: From the Hubbard tree $\mathcal{T}$ to the extended Böttcher conjugacy

$$h : \mathcal{D}(\infty) \cup \mathcal{J} \to \tilde{\mathcal{D}}(\infty) \cup \tilde{\mathcal{J}}. \tag{35.1}$$

By Theorem 25.58, the inverse Böttcher map $B^{-1} : \mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathcal{D}(\infty)$ admits an extension to a homeomorphism $((\mathbb{C}, \overline{\mathbb{D}})/\underset{\mathcal{K}}{\sim}) \to (\mathbb{C}, \mathcal{K})$ semi-conjugating the quotient of $f_0 : z \mapsto z^2$ on $\mathbb{C} \smallsetminus \mathbb{D}$ to f on $\mathcal{D}(\infty) \cup \mathcal{J}$. Moreover, its fibers are completely determined by the Hubbard tree $\mathcal{T}$ (by Thoerem 25.58 and Proposition 25.61). Hence the Böttcher conjugacy on the basins,

$$h = \tilde{B}^{-1} \circ B : \mathcal{D}(\infty) \to \tilde{\mathcal{D}}(\infty),$$

extends to a homeomorphism $h : \mathbb{C} \to \mathbb{C}$ that conjugates f to $\tilde{f}$ on their Julia sets. (Of course, the latter is automatic by continuity.)

Step 2: Conformal extension of conjugacy (35.1) *to* $\mathcal{D}(\mathbf{c})$. Let $\mathcal{D}_0$ be the immediate basin of 0 for f. We know that it is a Jordan disk (Corollary 25.5), and the Riemann mapping $\phi : (\mathrm{cl}\,\mathcal{D}_0, 0) \to (\overline{\mathbb{D}}, 0)$ (appropriately normalized) conjugates the return map $f^p : \mathrm{cl}\,\mathcal{D}_0 \to \mathrm{cl}\,\mathcal{D}_0$ to $f_0 : z \mapsto z^2$ on $\overline{\mathbb{D}}$.

The normalized Riemann mapping $\tilde{\phi} : \mathrm{cl}\,\mathcal{D}_0 \to \overline{\mathbb{D}}$ conjugates $\tilde{f}^p|\,\mathrm{cl}\,\tilde{\mathcal{D}}_0$ to the same map f_0 on $\overline{\mathbb{D}}$. Hence the composition $h_0 = \tilde{\phi}^{-1} \circ \phi : \mathrm{cl}\,\mathcal{D}_0 \to \mathrm{cl}\,\tilde{\mathcal{D}}_0$ conjugates $f^p \,|\, \mathrm{cl}\,\mathcal{D}_0$ to $\tilde{f}^p|\,\mathrm{cl}\,\tilde{\mathcal{D}}_0$.

We claim that this map h_0 continuously matches on $\partial \mathcal{D}_0$ with the conjugacy h from (35.1). Indeed, both of them conjugate $f^p|\,\partial\mathcal{D}_0$ to $\tilde{f}^p|\,\partial\tilde{\mathcal{D}}_0$. Hence the composition $h^{-1} \circ h_0 : \partial\mathcal{D}_0 \to \partial\mathcal{D}_0$ commutes with $f^p|\,\partial\mathcal{D}_0$. But the latter map is topologically equivalent to $z \mapsto z^2$ on $\mathbb{T}$, which has the trivial commutator (Proposition 19.58). Hence $h^{-1} \circ h_0 \,|\, \partial\mathcal{D}_0 = \mathrm{id}$, and the claim follows.

Let us now consider another component D of $\mathrm{int}\,\mathcal{K}$. Since $\mathrm{int}\,\mathcal{K}$ is equal to the basin of $\mathbf{0}$ (Theorem 25.2), there is $n = n_D \in \mathbb{Z}_+$ such that f^n conformally maps $\mathrm{cl}\,D$ onto $\mathrm{cl}\,\mathcal{D}_0$. Let $f_D^{-n} : \mathrm{cl}\,\mathcal{D}_0 \to \mathrm{cl}\,D$ stand for the inverse map. Then we let

$$h_D = \tilde{f}_D^{-n} \circ h_0 \circ f^n \,:\, \mathrm{cl}\,D \to \mathrm{cl}\,\tilde{D}. \tag{35.2}$$

Moreover, h_D matches continuously on ∂D with h. Indeed, since h is a conjugacy on the whole Julia set, we have

$$h|\,\partial D = \tilde{f}^{-n} \circ (h|\,\partial\mathcal{D}_0) \circ f^n \,:\, \partial D \to \partial\tilde{D}.$$

Comparing this with (35.2), taking into account that $h|\,\partial\mathcal{D}_0 = h_0$, yields $h|\,\partial D = h_D$.

Thus, we have extended h conformally and equivariantly to all the components D_i of $\mathrm{int}\,\mathcal{K}$. Since $\mathrm{diam}\,D_i \to 0$, this extension is a global homeomorphism (see Exercise 1.34), and Step 2 is accomplished.

Step 3: Dynamical qc removability of $\mathcal{J}$. We will now show that the conjugacy h just constructed is quasiconformal.

Lemma 35.2. *Let f and $\tilde{f}$ be two hyperbolic quadratic polynomials, let U and $\tilde{U}$ be neighborhoods of their Julia sets, and let $h : (U, \mathcal{J}) \to (\tilde{U}, \tilde{\mathcal{J}})$ be a homeomorphism conjugating f to $\tilde{f}$ near the Julia sets. If h is K-qc on $U \smallsetminus \mathcal{J}$ then h is K-qc on U .*

PROOF. We will use definition of quasiconformality in terms of the circular dilatation, see Proposition 12.14. It is enough to check that the image $h(D)$ of a sufficiently small disk $D := \mathbb{D}(z, \rho)$, $z \in \mathcal{J}$, has a bounded shape around $h(z)$. To this end, we will make use of the quasi-self-similarity of $\mathcal{J}$ and $\tilde{\mathcal{J}}$ (Exercise 25.17). According to that lemma, for all sufficiently small $\varepsilon \geq \rho > 0$ there exists an n such that f^n maps D univalently onto an oval V of size of order ε and bounded shape around $z_n = f^n z$. Since h is a homeomorphism, $h(V)$ is an oval whose inner and outer radii (around $h(z_n)$) are squeezed in between $r(\varepsilon) > 0$ and $R(\varepsilon) \to 0$ as $\varepsilon \to 0$. If $R(\varepsilon)$ is sufficiently small then there exists an inverse branch $\tilde{f}^{-n}$ on $h(V)$ with bounded Koebe distortion such that $\tilde{f}^{-n}(h(z_n)) = h(z)$. Hence $\tilde{f}^{-n}(h(V)) = h(D)$ has a bounded shape around $h(z)$, and quasiconformality of h follows.

By Proposition 25.23, the Julia set $\mathcal{J}$ has zero area. Hence it does not contribute to the dilatation of h, i.e., $\operatorname{Dil} h = \operatorname{Dil}(h|\,(U \smallsetminus \mathcal{J})) = K$. $\square$

Step 4: The conjugacy h is affine. As the conjugacy h is conformal on $\mathbb{C} \smallsetminus \mathcal{J}$, Lemma 35.2 implies that it is 1-qc on $\mathbb{C}$. By Weyl's Lemma, it is conformal on $\mathbb{C}$, so it is affine. Since no two different maps in the quadratic family (f_c) are affinely conjugate, Theorem 35.1 follows: $c = \tilde{c}$.

Thus, *superattracting parameters $c \in \mathcal{M}$ can be labeled by their Hubbard trees $\mathcal{T}$ or equivalently, by their characteristic angles θ_-.*

35.2. Multiplier Theorem and the centers of hyperbolic components.

35.2.1. *Statement.* Let us pick a favorite hyperbolic component Δ of the Mandelbrot set M. For $c \in \Delta$, the polynomial f_c has a unique attracting cycle $\boldsymbol{\alpha}_c = \{\alpha_k(c)\}_{k=0}^{p-1}$ of period p. By Lemma 33.3, the multiplier $\rho(c)$ of this cycle holomorphically depends on c, so that we obtain a holomorphic map $\rho : \Delta \to \mathbb{D}$. It is remarkable that this map gives an explicit uniformization of Δ by the unit disk:

THEOREM 35.3. *The multiplier map $\rho : \Delta \to \mathbb{D}$ is a conformal isomorphism.*

COROLLARY 35.4. *Any hyperbolic component Δ of $\operatorname{int} \mathcal{M}$ contains a unique superattracting parameter c_Δ.*

The superattracting parameter c_Δ is called the *center* of Δ.

COROLLARY 35.5. *For any real hyperbolic window $\Delta_{\mathbb{R}} \subset [-2, 1/4]$, the multiplier function $\rho : \Delta_{\mathbb{R}} \to (-1, 1)$ is a real analytic diffeomorphism.*

COROLLARY 35.6. *For any $p \in \mathbb{N}$, there exist 2^{p-1} hyperbolic components of $\operatorname{int} \mathcal{M}$ of some period $q|p$.*

PROOF. The centers of hyperbolic components of some period $q|p$ are roots of equation $f_c^p(0) = 0$, which is a polynomial equation of degree 2^{p-1}. All we need to check is that it has simple roots.

Consider the center $c_\circ$ of some hyperbolic component of period $q|p$. For c near $c_\circ$, let α_c be the attracting periodic point holomorphically moving with c such that $\alpha_\circ = 0$. Letting $\phi_n(c, z) := f_c^n(z)$, we have: $\alpha_c = \phi_q(c, \alpha_c) = \phi_p(c, \alpha_c)$. Differentiating at $c_\circ$, we obtain:

$$\frac{d\alpha_c}{dc}(c_\circ) = \partial_c \phi_p\,(c_\circ, 0) + (f_\circ^p)'(0) \cdot \frac{d\alpha_c}{dc}(c_\circ) = \partial_c \phi_p\,(c_\circ, 0). \tag{35.3}$$

Let us now consider dependence of the multiplier $\rho_c = (f_c^q)'(\alpha_c) \equiv \partial_z \phi_q\,(c, \alpha_c)$ on c. By the Multiplier Theorem, $\dfrac{d\rho_c}{dc}(c_\circ) \neq 0$. On the other hand,

$$\frac{d\rho_c}{dc}(c_\circ) = \partial_c \partial_z \phi_q\,(c_\circ, 0) + (f_\circ^q)''(0) \cdot \frac{d\alpha_c}{dc}(c_\circ) = (f_\circ^q)''(0) \cdot \frac{d\alpha_c}{dc}(c_\circ)$$

since $\partial_z \phi_q\,(c, 0) = (f_c^q)'(0) \equiv 0$. Hence $\dfrac{d\alpha_c}{dc}(c_\circ) \neq 0$.

Going back to (35.3), we conclude that $\partial_c \phi_p\,(c_\circ, 0) \neq 0$, implying that $c_\circ$ is a simple root of equation $\phi_p\,(c, 0) = 0$. $\square$

Exercise 35.7. *Calculate the number of hyperbolic components of* $\operatorname{int}\mathcal{M}$ *of exact period p.*

The Multiplier Theorem is in many respects analogous to Theorem 34.1 on connectivity of the Mandelbrot set. The latter gives an explicit dynamical uniformization of $\mathbb{C} \smallsetminus \mathcal{M}$, which the unbounded hyperbolic component in the quadratic family[3]; the former gives the dynamical uniformization for any bounded hyperbolic component in this family. The ideas of the proofs are also similar.

We already know that ρ is holomorphic, so for the Multiplier Theorem we need to verify that it is surjective and injective. The first statement is easy:

Exercise 35.8. *The multiplier map* $\rho : \Delta \to \mathbb{D}$ *is proper and hence surjective. In particular,* Δ *contains a superattracting parameter value.*

We will give several insights into the Multiplier Theorem. The first one is provided by the theory of holomorphic motions (see §17).

35.2.2. *Böttcher motion over* Δ.

Proposition 35.9. *Let* Δ *be a component of* $\operatorname{int}\mathcal{M}$ *with a base point* $c_\circ$. *Then there exists an equivariant biholomorphic motion* $h_c : \mathcal{D}_\circ(\infty) \to \mathcal{D}_c(\infty)$ *of the basin of infinity over* Δ.

Proof. Let $B_c : \mathcal{D}_c(\infty) \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ be the Böttcher-Riemann uniformization of the basin of infinity (see Theorem 23.25). It is a holomorphic function in two variables on the domain $\{(c, z) :\ c \in \Delta,\ z \in \mathcal{D}_c(\infty)\}$ (Corollary 23.37). It follows that the Böttcher conjugacy $h_c = B_c^{-1} \circ B_\circ$ is a biholomorphic motion of $\mathcal{D}_c$ over Δ. Since the maps B_c conjugate f_c to $z \mapsto z^2$, this motion is equivariant. $\square$

Exercise 35.10. *Show that an equivariant biholomorphic motion of the basin of* ∞ *over* Δ *is unique.*

Now the First λ-lemma implies:

Corollary 35.11. *For any component* Δ *of* $\operatorname{int}\mathcal{M}$*, there is a unique equivariant holomorphic motion* $h_c : \mathcal{D}_\circ(\infty) \cup \mathcal{J}_\circ \to \mathcal{D}_c(\infty) \cup \mathcal{J}_c$ *over* Δ *which is biholomorphic on* $\mathcal{D}_c(\infty)$.

Under the above circumstances, the biholomorphic motion $h_c : \mathcal{D}_\circ(\infty) \to \mathcal{D}_c(\infty)$ of the basin of infinity is called the *Böttcher motion*.

It follows that the filled Julia sets of all the maps f_c within one hyperbolic component ($c \in \Delta$) have the same pinched disk model (in the sense of §9.4). In

[3]Recall the definition of hyperbolicity given in §21.2.3.

particular, it is the same for a superattracting parameter $c_\circ \in \Delta$. As Theorem 25.58 shows, the latter can be explicitly described in terms of the characteristic angles $\theta_\pm$. Moreover, the conjugacy h_c respects the roots of components of $\operatorname{int}\mathcal{K}$ as well as the Böttcher coordinates on the basins $\mathcal{D}(\infty)$. Hence the rays landing at the valuable roots have the same angles for f_c and $f_\circ \equiv f_{c_\circ}$. Let us now define the *characteristic rays* $\mathcal{R}_c^{\theta_\pm}$ (and the corresponding *characteristic angles* $\theta_\pm$) for f_c in the same way as in the superattracting case (i.e., as the rays landing at the valuable root β_1 that enclose a sector containing the immediate valuable basin $\mathcal{D}_1$). We see that these rays are the same for all maps f_c, $c \in \Delta$.

Notice that at this stage we have already obtained Corollary 35.4. Indeed, as all superattracting parameters in Δ have the same characteristic angles, by the Rigidity Theorem, 35.1, there is only one such a parameter, the center c_Δ of Δ.

35.2.3. *Injectivity of the multiplier $\rho : \Delta \to \mathbb{D}$.* It is based on the following rigidity result (compare with Theorem 35.1):

THEOREM 35.12. *Let f_c and $f_{\tilde{c}}$ be two hyperbolic quadratic maps. Assume that the Böttcher conjugacy*

$$h : \mathcal{D}_c(\infty) \to \mathcal{D}_{\tilde{c}}(\infty), \quad h = B_{\tilde{c}}^{-1} \circ B_c,$$

extends to a homeomorphism $\mathcal{D}_c(\infty) \cup \mathcal{J}_c \to \mathcal{D}_{\tilde{c}}(\infty) \cup \mathcal{J}_{\tilde{c}}$. If the attracting cycles of these maps have the same multiplier $\rho \in \mathbb{D}$ then $c = \tilde{c}$.

PROOF. For $\rho = 0$, it is part of Theorem 35.1 (beginning Step 2 of the proof). For an arbitrary $\rho \in \mathbb{D}$, it follows the same lines, except that the model for the return map $f^p : \mathcal{D}(0) \to \mathcal{D}(0)$ is a general Blaschke map $g : \mathbb{D} \to \mathbb{D}$ (see Proposition 25.7) rather than the quadratic map $z \mapsto z^2$.

Continuous matching of h and h_0 on $\partial\mathcal{D}(0)$ in the proof of Step 2 is secured by the fact that the circle map $g|\,\mathbb{T}$, being expanding (Lemma 25.10), does not have non-trivial automorphisms (Corollary 19.63.).

We leave details to the reader. □

35.2.4. *Conclusion.* Together, Proposition 35.9 and Theorem 35.12 imply the Multiplier Theorem. So, for any hyperbolic parameter $c \in \operatorname{int} M$, there exists a unique superattracting parameter c_Δ that belongs to the same hyperbolic component Δ of $\operatorname{int} M$ as c. By definition, the *Hubbard tree* of f_c is the *abstract Hubbard tree* of $f_\Delta \equiv f_{c_\Delta}$.

COROLLARY 35.13. *A hyperbolic quadratic polynomial f_c, $c \in \mathcal{M}$, is uniquely determined by its Hubbard tree and the multiplier of the attracting cycle. The geodesic lamination that models $\mathcal{K}_c$ coincides with the lamination for $f_\circ$ described by* Theorem 25.58.

The uniformization $\rho : \Delta \to \mathbb{D}$ induces a foliation of Δ by (*parameter*) *internal rays* corresponding to straight radii $re(\theta)$, $0 \le r < 1$, in $\mathbb{D}$. For $c \in \bar{\Delta}$, $\arg \rho(c)$ is called the *internal angle* of c. Since Δ is a Jordan disk, the internal ray of angle $\theta \in [0, 1)$ lands at the neutral parameter $c(\theta) \in \partial\Delta$ such that $f \equiv f_{c(\theta)}$ has a *neutral* periodic point $\alpha \equiv \alpha_\theta$ fixed under f^p and such that $(f^p)'(\alpha) = e(\theta)$. Moreover, if $\theta \neq 0$ then p is the *smallest* period of α and θ is its rotation number. For $\theta = 0$, α may have a smaller period $p/k < p$ with rotation number j/k, $j \in \mathbb{Z}/k\mathbb{Z}$, $(j, k) = 1$ (see §33.5). This happens when Δ is a satellite hyperbolic component.

The parameter $\mathfrak{r}_\Delta := c(0) \in \partial\Delta$ for which $\theta = 0$ is called the *root* of Δ.

Since in the real case the Hubbard tree is reduced to the kneading sequence, we obtain:

COROLLARY 35.14. *Any real hyperbolic quadratic polynomial f_c, $c \in \mathcal{M}_{\mathbb{R}}$, is uniquely determined by its kneading sequence and the multiplier of the attracting cycle.*

35.3. Attracting multiplier as a conformal modulus. Let us select a hyperbolic component Δ of $\operatorname{int}\mathcal{M}$, and puncture it a the center: $\Delta^* := \Delta \smallsetminus \{c_\Delta\}$. Given a base point $c_\circ \in \Delta^*$, we will describe a qc deformation $f_c \equiv f_{c(\tau)}$ of the map $f_\circ \equiv f_{c_\circ}$ that provides us with an explicit interpretation of the attracting multiplier $\rho = \rho(c)$ of f_c as a conformal modulus in Δ. It gives us a new insight into the Multiplier Theorem, and in fact can be used to give an alternative proof for it.

We will produce this qc deformation by deforming the fundamental torus of $f_\circ$.

35.3.1. *Deformation of the fundamental torus.* Let us briefly recap the notion of fundamental torus from §23.1.3.

Take a little disk $D = \mathbb{D}(\alpha_\circ, \varepsilon)$ around an attracting periodic point $\alpha_\circ$ of $f_\circ$. It is invariant under $g_\circ \equiv f_\circ^p$ and the quotient of $D^* := D \smallsetminus \{\alpha_\circ\}$ under the action of $g_\circ$ is a conformal torus $T_\circ$. Its fundamental group has one marked generator corresponding to a little circle around $\alpha_\circ$.

By the Linearization Theorem, 23.4, the action of $g_\circ$ on D is conformally equivalent to the linear action of $\zeta \mapsto \rho_\circ \zeta$. Hence the partially marked torus $T_\circ$ is conformally equivalent to $\mathbb{T}^2_{\rho_\circ}$, so $\rho_\circ$ is the modulus of $T_\circ$. (In what follows we identify $T_\circ$ with $\mathbb{T}^2_{\rho_\circ}$.)

Let us select a holomorphic family of deformations $\psi_\tau : \mathbb{T}^2_{\rho_\circ} \to \mathbb{T}^2_\rho$ of $T_\circ$, where $\tau = \dfrac{1}{2\pi i} \log \rho \in \mathbb{H}$. For instance, ψ_τ can be chosen to be linear in the logarithmic coordinate $\log \zeta$:

$$x + y\tau_0 \mapsto x + y\tau; \quad (x, y) \in \mathbb{R}^2.$$

This gives us a complex one-parameter family of Beltrami differentials $\nu_\tau = \psi_\tau^*(\sigma)$ on T_0.

EXERCISE 35.15. *Calculate ν_τ explicitly* (*for the above linear deformation*).

35.3.2. *QC deformation of $f_\circ$.* We can lift ν_τ to the disk D and then pull it back by iterates of $f_\circ$. This gives us a family of $f_\circ-$invariant Beltrami differentials μ_τ on the attracting basin $\mathcal{D}(\boldsymbol{\alpha}_\circ)$. These Beltrami differentials have a bounded dilatation since the pullbacks under holomorphic maps preserve dilatation. Extend the μ_τ by 0 outside the attracting basin (keeping the notation). We obtain a family of measurable $f_\circ-$invariant conformal structures μ_τ on the Riemann sphere. Solving the Beltrami equation $(h_\tau)_*(\mu_\tau) = \sigma$ (with an appropriately normalization) we obtain a qc deformation of $f_\circ$:

$$f_{c(\tau)} = h_\tau \circ f_\circ \circ h_\tau^{-1} : \ z \mapsto z^2 + c(\tau), \tag{35.4}$$

where $c(\tau)$ depends holomorphically on τ (see Corollary 29.3). Moreover, this deformation is conformal on the basin of ∞.

EXERCISE 35.16. *Show that*

(i) *The attracting multiplier of $f_{c(\tau)}$ is equal to $\rho = e(\tau)$.*

(ii) *$c(\tau)$ is invariant under the translation $\tau \mapsto \tau + 1$ and hence it depends only on ρ.*

Thus, this qc deformation gives a natural uniformization

$$\mathbb{D}^* \to \Delta^*, \quad \lambda \mapsto c(\lambda)$$

which is inverse to the multiplier function $c \mapsto \rho(c)$. As a bi-product, we obtain (compare §36.5):

THEOREM 35.17. *All maps f_c, $c \in \Delta^*$, are qc equivalent, and the conjugacy is conformal on the basin of ∞.*

35.4. Attracting-superattracting surgery revisited. Let Δ be a hyperbolic component centered at c_Δ, and let $c \in \Delta$. Let $f \equiv f_c$ and $f_\Delta \equiv f_{c_\Delta}$.

Recall from §25.8 the surgery that turns a hyperbolic map f into a quasiregular map F coinciding with f on $\mathbb{C} \smallsetminus \mathcal{D}_0$. In turn, F is qc conjugate to a superattracting map $f_\circ$, and moreover, the conjugacy h is Böttcher on the basin of ∞. In this brief section, we will identify $f_\circ$ with f_Δ.

LEMMA 35.18. *Under the above circumstances, $f_\circ = f_\Delta$.*

PROOF. As we noticed in §25.8.3, the characteristic angles $\theta_\pm$ for f and $f_\circ$ coincide. But by Corollary 35.11, this is also true for f and f_Δ. Since the characteristic angles determine the superattracting parameter (by Theorem 35.1), the conclusion follows. □

35.5. Saddle-node bifurcation. The saddle-node bifurcation for a fixed point in a real one-parameter family of maps is represented on Figure 35.1. The simplest family exhibiting this scenario is

$$f_\varepsilon(x) = -\varepsilon + x + x^2. \tag{35.5}$$

For $\varepsilon > 0$, it has an attracting-repelling pair of real fixed points $\alpha = -\sqrt{\varepsilon}$ and $\beta = \sqrt{\varepsilon}$. Note also that the whole interval $[\alpha, \beta)$ is contained in the basin of α (see Figure 35.1).

For $\varepsilon = 0$, they merge into a parabolic fixed point ($\alpha = \beta = 0$) with multiplier 1, which then evaporates into the complex plane for $\varepsilon < 0$. (Compare with Exercise 20.15.)

In the complexified family, the attracting-repelling pair exists inside the cardioid obtained by applying $\varepsilon = \lambda^2$ to the disk λ-disk $\mathbb{D}(-1/2, 1/2)$. (Compare with the main cardioid of the Mandelbrot set, Exercise 33.2.) For $\varepsilon = 0$, these points merge into a parabolic fixed point with multiplier 1, which then split into two repelling fixed points outside the cardioid. Moreover, the cardioid has a 3/2-cusp at its root 0.

More generally, let us consider an analytic family of germs near 0,

$$f_\varepsilon(z) = f_0(z) - b\,\varepsilon(1 + \varepsilon\phi_\varepsilon(z)), \tag{35.6}$$

where $f_0(z) = z + z^2 + \dots$ is a non-degenerate parabolic germ, $\phi_\varepsilon(z)$ is holomorphic in two variables in some bidisk $\mathbb{D}_r^2$, and $b \neq 0$. We will call such a family *a generic saddle-node unfolding* (of a non-degenerate parabolic germ).

The germ f_ε has two fixed points

$$\alpha(\varepsilon) \sim -\sqrt{b\varepsilon} \quad \text{and} \quad \beta(\varepsilon) \sim \sqrt{b\varepsilon}.$$

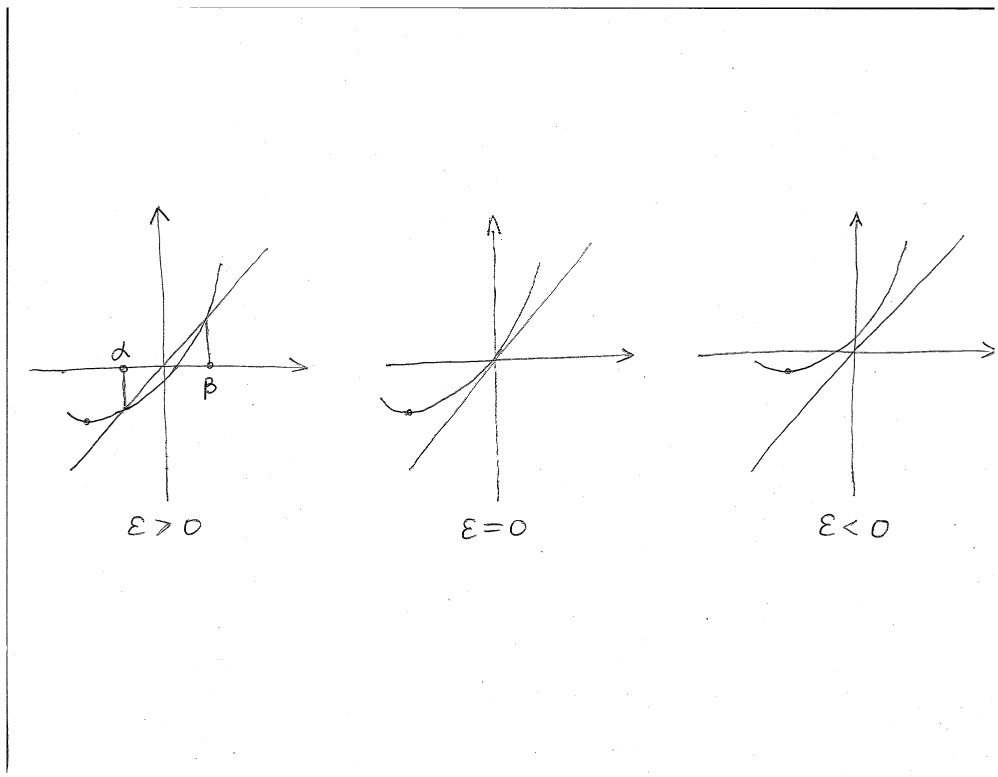

FIGURE 35.1. Real saddle-node bifurcation.

The multipliers of these points are equal to $1 \pm 2\sqrt{b\varepsilon} + O(\varepsilon)$ respectively, so if $b > 0$ then for small positive $\varepsilon > 0$, α is attracting while β is repelling (assuming the principal branch of $\sqrt{\cdot}$ is selected). The following statement shows that for these ε, the immediate attracting basin of α contains β on the boundary.

EXERCISE 35.19. *In a family* (35.6) *with $b > 0$, for $\varepsilon > 0$ sufficiently small, there is an invariant round disk contained in the basin of α and containing β on the boundary.*

35.6. Satellite bifurcation. A hyperbolic component Δ of $\operatorname{int} \mathcal{M}$ which is not primitive is called *satellite*. In this case, the parabolic cycle $\boldsymbol{\alpha}_{\mathfrak{r}}$ at the root point has a smaller period than the attracting cycle $\boldsymbol{\alpha}_c$ inside Δ. Let us study this scenario.

Let us consider a germ g_0 with a degenerate parabolic point,

$$g_0(z) = z + a_{\mathfrak{q}+1} z^{\mathfrak{q}+1} + \dots, \quad a_{\mathfrak{q}+1} \neq 0. \tag{35.7}$$

As usual, it can be normalized so that $a_{\mathfrak{q}+1} = 1$. Include it into a generic holomorphic one-parameter family fixing 0:

$$g_\varepsilon(z) = (1+b\varepsilon)z + z^{\mathfrak{q}+1} + \varepsilon(a_1(\varepsilon)z + a_2(\varepsilon)z^2 + \dots + a_{\mathfrak{q}+1}(\varepsilon)z^{\mathfrak{q}+1} + \dots), \tag{35.8}$$

where $b \neq 0$ and $a_1(\varepsilon) = O(\varepsilon)$. Then g_ε has $\mathfrak{q}$ fixed points besides 0: $\alpha_1(\varepsilon), \dots, \alpha_{\mathfrak{q}}(\varepsilon)$ – such that

$$\alpha_k(\varepsilon) \asymp \sqrt[\mathfrak{q}]{b\varepsilon} \tag{35.9}$$

near 0.

Exercise 35.20. *Show that these fixed points correspond to all* $\mathfrak{q}$*th roots of unity.*

Assume now that we have a holomorphic family f_ε of maps that have a holomorphically moving fixed point $\alpha(\varepsilon)$ such that for for $\varepsilon = 0$ it is parabolic with rotation number $\mathfrak{p}/\mathfrak{q} \neq 0 \mod 1$. Then the family

$$g_\varepsilon(z) := f_\varepsilon^{\mathfrak{q}}(z + \alpha(\varepsilon)) - \alpha(\varepsilon)$$

is of form (35.8). If it is generic then for ε near 0, the maps g_ε have $\mathfrak{q}$ extra fixed points (35.9).

Exercise 35.21. *Show that the extra g_ε-fixed points form an f_ε-cycle of period $\mathfrak{q}$. Moreover, the multiplier of this cycle is $\asymp \varepsilon$.*

Thus, the parabolic fixed point α "gives birth" to a cycle of order $\mathfrak{q}$. This scenario is called a *satellite bifurcation* of order $\mathfrak{q}$. In case $\mathfrak{q} = 2$ and $\mathfrak{q} = 3$ it is also referred to as a *period doubling* and *tripling bifurcation* respectively.

Exercise 35.22. *In a family* (35.8) *with $b > 0$, for $\varepsilon > 0$ sufficiently small, there is a round disk contained in the basin of the cycle $\boldsymbol{\alpha}$ and containing* 0 *on the boundary.*

35.7. Robustness of landing of local invariant curves. The following lemma shows that landing of a local invariant curve is a robust property under special parabolic-repelling perturbations:

Lemma 35.23. *Let us consider a parabolic germ*

$$f_0 : z \mapsto z + az^{k+1} + \dots \text{ with } k \geq 1,\ a \neq 0.$$

Let γ_0 be a topological arc contained in a local repelling petal P with wedge 1 at the origin (see Theorem 21.11) *whose endpoints are related by f_0. Let $f : z \mapsto \rho z + \dots$ be a perturbation of f with $\rho > 1$, and let γ be a nearby arc whose endpoints are related by f. Then $\bigcup_{n\in\mathbb{N}} f^{-n}(\gamma)$ is a curve landing at* 0 (*where f^{-n} are the iterates of the local inverse branch of f*).

Proof. By a power change of variable $Z = z^{-1/k}$, the parabolic germ f_0 can be brought to a Puiseux form (21.9)

$$G_0 : Z \mapsto Z + 1 + O(1/|Z|^{1/k})$$

near infinity. Applying the same change of variable to f, we bring it to a Puiseux form

$$G : Z \mapsto \rho^{-k} Z + O(1).$$

In this coordinate, the petal P becomes a half-plane $Q = \{Z : \operatorname{Re} Z < -t\}$. It is invariant under G^{-1} and the iterates G^{-n} converge to ∞ on Q. The conclusion follows. $\square$

35.8. Roots of hyperbolic components. Let Δ be a hyperbolic component of $\operatorname{int} \mathcal{M}$ of period p. Let $\mathfrak{r} \equiv \mathfrak{r}_\Delta \in \partial\Delta$ be the root of Δ. Let us consider the dynamical characteristic rays $\mathcal{R}_\mathfrak{r}^{\theta_\pm}$ at $\mathfrak{r}$. The following lemma shows that these rays are stable under perturbations *into* Δ:

LEMMA 35.24. *There exists a parameter $c \in \Delta$ near $\mathfrak{r}$ such that the rays $\mathcal{R}_c^{\theta_\pm}$ are the characteristic rays for f_c.*

PROOF. For any level $t > 0$, if c is sufficiently close to $\mathfrak{r}$ then the rays $\mathcal{R}_c^{\theta_\pm}[t, \infty)$ closely follow the $\mathcal{R}_\mathfrak{r}^{\theta_\pm}[t, , \infty)$. Hence for t small enough, the fundamental arc $\mathcal{R}_c^{\theta_\pm}[t, 2^{p_0}t]$ is trapped in the repelling petal of $f_\mathfrak{r}$ (where p_0 is the period of the parabolic point of $f_\mathfrak{r}$ (which can be smaller than p). By Lemma 35.23, the rays $\mathcal{R}_c^{\theta_\pm}[t, 2^{p_0}t]$ land at the repelling point β_c that bifurcated from the parabolic point.

For the same reason, the whole ray portrait $\mathfrak{R}_\mathfrak{r}$ persists under this perturbation, so we obtain a ray portrait $\mathfrak{R}_c$ with the same combinatorics. In particular, by Lemma 24.17 we obtain the characteristic strip Π^c_{ch} mapped by f_c^p with degree two onto a bigger strip. Such a map has two fixed points in $\bar{\Pi}^c_{\text{ch}}$, which we identify with the attracting periodic point $\alpha_c \in \Pi^c_{\text{ch}}$ and the repelling periodic point $\beta_c \in \partial\Pi^c_{\text{ch}}$. Since the immediate basin of α_c is contained in $\overline{\Pi}^c_{\text{ch}}$, the f_c^p-fixed point on its boundary must be β_c.

We conclude that the rays $\mathcal{R}_c^{\theta_\pm}$ are characteristic for f_c. $\square$

Together with the results of §35.2.2, this implies:

COROLLARY 35.25. *Under the above circumstances, the angles $\theta_\pm$ are characteristic for any $c \in \Delta$.*

We say that two hyperbolic components, Δ_1 and Δ_2, *touch*[4] at some point $c_\circ$ if $c_\circ \in \bar{\Delta}_1 \cap \bar{\Delta}_2$.

COROLLARY 35.26. *No two hyperbolic components can touch at their roots.*

PROOF. Otherwise, by Corollary 35.25 there would be two hyperbolic components with the same characteristic angles at their centers. But by the Rigidity Theorem 35.1 the characteristic angles at the center determine the hyperbolic component – contradiction. $\square$

As the characteristic angles determine the dynamical model for the Julia set (Theorems 25.58 and 26.28), we conclude:

THEOREM 35.27. *For the root $\mathfrak{r}$ of a hyperbolic component Δ and any $c \in \Delta$, the Böttcher conjugacy*

$$h_c \equiv B_c^{-1} \circ B_\mathfrak{r} : \mathcal{D}_\mathfrak{r} \to \mathcal{D}_c$$

extends to a conjugacy $\mathcal{J}_\mathfrak{r} \to \mathcal{J}_c$.

In fact, up to a modification on the central component $\mathcal{D}_0$, the above conjugacy is globally dynamical:

[4]It would be more consistent, but less concise, to talk about touching of the *closures* of these components.

Exercise 35.28. *Under the circumstances of the above theorem, the Böttcher conjugacy h_c extends to a conjugacy between the superattracting quasiregular map F constructed in* §26.6 *and the polynomial f_{c_Δ} centered at Δ. However, this conjugacy is not quasiconformal.*

35.9. Bifurcations in the quadratic family.

35.9.1. *Saddle-node bifurcations.* Let us consider a parabolic parameter $c_\circ \in \mathcal{M}$ of period p with multiplier 1. Let $\alpha_\circ$ be one of its parabolic point. By Corollary 21.26, the parabolic map $f_\circ^p$ has only one attracting petal at $\alpha_\circ$, and hence it is non-degenerate at $\alpha_\circ$.

Lemma 35.29. *Under the above circumstances, the family $f_c^p(z)$ is a generic saddle-node unfolding of the parabolic map $f_\circ^p$.*

Proof. Let us translate the parameter $c_\circ$ and the parabolic point $\alpha_\circ$ at 0, and then normalize the quadratic term to be 1. The family $f_c^p(z)$ assumes a form

$$g_\varepsilon(\zeta) = a_0(\varepsilon) + (1 + a_1(\varepsilon))\zeta + (1 + a_2(\varepsilon))\zeta^2 + \dots,$$

where $a_i(0) = 0$. We need to show that $a_0(\varepsilon) \asymp \varepsilon$.

Note first that both a_0 and a_1 cannot identically vanish, for otherwise the fixed points $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ would blow up as $\varepsilon \to 0$. It follows that the fixed points behave as $\varepsilon^{n/2}$ with some $n \in \mathbb{N}$. If n is even then α and β are two regular local branches of the fixed point locus. These points would become attracting at some hyperbolic components Λ_α and Λ_β rooted to 0, contradicting Corollary 35.26.

If $n > 1$ is odd then the multiplier of these fixed points would also behave $1 + q\varepsilon^{k/2}$ with some odd $k > 1$. In this case, there are two sectors of wedge $2\pi/k$ centered at 0 where the multiplier would become < 1 in absolute value. These sectors would be local germs of two hyperbolic components rooted at 0, again contradicting Corollary 35.26.

Hence the fixed points behave as $\sqrt{\varepsilon}$, which is possible only when $a_0(\varepsilon) \asymp \varepsilon$. $\square$

Corollary 35.30. *Under the above circumstances, $\partial\Delta$ has a 3/2-cusp at the root $\mathfrak{r}$.*

So, any parabolic parameter $\mathfrak{r}$ with multiplier 1 is the root point of exactly one hyperbolic component Λ of $\operatorname{int}\mathcal{M}$. Moreover, the period p of the attracting cycle $\boldsymbol{\alpha}_c$ for $c \in \Lambda$ is equal to the period of the parabolic cycle $\boldsymbol{\alpha}_\mathfrak{r}$ at the root. Such a hyperbolic component is called *primitive*.

35.9.2. *Satellite bifurcations.* Let us go back to the quadratic family f_c. Assume $f_\circ \equiv f_{c_\circ}$ has a parabolic periodic point $\alpha_\circ$ of period p with rotation number $\mathfrak{p}/\mathfrak{q} \neq 0 \mod 1$.

Lemma 35.31. *Under the above circumstances, the family f_c^p entertains a generic satellite bifurcation at $c_\circ$ of order $\mathfrak{q}$.*

Proof. By the Implicit Function Theorem, the periodic point $\alpha(c)$ moves holomorphically with c near c_0, so we can consider a local holomorphic family

$$g_c(z) := f_c^p(z + \alpha(c)) - \alpha(c)$$

near 0. The origin is a parabolic fixed point for $g_\circ$ with rotation number $\mathfrak{p}/\mathfrak{q}$, so by Corollary 21.26, it has $\mathfrak{q}$ petals attached to 0. Hence $g^\mathfrak{q}$ has a local expansion (35.7). Moreover, $\partial_c g|_{c=c_\circ} \neq 0$ since otherwise there would be two hyperbolic components that touch at their roots, contradicting Corollary 35.26. Thus, g_c is a generic family of (35.8) kind to which the discussion of §35.6 is applicable. □

36. Structural stability

In §33.6.3 we showed that structural stability is generic in the real quadratic family. Here we will establish a similar result for the complex quadratic family.

36.1. Statement of the result. Let us give the definition of structural stability in our setting. A map $f_\circ : z \mapsto z^2 + c_\circ$ (and the corresponding parameter $c_\circ \in \mathbb{C}$) is called *structurally stable* if for any $c \in \mathbb{C}$ sufficiently close to $c_\circ$, the map f_c is topologically conjugate to $f_\circ$, and moreover, the conjugacy $h_c : \mathbb{C} \to \mathbb{C}$ can be selected continuously in c (in the uniform topology on the Riemann sphere $\hat{\mathbb{C}}$). By definition, the set of structurally stable parameters is open. In this section we will prove that it is dense:

THEOREM 36.1. *The set of structurally unstable parameters is equal to the boundary of the Mandelbrot set together with the centers of hyperbolic components. Hence the set of structurally stable parameters is dense in $\mathbb{C}$. Moreover, any structurally stable map $f_\circ$ is quasiconformally conjugate to all nearby maps f_c.*

Notice that parameters $c_\circ \in \partial\mathcal{M}$ are obviously unstable since the Julia set $\mathcal{J}_\circ$ is connected, while the Julia sets $\mathcal{J}_c$ for nearby $c \in \mathbb{C} \smallsetminus \mathcal{M}$ are disconnected. The centers of hyperbolic components are also unstable since the topological dynamics near a superattracting cycle is different from the topological dynamics near an attracting cycle (the grand orbits on the basin of attraction are discrete in the latter case and are not in the former).

The proof of stability of other parameters will occupy §36.2 – §36.6. The desired conjugacies will be constructed as equivariant holomorphic motions. Since by the Second λ-lemma, holomorphic motions are quasiconformal in the dynamical variable, the last assertion of Theorem 36.1 will follow automatically.

36.2. $\mathcal{J}$-stability. Let us first show that the Julia set $\mathcal{J}_c$ moves holomorphically outside the boundary of $\mathcal{M}$. (Strictly speaking, this step is not needed for the proof of Theorem 36.1 given below, but it gives a good illustration of the method.)

A map $f_\circ : z \mapsto z^2 + c_\circ$ (and the corresponding parameter $c_\circ \in \mathbb{C}$) is called *$\mathcal{J}$-stable* if for any $c \in \mathbb{C}$ sufficiently close to $c_\circ$, the map $f_c|\,\mathcal{J}_c$ is topologically conjugate to $f_\circ|\,\mathcal{J}_\circ$, and moreover the conjugacy $h_c : \mathcal{J}_\circ \to \mathcal{J}_c$ depends continuously on c.

THEOREM 36.2. *The set of $\mathcal{J}$-stable parameters is equal to $\mathbb{C} \smallsetminus \partial\mathcal{M}$ and hence is dense in $\mathbb{C}$. Moreover, the corresponding conjugacies $h_c : \mathcal{J}_\circ \to \mathcal{J}_c$ form a holomorphic motion of the Julia set over the component of $\mathbb{C} \smallsetminus \partial\mathcal{M}$ containing $c_\circ$.*

PROOF. Parmeters $c_\circ \in \partial\mathcal{M}$ are obviously $\mathcal{J}$-unstable as the Julia set $\mathcal{J}_\circ$ is connected while the Julia sets $\mathcal{J}_c$ get disconnected for nearby $c \in \mathbb{C} \smallsetminus \mathcal{M}$.

Let C be the component of $\operatorname{int} \mathcal{M}$ containing $c_\circ$. By Corollary 33.5, C does not contain neutral parameters, and hence all periodic points are persistently hyperbolic over C, either repelling or attracting. Hence they depend holomorphically on $c \in C$. Since C is simply connected (Exercise 33.1 (iii)), these holomorphic functions $c \mapsto \alpha(c)$ are single valued. Moreover, they cannot collide since collisions could occur only at parabolic parameters. Thus, they provide us with a holomorphic motion $h_c : \mathrm{Per}_\circ \to \mathrm{Per}_c$ of the set of periodic points.

This holomorphic motion is equivariant. Indeed, if

$$c \mapsto \alpha(c) = h_c(\alpha)$$

is a holomorphically moving periodic point then $c \mapsto f_c(\alpha(c))$ is also a holomorphically moving periodic point. Hence $f_c(\alpha(c)) = h_c(f_\circ(\alpha))$ and we obtain:

$$f_c(h_c(\alpha)) = f_c(\alpha(c)) = h_c(f_\circ(\alpha)).$$

By the First λ-lemma, this holomorphic motion extends to a continuous equivariant holomorphic motion of the closure of periodic points, which contains the Julia set. Moreover, this motion is automatically continuous in both variables (λ, z), and hence provides us with a family of topological conjugacies between $\mathcal{J}_\circ$ and $\mathcal{J}_c$ continuously depending on c. □

EXERCISE 36.3. *An equivariant holomorphic motion of the Julia set is unique.*

36.3. Bifurcation locus. The set of $\mathcal{J}$-unstable parameters is called the *bifurcation locus*. So, for the quadratic family, it is equal to the boundary of the Mandelbrot set, $\partial\mathcal{M}$. It can also be described in several other ways:

PROPOSITION 36.4. *The bifurcation locus $\partial\mathcal{M}$ coincides with the following sets:*

(i) *The set parameters for which the critical point* 0 *is active;*

(ii) *The closure of parabolic parameters;*

(iiI) *The accumulation set for superattracting parameters;*

(iv) *The closure of Misiurewicz parameters.*

PROOF. (i) This follows from Exercise 33.1 (iv). However, the reader can entertain himself by relating directly the $\mathcal{J}$-instability to activity of the critical point (rather than checking that both properties occur on $\partial\mathcal{M}$).

(ii) We know by Corollary 33.5 that the parabolic parameters belong to $\partial\mathcal{M}$. On the other hand, if $c_\circ \in \partial\mathcal{M}$ is not approximated by neutral parameters, then the proof of Theorem 36.2 shows that $c_\circ$ is $\mathcal{J}$-stable.

Properties (iii) and (iv) follow from Exercise 33.6. □

PROBLEM 36.5. *For a generic parameter $c \in \partial\mathcal{M}$, the orbit of the critical point is dense in the Julia set $\mathcal{J}_c$.*

36.4. Motion over a queer component. Recall Proposition 35.9 asserting that the basin of infinity $\mathcal{D}_c(\infty)$ moves biholomorphically over any component Δ of $\operatorname{int}\mathcal{M}$.

If $\Delta \equiv Q$ is a queer component then $\mathbb{C} = \overline{\mathcal{D}}_c(\infty)$ for any $c \in Q$, and so, we obtain the Structural Stability Theorem in this case:

COROLLARY 36.6. *For a queer component Q of $\operatorname{int}\mathcal{M}$, there is a unique equivariant holomorphic motion $h_c : \mathbb{C} \to \mathbb{C}$ over Q which is biholomorphic on $\mathcal{D}_c(\infty)$. Hence all parameters $c \in Q$ are structurally stable.*

36.5. Motion of an attracting basin. For a hyperbolic parameter c, let $\boldsymbol{\alpha}_c = (f_c^k(\alpha_c))_{k=0}^{p-1}$ be the corresponding attracting cycle. Recall that $\mathcal{D}(\boldsymbol{\alpha}_c)$ stands for its basin. Also, we use notation Δ^* for a punctured hyperbolic component (see §35.3).

PROPOSITION 36.7. *Let Δ be a hyperbolic component of $\operatorname{int}\mathcal{M}$, and let $c_\circ \in \Delta^*$. Then there is an equivariant smooth holomorphic motion of the attracting basin $\mathcal{D}(\boldsymbol{\alpha}_c)$ over some neighborhood of $c_\circ$.*

PROOF. Let us consider the maps f_c^p near their fixed points α_c. Lemma 23.7 implies that there is a neighborhood $\Lambda \subset \Delta^*$ of $c_\circ$ and an $\varepsilon > 0$ such that the inverse linearizing coordinate $\phi_c^{-1}(z)$ for f_c^p is holomorphic on $\Lambda \times \mathbb{D}_\varepsilon$. Let $V_c = \phi_c^{-1}(\mathbb{D}_\varepsilon) \ni \alpha_c$, and let us consider a fundamental annulus $A_c = \operatorname{cl}(V_c \smallsetminus f_c(V_c))$.

By Theorem 21.4, the critical orbit $\operatorname{orb}_c(0)$ must cross A_c. By adjusting ε and shrinking Λ if needed, we can ensure that it does not cross ∂A_c. Then it crosses A_c at a single point $\upsilon_n(c) = f_c^n(0) \in \operatorname{int} A_c$, where $n \in \mathbb{N}$ is independent of c. Its position in the linearizing coordinate, $a_c = \phi_c(\upsilon_n(c)) \in \mathbb{A}(\varepsilon, \rho_c\varepsilon) \equiv \mathbb{A}_c$, depends holomorphically on c (here ρ_c is the multiplier of $\boldsymbol{\alpha}_c$).

Let $Q_c = \partial\mathbb{A}_c \cup \{a_c\}$. Let us define a smooth equivariant holomorphic motion $\mathbf{h}$ of a small neighborhood of Q_c over Λ as follows: $h_c = \mathrm{id}$ near the outer boundary of $\mathbb{A}_\circ$, $h_c : z \mapsto \rho_c z/\rho_\circ$ near the inner boundary of $\mathbb{A}_\circ$, and $h_c : z \mapsto a_c z/a_\circ$ near $a_\circ$. By Lemma 17.1, this motion extends to a smooth motion of the whole plane over some neighborhood of $c_\circ$ (we will keep the same notation Λ for this neighborhood). Let us restrict the motion to the fundamental annulus $\mathbb{A}_c$ (keeping the same notation h_c for it). By Lemma 17.10 (in the simple case when there are no critical points), this motion can be first extended to the forward orbit of $\mathbb{A}_c$, (providing us with an equivariant holomorphic motion of $\mathbb{D}_\varepsilon$).

We can now transfer it using the linearizing coordinates to a holomorphic motion of V_c, then extend it to an invariant neighborhood $\mathbf{V}_c = \bigcup_{k=0}^{p-1} f_c^k(V_c)$ of $\boldsymbol{\alpha}$, and finally we can use Lemma 17.10 to pull this motion back to all preimages of $\mathbf{V_c}$ (the assumption of Lemma 17.10 on the critical values is secured by the property that a_c is a leaf of the motion $\mathbf{h}$). It provides us with the desired equivariant holomorphic motion of the basin $\mathcal{D}(\boldsymbol{\alpha}_c)$. $\square$

COROLLARY 36.8. *Let Δ be a hyperbolic component of $\operatorname{int}\mathcal{M}$, and let $c_\circ \in \Delta^*$. Then there is an equivariant holomorphic motion of the whole plane $\mathbb{C}$ over some neighborhood of $c_\circ$ which is biholomorphic on $\mathcal{D}_c(\infty)$. Hence all parameters $c \in \Delta^*$ are structurally stable.*

PROOF. Since for $\mathbb{C} = \operatorname{cl}(\mathcal{D}_c(\infty) \cup \mathcal{D}(\boldsymbol{\alpha}_c))$ for $c \in \Delta$, Propositions 35.9 and 36.7, together with the First λ-lemma yield the desired. $\square$

36.6. Motion of the Cantor set. Let us finally deal with the complement of $\mathcal{M}$.

PROPOSITION 36.9. *Let $c_\circ \in \mathbb{C} \smallsetminus \mathcal{M}$. Then there is an equivariant smooth holomorphic motion of the basin of infinity, $\mathcal{D}_c(\infty)$, over some neighborhood of $c_\circ$.*

The proof is similar to the one given in the attracting case, using the Böttcher coordinate in place of the linearizing coordinate. To implement it, we need a rotationally equivariant Extension Lemma:

LEMMA 36.10. *Let $R > r > 1$ and let $z \in \mathbb{A}(r, R)$. Let ϕ be a holomorphic function on a domain $(\Lambda, \lambda_\circ)$ with $\phi(\lambda_\circ) = z$. Then there is a smooth holomorphic motion H_λ of the whole complex plane $\mathbb{C}$ over some neighborhood Λ' of $\lambda_\circ$ such that*

(i) $H_\lambda(z) = \phi(\lambda)$;
(ii) $H_\lambda = \mathrm{id}$ *on* $\mathbb{C} \smallsetminus \mathbb{A}(r, R) = \mathrm{id}$;
(iii) *The H_λ commute with the rotation group $\zeta \mapsto e(\theta)\,\zeta$.*

PROOF. Let $\tau(\lambda) = \phi(\lambda)/z$, and let $h_\lambda(\zeta) = \tau(\lambda)\zeta$. This motion satisfies requirements (i) and (iii). To make it satisfy (ii) as well, we will use a smooth cut-off function $\phi : \mathbb{R} \to \mathbb{R}$ supported on a small neighborhood of $|z|$. Then the motion

$$H_\lambda(\zeta) = \phi(|\zeta|)\, h_\lambda(\zeta) + (1 - \phi(|\zeta|))\, \zeta$$

satisfies all the requirements. □

Proof of Proposition 36.9. Let us consider the Böttcher coordinate B_c of f_c near ∞. Since it depends holomorphically on c, there is a neighborhood $U \subset \mathbb{C} \smallsetminus \mathcal{M}$ of $c_\circ$ and an $R > 1$ such that the function $(c, z) \mapsto B_c^{-1}(z)$ is holomorphic on $U \times (\mathbb{C} \smallsetminus \overline{\mathbb{D}}_R)$.

Let $V_c = B_c^{-1}(\mathbb{C} \smallsetminus \overline{\mathbb{D}}_R)$. By adjusting R and U if necessary, we can ensure that the $\mathrm{orb}_c(0)$ does not cross the boundary of the fundamental annulus $A_c = V_c \smallsetminus f_c(V_c)$. Then there is a unique $n > 0$ such that $\upsilon_n(c) = f_c^n(0) \in \operatorname{int} A_c$. Let us mark the corresponding point $a_c = B_c(\upsilon_n(c))$ in the annulus $\mathbb{A} \equiv \mathbb{A}(R, R^2)$.

Applying lemma 36.10, we find a rotationally equivariant holomorphic motion $H_c : \mathbb{A} \to \mathbb{A}$ such that $H_c(a_\circ) = a_c$ and $H_c = \mathrm{id}$ on $\partial\mathbb{A}$.

EXERCISE 36.11. *Show that this holomorphic motion extends to a holomorphic motion $H_c : \mathbb{C} \smallsetminus \mathbb{D}_R \to \mathbb{C} \smallsetminus \mathbb{D}_R$ commuting with $z \mapsto z^2$.*

Let us now transfer H_c by means of the Böttcher coordinate to a holomorphic motion $h_c : V_c \to V_c$, $h_c = B_c^{-1} \circ H_c \circ B_\circ$. This motion is equivariant, $B_c \circ f_\circ = f_c \circ B_c$, and has $\upsilon_n(c)$ as one of its leaves. By Lemma 17.10, it can be lifted to a holomorphic motion of $f_c^{-1}(V_c)$ that has $\upsilon_{n-1}(c)$ as its leaf. Moreover, by the equivariance of the original motion h_c, the new motion coincides with h_c on $V_\circ$ which implies that it is equivariant. Then we can lift it further to $f^{-2}(V_c)$, and so on: in this way we will exhaust the whole basin of ∞. □

Since $\mathbb{C} = \overline{\mathcal{D}}_c(\infty)$ for $c \in \mathbb{C} \smallsetminus \mathcal{M}$, Proposition 36.9 (together with the First λ-lemma) yields:

COROLLARY 36.12. *Let $c_\circ \in \mathbb{C} \smallsetminus \mathcal{M}$. Then there is an equivariant holomorphic motion of the whole plane $\mathbb{C}$ over some neighborhood of $c_\circ$. Hence all parameters $c \in \mathbb{C} \smallsetminus \mathcal{M}$ are structurally stable.*

Corollaries 36.6, 36.8 and 36.12 cover all types of components of $\mathbb{C} \smallsetminus \partial\mathcal{M}$, and together prove the Structural Stability Theorem, 36.1.

36.7. Invariant line fields and queer components.

36.7.1. *Definition.* Informally speaking, a line field on $\mathbb{C}$ is a family of tangent lines $l(z) \in \mathrm{T}_z\mathbb{C}$ depending measurably on $z \in \mathbb{C}$.

Here is a precise definition. Any line $l \in \mathbb{C}$ passing through the origin is uniquely represented by a pair of centrally symmetric points $e(\pm\theta) \in \mathbb{T}$ in the unit circle, or by a single number

$$\nu = e(2\theta) \in \mathbb{T}, \quad \theta \in \mathbb{R}/(\mathbb{Z}/2). \tag{36.1}$$

The space of these lines form, by definition, the one-dimensional projective line $\mathbb{PR}^1$, and (36.1) provides us with its parametrization by the angular coordinate (and shows that $\mathbb{PR}^1 \approx \mathbb{T}$).

Let us now consider the projective tangent bundle over $\mathbb{C}$,

$$\mathrm{PT}(\mathbb{C}) = \mathbb{C} \times \mathbb{PR}^1$$

parametrized by $\mathbb{C} \times (\mathbb{R}/(\mathbb{Z}/2))$. A *line field* on $\mathbb{C}$ is a measurable section of $\mathrm{PT}(\mathbb{C})$ defined on some set $X \subset \mathbb{C}$ of positive area called its *(measurable) support*. In terms of the angular coordinate, we obtain a measurable function $X \to \mathbb{R}/(\mathbb{Z}/2)$, $z \mapsto \theta(z)$.[5] In the circular coordinate ν, we obtain a measurable function $X \to \mathbb{T}$. *In what follows, we will always extend ν by 0 to the whole plane.*

Exercise 36.13. *Show that a line field on a Riemann surface S is given by a Beltrami differential $\nu(z)\dfrac{d\bar{z}}{dz}$ with $|\nu(z)| \in \{0, 1\}$.*

A line field on a set $J \subset \mathbb{C}$ is a line field on $\mathbb{C}$ whose support is contained in J. If such a non-trivial line field exists then area $J > 0$.

A line field is called *invariant* (under a holomorphic map f) if it is invariant under the natural action of f on the projective line bundle: $l(fz) = Df(z)\, l(z)$, or in the angular coordinate, $\theta(fz) = \theta(z) + \arg f'(z) \mod 1/2$, or in the Beltrami coordinate, $f^*\nu = \nu$ (where the pullback is understood in the sense of Beltrami differentials).

If an invariant line field l is supported on a set X then we can pull it back by the dynamics[6] to obtain an invariant line field supported on the set $\tilde{X} = \bigcup_{n=0}^{\infty} f^{-n}(X)$. Hence we can assume in the first place that l *is supported on a completely invariant set*: this will be our standing assumption.

36.7.2. *Existence criterion.*

Proposition 36.14. *Let Q be a queer component of* $\operatorname{int}\mathcal{M}$. *Then any map f_c, $c \in Q$, has an invariant line field on its Julia set. In particular,* area $\mathcal{J}(f_c) > 0$.

Vice versa, if f_c has an invariant line field on its Julia set then c belongs to a queer component of $\operatorname{int}\mathcal{M}$.

Proof. Take some $c_\circ \in Q$. By Corollary 36.6, there is an equivariant holomorphic motion h_c over $(Q, c_\circ)$ which is biholomorphic on $\mathcal{D}_c(\infty)$. Let us consider the corresponding Beltrami differentials $\mu_c = \bar{\partial} h_c / \partial h_c$, $c \in Q$. Each μ_c vanishes on $D_\circ(\infty)$, however $\mu_c \neq 0$ for $c \neq c_\circ$ (for otherwise, by Weyl's Lemma the map h_c would be affine, contrary to the fact the quadratic maps f_c and $f_\circ$ are not affinely conjugate). Hence area(supp μ_c) > 0 for any $c \neq c_\circ$, and all the more, area $\mathcal{J}_\circ > 0$. Moreover, since μ_c is $f_\circ$-invariant, the normalized Beltrami differential $\nu_c = \mu_c/|\mu_c|$ (where we let $\nu_c = 0$ outside supp μ_c) is also $f_\circ$-invariant, and hence determines an invariant line field on the Julia set $\mathcal{J}_\circ$.

Vice versa, assume $f_\circ$ has an invariant line field on $\mathcal{J}_\circ$ given by an invariant Beltrami differential $\nu_\circ$. For any $\lambda \in \mathbb{D}$, the Beltrami differential $\lambda\nu_\circ$ is also f-invariant. Let $h_\lambda : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ be the solution of the corresponding Beltrami

[5] As always, a measurable function is considered up to an arbitrary change on null-sets.

[6] The pullback would fail at the critical point but we can always remove its grand orbit (as any other completely invariant null-set) from $\tilde{X}$.

equation tangent to the identity at infinity. Then the map $h_\lambda \circ f_\circ \circ h_\lambda^{-1}$ is a quadratic polynomial $f_{\sigma(\lambda)} : z \mapsto z^2 + \sigma(\lambda)$ (see §29.1.2). By Corollary 29.3, the map $\sigma : \mathbb{D} \to \mathbb{C}$ is holomorphic. Since the line field is non-trivial, σ is not identically constant. Hence its image covers a neighborhood of $c_\circ$ contained in $\operatorname{int}\mathcal{M}$. So, it is contained in some component Q of $\operatorname{int}\mathcal{M}$. Since $\operatorname{area}\mathcal{J}(f_\circ) > 0$, the map $f_\circ$ is not hyperbolic (see Proposition 25.23), so Q must be queer. $\square$

Thus, the Fatou Conjecture, 33.7, is equivalent to the following one:

CONJECTURE 36.15 (No Invariant Line Fields). *No quadratic polynomial has an invariant line field on its Julia set.*

36.7.3. *Uniqueness and ergodicity.* As a line field $l(z)$ is rotated by angle $2\pi\alpha$ with $\alpha \in \mathbb{R}/(\mathbb{Z}/2)$, the corresponding Beltrami differential is multiplied by $\lambda = e(2\alpha) \in \mathbb{T}$. Of course, if the original line field was f-invariant then so is the rotated one.

LEMMA 36.16. *A quadratic polynomial can have at most one, up to rotation, invariant line field on its Julia set.*

This will follow from the ergodicity (recall §19.6.5) of the action of f on the measurable support of any invariant line field.

LEMMA 36.17. *Let f be a quadratic polynomial, and let $l(z)$ be an invariant line field on $J(f)$. Then the action of f on* $\operatorname{supp} l$ *is ergodic.*

PROOF. Assume that $\operatorname{supp} l$ admits a disjoint decomposition $X_1 \sqcup X_2$ into two measurable invariant subsets of positive measure. Then the restriction of l to these sets gives us two invariant line fields l_i with disjoint supports. Let ν_i be the corresponding Beltrami differentials. Then we can consider a complex two-parameter family of Beltrami differentials $\nu_\lambda = \lambda_1\nu_1 + \lambda_2\nu_2$, where $\lambda = (\lambda_1, \lambda_2) \in \mathbb{D}^2$. Since $\|\nu_\lambda\|_\infty < 1$ for each λ, we can solve the corresponding Beltrami equations and obtain a two parameter family of qc maps $h_\lambda : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ tangent to the identity at infinity. Then the maps $h_\lambda \circ f_\circ \circ h_\lambda^{-1}$ form a family of quadratic polynomials $f_{\sigma(\lambda)} : z \mapsto z^2 + \sigma(\lambda)$ (see §29.1.2).

By Proposition 14.5, the map $\sigma : \mathbb{D}^2 \to \mathbb{C}$ we have obtained this way is continuous (in fact, by Corollary 29.3, it is holomorphic). Hence it cannot be injective: there exist $\lambda \neq \kappa$ in $\mathbb{D}^2$ such that $\sigma(\lambda) = \sigma(\kappa)$. Then the map $\phi = h_\kappa^{-1} \circ h_\lambda$ commutes with $f_\circ$. But the only conformal automorphism of $D_\circ(\infty)$ commuting with $f_\circ$ is the identity (see Exercise 23.24). Hence $h_\lambda = h_\kappa$ implying that $\lambda = \kappa$ – contradiction. $\square$

Proof of Lemma 36.16. Assume we have two invariant line fields given by Beltrami differentials ν_i. Let $X_i = \operatorname{supp} \nu_i$. Notice that due to our convention, both differences, $X_1 \smallsetminus X_2$ and $X_2 \smallsetminus X_1$, are completely invariant sets. If $\operatorname{area}(X_2 \smallsetminus X_1) > 0$ then an invariant Beltrami differential ν which is equal to ν_1 on X_1 and is equal to ν_2 on $X_2 \smallsetminus X_1$ has a non-ergodic support, contradicting Lemma 36.17. Hence $\operatorname{area}(X_2 \smallsetminus X_1) = 0$, and for the same reason $\operatorname{area}(X_1 \smallsetminus X_2) = 0$, so that the set $Y = X_1 \cap X_2$ can be taken as a measurable support for both differentials.

By Lemma 36.17, f acts ergodically on Y. But the ratio ν_2/ν_1 is an invariant function on Y. By ergodicity, it is equal to const a.e. on Y, and we are done. $\square$

36.7.4. *Dynamical uniformization of queer components.* We can now construct a dynamical uniformization of any queer component Q by a Beltrami disk. (Compare with the uniformizations of hyperbolic components of $\mathbb{C} \setminus \partial\mathcal{M}$ given by Theorems 34.1 and 35.3.)

For a base map $f_\circ$, let us select an invariant line field on $\mathcal{J}_\circ$ given by an f-invariant Beltrami differential $\nu_\circ$. Then the Beltrami disk $\{\lambda\nu_\circ\}_{\lambda\in\mathbb{D}}$ generates a holomorphic family of quadratic polynomials $f_{\sigma(\lambda)} : z \mapsto z^2 + \sigma(\lambda)$ (see the proof of the second part of Lemma 36.14). This is the desired uniformization:

Queer Theorem. *The map $\sigma : (\mathbb{D}, 0) \to (Q, c_\circ)$ is the Riemann mapping.*

Proof. The map σ is a holomorphic embedding for the same reason as in the proof of Lemma 36.17. Let us show that it is surjective. Let $c \in Q$. By Corollary 36.6, the map f_c is conjugate to $f_\circ$ by a qc homeomorphism h_c which is conformal outside $\mathcal{J}_\circ$. Let $\mu_c = \bar\partial h_c/\partial h_c$ be the Beltrami differential of h_c, and let $\nu_c = \mu_c/|\mu_c|$. Since the latter differential determines an invariant line field on $\mathcal{J}_\circ$, Lemma 36.16 yields:

$$\operatorname{supp}\mu_c = \operatorname{supp}\nu_c = \operatorname{supp}\nu_\circ.$$

Since the differential μ_c is $f_\circ$-invariant, the ratio $\mu_c/\nu_\circ$ is an f-invariant function. By ergodicity, it is const a.e., so that $\mu_c = \lambda\nu_\circ$ for some $\lambda \in \mathbb{D}$. It follows that $c = \sigma(\lambda)$, and we are done. $\square$

36.8. Quasiconformal classification of the quadratic maps. We can now give a complete classification of the quadratic maps up to qc conjugacy:

Theorem 36.18. *Any qc class in the parameter plane $\mathbb{C}$ of the quadratic family is one on the following list:*

- *the complement of the Mandelbrot set;*
- *a hyperbolic component of* $\operatorname{int}\mathcal{M}$ *punctured at the center;*
- *a queer component of* $\operatorname{int}\mathcal{M}$*;*
- *the center of a hyperbolic component;*
- *a single point of the boundary of* $\mathcal{M}$*.*

The first three types of maps are *qc deformable*, the last two are *qc rigid*.

Proof. By the Structural Stability Theorem, 36.1, each of the above listed sets is contained in some qc class. What we need to show that they belong to different qc classes.

Assume it is not the case: let $c_\circ$ and c be two parameters in different sets but in the same qc class. Then the quadratic polynomials $f_\circ$ and f_c are conjugate by a qc map h. Let $\mu = \bar\partial h/\partial h$ be the Beltrami differential of h, and let $r = 1/\|\mu\|_\infty$. Let us consider the Beltrami disk $\{\lambda\mu : \ |\lambda| < r\}$ and the corresponding qc deformation

$$f_{\sigma(\lambda)} : z \mapsto z^2 + \sigma(\lambda)$$

of $f_\circ$ (see Corollary 29.3). Then $\sigma : \mathbb{D}_r \to \mathbb{C}$ is a holomorphic map such that $\sigma(0) = c_\circ$ and $\sigma(1) = c$. In particular, it is not identically constant and hence its image U is a domain in $\mathbb{C}$. But U is not contained in a single component of $\operatorname{int}\mathcal{M}$, so it must intersect $\partial\mathcal{M}$, and hence it must intersect $\mathbb{C} \setminus \mathcal{M}$. Thus, U contains quadratic maps of both dichotomy types: with connected as well as Cantor Julia sets, which is impossible as all the maps in U are topologically conjugate. $\square$

37. Limbs and wakes of the Mandelbrot set

37.1. Stability of landing.

37.1.1. *General condition.* Recall that $\theta(c)$ stands for the external argument of c, which can be viewed both dynamically (as the critical value) and parametrically (as the parameter of a quadratic map).

LEMMA 37.1. *Assume for some $\theta \in \mathbb{R}/\mathbb{Z}$ and some parameter domain $\Lambda \subset \mathbb{C}$, we have*

$$\theta(c) \neq T^n\theta \text{ for any } n \geq 1 \text{ and } c \in \Lambda \smallsetminus \mathcal{M}.$$

Assume also that for some $c_\circ \in \Lambda$, the ray $\mathcal{R}^\theta_\circ$ lands at some point $a_\circ \in \mathcal{J}_\circ$. Then for any $c \in \Lambda$, the ray $\mathcal{R}^\theta_c$ lands at some point $a_c \in \mathcal{J}_c$ holomorphically depending on $c \in \Lambda$.

PROOF. By Lemma 34.7, the ray $\mathcal{R}^\theta_c$ moves holomorphically under the Böttcher motion h_c over Λ. By the λ-lemma, h_c extends to $\operatorname{cl}\mathcal{R}^\theta_c$. Since $\operatorname{cl}\mathcal{R}^\theta_\circ = \mathcal{R}^\theta_\circ \cup \{a_\circ\}$, we have: $\operatorname{cl}\mathcal{R}^\theta_c = h_c(\operatorname{cl}\mathcal{R}^\theta_\circ) = \mathcal{R}^\theta_c \cup \{h_c(a_\circ)\}$. Finally, let $a_c := h_c(a_\circ)$. □

37.1.2. *Rational case: persistence of being repelling.*

LEMMA 37.2. *Assume for some rational $\theta = \mathfrak{p}/\mathfrak{q} \in \mathbb{Q}_{\text{odd}}/\mathbb{Z}$ and some parameter domain $\Lambda \subset \mathbb{C}$ intersecting $\mathcal{M}$, we have:*

$$\theta(c) \notin \operatorname{orb}_T \theta \text{ for any } c \in \Lambda \smallsetminus \mathcal{M}.$$

Then for any $c \in \Lambda$, the landing point a_c of $\mathcal{R}^\theta_c$ is a repelling periodic point depending holomorphically on c.

PROOF. By assumption, there exists $c_\circ \in \Lambda \cap \mathcal{M}$. By Theorem 24.3, the dynamical ray $\mathcal{R}^\theta_\circ$ lands at some periodic point $a_\circ \in \mathcal{J}_\circ$. Then by Lemma 37.1, the ray $\mathcal{R}^\theta_c$ lands at some point a_c holomorphically depending on c. Since $f^p_c(\mathcal{R}^\theta_c) = \mathcal{R}^\theta_c$ for some p depending only on $\mathfrak{q}$, the point a_c is necessarily periodic. If somewhere it stops being repelling then nearby it becomes attracting. But an attracting periodic point cannot be a landing point of any ray – contradiction. □

EXERCISE 37.3. *Assume for some rational $\theta = \mathfrak{p}/\mathfrak{q} \in \mathbb{Q}_{\text{ev}}/\mathbb{Z}$ and some parameter domain $\Lambda \subset \mathbb{C}$ intersecting $\mathcal{M}$, we have:*

$$\theta(c) \notin \operatorname{orb}_T T(\theta) \text{ for any } c \in \Lambda \smallsetminus \mathcal{M}.$$

Then the ray $\mathcal{R}^\theta_c$ moves biholomorphically over Λ, and for any $c \in \Lambda$, its landing point a_c is a repelling preperiodic point depending holomorphically on c.

37.1.3. *Stability under small perturbation* (*repelling case*). If a is a repelling periodic point of period p for a polynomial f then by the Implicit Function Theorem, a nearby polynomial $\tilde{f}$ has a unique repelling periodic point $\tilde{a}$ near a. We will refer to $\tilde{a}$ as the *perturbed* a.

STABILITY LEMMA. *Assume that a periodic ray $\mathcal{R} \equiv \mathcal{R}^\theta(f)$ of a polynomial f lands at a repelling periodic point a. Then for $\tilde{f}$ sufficiently close to f, the corresponding ray $\tilde{\mathcal{R}} \equiv \mathcal{R}^\theta(\tilde{f})$ lands at the perturbed repelling periodic point $\tilde{a}$.*

REMARK 37.4. Let us emphasize that this lemma applies to both connected and disconnected cases.

Proof. Without loss of generality we can assume that the point a is fixed and the ray $\mathcal{R}$ is invariant. Let $d \geq 2$ be the degree of f.

Let us take a small disk $D \equiv \mathbb{D}(a, 2\varepsilon)$ such that the local inverse branch g of f^{-1} is well defined in D and $g(D) \Subset D$. Then the same is true for $\tilde{f}$ sufficiently close to f.

Let us fix some equipotential level $t > 0$ such that $\mathcal{R}(0, d\,t] \subset \mathbb{D}(a, \varepsilon)$, and let $\gamma = \mathcal{R}[t, d\,t] \subset \mathbb{D}(a, \varepsilon)$.

Let us consider the inverse Böttcher functions $B_{\tilde{f}}^{-1}$. Note that even in the disconnected case, these functions can be analytically extended along any (non-singular) ray (see Problem 23.30). Moreover, $B_{\tilde{f}}^{-1}$ depends continuously on $\tilde{f}$ in the closed-open topology. It follows that if $\tilde{f}$ is sufficiently close to f, then the ray $\tilde{\mathcal{R}}[t, \infty)$ is well defined and ε-close to the ray $\mathcal{R}[t, \infty)$. Hence the ray arc $\tilde{\gamma} := \tilde{\mathcal{R}}[t, d\,t]$ is contained in D for $\tilde{f}$ sufficiently close to f, so the inverse branch $\tilde{g}$ is well defined on $\tilde{\gamma}$.

Let $\tilde{e} := \tilde{\mathcal{R}}(t)$, $\tilde{b} := \tilde{\mathcal{R}}(d\,t) = \tilde{f}(\tilde{e})$ Since both $\tilde{e}$ and $\tilde{b}$ lie in D, we have: $\tilde{e} = g(\tilde{b})$. Hence the arc $\tilde{g}(\tilde{\gamma}) \subset D$ gives an extension of the ray $\tilde{\mathcal{R}}[t, \infty)$ to the ray $\tilde{\mathcal{R}}[t/d, \infty)$. Repeating this argument, we see that the arcs $\tilde{g}^n(\tilde{\gamma})$ give an extension of $\tilde{\mathcal{R}}[0, \infty)$ to the full ray $\tilde{\mathcal{R}}$ landing at $\tilde{a}$. $\square$

37.1.4. *Parabolic case.*

Lemma 37.5. *Let $\mathfrak{r}$ be the root of a hyperbolic component Δ of $\operatorname{int} \mathcal{M}$, and let $\mathcal{R}^\theta$ be a dynamical ray landing at the corresponding parabolic root β of $f \equiv f_{\mathfrak{r}}$. Then for any $c \in \Delta$, the ray $\mathcal{R}_c^\theta$ lands at the corresponding repelling root β_c of f_c.*

Proof. By Corollary 35.11, it is sufficient to prove this for some parameter $c \in \Delta$, so we can consider a small perturbation of $\mathfrak{r}$ whose attracting cycle $\boldsymbol{\alpha}_c = (f_c^n \alpha_c)_{n=0}^{p-1}$ has a real multiplier. Let $d = 2^m$ where m is the period of θ under the doubling map. We can now proceed as in the Stability Lemma. Namely, the ray $\mathcal{R}_c^\theta$ follows $\mathcal{R}^\theta$ for long time, so eventually some fundamental arc $\gamma_c := \mathcal{R}_c^\theta[t, d\,t]$. is trapped inside the repelling petal of β_c. Then it extends to a ray landing at the repelling periodic point β_c which is a perturbation of β. By Exercises 35.19 and 35.22, β_c lies on the boundary of the immediate basin $\mathcal{D}^\bullet(\boldsymbol{\alpha}_c)$. Moreover, it is fixed under the return map to $\mathcal{D}(\alpha_c)$, so it is its root. $\square$

Corollary 37.6. *Let Δ be a hyperbolic component rooted at $\mathfrak{r}$, and let $\mathfrak{p}/\mathfrak{q}$ be the rotation number of the parabolic point of $f_{\mathfrak{r}}$. Then for any $c \in \Delta$, the combinatorial rotation number of the repelling root point β_c is equal to $\mathfrak{p}/\mathfrak{q}$.*

Proof. By Lemma 37.5, the combinatorial rotation number of the repelling root β_c of f_c, $c \in \Delta$, is equal to the combinatorial rotation number of the parabolic root β of $f_{\mathfrak{r}}$. But the latter is equal to the rotation number of this parabolic point (see Theorem 24.6). $\square$

37.2. Landing of rational parameter rays (with odd denominator). Let $\mathfrak{P}_p \subset \mathcal{M}$ be the set of parabolic parameters of period p.

Lemma 37.7. *Let $\theta \in \mathbb{Q}_{\text{odd}}/\mathbb{Z}$, and let p be its period under the doubling map. Then the rational parameter ray $\mathcal{R}_{\text{par}}^\theta$ lands at some parabolic parameter $\mathfrak{r} \in \mathfrak{P}_q$ with $q|p$.*

PROOF. Let $\Theta = (\theta_i)_{i=0}^{\mathfrak{q}-1} \subset \mathbb{T}$, $i \in \mathbb{Z}/p\mathbb{Z}$, be the cycle of $\theta \equiv \theta_1$ under the doubling map.

Observe first that *for any $c \in \mathcal{R}^\theta_{\mathrm{par}}$, the corresponding dynamical ray $\mathcal{R}^\theta_c$ crashes at some pre-critical point.* Indeed, by the Basic Phase-Parameter relation (§34.2), for $c \in \mathcal{R}^\theta_{\mathrm{par}}$ we have $c \in \mathcal{R}^\theta_c$. Then $0 \in \mathcal{R}^{\theta_0}_c$, so, the ray $\mathcal{R}^{\theta_0}_c$ crashes at the critical point 0. Going backwards along the cycle of rays $\mathcal{R}^{\theta_i}_c$, we see that all the dynamical rays of this cycle crash at some precritical points. In particular, the ray $\mathcal{R}^{\theta_1}_c$ does.

Together with the Stability Lemma, this observation implies that *any accumulation point $\mathfrak{r} \in \mathcal{M}$ of the parameter ray $\mathcal{R}^\theta_{\mathrm{par}}$ belongs to $\bigcup_{q|p} \mathfrak{P}_q$.* Indeed, by Theorem 24.3, the corresponding dynamical ray $\mathcal{R}^\theta_{\mathfrak{r}}$ must land at a repelling or parabolic periodic point $\beta_{\mathfrak{r}}$ (with a period q dividing p). If $\beta_{\mathfrak{r}}$ were repelling, then by the Stability Lemma, for c near $\mathfrak{r}$ the ray $\mathcal{R}^\theta_c$ would also safely land at a repelling periodic point β_c, contradicting the above observation.

Finally, the accumulation set ω of the ray $\mathcal{R}^\theta_{\mathrm{par}}$ is either a single point or a continuum. Since $\bigcup_{q|p} \mathfrak{P}_q$ is finite, ω must be a single point $\{\mathfrak{r}\}$, so the ray in question lands at $\mathfrak{r}$. □

37.3. Wake Theorem. Let us consider a parabolic parameter $\mathfrak{r}$ which is not the main cusp of $\mathcal{M}$ (i.e., $\mathfrak{r} \neq 1/4$). Then the corresponding parabolic polynomial has two dynamical characteristic rays $\mathcal{R}^\pm_{\mathfrak{r}} \equiv \mathcal{R}^{\theta_\pm}_{\mathfrak{r}}$ (see §26.7). These rays are periodic with the same period, p. Let $\Delta \equiv \Delta_{\mathfrak{r}}$ be the hyperbolic component rooted at $\mathfrak{r}$ (which is unique by Corollary 35.26). For $c \in \Delta$, we let:

- $\alpha_c \equiv \alpha^{\mathrm{ch}}_c$ and $\beta_c \equiv \beta^{\mathrm{ch}}_c$ be the corresponding characteristic attracting periodic point and the repelling root of f_c.
- $\Theta_\pm \equiv \Theta_\pm(\mathfrak{r}) := \mathrm{orb}_T\, \theta_\pm$ be the corresponding cycles of rays (which can coincide), and $\Theta \equiv \Theta(\mathfrak{r}) := \Theta_+ \cup \Theta_-$. By Lemma 37.5, we know that for $c \in \Lambda$, the rays $\mathcal{R}^\pm_c := \mathcal{R}^{\theta_\pm}_c$ are characteristic rays landing at the root point β_c.
- $\mathcal{W}^{\mathrm{ch}}_c \ni c$ be the characteristic dynamical sector bounded by $\mathcal{R}^\pm_c$.

Let is also consider the following parameter objects:

- $\mathcal{R}^\pm_{\mathrm{par}} \equiv \mathcal{R}^{\theta_\pm}_{\mathrm{par}}$ be the corresponding *parameter rays.*

THEOREM 37.8. *For any parabolic parameter $\mathfrak{r} \neq 1/4$, the parameter rays $\mathcal{R}^\pm_{\mathrm{par}}$ land at $\mathfrak{r}$. (The sector $\mathcal{W}^{\mathrm{par}}_{\mathfrak{r}}$ bounded by $\mathcal{R}^+_{\mathrm{par}} \cup \mathcal{R}^-_{\mathrm{par}} \cup \{\mathfrak{r}\}$ and containing $\Delta_{\mathfrak{r}}$ is called the parameter wake rooted at $\mathfrak{r}$.) The root point β_c, and the dynamical rays $\mathcal{R}^\pm_c$ persist as holomorphically moving objects over the whole wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{r}}$ (and β_c remains repelling throughout). Moreover, $\mathcal{W}^{\mathrm{par}}_{\mathfrak{r}} \not\ni 0$.*

Vice versa, if the ray configuration $\mathfrak{R}_{\mathfrak{r}}$ is realized for some parameter c, then $c \in \mathcal{W}^{\mathrm{par}}_{\mathfrak{r}} \cup \{\mathfrak{r}\}$.

PROOF. • First, define the *wake* $\mathcal{W}^{\mathrm{par}} \equiv \mathcal{W}^{\mathrm{par}}_{\mathfrak{r}}$ as the component of

$$\mathbb{C} \smallsetminus \bigcup_{\theta \in \Theta} \mathrm{cl}(\mathcal{R}^\theta_{\mathrm{par}}) \tag{37.1}$$

containing $\Delta_{\mathfrak{r}}$. Then

$$\theta(c) \notin \Theta \text{ for any } c \in \mathcal{W}^{\mathrm{par}} \smallsetminus \mathcal{M}. \tag{37.2}$$

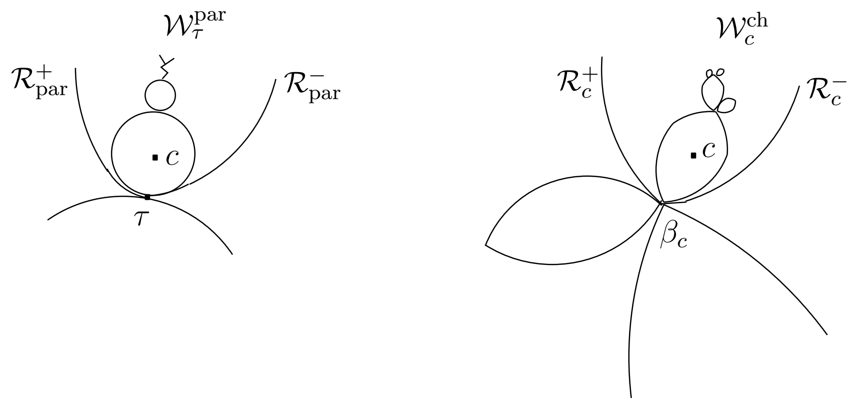

FIGURE 37.1. A parabolic parameter wake $\mathcal{W}_\tau^{\mathrm{par}}$ and the corresponding dynamical characteristic wake $\mathcal{W}_c^{\mathrm{ch}}$ for $c \in \Delta$ (in the satellite case).

By Lemmas 34.7 and 37.2, all the dynamical rays $\mathcal{R}_c^\theta$, $\theta \in \Theta$, and their landing points, move holomorphically over $\mathcal{W}^{\mathrm{par}}$. Hence *for any $c \in \mathcal{W}^{\mathrm{par}}$, we have naturally defined "characteristic rays" $\mathcal{R}_c^\pm$ landing at the "characteristic root" β_c that bound the "characteristic sector" $\mathcal{W}_c^{\mathrm{ch}}$.* (Here all quotation marks can be removed for $c \in \Delta$.) Moreover, the root β_c of this sector remains repelling throughout the wake.

• By (37.2), for $c \in \mathcal{W}^{\mathrm{par}} \setminus \mathcal{M}$, none of the rays $\mathcal{R}_c^\theta$, $\theta \in \Theta$, lands at the critical value $v = c$. Hence *for all $c \in \mathcal{W}^{\mathrm{par}}$, the critical value v remains inside the dynamical characteristic sector $\mathcal{W}_c^{\mathrm{ch}}$.*

• *The wake $\mathcal{W}^{\mathrm{par}}$ is bounded by the parameter characteristic rays $\mathcal{R}_{\mathrm{par}}^\pm$.* Indeed, let $c_\circ \in \partial\mathcal{W}^{\mathrm{par}} \setminus \mathcal{M}$ and $t(c_\circ)$ be the equipotential height of $c_\circ$ (both: dynamical and parameter). Let $\tau = (3/4)t(c_\circ)$. Then the dynamical subrays $\mathcal{R}_c^\theta[\tau, \infty)$, $\theta \in \Theta$, move holomorphically over some neighborhood U of $c_\circ$. It follows that they stay some definite distance $d > 0$ apart for $c \in U$. Hence, if $\theta \in \Theta$ were not a characteristic angle then for $c \in U \cap \mathcal{W}^{\mathrm{par}}$ sufficiently close to $c_\circ$, the critical value $v = c$ would be closer to $\mathcal{R}_c^\theta$ than to the characteristic rays $\mathcal{R}_c^\pm$. But this is impossible since such a v belongs to the dynamical characteristic sector $\mathcal{W}_c^{\mathrm{ch}}$ bounded by the rays $\mathcal{R}_c^\pm$, while other rays $\mathcal{R}_c^\theta$, $\theta \in \Theta$, lie outside this sector (see Lemma 24.17).

Hence *the parameter rays $\mathcal{R}_{\mathrm{par}}^\pm$ land at the same point, and bound the wake $\mathcal{W}^{\mathrm{par}}$.*

• *The parameter rays $\mathcal{R}_{\mathrm{par}}^\pm$ land at $\mathfrak{r}$.* Otherwise, $\mathfrak{r}$ would belong to $\operatorname{int} \mathcal{W}^{\mathrm{par}}$. But then the dynamical root β_c would become attracting for some c near $\mathfrak{r}$, contradicting Lemma 37.2.

• Moreover, $0 \notin \mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}$ since for f_0, no two rays land at the same point. This completes the proof of the first part of the Theorem.

• Vice versa, assume the ray portrait $\mathfrak{R} \equiv \mathfrak{R}_{\mathfrak{r}}$ is realized for some parameter c. If the corresponding landing cycle $\boldsymbol{\beta}_c$ is parabolic then by Theorem 35.27 it can be perturbed to an attracting-repelling pair without changing the ray portrait. It is sufficient to show that the perturbed c belongs to $\mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}$, so we can assume that $\boldsymbol{\beta}_c$ is repelling to start with.

As we know, the ray portrait $\mathfrak{R}$ persists over the full component W of the set (37.1) containing c. Also, the critical value $v = c$ belongs to the characteristic sector S_c^{ch} for all $c \in W$. But for $c \in W \smallsetminus \mathcal{M}$, the Phase-Parameter relation implies that c must belong to the parameter characteristic sector $\mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}$. Hence $W = \mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}$. $\square$

The piece of the Mandelbrot set contained in a wake $\mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}$ is called an *unrooted limb of* $\mathcal{M}$,

$$\mathcal{L}_{\mathfrak{r}}^* \equiv \mathcal{L}_{\mathfrak{r}}^*(\mathcal{M}) := \mathcal{M} \cap \mathcal{W}_{\mathfrak{r}}^{\mathrm{par}}.$$

Its closure

$$\mathcal{L}_{\mathfrak{r}} \equiv \mathcal{L}_{\mathfrak{r}}(\mathcal{M}) := \mathrm{cl}(\mathcal{L}_{\mathfrak{r}}^*) = \mathcal{L}_{\mathfrak{r}}^* \cup \{\mathfrak{r}\}$$

is the *limb of* $\mathcal{M}$ *rooted at* $\mathfrak{r}$ (compare §9.1.5).

We also say that the above wake and the limb are *centered at* $c_\circ$, where $c_\circ$ is the center of the hyperbolic component Δ, and will sometimes label them accordingly: $\mathcal{W}_\circ \equiv \mathcal{W}_{c_\circ}$, $\mathcal{L}_\circ \equiv \mathcal{L}_{c_\circ}$.

The origin is the natural base point (the *center*) of $\mathcal{M}$. Following §9.1.5, the branch of $\mathcal{M}$ at $\mathfrak{r}$ containing 0 is called the *body* $\mathcal{B}_{\mathfrak{r}}$ of $\mathcal{M}$ at $\mathfrak{r}$.

Let us now show that any rational parameter ray with odd denominator is characteristic for its landing point:

PROPOSITION 37.9. *For any* $\theta \in (\mathbb{Q}_{\mathrm{odd}}/\mathbb{Z})^*$, *the parameter ray* $\mathcal{R}^\theta$ *lands at the parabolic parameter* $\mathfrak{r}$ *for which* θ *is one of two characteristic angles.*

PROOF. Let us use a counting argument. Take any period $p \in \mathbb{Z}_+$. By Corollary 35.6, there are 2^{p-1} hyperbolic components of some period $q|p$.

Since no two hyperbolic components touch at their roots (Corollary 35.26), there exists 2^{p-1} distinct roots of the above hyperbolic components. Each root $\mathfrak{r} \neq 1/4$ has two associated characteristic angles $\theta_\pm(\mathfrak{r})$. Altogether, we account for $2^p - 2$ characteristic angles associated with these roots.

On the other hand, the doubling map $T : \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$ has $2^p - 1$ periodic angles of some period $q|p$ (which are fixed points of T^p). Disregarding $\theta = 0$, we obtain exactly the same number, $2^p - 2$, of such periodic points. Hence, each of them must serve as a characteristic ray for some parabolic point. $\square$

37.4. Main Wakes Decomposition. Let us now specify the above discussion to the *main wakes*, i.e., the satellite wakes attached to the main cardioid. Let

- α_c be the fixed point of f_c holomorphically depending on $c \in \mathbb{C} \smallsetminus [1/4, +\infty)$ and such that $\alpha_0 = 0$ (compare §24.4);
- Δ_0 stand for the main hyperbolic component of $\operatorname{int} \mathcal{M}$ bounded by the main cardioid $\mathcal{N}_0 \equiv \partial\Delta_0$;
- For $\gamma \in \mathbb{R}/\mathbb{Z}$, $c(\gamma) \in \partial\Delta_0$ be the boundary point of Δ_0 with *internal angle* γ.

For $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, we let

- $\Delta_{\mathfrak{p}/\mathfrak{q}}$ be the satellite hyperbolic component rooted at $c(\mathfrak{p}/\mathfrak{q})$ (by Corollary 35.26, it is unique);
- $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}^{\mathrm{par}}$ be the *main* $\mathfrak{p}/\mathfrak{q}$*-satellite wake* rooted at $c(\mathfrak{p}/\mathfrak{q})$ (we also let $\mathcal{W}_0^{\mathrm{par}} := \mathbb{C} \smallsetminus \mathcal{R}^0$);
- $\Theta(\mathfrak{p}/\mathfrak{q}) \subset \mathbb{R}/\mathbb{Z}$ be the rotation set for the doubling map T with rotation number $\mathfrak{p}/\mathfrak{q}$ (see Appendix 24.7), and $\theta_\pm(\mathfrak{p}/\mathfrak{q}) \in \Theta(\mathfrak{p}/\mathfrak{q})$ be the corresponding characteristic angles.

Theorem 37.8, together with Lemma 24.8 (ii), imply:

COROLLARY 37.10. *The wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$, together with its root $c(\mathfrak{p}/\mathfrak{q})$, is equal to the set of parameters c for which the combinatorial rotation number of α_c is equal to $\mathfrak{p}/\mathfrak{q}$. The dynamical rays $\mathcal{R}^\theta_c$ with $\theta \in \Theta(\mathfrak{p}/\mathfrak{q})$, together with their landing point α_c, move holomorphically under the Böttcher motion over $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$.*

For $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, let us define the *main unrooted limb* $\mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}}$ of $\mathcal{M}$ as $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}} \cap \mathcal{M}$ and the *main limb* $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ as $\mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}} \cup \{c(\mathfrak{p}/\mathfrak{q})\}$.

THEOREM 37.11. *The Mandelbrot set admits the following decomposition:*

$$\mathcal{M} = \overline{\Delta}_0 \sqcup \bigsqcup_{\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*} \mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}}.$$

PROOF. For any parameter $c \in \mathcal{M} \smallsetminus \overline{\Delta}_0$, the fixed point α_c has a well defined rotation number $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$. By Corollary 37.10, $c \in \mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$, and the conclusion follows. □

For $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, let us consider the *shadow*

$$\mathrm{Sh}_{\mathfrak{p}/\mathfrak{q}} \equiv \mathrm{Sh}(\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}) = [\theta_-(\mathfrak{p}/\mathfrak{q}), \theta_+(\mathfrak{p}/\mathfrak{q})]$$

of the corresponding wake at infinity (as defined in §9.1.1). Let $\mathrm{Sh}_{\mathbb{Q}^*}$ be the union of all these shadows. (We can also define Sh_0 as $(\mathbb{R}/\mathbb{Z})^*$.)

PROPOSITION 37.12. *The union of the shadows, $\mathrm{Sh}_{\mathbb{Q}^*}$, is dense in $\mathbb{R}/\mathbb{Z}$. For any $\theta \in (\mathbb{R}/\mathbb{Z})^* \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}$, the parameter ray $\mathcal{R}^\theta_{\mathrm{par}}$ lands at some irrational point $c(\gamma)$, $\gamma \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$ of the main cardioid $\mathcal{N}_0$. Moreover, every irrational point $c(\gamma) \in \mathcal{N}_0$ is the landing point of exactly one ray $\mathcal{R}^\theta$, and the corresponding θ lies in $(\mathbb{R}/\mathbb{Z})^* \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}$.*

PROOF. Let $\theta \in (\mathbb{R}/\mathbb{Z})^* \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}$. Theorem 37.11 implies that the ray $\mathcal{R}^\theta_{\mathrm{par}}$ accumulates on some connected subset $\omega \subset \mathcal{N}_0$ which contains at most one parabolic point. Hence $\omega = \{c_\gamma\}$ is a single point, so $\mathcal{R}^\theta_{\mathrm{par}}$ lands at c_γ. In fact, c_γ is not parabolic for otherwise it would be a landing point of two rays ($\mathcal{R}^\theta$ and one of the characteristic rays landing at c_γ) that bound an (open) sector S that does not intersect $\mathcal{M}$ (contradicting Lemma 8.18). Furthermore, if $(\mathbb{R}/\mathbb{Z}) \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}$ contained an interval J then all the rays $\mathcal{R}^\theta$, $\theta \in J$, would land at the same point c_γ, which is impossible. The conclusions follow. □

Thus, we obtain a natural *Devil Staircase projection* $\mathbb{R}/\mathbb{Z} \to \mathcal{N}_0$ that collapses the shadows $\mathrm{Sh}_{\mathfrak{p}/\mathfrak{q}}$ to the corresponding roots $c(\mathfrak{p}/\mathfrak{q})$.

37.5. General Wakes Decomposition. The previous discussion can be generalized to the satellite limbs attached to any hyperbolic component Δ of $\operatorname{int} \mathcal{M}$. We will use a similar notation as above, usually keeping Δ implicit. So, let

• For $\gamma \in \mathbb{R}/\mathbb{Z}$, $c(\gamma) \equiv c_\Delta(\gamma) \in \partial\Delta$ be the boundary point of Δ with *internal angle* γ;

• $\mathfrak{r} \equiv \mathfrak{r}_\Delta = c(0)$ be the root of Δ, and $\mathcal{W}^{\mathrm{par}}_0 \equiv \mathcal{W}^{\mathrm{par}}_0(\Delta)$ be the parabolic wake rooted at $\mathfrak{r}$;

• For $c \in \Delta$, $\alpha_c \equiv \alpha_c(\Delta)$ be the valuable attracting periodic point of f_c (of period p); (the same notation will be used for its continuous extension to $\overline{\Delta}$ and holomorphic continuation to bigger parameter regions);

- $\alpha_{\mathfrak{p}/\mathfrak{q}} \equiv \alpha_{c(\mathfrak{p}/\mathfrak{q})}$ be the valuable parabolic point for $f_{c(\mathfrak{p}/\mathfrak{q})}$;
- $\Delta_{\mathfrak{p}/\mathfrak{q}}$ be the satellite hyperbolic component rooted at $c(\mathfrak{p}/\mathfrak{q})$ (by Corollary 35.26, it is unique);
- $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}} \equiv \mathcal{W}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ be the $\mathfrak{p}/\mathfrak{q}$-*parabolic wake* rooted at $c(\mathfrak{p}/\mathfrak{q})$;
- $\Theta(\mathfrak{p}/\mathfrak{q};\Delta) \subset \mathbb{R}/\mathbb{Z}$ be the set of angles θ whose rays $\mathcal{R}^{\theta}_{c(\mathfrak{p}/\mathfrak{q})}$ land at $\alpha_{\mathfrak{p}/\mathfrak{q}}$. It is a rotation set under T^p.

At the center $c_\circ$ of Δ, we can consider characteristic strip $\Pi^{\circ}_{\mathrm{ch}}$ (see §25.7.1) and the corresponding characteristic intervals I^{ch}_1 and I^{ch}_2 at the circle of infinity. Since the ray portrait $\mathfrak{R}_c$ persists over the wake $\mathcal{W}^{\mathrm{par}}_0 \cup \{\mathfrak{r}\}$, so does the strip; call it Π^{c}_{ch}. Note that the characteristic intervals do not depend on $c \in \mathcal{W}^{\mathrm{par}}_0 \cup \{\mathfrak{r}\}$.

Exercise 37.13. *Show that for $c \in \mathcal{W}^{\mathrm{par}}_0$, the map f^p has two fixed points, β_c and α_c, in $\overline{\Pi^{c}_{\mathrm{ch}}}$. If α_c is repelling or parabolic with combinatorial rotation number $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$ then the external angles of α_c form a tuned rotation set under $f^p|\,(I^{\mathrm{ch}}_1 \cup I^{\mathrm{ch}}_2)$ with rotation number $\mathfrak{p}/\mathfrak{q}$.*

Thus, *the combinatorial rotation number of α_c determines the ray portrait of its cycle.* Now, Theorem 37.8, together with Lemma 24.8 (ii), imply:

Corollary 37.14. *For any hyperbolic component Δ, the dynamical rays $\mathcal{R}^{\theta}_c$ with $\theta \in \Theta(\mathfrak{p}/\mathfrak{q};\Delta)$, together with their landing point α_c, move holomorphically under the Böttcher motion over the satellite wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$. The wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$, together with its root $\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}}$, is equal to the set of parameters $c \in \mathcal{W}^{\mathrm{par}}_0$ for which the combinatorial rotation number of α_c is equal to $\mathfrak{p}/\mathfrak{q}$. The characteristic angles $\theta_\pm$ of the wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}$ have tuned rotation number $\mathfrak{p}/\mathfrak{q}$.*

Let us now define general limbs by letting $\mathcal{L}^*_0 \equiv \mathcal{L}^*_0(\Delta) := \mathcal{W}^{\mathrm{par}}_0(\Delta) \cap \mathcal{M}$ and $\mathcal{L}_0 \equiv \mathcal{L}_0(\Delta) := \mathcal{L}^*_0 \cup \{\mathfrak{r}\}$. We will also use notation $\mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}} \equiv \mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ and $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}} \equiv \mathcal{L}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ for the satellite limbs attached to Δ.

Theorem 37.15. *Any limb $\mathcal{L}_0 \equiv \mathcal{L}_0(\Delta)$ admits the following decomposition:*

$$\mathcal{L}_0 = \overline{\Delta} \sqcup \bigsqcup_{\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*} \mathcal{L}^*_{\mathfrak{p}/\mathfrak{q}}.$$

Proof. Let $c \in \mathcal{L}^*_0$. If the corresponding periodic point α_c is repelling then it has some combinatorial rotation number $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, so by Corollary 37.14 it must belong to $\mathcal{W}^{\mathrm{par}}_{\mathbb{Q}^*}$. If α_c is non-repelling then $c \in \overline{\Delta}$. □

Corollary 37.16. *For any hyperbolic component Δ, the unrooted limb $\mathcal{L}^*_0(\Delta)$ is connected.*

Proof. The unrooted limb $\mathcal{L}^*_0$ is obtained by attaching to the unrooted component $\Delta \smallsetminus \{\mathfrak{r}\}$ the limbs $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$, $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$. As all these sets are connected, the limb $\mathcal{L}^*_0$ is connected as well. □

Let $\mathrm{Sh}_0 \equiv \mathrm{Sh}(\mathcal{W}^{\mathrm{par}}_0)$ be the *shadow* of the the wake $\mathcal{W}^{\mathrm{par}}_0$ at infinity. As in the fixed point case, for $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, let $\mathrm{Sh}_{\mathfrak{p}/\mathfrak{q}} \equiv \mathrm{Sh}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ be the shadow of the satellite wake $\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ at infinity, and let $\mathrm{Sh}_{\mathbb{Q}^*} \equiv \mathrm{Sh}_{\mathbb{Q}^*}(\Delta)$ be the union of all these shadows.

Proposition 37.17. *For any hyperbolic component Δ, we have:*

(i) *The shadow* $\mathrm{Sh}_{\mathbb{Q}^*}$ *is dense in* Sh_0.

(ii) *For any rotation number* $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$*, the parabolic point* $c(\mathfrak{p}/\mathfrak{q}) \in \partial\Delta$ *is the landing point of exactly two external rays* (*its characteristic rays*).

(iii) *For any* $\theta \in \mathrm{Sh}_0 \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}$*, the parameter ray* $\mathcal{R}^\theta_{\mathrm{par}}$ *lands at some point* $c(\gamma) \in \partial\Delta$ *with irrational* $\gamma \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$*. Moreover, every irrational point* $c(\gamma) \in \partial\Delta$ *is the landing point of exactly one ray* $\mathcal{R}^\theta$*, and the corresponding* θ *lies in* $\mathrm{Sh}_0 \smallsetminus \mathrm{Sh}_{\mathbb{Q}^*}(\Delta)$.

(iv) *Only the boundary external rays of* Sh_0 *land at the root* $\mathfrak{r}$.

Proof. Assertion (ii) follows from Corollary 37.16 (applied to $\Delta_{\mathfrak{p}/\mathfrak{q}}$). All the rest is similar to Proposition 37.12. □

Again, we obtain a natural *Devil Staircase projection* $\mathrm{Sh}_0(\Delta) \to \partial\Delta$ that collapses the shadows $\mathrm{Sh}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ to the corresponding roots $c_\Delta(\mathfrak{p}/\mathfrak{q})$.

37.6. Boundaries of hyperbolic components.

37.6.1. *Only primitive roots are singular.*

Proposition 37.18. *If two hyperbolic components of* $\mathrm{int}\, M$ *touch, then one of them is obtained from the other by a satellite bifurcation. In the latter case, there are no other components attached to the bifurcation point.*

Proof. Let us consider two hyperbolic components, Δ and $\tilde{\Delta}$, touching at some point c. By Corollary 35.26, c cannot be the root for both of them, so assume for definiteness that c is not the root of Δ. Then we have for the corresponding multiplier: $\rho(c) = e^{2\pi i\theta} \neq 1$. Now Theorem 37.15 implies:

• If θ is irrational then nothing can be attached to Δ at c.

• If $\theta = \mathfrak{p}/\mathfrak{q}$ is rational then the only hyperbolic component attached to Δ at c is the one obtained by the satellite $\mathfrak{p}/\mathfrak{q}$-bifurcation, i.e., $\tilde{\Delta} = \Delta_{\mathfrak{p}/\mathfrak{q}}$. □

We are now prepared to show that the boundaries of hyperbolic components are non-singular except at primitive roots.

Proposition 37.19. *Let* Δ *be a hyperbolic component of the Mandelbrot set rooted at* $\mathfrak{r}$*. Then:*

(i) *If* Δ *is satellite then* $\partial\Delta$ *is a non-singular real analytic Jordan curve;*

(ii) *If* Δ *is primitive then* $\partial\Delta$ *is a real analytic Jordan curve with the only singularities at the root* $\mathfrak{r}$*, where it has a* $(3/2)$*-cusp.*

Proof. Let us consider the multiplier function $\rho : \Delta \to \mathbb{D}$. Since it is a branch of an algebraic function, the boundary $\partial\Delta$ is a piece of a real algebraic curve $\{|\rho(z)| = 1\}$ (maybe, singular). Since Δ is a topological disk, $\partial\Delta$ is a Jordan curve.

Assume $c_0 \in \partial\Delta$ is not the root. Then ρ analytically extends to a neighborhood of c_0. If $\rho'(c_0) = 0$ then $\rho^{-1}(\mathbb{D})$ would consist of several components attached to c_0. Each of them would be a hyperbolic component of the same period as Δ, contradicting Proposition 37.18.

Hence $\rho'(c_0) \neq 0$, and the IFT implies that near c_0 the boundary $\partial\Delta = \rho^{-1}(\mathbb{T})$ is non-singular.

If Δ is satellite and $c_0 = \mathfrak{r}$ is its root, then Lemma 35.31 and Exercise 35.21 imply that the multiplier ρ analytically extends to a neighborhood of c_0, with $\rho'(c_0) \neq 0$. By the IFT, $\partial\Delta$ is non-singular near c_0.

Finally, if Δ is primitive and $c_0 = \mathfrak{r}$ is its root, then $\partial\Delta$ has a (3/2)-cusp at c_0 by Corollary35.30. $\square$

37.6.2. *The inverse multiplier.*

LEMMA 37.20. *Let Δ be a hyperbolic component of $\mathcal{M}$, and let $\rho : \Delta \to \mathbb{D}$ be the corresponding multiplier function. Then the inverse function $\tau := \rho^{-1}$ extends to a holomorphic function in a neighborhood of $\overline{\mathbb{D}}$. Moreover,*

- *If Δ is satellite then τ is univalent.*
- *If Δ is primitive then τ is univalent outside a neighborhood of $1 \in \mathbb{T}$ and has a simple critical point at $\rho = 1$.*

PROOF. Let p be the period of Δ, and let $c_\circ \in \partial\Delta$. Assume first that $c_\circ$ is not the root of Δ. Then $\rho(c)$ extends holomorphically to a neighborhood of $c_\circ$. If c were critical for ρ then there would be another hyperbolic component Δ' with the same period p attached to $c_\circ$ (such that $\rho(\Delta') = \rho(\Delta)$ locally), which would contradict Corollary 37.18. Hence ρ extends univalently to a neighborhood of $c_\circ$.

Let us now consider the root point $c_\circ \in \partial\Delta$. If Δ is satellite then again, ρ extends univalently to a neighborhood of $c_\circ$ (Exercise 35.21). If Δ is primitive then ρ extends to a univalent function of $\sqrt{c - c_\circ}$ (Lemma 35.29).

The conclusion follows. $\square$

COROLLARY 37.21. *If Δ is satellite then $\partial\Delta$ is a non-singular real analytic Jordan curve. If Δ is primitive then Δ is a real analytic Jordan curve with the only singularity at the root, where it has a 3/2-cusp.*

37.6.3. *MLC outside primitive roots.* We are now ready to prove perfect rigidity (see §9.2.1), and hence local connectivity, of $\mathcal{M}$ on the main cardioid.

LEMMA 37.22. *The limbs $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}(\Delta)$ attached to any hyperbolic component Δ of* $\operatorname{int}\mathcal{M}$ *shrink as $\mathfrak{q} \to \infty$.*

PROOF. Follows from the General Limbs Decomposition (Theorem 37.15) together with Lemma 9.20. $\square$

LEMMA 37.23. (i) *The Mandelbrot set is perfectly rigid at its root 1/4.*

(ii) *Any satellite limb $\mathcal{L}_\mathfrak{r}$ of $\mathcal{M}$ is perfectly rigid at its root $\mathfrak{r}$.*

(iii) *Let Δ be a hyperbolic component rooted at $\mathfrak{r}$, and let c be a boundary point of Δ. Assume $c \neq \mathfrak{r}$ in case when Δ is a primitive component different from the main cardioid. Then $\mathcal{M}$ is perfectly rigid at c.*

PROOF. By Lemmas 9.19 and 37.22, $\mathcal{M}$ is perfectly rigid at the root 1/4 and at any irrational point $c(\theta)$, $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$, of the main cardioid $\mathcal{N}_0$.

Let $\Delta_{\mathfrak{p}/\mathfrak{q}}$ be a satellite hyperbolic component rooted at a parabolic point $c(\mathfrak{p}/\mathfrak{q})$, $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, and let $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ be the corresponding limb. As above, Lemmas 9.19 and 37.22 imply that $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ is perfectly rigid at $c(\mathfrak{p}/\mathfrak{q})$. By Lemma 9.20 (iii), $\mathcal{M}$ is perfectly rigid at $c(\mathfrak{p}/\mathfrak{q})$.

Thus, $\mathcal{M}$ is perfectly rigid at any point of the main cardioid.

More general assertions of (ii) and (iii) are handled similarly. $\square$

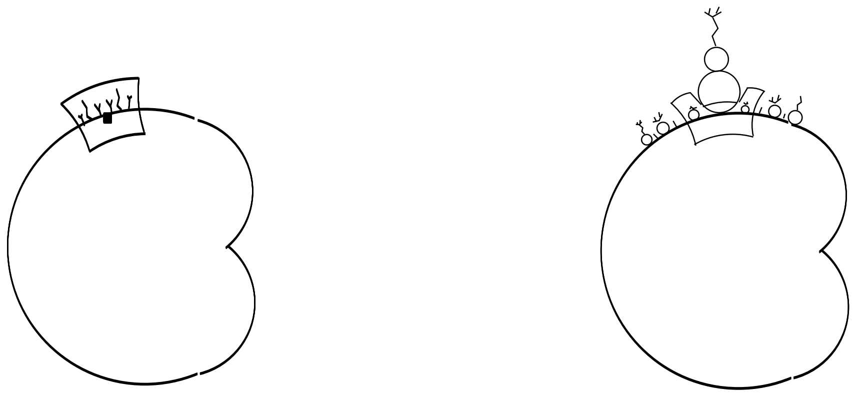

FIGURE 37.2. MLC on the main cardioid at a parabolic (right) and an irrational (left) parameters.

37.6.4. *Size of the limbs.*

PROPOSITION 37.24. *For any hyperbolic component* Δ *of* $\operatorname{int}\mathcal{M}$*, there exists a constant* $C = C_\Delta$ *such that for any limb* $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ *attached to* Δ*, we have:*

$$\operatorname{diam}\mathcal{L}_{\mathfrak{p}/\mathfrak{q}} \leq \frac{C}{\mathfrak{q}}.$$

PROOF. Let $\rho : \Delta \to \mathbb{D}$ be the multiplier function on Δ. By Lemma 37.20, the inverse function $\tau = \rho^{-1}$ extends to a holomorphic function in a neighborhood of $\overline{\mathbb{D}}$.

Let us pull the limbs $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ back by τ to an outer neighborhood of $\mathbb{T}$, i.e., consider components $\tilde{\mathcal{L}}_{\mathfrak{p}/\mathfrak{q}}$ of $\tau^{-1}(\mathcal{L}_{\mathfrak{p}/\mathfrak{q}})$ attached to $e^{2\pi i\mathfrak{p}/\mathfrak{q}} \in \mathbb{T}$. By the Yoccoz Inequality, these sets have diameters $O_\Delta(1/\mathfrak{q})$. Hence, so do their images under ρ. □

EXERCISE 37.25. *If* Δ *is primitive then* $\operatorname{diam}\Delta_{1/\mathfrak{q}} = O(1/\mathfrak{q}^2)$.

PROPOSITION 37.26. *For any irrational* $\theta \in \mathbb{R}/\mathbb{Z}$*, there is a single parameter ray* $\mathcal{R}^\eta$ *landing at the point* $c(\theta) \in \mathcal{C}$ *with internal angle* θ.

37.7. Growth of entropy.

37.7.1. *Real cycles settle forever.* Let us now consider a real superattracting quadratic polynomial $f = f_c$, $c \in [-2, 1/4]$ of period $p > 0$. Its Hubbard tree $\mathcal{T}$ is the interval $[v, f(v)]$ with the marked postcritical set $\mathcal{O}$. As this information is equivalent to prescribing the (finite) kneading sequence of f, the Rigidity Theorem 35.1 implies:

COROLLARY 37.27. *A finite kneading sequence uniquely determines a real superattracting parameter.*

For $c > 1/4$, the map f_c does not have any real periodic points, while for $c < -2$, all periodic points are real (see §20.4). So, in between they gradually come to the real line from the complex plane (by means of saddle-node or doubling bifurcations). However, theoretically some of them can then disappear before coming back again, and eventually settling down permanently on the real line. We are now prepared to show that in reality this scenario does not happen.

Let us consider a real hyperbolic window $(c_1, c_2) \subset [-2, 1/4]$, endowed with the multiplier parametrization $\rho(c)$. According to the Multiplier Theorem, $\rho : (c_1, c_2) \to (-1, 1)$ is a real analytic diffeomorphism, so it is monotonic. However, the Multiplier Theorem does not tell us whether ρ is increasing or decreasing. The Wake Theorem does:

LEMMA 37.28. *For any real hyperbolic window (c_1, c_2), the multiplier coordinate $\rho : (c_1, c_2) \to (-1, 1)$ is monotonically increasing.*

PROOF. Otherwise $\rho(c_1) = 1$, so c_1 is the root of the corresponding hyperbolic component $\Delta \supset (c_1, c_2)$. Let us consider the corresponding parabolic wake $\mathcal{W} \supset \Delta$ rooted at c_1. It is bounded by an $\mathbb{R}$-symmetric topological line L comprising two unreal rays. Hence $L \cap \mathbb{R} = \{c_1\}$, implying that $\mathcal{W} \cap \mathbb{R} = (c_1, +\infty) \ni 0$. But 0 does not belong to any wake – contradiction. $\square$

PROPOSITION 37.29. *Once a periodic point α_c appears on the real line for some $\mathfrak{r} \in [-2, 1/4]$, it stays real for all $c < \mathfrak{r}$. If it becomes repelling for $c < \mathfrak{r}$ near $\mathfrak{r}$, it will stay repelling forever. If it becomes attracting, it stays attracting over the corresponding hyperbolic window $(t, \mathfrak{r})$, and then turns into repelling forever.*

PROOF. Let α_c has period p in the punctured neighborhood of $\mathfrak{r}$. For $c = \mathfrak{r}$, it must merge with another periodic point β_c (of period dividing p). So, $\alpha_{\mathfrak{r}} = \beta_{\mathfrak{r}}$ is a parabolic periodic point for $f_{\mathfrak{r}}$ with $(f_{\mathfrak{r}}^p)'(\alpha_{\mathfrak{r}}) = 1$.

By Proposition 37.18, there is exactly one hyperbolic component Δ of period p rooted at $\mathfrak{r}$. Since $\mathcal{M}$ is $\mathbb{R}$-symmetric, so is Δ. Hence the real slice $\Delta \cap \mathbb{R}$ is a real hyperbolic window rooted at $\mathfrak{r}$. By Lemma 37.28, $\Delta \cap \mathbb{R} = (t, \mathfrak{r})$, where $\rho(t) = -1$.

Let $\mathcal{W}_{\mathfrak{r}}$ and $\mathcal{W}_t$ be the corresponding hyperbolic wakes. The same argument as for Lemma 37.28 shows that $\mathcal{W}_{\mathfrak{r}} \cap \mathbb{R} = (-\infty, \mathfrak{r})$ and $\mathcal{W}_t \cap \mathbb{R} = (-\infty, t)$.

If the component Δ is primitive then for $c \in (t, \mathfrak{r})$, one of the points α_c, β_c is attracting (say, α_c) while the other is repelling. By Theorem 37.8, β_c stays repelling throughout the whole wake $\mathcal{W}_{\mathfrak{r}} \supset (-\infty, \mathfrak{r})$, while α_c becomes repelling as c crosses t and stays repelling throughout the wake $\mathcal{W}_t \supset (-\infty, t)$.

If the component Δ is satellite, then for $c \in (t, \mathfrak{r})$, α_c and β_c belong to one attracting cycle. Applying Theorem 37.8 again, we conclude that this cycle becomes repelling as c crosses t and stays repelling throughout the wake $\mathcal{W}_t \supset (-\infty, t)$. $\square$

PROBLEM 37.30. *Given a period p, describe the first moment when a periodic point of period p appears in the real quadratic family.*

37.7.2. *Entropy.* For real maps, *topological entropy* can be defined as the growth rate for the number of periodic points (see Theorem 48.7):

$$h(f) = \lim \frac{1}{n} \log |\mathrm{Per}_n|,$$

where $\mathrm{Per}_n = \{x : f^n x = x\}$.

Proposition 37.29 immediately implies:

THEOREM 37.31. *As c moves from $1/4$ to -2, the topological entropy $h(f_c)$ monotonically changes from 0 to $\log 2$.*

REMARK 37.32. The entropy function $c \mapsto h(f_c)$ is an example of the Devil Staircase: it is constant on the hyperbolic windows, and grows on the complementary Cantor set (of positive length).

37.8. Monotonicity of the kneading function $c \mapsto \mathrm{Kn}(c)$. The space of kneading sequences is ordered with the twisted lexicographic order defined in §32.4.2. This allows us to formulate a more complete version of Theorem 37.31 on the entropy growth (and of the preceding Proposition 37.29):

THEOREM 37.33. (i) *The level sets of the kneading function $c \mapsto \mathrm{Kn}(c)$ are connected (and hence, are either intervals or singletons).*
(ii) *The kneading sequence $\mathrm{Kn}(c)$ depends monotonically on $c \in \mathcal{M}_{\mathbb{R}}$.*

PROOF. (i) Rigidity of superattracting parameters (Corollary 38.13) implies that the level set of a finite kneading sequences is a single point.

The level set of a periodic kneading sequence is an open interval with superattracting endpoints (of period p and $2p$). It comprises two hyperbolic half-intervals meeting at a doubling bifurcation point. (There is one exception: the level set of $\bar{\varepsilon} = (+ + + \dots)$ is the interval $(0, 1/4]$.)

Assume $\mathrm{Kn}(c_0) = \mathrm{Kn}(c_1)$ for some non-hyperbolic parameters $c_1 < c_0$, but $\mathrm{Kn}(c)$ is not constant on the interval $[c_1, c_0]$. For the kneading sequence to change, one of the postcritical points 0_p must cross 0, creating a superattracting parameter $c' \in [c_1, c_0]$ (compare Theorem 33.12).

Let us consider the first event like this with period $\leq p$, so $\mathrm{Kn}_p(c) = \mathrm{Kn}_p(c_0)$ for $c \in (c', c_0)$. By the Multiplier Theorem, after the crossing the kneading sequence must change: $\mathrm{Kn}_p(c) \neq \mathrm{Kn}_p(c_0)$ for $c < c'$ near c'.

But then the points 0_n, $n = 0, \dots, p$ should be re-organized back creating another superattracting parameter $c'' \in (c_1, c')$ with the same kneading sequence as for c'. As we already know, this is impossible.

□

37.9. Realization of abstract Hubbard trees.

THEOREM 37.34. *Any abstract Hubbard tree with a periodic critical point is realized by a unique superattracting quadratic polynomial f_c.*

PROOF. Uniqueness part is taken care of by Theorem 35.1, so let us deal with the existence.

The idea for the existence is to extend f to a topological double branched covering $\mathbb{C} \to \mathbb{C}$ and to read from it the charcateristic angles of the desired superattracting parameter.

Let $\mathcal{T} \subset \mathbb{C}$ be an abstract Hubbard tree endowed with with an F-action. First, let us extend F to a topological double branched covering $\hat{F} : \mathbb{C} \to \mathbb{C}$ that looks like a quadratic polynomial.

To this end, let us embed $\mathcal{T}$ into a symmetric tree $\mathcal{T}' \subset \mathbb{C}$ and extend F to a double branched covering $\mathcal{T}' \to \mathcal{T}$ (which we denote in the same way). Mark on $\mathcal{T}'$ all the points $F^{-1}(\mathbf{0})$ and all the branch points.

In this context, a "ray" $\mathcal{R}$ for $\mathcal{T}$ (resp., for $\mathcal{T}'$) means a simple path $\mathbb{R}_+ \to \mathbb{C} \smallsetminus \mathcal{T}$ landing at some point of $\mathcal{T}$ (resp., of $\mathcal{T}'$). (Similarly, we can consider "closed rays".) A "ray configuration" is a disjoint family of rays.

Let us consider two ray configurations $\mathfrak{R} \subset \mathfrak{R}'$, for $\mathcal{T}$ and $\mathcal{T}'$ respectively, landing at the marked points of the corresponding tree and uniquely representing every access to them (with respect to the corresponding tree). Together with the corresponding Hubbard trees, they provides us with two cell decompositions $\mathcal{C}$ and $\mathcal{C}'$ of $\mathbb{C}$ such that $\mathcal{C}'$ is a refinement of $\mathcal{C}$. (Thus, vertices for $\mathcal{C}$ are the marked points

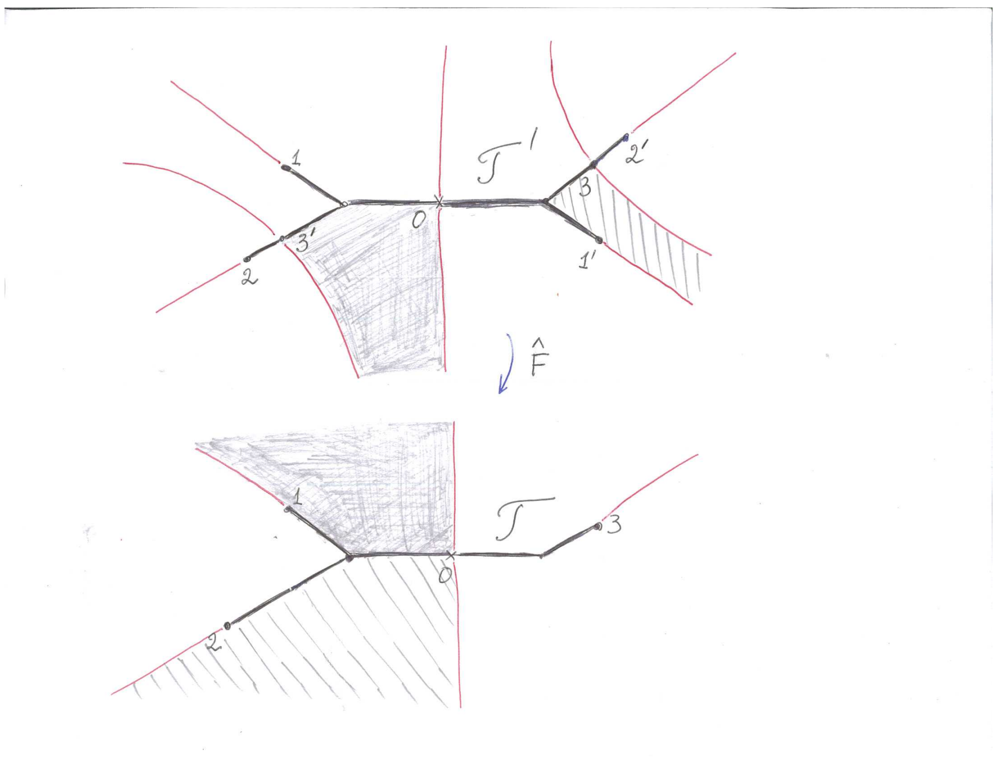

FIGURE 37.3. Extension of a Hubbard map to a double branched covering $\hat{F}:\mathbb{C}\to\mathbb{C}$. Two cells and their respective images are shadowed.

of $\mathcal{T}$ and ∞; its edges are rays of $\mathfrak{R}$ and arcs of $\mathcal{T}$ connecting two neighboring vertices; and similarly for $\mathcal{T}'$).

Since $F:\mathcal{T}'\to\mathcal{T}$ maps the accesses to $\mathcal{T}'$ to accesses to $\mathcal{T}$, it can be extended to a map $\hat{F}:\mathfrak{R}'\to\mathfrak{R}$ mapping each closed ray of $\mathfrak{R}'$ homeomorphically to a closed ray of $\mathfrak{R}$. This map further extends to a cell maps $\hat{F}:(\mathbb{C},\mathcal{C}')\to(\mathbb{C},\mathcal{C})$ acting homeomorphically on the cells and being a global double branched covering.

Let us further adjust $\hat{F}$ to supply it with "characteristic rays". To this end, let us select a point $\beta\in\mathcal{T}$ near c_1 and adjust the original f on $\mathcal{T}$ to make β periodic. Let us consider two "rays" $\mathcal{R}^{\pm}_{\text{ch}}$ landing at β on the opposite sides of $\mathcal{T}$ (selected so that they are disjoint from the previously constructed rays). The union $\mathcal{R}^{+}_{\text{ch}}\cup\mathcal{R}^{-}_{\text{ch}}$ is homotopically periodic under $\hat{F}$ rel $\mathcal{T}$. Reading off their itineraries with respect to the extended Hubbard tree, we obtain the desired chractecteritic angles.

□

37.10. Misiurewicz rays, wakes and decorations.

37.10.1. *Parameter rays.* The following result is an analogue of Theorem 37.8, but it is much easier:

THEOREM 37.35. *Let $c_\circ$ be a Misiurewicz parameter, and let η_j, $j = 1, \dots, s$, be the corresponding dynamical valuable angles. Then the parameter rays $\mathcal{R}_{\mathrm{par}}^{\eta_j}$ land at $c_\circ$.*

Vice versa, any rational angle $\eta \in (\mathbb{Q}_{\mathrm{ev}}/\mathbb{Z})$ with even denominator is valuable for some Misiurewicz parameter $c_\circ$, so the corresponding parameter ray $\mathcal{R}_{\mathrm{par}}^{\eta}$ lands at $c_\circ$.

PROOF. Let $f_\circ \equiv f_{c_\circ}$, $\boldsymbol{\alpha}_\circ$ be the postcritical cycle of $f_\circ$, and let $l \in \mathbb{N}$ be the smallest moment when $f_\circ^{l+1}(0) \in \boldsymbol{\alpha}_\circ$. Let θ_i be the external angles of the rays landing on $\boldsymbol{\alpha}_\circ$. By the Stability Lemma, the configuration of rays $\mathcal{R}_c^{\theta_i}$, together with their landing cycle $\boldsymbol{\alpha}_c$, moves holomorphically under the (extended to $\boldsymbol{\alpha}_c$) Böttcher motion near $c_\circ$.

By continuity, for c near $c_\circ$,

$$f_c^n(0) \notin \boldsymbol{\alpha}_c \cup \bigcup_j \mathcal{R}_c^{\theta_i}, \quad n = 0, \dots, l.$$

Hence the above motion can be pulled back to a motion of the rays $\mathcal{R}_c^{\eta_j}$, together with their landing point a_c, where $a_{c_\circ} = c_\circ$ (compare Proposition 34.6).

Lemma 34.4 implies that the diagonal $\Gamma = \{(c, c)\}$ is transverse to these rays motion intersecting once every leaf. By the Phase-Parameter relation (Lemma 34.9), we obtain a configuration of parameter rays $\mathcal{R}_{\mathrm{par}}^{\eta_i}$ landing at $c_\circ$.

Vice versa, given an angle $\eta \in \mathbb{Q}_{\mathrm{ev}}/\mathbb{Z}$, let us consider its orbit $\Theta := \mathrm{orb}_T(\eta)$ under the doubling map. Let

$$\mathfrak{R}_c := \bigcup_{\eta_i \in \Theta} \mathcal{R}_c^{\eta_i}, \quad \mathfrak{R}_{\mathrm{par}} := \bigcup_{\eta_i \in \Theta} \mathcal{R}_{\mathrm{par}}^{\eta_i}$$

be the corresponding dynamical and parameter ray portraits. For $c \in \mathbb{C} \smallsetminus \mathrm{cl}\, \mathfrak{R}_{\mathrm{par}}$, the dynamical portrait $\mathfrak{R}_c$ is well defined (i.e., the rays do not crash) and moves biholomorphically under the Böttcher motion.

Let $c_\circ \in \mathcal{M}$ be any accumulation point for $\mathfrak{R}_{\mathrm{par}}^{\eta}$. Let us show that this parameter is not parabolic.[7] Otherwise, let α_v be the valuable parabolic point for $f_\circ \equiv f_{c_\circ}$ and let $\mathcal{D}_v$ be its immediate basin. By Theorem 24.6, the dynamical rays landing at α_v are rational rays with odd denominator, so $\mathcal{R}_\circ^{\eta}$ does not land at α_v. Hence it is separated from the critical value $v_\circ = c_\circ \in \mathcal{D}_v$ by some cut-line $L_\circ$ through a repelling periodic point. By the Stability Lemma, this cut-line persists under perturbations: for c near $c_\circ$, there is a cut-line L_c near $L_\circ$. By continuity, L_c separates $v = c$ from $\mathcal{R}_c^{\eta}$, so this ray cannot land at v. On the other hand, it does for $c \in \mathcal{R}_{\mathrm{par}}^{\eta}$, which can be selected arbitrary close to $c_\circ$ – contradiction.

By Corollary 24.4, the ray $\mathcal{R}_\circ^{\eta}$ lands at some repelling preperiodic point $a \in \mathcal{J}_c$.

If a is not pre-valuable then by the Stability Lemma, the pullback-rays $\mathcal{R}_c^{\eta/2}$ and $\mathcal{R}_c^{\eta+1/2}$ do not crash for c near $c_\circ$. On the other hand, they do for any $c \in \mathcal{R}_{\mathrm{par}}^{\eta}$,

[7] Note that in the satellite case, it also follows from Proposition 37.17.

which accumulate on $c_\circ$. This contradiction shows that a is pre-valuable, so $f_\circ$ is postcritically preperiodic.

In fact, η is valuable for $f_\circ$. Otherwise $\mathcal{R}_\circ^\eta$ can be separated from the critical value $v_\circ = c_\circ$ by a cut-line $L_\circ$ through an α-prefixed point (since the Julia set $\mathcal{J}_\circ$ is a dendrite, see Exercise 27.5). By the Stability Lemma, this cut-line persists under a perturbation separating the critical value $v_c = c$ from the ray $\mathcal{R}_c^\eta$. Hence $v_c \notin \mathcal{R}_c^\eta$. $\square$

As an immediate Corollary, we obtain the following Rigidity result:

COROLLARY 37.36. *A preperiodic parameter c is determined by any valuable angle η_j of f_c.*

In particular, for any $\mathbf{n} \in \mathbb{Z}_+$, the Mandelbrot set has $2^{\mathbf{n}-1}$ *dyadic tips* $\mathbf{t}_{\mathbf{n},k}^{\mathrm{Mis}}$ which are the landing points of the dyadic rays $\mathcal{R}_{\mathrm{par}}^{2k-1/2^{\mathbf{n}}}$ $k = 1, \dots, 2^{\mathbf{n}-1}$. Moreover, since the β-fixed point has rotation number 0, this is the only ray landing at the corresponding tip. Note that the first of these tips (corresponding to $\mathbf{n} = 1$) is the landing point of $\mathcal{R}_{\mathrm{par}}^{1/2}$, which is the *Chebyshev* parameter -2.

37.10.2. *Wakes and decorations.* The parameter rays $\mathcal{R}_{\mathrm{par}}^{\eta_j}$ landing at a Misiurewicz (preperiodic) parameter $c_\circ$ divide the plane into s sectors. The $s-1$ sectors that do not contain 0 are called *Misiurewicz (or preperiodic) wakes* $\mathcal{V}_k$ rooted at $c_\circ$. The corresponding limbs of the Mandelbrot set, $\mathcal{T}_i^{\mathrm{Mis}} := (\mathcal{M} \cap \mathcal{V}_k) \cup \{c_\circ\}$, are called *Misiurewicz (or preperiodic) decorations* rooted at $c_\circ$.

Let $\boldsymbol{\alpha}_\circ$ be the postcritical repelling cycle for a Misiurewicz map $f_\circ \equiv f_{c_\circ}$, and let $\mathfrak{R}_\circ$ be the corresponding periodic ray portrait. Let l be the smallest natural number such that $f_\circ^{l+1}(0) \in \boldsymbol{\alpha}_\circ$.

EXERCISE 37.37. *Under the above circumstances, let $\mathcal{R}_\circ^{\eta\pm}$ be two consecutive dynamical rays landing at the critical value $v_\circ = c_\circ$ and bounding a dynamical wake $\mathcal{W}_\circ^{\mathrm{Mis}}$. Let $\mathcal{W}^{\mathrm{Mis}}$ be the corresponding parameter wake bounded by $\mathcal{R}_{\mathrm{par}}^{\eta\pm}$. Assume $f^n(0) \notin \mathcal{W}_\circ^{\mathrm{Mis}}$ for $n = 0, \dots, l$. Then the corresponding ray portrait $\mathfrak{R}_c$, and its preimages up to order l, move biholomorphically under the Böttcher motion over $\mathcal{W}_{\mathrm{par}}^{\mathrm{Mis}}$.*

37.10.3. *Well branching.*

PROPOSITION 37.38. *The Mandelbrot set is well branched at any preperiodic parameter $c_\circ$.*

PROOF. Let $\boldsymbol{\alpha}_\circ = (f_\circ^n \alpha_\circ^{\mathrm{ch}})_{n=0}^{p-1}$ be the postcritical cycle for $f_\circ \equiv f_{c_\circ}$, $f_\circ^{l+1}(0) = \alpha_\circ^{\mathrm{ch}}$, $l \in \mathbb{N}$. Let S_j° be the sectors rooted at $\alpha_\circ^{\mathrm{ch}}$, $j \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$. For any j, approximate $\alpha_\circ^{\mathrm{ch}}$ with pre-$\boldsymbol{\alpha}$ points a_k° in such a way that there are cut-lines L_k° through a_k° converging to ∂S_k°. These cut lines persist in some neighborhood of $c_\circ$. Since $v_l = f^l(c)$ move transversely to $\alpha_{\mathrm{ch}}(c)$, we can find nearby parameters c_k such that $f_{c_k}^l(c_k) = a_k(c)$. Since the corresponding parameter cut-lines L_k^{par} converges to the $\partial S_j^{\mathrm{par}}$, there is only one branch of $\mathcal{M}$ in S_j^{par}. $\square$

37.11. Renormalization windows.

37.11.1. *Complex windows.* Let us start with a hyperbolic map $f_\circ = f_{c_\circ}$ of period $p > 1$ with the characteristic ray portrait Θ. Let $\mathfrak{r}_\circ$ be the root of the corresponding hyperbolic component $\Delta_\circ \equiv \Delta_{\mathfrak{r}_\circ} \subset \operatorname{int}\mathcal{M}$, and let $\mathcal{W}_\circ^{\mathrm{par}} \equiv \mathcal{W}_{\mathfrak{r}_\circ}^{\mathrm{par}} \supset \Delta_\circ$ be the corresponding parameter wake.

In §28.4.3 we constructed, for any $c \in \mathcal{W}_\circ$, a valuable degenerate ql map $f_c^p : W_c \to W_c'$, where $W_c' \ni c$ is a topological triangle bounded by two characteristic rays and an equipotential of level t'. Let $\mathcal{W}_\circ^{\mathrm{par}}(t') \equiv \mathcal{W}_{\mathfrak{r}_\circ}^{\mathrm{par}}(t')$ be the corresponding *truncated parameter wake* bounded by the parameter rays landing at $\mathfrak{r}$ and the equipotential of level t' (see Figure 28.3).

Lemma 37.39. *The boundaries of W_c, W_c' move biholomorphically under the Böttcher motion over the truncated parameter wake $\mathcal{W}_\circ^{\mathrm{par}}(t')$.*

Proof. By Theorem 37.8, the characteristic rays $\mathcal{R}_c^{\pm}$ move bi-holomorphically under the Böttcher motion over the wake $\mathcal{W}_\circ^{\mathrm{par}}$. Moreover, $f^p(0)$ lands in W_c' for $c \in \mathcal{W}_\circ^{\mathrm{par}}$, so, it does not hit the boundary rays. By Lemma 17.10, two additional boundary rays of W_c also move holomorphically over $\mathcal{W}_\circ^{\mathrm{par}}$.

The equipotential story is similar. By Lemma 34.4 (ii), the equipotential $\mathcal{E}_c(t')$ moves bi-holomorphically over $\mathcal{V}^{\mathrm{par}}(t')$. Since $f_c^p(0) \notin \mathcal{E}_c(t')$ for $c \in \mathcal{W}_\circ^{\mathrm{par}}(t')$, so does its f_c^p-pullback $\mathcal{E}_c(t)$. □

Thus, we obtain a holomorphic family $f_c^p : W_c \to W_c'$ of degenerate ql maps over the truncated parameter wake $\mathcal{W}_\circ^{\mathrm{par}}(t')$. By the Phase-Parameter Relation, it yields the parameter domain $\mathcal{V}_\circ \equiv \mathcal{V}_\circ^{\mathrm{par}}(t')$ corresponding to W_c. It is called a *renormalization window* (*rooted at* $\mathfrak{r}_\circ$*, or centered at* $c_\circ$). Note that the extra boundary $\partial\mathcal{V}_\circ \smallsetminus \partial\mathcal{W}_\circ^{\mathrm{par}}(t')$ comprises two arcs of rays landing at the Misiurewicz point $\mathfrak{t}$ (for which $f(0) = \alpha'_{\mathrm{ch}}$ where $f^p(\alpha'_{\mathrm{ch}}) = \alpha_{\mathrm{ch}}$) [compare Theorem 37.35] and two equipotential arcs of level $t = t'/2^p$. The landing point $\mathfrak{t}$ is called the *tip* of $\mathcal{V}_\circ$,

Proposition 37.40. (i) *A map f_c, $c \in \mathcal{V}_\circ$, is renormalizable with period p if and only if*

$$f_c^{pn}(0) \in \overline{W}_c, \quad n = 0, 1, \dots \tag{37.3}$$

The tip map $f_{\mathfrak{t}}$ is renormalizable as well.

(ii) *The root map $f_{\mathfrak{r}_\circ}$ is renormalizable with period p if and only if the hyperbolic component $\Delta_\circ$ is primitive.*

(iii) *The above parameters constitute all parameters c for which f_c is renormalizable* (*or almost renormalizable*) *with combinatorics Θ* (*as defined in* §28.4.5).

Proof. Assertion (i) follows from Lemma 28.20 for the "if" part (including the tip point) and from Exercise 28.28 (iii) for the "only if" part.

Assertion (ii) follows from Exercise 28.24.

(iii) Since the wake $\mathcal{W}_\circ^{\mathrm{par}}$ (together with its root) consists of all parameters for which Θ is realizable, our parameters belong to this wake. The previous assertions tell us which of them are renormalizable (resp.: almost renormalizable). □

Let $M_\circ$ be the above set of almost renormalizable parameters (which are genuinely renormalizable, except for the root parameter $\mathfrak{r}_\circ$ in the case when it is satellite). It is called a *little M-copy* (*centered at* $c_\circ$ *or rooted at* $\mathfrak{r}_\circ$). In §43, we will see that $M_\circ$ is indeed, a homeomorphic copy of the big Mandelbrot set $\mathcal{M}$.

EXERCISE 37.41. *Show that any little copy $M_\circ$ is a hull.*

37.11.2. *Renormalization combinatorics revisited.* Recall from §28.4.5 that the DH renormalization combinatorics is recorded by the ray portrait Θ of the cut-cylce $\boldsymbol{\alpha}$ defining the renormalization. By Proposition 37.9, the charactreristic parameter rays of Θ land at the root of some hyperbolic component Δ. So, we can record the combinatorial data by the center c of this component, or else by the corresponding *Hubbard tree* $\mathcal{T}$. [Notice that $\mathcal{T}$ can be directly read off from Θ by generating the completely invariant geodesic lamination $\mathcal{L}_\Theta$ and taking its Hubbard tree (compare Proposition 25.61).]

EXERCISE 37.42. (i) *Assume f is renormalizable with combinatorics $\mathcal{T}_\circ$ and its renormalization $g = Rf$ is superattracting with Hubbard tree $\mathcal{T}$.* (*Work out the notion of the Hubbard tree for a hyperbolic ql map.*) *Describe the Hubbard tree of f. It is called the* tuning *of the Hubbard tree $\mathcal{T}_\circ$ by $\mathcal{T}$.*
(ii) *Show that the tuning on the level of abstract Hubbard trees is realizable by quadratic polynomials.*

Yet another way of recording the renormalization combinatorics is by the little M-copy introduced above (§37.11.1). Indeed, the hyperbolic center c determines such an M, and the other way around. In this way we are reminded of all the maps that are renormalizable with a given combinatorics.

Accordingly, the *full renormalization combinatorics* $(\Theta^{[0]}, \Theta^{[1]}, \dots)$ (see §28.4.7) can be alternatively recorded by a sequence of superattracting parameters, $(c^{[0]}, c^{[1]}, \dots)$ (such that $c^{[n+1]}$ belongs to the little $M-$copy centered at $c^{[n]}$), or by the sequence of the corresponding Hubbard trees $(\mathcal{T}^{[0]}, \mathcal{T}^{[1]}, \dots)$ (such that $\mathcal{T}^{[n+1]}$ is a "tuning" of $\mathcal{T}^{[n]}$ by some *relative* Hubbard tree $\mathcal{T}_n$, see Exercise 37.42) or else, by a nest of little M-copies $M^{[0]} \supset M^{[1]} \supset \dots$

REMARK 37.43. It is often more instructive to describe the full renormalization combinatorics as the string of *relative* Hubbard trees $\mathcal{T}_n$ recording types of the renormalizations $g_{n+1} = R_n g_n$. Once we know that the "little M-copies" are actually homeomorphic to $\mathcal{M}$, it will be easy to show that any string $(\mathcal{T}_n)_{n=0}^\infty$ of relative renormalization combinatorics is realizable by some quadratic polynomial (see Proposition 43.10).

37.11.3. *Real windows.* Let us take the real slice $\mathcal{V}_\circ^\mathbb{R} := \mathcal{V}_\circ \cap \mathbb{R}$ of a complex renormalization window. It is an open interval bounded be the root and tip of $\mathcal{V}_\circ$, naturally called an (*open*) *real renormalization window.* [We will refer to its closure $\overline{\mathcal{V}}_\circ^\mathbb{R}$ as a *closed real renormalization window.*] Proposition 37.39 implies:

COROLLARY 37.44. *Let $f_\circ$ be a real hyperbolic map of period $p > 1$. Then the configuration of rays and equipotentials that creates the canonical almost renormalization moves continuously over the real renormalization window $\mathcal{V}_\circ^\mathbb{R}$.*

REMARK 37.45. In fact, the rays still move continuously up to the root point. However, pinching occurs at the tip.

Similarly, we have the real counterpart of Proposition 37.40:

COROLLARY 37.46. *All the maps f_c, $c \in \overline{\mathcal{V}}_\circ^\mathbb{R}$, are really renormalizable with period p, where the renormalization can be degenerate at the root parameter.*

REMARK 37.47. Recall from §30.7 that even in the real case we require *repulsiveness* for the ("non-degenerate") renormalization.

PROPOSITION 37.48. *To any admissible superattracting kneading sequence* $\mathrm{Kn}_\circ$*, corresponds a unique real renormalization window* $\overline{\mathcal{V}}_\circ^{\mathbb{R}}$ *of parameters renormalizable with combinatorics* $\mathrm{Kn}_\circ$*. Moreover, the corresponding renormalizations form a full unimodal family.*

PROOF. Since the little copy $M_\circ$ is a real symmetric hull, its slice $M_\circ^{\mathbb{R}}$ is a closed interval. As both endpoints of $\mathcal{V}_\circ^{\mathbb{R}}$ belong to $M_\circ^{\mathbb{R}}$, we conclude that $M_\circ^{\mathbb{R}} = \overline{\mathcal{V}}_\circ^{\mathbb{R}}$. But all maps in $M_\circ$ are renormalizable with the same combinatorics, which for real maps amounts to the same kneading sequence, $\mathrm{Kn}_\circ$.

Now, let $f_\circ : \mathcal{I} \to \mathcal{I}$ be a superattracting quadratic polynomial with kneading sequence $\mathrm{Kn}_\circ$, and let g be any unimodal map. Then we can *tune* $f_\circ$ by g in the following way.

Let $(\mathcal{I}_n)_{n=0}^{p-1}$ be the immediate attracting basin for $f_\circ$, let β_0 be the root of $\mathcal{I}_0$, and let $J := \operatorname{Im} f^p(\mathcal{I}_0) = [0, \beta_0]_\#$. By rescaling, we can realize g as a unimodal map $\mathcal{I}_0 \to \mathcal{I}_0$ fixing β_0; let $L := \operatorname{Im} g \ni \beta_0$. Then there exists an orientation preserving diffeomorphism $h : (\mathcal{I}_0, J) \to (\mathcal{I}_0, L)$ fixing β_0 and such that $g = h \circ f_\circ^p \,|\, \mathcal{I}_0$.

Let us now define $f : \mathcal{I} \to \mathcal{I}$ so that it coincides with $f_\circ$ on $\mathcal{I} \smallsetminus \mathcal{I}_{p-1}$ and equal to $h \circ f$ on $\mathcal{I}_{p-1}$. It is a (continuous) unimodal map whose renormalization is g.

Since the quadratic family is full (Theorem 33.9), there is a quadratic polynomial f_c with the same kneading sequence as the map f just constucted.

The uniqueness part follows from Corollary 37.27. □

REMARK 37.49. a) If g is selected as a map of class $\mathfrak{G}$ then one can show that the quadratic polynomial f_c is actually *topologically conjugate* to f. [Though the latter is not in class $\mathfrak{G}$, the results from §32.4.1 (based on §30 and, in particular, on the No Wandering Intervals Theorem) are still valid.]

b) The Straightening Theorem, 43.1, will imply that in fact, there is a natural homeomorphism $\chi : \overline{\mathcal{V}}_\circ^{\mathbb{R}} \to \mathcal{M}_{\mathbb{R}}$ respecting combinatorics.

c) The result is obviously valid (except for the uniqueness part) for *any full unimodal family*, not just for the quadratic one.

EXERCISE 37.50. *Specialize the discussion of tuning from* Exercise 37.42 *to the real case. How does tuning act on the kneading sequences?*

EXERCISE 37.51. *Let us consider a renormalization window centered at* $c_\circ \in \overline{\mathcal{V}}_\circ^{\mathbb{R}}$*, and let* Kn_0 *be its kneading sequence. For* $c \in \overline{\mathcal{V}}_\circ^{\mathbb{R}}$*, let* $g = Rf_c$ *be the corresponing renormalization, and let* Kn_g *be its kneading sequence.*

(i) *Describe the kneading sequence* Kn_c *of* f_c *in terms of* $\mathrm{Kn}_\circ$ *and* Kn_g*.*

(ii) *Describe the kneading model for* f_c *in terms of the model for* $f_\circ$ *and* g*.*

Under the above circumstances, Kn_c is called the *tuning* of $\mathrm{Kn}_\circ$ by Kn_g,

$$\mathrm{Kn}_c = \mathrm{Kn}_\circ * \mathrm{Kn}_g.$$

If a superattracting kneading sequence Kn_c is indecomposible in such a way (with Kn_g being necessarily superattracting as well), it is called *prime*.

Adapting further the above complex discussion to the real case, we see that the *full real renormalization combinatorics* can be recorded by a sequence of real superattracting parameters, $(c^{[0]}, c^{[1]}, \dots)$ (such that $c^{[n+1]}$ belongs to the real renormalization window centered at $c^{[n]}$), or by the sequence of the corresponding superattracting kneading sequences $(\mathrm{Kn}^{[0]}, \mathrm{Kn}^{[1]}, \dots)$ (such that $\mathrm{Kn}^{[n+1]}$ is a tuning of $\mathrm{Kn}^{[n]}$ by some *relative* prime superattracting kneading sequence Kn_n). (Recall that kneading sequences are real counterparts of Hubbard trees.) Moreover, the full renormalization combinatorics can be described as the string of relative prime superattracting kneading sequences Kn_n recording types of the consecutive renormalizations $g_{n+1} = R_n g_n$.

We are now ready to give a real counterpart of Remark 37.43:

COROLLARY 37.52. *Any sequence* $(\mathrm{Kn}_n)_{n=0}^{\infty}$ *of admissible prime superattracting kneading sequences is realizable as a full string of relative renormalizaton combinatorics for some infinitely renormalizable quadratic polynomials* f_c*,* $c \in \mathcal{M}_{\mathbb{R}}$*.*

PROOF. Proposition 37.48 provides us with a full unimodal family $R_1 f_c$ over a renormalization window $\overline{\mathcal{V}}_1^{\mathbb{R}}$ corresponding to the first kneading sequence, Kn_1. Inside this family, we find (see Remark 37.49c) a second window, $\overline{\mathcal{V}}_2^{\mathbb{R}}$, corresponding to the relative kneading sequence Kn_2, with a full family of renormalizations $R_2 \circ R_1$. Proceeding indiuctively, we obtain a nest of renormalization windows $\overline{\mathcal{V}}_n^{\mathbb{R}}$, $n \in \mathbb{N}$, corresponding to finite strings of the kneading sequences under consideration. Their intersection provides us with a desired infinitely renormalizable polynomial. ☐

EXERCISE 37.53. (i) *What is the keanding sequence of the classical Feigenbaum map corresponding to the doubling renormalization?*

(ii) *Generalize it to an arbitrary real Feigenbaum map.*

38. Combinatorial Rigidity, MLC, and Density of Hyperbolicity

In this section we will formulate two equivalent versions of the main open problem in the field: the *MLC Conjecture* (on the local connectivity of the Mandelbrot set) and the *Combinatorial Rigidity Conjecture*. It would imply the *Fatou Conjecture* on the density of hyperbolicity. *Real* counterparts of these Conjectures have been established; they will be among main themes of vol. III.

In this section we will overview various relations, reductions, and advances in these problems.

38.1. Combinatorial classes. Given a point $c_\circ \in \mathcal{M}$, the *combinatorial class* $\mathcal{C}(c_\circ) \subset \mathcal{M}$ is the equivalence class of quadratic polynomials f_c that are combinatorially equivalent to $f_\circ \equiv f_{c_\circ}$. By definition, it is equivalent to saying that $f_\circ$ and f_c have the same combinatorial laminations, $\mathcal{L}_{\mathrm{com}}(f_c) = \mathcal{L}_{\mathrm{com}}(f_\circ)$, and hence the same combinatorial models (see §32.1.3). Green puzzle pieces (in the sense of §9.1) for the Mandelbrot set are called *parameter puzzle pieces*. Such a piece is called *rational* if its vertical boundary consists of arcs of parameter rays landing at parabolic or Misiurewicz points (which is equivalent to saying that the rays are rational but we have not yet given a complete proof of this fact (see Lemma 37.7, Theorems 37.8, 37.35, and Theorem 47.11 below.)

EXERCISE 38.1. *Assume* $f_\circ$*,* $c_\circ \in \mathcal{M}$*, is periodically repelling. Then:*

(i) *The combinatorial class* $\mathcal{C}(c_\circ)$ *consists of those parameters* c *that cannot be separated from* $c_\circ$ *by a cut-line through a parabolic or preperiodic cut-point;*

(ii) $\mathcal{C}(c_\circ)$ *is closed and* $\partial\mathcal{C}(c_\circ) \subset \partial\mathcal{M}$*;*

(iii) $\mathcal{C}(c_\circ)$ *is the impression of the rational puzzle end* $E_{\mathbb{Q}}(c_\circ)$ *of* $c_\circ$ *(in other words,* $\mathcal{C}(c_\circ)$ *is the intersection of all rational puzzle pieces containing* $c_\circ$*).*

Exercise 38.2. *The combinatorial class* $\mathcal{C}(c_\circ)$ *of an attracting parameter* $c_\circ$ *is the corresponding hyperbolic component* $\Delta_\circ$ *together with its root* $\tau_\circ$ *and all irrational parameters on the boundary.*[8]

Lemma 38.3. *Assume* $f_\circ$, $c_\circ \in \mathcal{M}$, *is periodically repelling. Then:*

(i) *For any parameter ray* $\mathcal{R}^\theta_{\mathrm{par}}$ *accumulating on some point of* $\mathcal{C}(c_\circ)$, *i.e.,*

$$\omega(\mathcal{R}^\theta_{\mathrm{par}}) \cap \mathcal{C}(c_\circ) \neq \emptyset,$$

the lamination $\mathcal{L}_\theta$ (see §32.5.1) *is equal to the combinatorial lamination* $\mathcal{L}_{\mathrm{com}}(c_\circ)$.

(ii) *If* $c_\circ$ *is not Misiurewicz then there exist at most two parameter rays converging to* $E_{\mathbb{Q}}(c_\circ)$.

Proof. Since the odd rational rays land at parabolic parameters (Proposition 37.9), $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q}_{\mathrm{odd}})/\mathbb{Z}$. Let us show that the diameter $\gamma_\theta = (e(\theta/2), -e(\theta/2))$ is unlinked with all periodic ray portraits of $f_\circ$.

By definition, periodic ray portraits are the same over the combinatorial class $\mathcal{C}(f_\circ)$, so we can replace $f_\circ$ with any other map in $\mathcal{C}(f_\circ)$. Since $\mathcal{R}^\theta_{\mathrm{par}}$ accumulates on some point of $\mathcal{C}(f_\circ)$, we can assume that $c_\circ \in \omega(\mathcal{R}^\theta_{\mathrm{par}})$. Since $f_\circ$ is periodically repelling, the Stability Lemma ensures that any periodic ray portrait $\Theta \subset \mathbb{Q}_{\mathrm{odd}}/\mathbb{Z}$ of $f_\circ$ is robust under perturbations. But $c_\circ$ can be perturbed to a parameter $c \in \mathcal{R}^\theta_{\mathrm{par}}$, and we know that periodic ray portraits for such a parameter are unlinked with γ_θ (by Proposition 32.15, using that $\theta \notin \mathbb{Q}_{\mathrm{odd}}/\mathbb{Z}$). It follows that Θ is unlinked with γ_θ. The conclusions follow from Theorem 32.10. $\square$

38.2. Main Conjectures. A map $f_\circ$ (and the corresponding parameter $c_\circ \in \mathcal{M}$) is called *combinatorially rigid* if $\mathcal{C}(c_\circ) = \{c_\circ\}$.

Combinatorial Rigidity Conjecture. *Any periodically repelling quadratic polynomial* f_c *is combinatorially rigid.*

MLC Conjecture. *Mandelbrot set is locally connected.*

Lemma 38.4. *Combinatorial rigidity of a periodically repelling parameter* $c_\circ \in \mathcal{M}$ *implies weak local connectivity of* $\mathcal{M}$ *at* $c_\circ$.

Proof. Assume $c_\circ$ is rigid. Then By Lemma 38.1 (iii), the rational puzzle pieces around $c_\circ$ shrink. By Corollary 9.9, this implies the desired. $\square$

Exercise 38.5. *A more general assertion holds: any prime-end impression* $I(E)$ *for* $\mathcal{M}$ *that intersects a combinatorial class* $\mathcal{C}(c_\circ)$ *is contained in* $\mathcal{C}(c_\circ)$.

We will see later (Theorem 47.13) that the above two conjectures are actually equivalent. Moreover, they imply the Fatou Conjecture:

---

[8] The latter shows that the notion of combinatorial class is not so meaningful for attracting and neutral parameters. To define it in terms of the critical laminations $\mathcal{L}_\theta$ would make a better sense.

Lemma 38.6. *The Combinatorial Rigidity Conjecture implies the Fatou Conjecture on the density of hyperbolicity.*

Proof. If the Fatou Conjecture fails, then there is a queer component Δ of $\operatorname{int} \mathcal{M}$. By Theorem 36.2, any $c \in \Delta$ is $\mathcal{J}$-stable and hence not rigid. $\square$

Let us finish with one measure-theoretic conjecture:

ABM Conjecture. The boundary of the Mandelbrot set has zero area:

$$\operatorname{area} \partial \mathcal{M} = 0.$$

We will give below the reason why it is conceivable (see §47.10).

38.3. Density of Real Hyperbolicity. The *real combinatorial class* of a point $c_\circ \in \mathcal{M}_\mathbb{R}$ is naturally defined as $\mathcal{C}_\mathbb{R}(c_\circ) := \mathcal{C}(c_\circ) \cap \mathbb{R}$. According to Proposition 32.30, it comprises real parameters with the same kneading data.

The *Real Combinatorial Rigidity Conjecture* asserted that all real periodically repelling combinatorial classes are singletons. A *Real Fatou Conjecture* asserted that hyperbolic parameters are dense in $\mathcal{M}_\mathbb{R}$. Both assertions have been confirmed:

Density of Real Hyperbolicity. *The real combinatorial class of any peridically repelling parameter $c \in \mathcal{M}_\mathbb{R}$ is a singleton. Real hyperbolic maps are dense in $\mathcal{M}_\mathbb{R}$.*

Later in this section (§38.8) we will show that these two assertions are equivalent. The whole Theorem will be proved in vol III.

38.4. Thurston equivalence. Let us consider two quadratic polynomials, $f = f_c$ and $\tilde{f} = f_{\tilde{c}}$, with post-valuable sets $\overline{\mathcal{P}} \equiv \operatorname{cl} \mathcal{P}$ and $\operatorname{cl} \tilde{\mathcal{P}}$. They are called *Thurston equivalent* if there exists a homeomorphism $h : (\mathbb{C}, \operatorname{cl} \mathcal{P}) \to (\mathbb{C}, \operatorname{cl} \tilde{\mathcal{P}})$ which is a conjugacy on the post-valuable sets and that can be lifted (via f and $\tilde{f}$) to a homeomorphism $h_1 : (\mathbb{C}, \operatorname{cl} \mathcal{P}) \to (\mathbb{C}, \operatorname{cl} \tilde{\mathcal{P}})$ homotopic to h rel $\overline{\mathcal{P}}$. So, we have a commjutative diagram

$$\begin{array}{ccc}
(\mathbb{C}, \operatorname{cl} \mathcal{P}) & \xrightarrow{H} & (\mathbb{C}, \operatorname{cl} \tilde{\mathcal{P}}) \\
f \downarrow & & \downarrow \tilde{f} \\
(\mathbb{C}, \operatorname{cl} \mathcal{P}) & \xrightarrow[h]{} & (\mathbb{C}, \operatorname{cl} \tilde{\mathcal{P}})
\end{array}$$

with h and h_1 homotopic rel $\overline{\mathcal{P}}$.

A couple of remarks are due. First, under our circumstances, the Lifting Criterion implies that a homeomorphism $h : \mathbb{C} \to \mathbb{C}$ is liftable if and only if $h(c) = \tilde{c}$. If so, there are two lifts determined by whether $h_1(c) = \tilde{c}$ or $h_1(c) = -\tilde{c}$.

Next, since $h : \operatorname{cl} \mathcal{P} \to \operatorname{cl} \tilde{\mathcal{P}}$ is a conjugacy, we have $h(c) = \tilde{c}$, hence h is liftable in two ways. The above definition requires that if the lift is selected so that $h_1(c) = \tilde{c}$, then $h_1 | \overline{\mathcal{P}} = h | \overline{\mathcal{P}}$, and h_1 is homotopic to h rel $\overline{\mathcal{P}}$.

Finally, we can specify an extra regularity (qc, smooth, etc.) of a Thurston equivalence h.

38.5. Pullback Argument. Here we will introduce a very useful tool that allows one to promote a qc map equivariant on the postcritical set to a global qc conjugacy.

LEMMA 38.7. *If two quadratic polynomials $f = f_c$ and $\tilde{f} = f_{\tilde{c}}$ are Thurston $K-$qc equivalent, then the Böttcher conjugacy $H : \mathcal{D}(\infty) \to \tilde{\mathcal{D}}(\infty)$ admits a $K-$qc extension to the whole complex plane which is a conjugacy on the postcritical sets (and is automatically a conjugacy on the Julia sets).*

PROOF. For a big level $t > 0$, let us consider the super- and sup-potential domains $\Omega(t)$ and $\Sigma(t)$ for f (see §23.5.4). As usual, the corresponding objects for $\tilde{f}$ are marked with tilde.

Take a quasidisk Δ containing $\overline{\mathcal{P}}$, let $\tilde{\Delta} = h(\Delta)$, and select an equipotential level $t > 0$ so that $\Delta \Subset \Sigma(t)$ and $\tilde{\Delta} \Subset \tilde{\Sigma}(t)$. We can modify h on $\mathbb{C} \smallsetminus \Delta$ so that on $\Omega(t)$ it becomes equal to the Böttcher conjugacy $H : \Omega(t) \to \tilde{\Omega}(t)$ (by means of the quasiconformal interpolation through the annulus $\Sigma(t) \smallsetminus \Delta$, see Lemma 15.19). Let h' be the modified map, and let h'_1 be its $(f, \tilde{f})-$lift coinciding with h_1 on $f^{-1}(\Delta) \supset f^{-1}(\overline{\mathcal{P}}) \supset \overline{\mathcal{P}}$. Since $h' | \Omega(t) = H$, its lift $h'_1 | \Omega(t/2)$ is equal to either H or $-H$. In the latter case, we can twist h on the annulus $\Omega(t) \smallsetminus \Delta$ rel its boundary (i.e., compose h with the twist τ from §1.38) to get the correct lift. Then h'_1 becomes homotopic to h' rel $\overline{\mathcal{P}} \cup \Omega(t)$.

So, we can assume in the first place that $h | \Omega(t) = H | \Omega(t)$, $h_1 | \Omega(t/2) = H | \Omega(t/2)$, and h_1 is homotopic to h rel $\overline{\mathcal{P}} \cup \Omega(t)$ (and in particular, h_1 coincides with h on $\overline{\mathcal{P}} \cup \Omega(t)$). Moreover, since holomorphic lifts preserve dilatation, $\operatorname{Dil} h_1 = \operatorname{Dil} h =: K$.

By the Lift Homotopy Theorem, h_1 admits a lift h_2 homotopic to h_1 (which is a lift of h) rel $f^{-1}(\overline{\mathcal{P}}) \cup \Omega(t/2) \supset \overline{\mathcal{P}} \cup \Omega)t/2)$. In particular, $h_2 | f^{-1}(\overline{\mathcal{P}}) = h_1 | f^{-1}(\overline{\mathcal{P}})$. Moreover: h_2 is globally K-qc and conformal on $\Omega(t/4)$.

Repeating this lifting procedure we obtain a sequence of K-qc homeomorphisms $h_n : (\mathbb{C}, f^{-n}(\operatorname{cl} \mathcal{P})) \to (\mathbb{C}, \tilde{f}^{-n}(\operatorname{cl} \tilde{\mathcal{P}}))$ in the same homotopy class rel $\overline{\mathcal{P}}$ and such that h_n coincides with the Böttcher conjugacy on $D(t/2^n)$. In particular, these maps are normalized at any two points of $\overline{\mathcal{P}}$. (We can assume that $|\overline{\mathcal{P}}| \geq 2$ since in the case when $\overline{\mathcal{P}}$ is a singleton, $c = \tilde{c} = 0$, and there is nothing to prove.)

By compactness of the space of normalized K-qc maps, there exists a subsequence $h_{n(k)}$ uniformly converging to a K-qc map coinciding with h on $\overline{\mathcal{P}}$ and coinciding with the Böttcher conjugacy on the whole basin of infinity. □

COROLLARY 38.8. *Let $f = f_c$ and $\tilde{f} = f_{\tilde{c}}$ be two quadratic polynomials with nowhere dense filled Julia sets. If they are Thurston $K-$qc equivalent then the Böttcher conjugacy $B : \mathcal{D}(\infty) \to \tilde{\mathcal{D}}(\infty)$ admits an extension to a global K-qc conjugacy $\Phi : \mathbb{C} \to \mathbb{C}$ between f and $\tilde{f}$. Moreover, if f does not have invariant line fields on its Julia set (in particular, if* $\operatorname{area} \mathcal{J}(f) = 0$) *then* $c = \tilde{c}$.

PROOF. The main assertion is the direct consequence of the lemma. Moreover, in the absense of invariant line fields on $\mathcal{J}(f)$, we have $\bar{\partial}\, \Phi = 0$ a.e., so by Weyl's Lemma, Φ is conformal on the whole complex plane. Hence it is affine, and the conclusion follows. □

The above results can be easily adapted to the quadratic-like case. Let us consider a ql map $f : U \to V$ with connected Julia set, nicely adjusted domains,

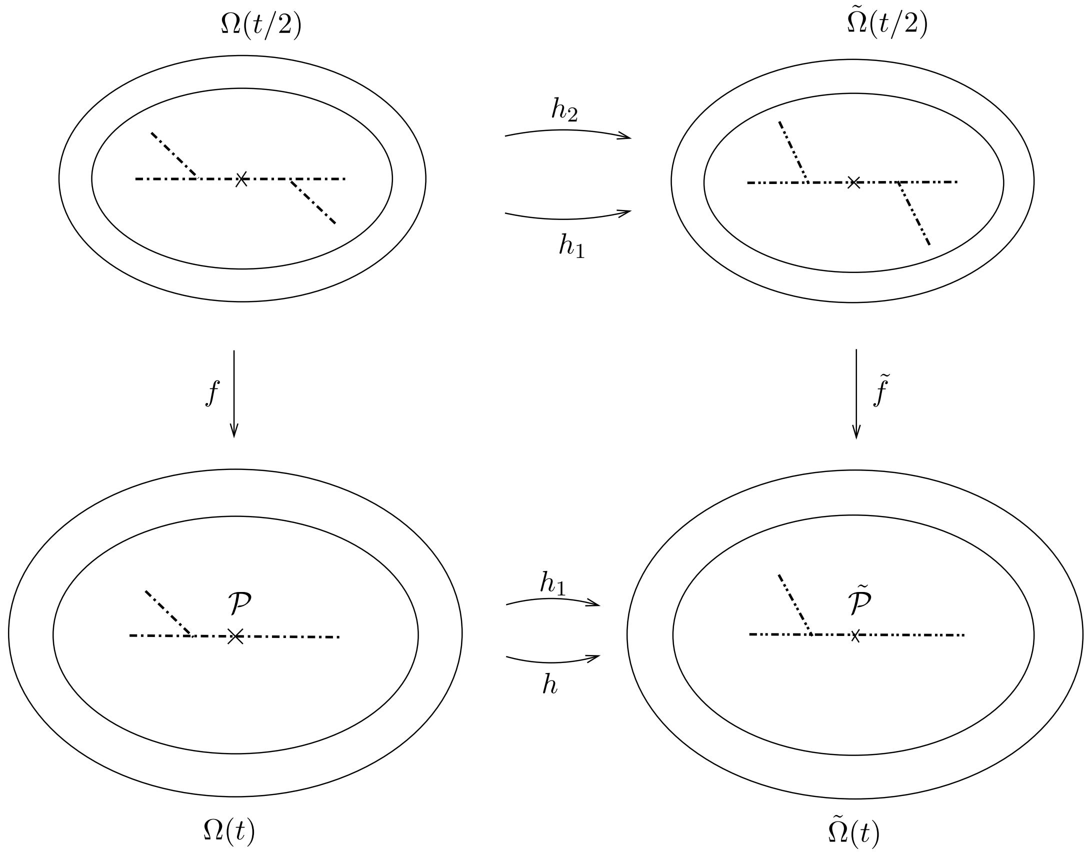

FIGURE 38.1. One step of the Pullback Argument.

the fundamental annulus A, and the critical value v. Let $\tilde{f} : \tilde{U} \to \tilde{V}$ be a similar map, with the corresponding objects marked by tilde. Let us define a Thurston $K-$qc equivalence between f and $\tilde{f}$ as a global K-qc map

$$h : (\mathbb{C}, \mathcal{P}, v, \partial A) \to (\mathbb{C}, \tilde{\mathcal{P}}, \tilde{v}, \partial \tilde{A})$$

which is equivariant on $\mathcal{P}$ and whose $(f, \tilde{f})-$lift $h' : U \to \tilde{U}$ is homotopic to $h|U$ rel $\overline{\mathcal{P}}$.

EXERCISE 38.9. *Let f and $\tilde{f}$ be two K-qc equivalent ql maps as above. Then:*
(i) *h can be twisted so that it is homotopic to its lift h' rel $\overline{\mathcal{P}} \cup \partial A$ (with dilatation K_0 depending only on K and $\operatorname{mod} A$).*
(ii) *There exists a K_0-qc map $H : V \to \tilde{V}$ equivariant on $\overline{\mathcal{P}} \cup (V \smallsetminus \operatorname{int} \mathcal{K})$.*

In particular, in the periodically repelling case, we have:
(iii) *$H : V \to \tilde{V}$ is a conjugacy.*
(iv) *If f does not have invariant line fields on the Julia set, then*

$$\operatorname{Dil}(H|\,V) = \operatorname{Dil}(h|\,A).$$

38.6. Rigidity of superattracting polynomials (revisited). As a first illustration of the Pullback Argument, let us give a different proof of Theorem 35.1.

LEMMA 38.10. *Let $f = f_c$ and $\tilde{f} = f_{\tilde{c}}$ be two superattracting quadratic polynomials. If they are Thurston equivalent then $c = \tilde{c}$.*

PROOF. Since the postcritical set is finite, the Thurston equivalence can be assumed smooth, and hence qc, on the whole Riemann sphere. By the Pullback Argument, the Böttcher conjugacy between f and $\tilde{f}$ extends continuously to the Julia sets. Since the attracting cycles of our maps have the same multipliers (equal to 0), the conclusion follows from Theorem 35.12. □

REMARK 38.11. Instead of using Theorem 35.12, one could adjust the Pullback Argument so that it would directly imply the statement. Namely, one can modify the Thurston equivalence so that is becomes a conformal conjugacy near the superattracting cycles (similarly to the adjustment near ∞ carried in the proof of Lemma 38.7). Then the Pullback Argument will turn it into a qc conjugacy which is conformal outside the Julia set. Since the latter has zero area, it is conformal on the whole plane.

LEMMA 38.12. *If two superattracting quadratic polynomials have the same Hubbard trees, then they are Thurston equivalent.*

PROOF. Let us partition the plane by the rays $\mathcal{R}_i$ landing at the marked points of the Hubbard trees, and let $\mathcal{R} = \bigcup \mathcal{R}_i$. Then the conjugacy $h_{\mathcal{T}} : \mathcal{T} \to \tilde{\mathcal{T}}$ can be extended to the whole plane so that it is the Böttcher conjugacy on $\mathcal{R}$. It further lifts to a homeomorphism

$$h_1 : (\mathbb{C}, f^{-1}(\mathcal{T} \cup \mathcal{R}) \to (\mathbb{C}, \tilde{f}^{-1}(\tilde{\mathcal{T}} \cup \tilde{\mathcal{R}})$$

which is Böttcher on $f^{-1}(\mathcal{R})$. As $h_1 = h$ on $\mathcal{R}$, these two maps are homotopic rel $\mathcal{R}$, all the more rel $\mathcal{O}$. It gives us a desired Thurston equivalence. □

Putting this together with Lemma 38.10, we obtain:

COROLLARY 38.13. *If two superattracting quadratic polynomials, f_c and $f_{\tilde{c}}$, have the same Hubbard tree then $c = \tilde{c}$.*

38.7. Combinatorial vs Thurston qc rigidity. Let us say that a parameter $c \in \mathcal{M}$ is *Thurston qc rigid* if it is not Thurston qc equivalent to any other parameter c'.

LEMMA 38.14. *Any boundary parameter $c \in \partial\mathcal{C}(c_\circ)$ is Thurston qc rigid.*

PROOF. Otherwise, by Corollary 38.8 c would have a non-trivial qc class. Then by Theorem 36.18, c would belong to either a punctured hyperbolic component or to a queer component of $\mathcal{M}$. In either case, c would belong to the interior of its combinatorial class, contradicting the assumption. □

COROLLARY 38.15. *Assume $f_{c_\circ}$ is periodically repelling. If any two parameters $c, c' \in \mathcal{C}(c_\circ)$ are Thurston qc equivalent then $c_\circ$ is combinatorially rigid.*

PROOF. For a periodically repelling parameter, the combinatorial class $\mathcal{C}(c_\circ)$ is closed, so it contains a boundary parameter. □

REMARK 38.16. The above proof can be presented in a a more direct fashion as an *open-closed* argument. Indeed, $\mathcal{C}(c_\circ)$ is closed while $\mathcal{QC}(c_\circ)$ is open unless it is a singleton (by the *Beltrami deformation* argument from the proof of Theorem 36.18). So, if $\mathcal{C}(c_\circ) = \mathcal{QC}(c_\circ)$ then $\mathcal{C}(c_\circ)$ must be a singleton.

EXERCISE 38.17. *Show that the "periodically repelling assumption" can be dropped in the above Corollary.*

Due to Corollary 38.15, the Combinatorial Rigidity Conjecture is reduced to the following one:

CONJECTURE 38.18. *If two periodically repelling quadratic polynomials are combinatorially equivalent then they are Thurston qc equivalent.*

38.8. Real Combinatorial Rigidity and Density of Hyperbolicity. Let us start with a criterion for the Density of Real Hyperbolicity:

LEMMA 38.19. *The Real Combinatorial Rigidity is equivalent to the Density of Hyperbolicity in* $\mathcal{M}_\mathbb{R} = [-2.1/4]$.

PROOF. Assume the Real Combinatorial Rigidity. Take any non-hyperbolic parameter $c_\circ \in \mathcal{M}_\mathbb{R}$. If it is parabolic then it can obviously perturbed to a hyperbolic one. Otherwise $\mathcal{C}_\mathbb{R}(c_\circ) = \{c_\circ\}$ by assumption, so there are nearby parameters c with different kneading sequences. To pass from $\mathrm{Kn}(c_\circ)$ to $\mathrm{Kn}(c)$, one has to pass through a superattracting parameter.

Vice versa, if the Real Combinatorial Rigidity fails, then there is a non-singleton non-hyperbolic real combinatorial class $\mathcal{C}_\mathbb{R}$. By Theorem 37.33, $\mathcal{C}_\mathbb{R}$ is an interval. Non-density of hyperbolicity follows. $\square$

Let us now consider two real periodically repelling quadratic polynomials, $f \equiv f_c$ and $\tilde{f} \equiv f_{\tilde{c}}$, $c, \tilde{c} \in [-2, 1/4]$, with postcritical sets $\mathrm{cl}\,\mathcal{P}_0$ and $\mathrm{cl}\,\tilde{\mathcal{P}}_0$. According to Proposition 32.16, if they are combinatorially equivalent then they are topologically conjugate on the real line. All the more, there is a homeomorphism $h : (\mathbb{R}, \mathrm{cl}\,\mathcal{P}_0) \to (\mathbb{R}, \mathrm{cl}\,\tilde{\mathcal{P}}_0)$ that restricts to a conjugacy $(\mathrm{cl}\,\mathcal{P}_0, 0) \to (\mathrm{cl}\,\tilde{\mathcal{P}}_0, 0)$. If h can be selected quasisymmetric, we say that the maps are *combinatorially qs equivalent*.

LEMMA 38.20. *Let f and $\tilde{f}$ be two periodically repelling quadratic polynomials.*

(i) *If f and $\tilde{f}$ are combinatorially equivalent then they are Thurston equivalent.*

(ii) *If the Thurston equivalence $h : \mathbb{R} \to \mathbb{R}$ can be selected κ-quasisymmetric then f and $\tilde{f}$ are Thurston $K(\kappa)$-qc equivalent on $\mathbb{C}$.*

PROOF. As always, we can assume that 0 is the minimum for both f and $\tilde{f}$, so our combinatorial equivalence $h : \mathbb{R} \to \mathbb{R}$ is orientation preserving.

(i) First, let us extend h to a homeomorphism $\mathbb{C} \to \mathbb{C}$ (in an arbitrary way). Since $h[v, +\infty) = [\tilde{v}, +\infty)$, h can be lifted to a $\mathbb{C}$-homeomorphism in two ways, $\pm h_1 : (\mathbb{C}, \mathbb{R}, 0) \to (\mathbb{C}, \mathbb{R}, 0)$. Let h_1 be the lift preserving orientation of $\mathbb{R}$.

Then $h_1 \,|\, \overline{\mathcal{P}} = h \,|\, \overline{\mathcal{P}}$. Indeed, since h is a conjugacy on $\overline{\mathcal{P}}_0$, $h(x) = \pm h_1(x)$ for any $x \in \overline{\mathcal{P}}$ (where the sign may *a priori* depend on x). However, since both h and h_1 are orientation preserving on $\mathbb{R}$ and $h(0) = 0 = h_1(0)$, we have:

$$\operatorname{sign} h(x) = \operatorname{sign} x = \operatorname{sign} h_1(x),$$

and hence $h(x) = h_1(x)$.

By Exercise 1.37, h_1 is $\mathbb{R}$-symmetrically isotopic rel $\overline{\mathcal{P}}$ to some homeomorphism $h' : \mathbb{C} \to \mathbb{C}$ coinciding with h on $\mathbb{R}$. By the Alexander Trick, h' is isotopic to h rel $\mathbb{R}$, so the pair (h, h_1) provides us with a Thurston equivalence between f and $\tilde{f}$.

(ii) If h is κ-qs then it admits a K-qs extension to $\mathbb{C}$, where $K = K(\kappa)$. Its lift h_1 by means of holomorphic maps $(f,\ \tilde{f})$ will be K-qc as well. Thus, we obtain a a Thurston K-qc equivalence. ☐

Corollary 38.21. *Let f and $\tilde{f}$ be two real periodically repelling quadratic polynomials* (*with connected Julia sets*)*. If they are combinatorially κ-qs equivalent then the are $K(\kappa)$-qc conjugate* (*with the conjugacy being Böttcher on the basin of ∞*)*.*

Proof. Under these circumstances, f and $\tilde{f}$ have nowhere dense filled Julia sets, so Corollary 38.8 applies. ☐

Lemma 38.22. *Let $\mathcal{C}_{\mathbb{R}} \subset \mathcal{M}_{\mathbb{R}}$ be a periodically repelling real combinatorial class of quadratic polynomials. If for any two $c, \tilde{c} \in \mathcal{C}_{\mathbb{R}}$, the polynomials f_c and $f_{\tilde{c}}$ are combinatorially qs equivalent then $\mathcal{C}_{\mathbb{R}}$ is a singleton.*

Proof. Assume otherwise. By Corollary 38.21, any two polynomials f_c and $f_{\tilde{c}}$, $(c, \tilde{c} \in \mathcal{C}_{\mathbb{R}})$ are qc conjugate. So, they are not qc rigid.

As peridically repelling combinatorial classes are closed, we can take a boundary parameter $c \in \partial\mathcal{C}_{\mathbb{R}} \subset \partial\mathcal{M}$ (see Exercise 38.1(ii)). By the last item of Theorem 36.18, c is qc rigid, and we have arrived at a contradiction. ☐

38.9. Rigidity for real Feigenbaum maps. According to Proposition 32.30 the renormalization combinatorics of a real Feigenbaum map is determined by its kneading sequence (described in Exercises 37.50 and 37.53).

Theorem 38.23. *There is a unique real Feigenbaum map with a given admissible combinatorics.*

Proof. Corollary 37.52 secures existence of such a map. Let $\mathcal{C}_{\mathbb{R}}$ be the corresponding real combinatorial class. By Theorem 30.49, the postcritical sets of all maps in $\mathcal{C}_{\mathbb{R}}$ are Cantor sets with bounded geometry. It follows from Exercise 15.8 that all of them are combinatorially qs equivalent, and the conclusion follows from Lemma 38.22. ☐

38.10. Notes. The MLC Conjecture was formulated by Douady and Hubbard [**DH2**]. It was motivated by a perspective to obtain a precise topological model for $\mathcal{M}$. The potential model was described by Thurston, in terms of a geodesic lamination [**Th1**], and Douady [**D3**].

The Fatou conjecture was reduced to MLC by Douady and Hubbard [**DH2**] (in a different way from presentred above).

The Combinatorial Rigidity Conjecture is a deep dynamical analogue of the *Mostow Rigidity phenomenon* in 3D hyperbolic geometry. In fact, the conjecture has a precise geometric counterpart, known as the *Thurston Ending Laminations Conjecture*, now is a theorem, due to Brock, Canary and Minsky [**BCM**].

The equivalence between the MLC and the Combinatorial Rigidity Conjecture was a folklore in the 1990s. The first detailed written account was given by Schleicher [**Sc4**].

According to Dennis Sullivan's oral communication, the Pullback Argument was designed by Thurston.

Rigidity Theorem 38.23 for real Feigenbaum maps is due to Sullivan (see [**MvS**]).

39. Thurston Realization Theorem

39.1. Statement. Let $g : \mathbb{R}^2 \to \mathbb{R}^2$ be a topological double branched covering of the plane with the critical point at 0. We say that it is *critically periodic* if $g^p(0) = 0$ for some $p \geq 1$. (As we know, quadratic polynomials with this property are called "superattracting", but this term could be misleading in the topological setting since a periodic critical point can be even repelling.) Let $\mathcal{P} = (g^n(0))_{n=0}^{p-1}$.

Similarly to the actual quadratic maps, we can define *Thurston equivalence* between such maps (see §38.4). We say that a map g in question is *realizable* if there exists a quadratic polynomial f_c in the Thurston class of g. By Lemma 38.10, this realization is unique.

Thurston Realization Theorem. *Any critically periodic topological double branched covering $g : \mathbb{C} \to \mathbb{C}$ is realizable.*

39.2. Proof.

39.2.1. *The Teichmüller and Moduli spaces.* Without loss of generality, we can assume that $\mathbb{R}^2 \approx \mathbb{C}$ is endowed with a qc structure and that g is quasiregular. All the conformal structures on $\mathbb{R}^2$ (or, on the punctured $\mathbb{R}^2$) will be assumed to have a bounded dilatation.

Let us consider the *Teichmüller space* $\mathcal{T} = \mathcal{T}_g \approx \mathcal{T}_{p+1}$ of the punctured plane $\mathbb{R}^2 \smallsetminus \mathcal{P}$. By definition, it is the space of conformal structures μ on $\mathbb{R}^2 \smallsetminus \mathcal{P}$ up to homotopy. More precisely, let $h_\mu : (\mathbb{R}^2, \mathcal{P}) \to (\mathbb{C}, \mathcal{O}_\mu)$ be the solution of the Beltrami equation for the structure μ, where $\mathcal{O}_\mu = h_\mu(\mathcal{P})$. Two structures μ and μ' are equivalent if there is a complex affine transformation $\phi : (\mathbb{C}, \mathcal{O}_\mu) \to (\mathbb{C}, \mathcal{O}_{\mu'})$ such that $\phi \circ h_\mu$ is homotopic to $h_{\mu'}$ rel $\mathcal{P}$. A class $\tau = [h_\mu]$ of equivalent maps represents a point of $\mathcal{T}$.

The *Moduli space* $\mathcal{M} = \mathcal{M}_g$ is the space of embeddings $\mathcal{P} \to \mathbb{C}$ up to affine transformation. The natural projection $\pi : \mathcal{T} \to \mathcal{M}$ associates to a class $[h_\mu] \in \mathcal{T}$ the embedding $h_\mu | \mathcal{P} \in \mathcal{M}$ (up to affine transformation).

39.2.2. *Pullback operator.* By an affine conjugacy and homotopy rel $\mathcal{P}$, we can normalize g so that $c_0 = 0$ is its critical point and $g(z) = z^2$ near ∞. Recall that $0_k = g^k(0)$, $k = 0, \dots, p-1$.

Let us now define the pullback operator $g^* : \mathcal{T} \to \mathcal{T}$ induced by the pullback $\mu \mapsto \mu' = g^*(\mu)$ of the complex structures on $\mathbb{C}$. More precisely, let a point $\tau \in \mathcal{T}$ be represented by a homeomorphism $h : (\mathbb{C}, \mathcal{P}, 0) \to (\mathbb{C}, \mathcal{O}_\tau, 0)$. Then $\mu = h^*(\sigma)$ and $\mu' = (h \circ g)^*(\sigma)$. Let $h' : (\mathbb{C}, \mathcal{P}, 0) \to (\mathbb{C}, h'(\mathcal{P}, 0))$ be the solution of the Beltrami equation for the conformal structure μ'. Then h' represents the point $\tau' = g^*(\tau)$. If $\tilde{h}$ is homotopic to h rel $\mathcal{P}$ then the Lift Homotopy Theorem ensures that $\tilde{h}'$ is homotopic to $\tilde{h}$ rel $\mathcal{P}$, so the operator is well defined. In particular, $h'(\mathcal{P})$ (up to rescaling) depends only on τ', and it can be called $\mathcal{O}_{\tau'}$.

Let $h(\mathcal{P}) = (z_0 = 0, z_1, \dots, z_{p-1})$, $h'(\mathcal{P}) = (z'_0 = 0, z'_1, \dots, z'_{p-1})$, where

$$z_i = h(0_i), \quad z'_i = h'(0_i).$$

Composition $h \circ g \circ (h')^{-1}$ is a holomorphic double branched covering

$$f : (\mathbb{C}, \mathcal{O}_{\tau'}) \to (\mathbb{C}, \mathcal{O}_{\tau}),$$

so it is a quadratic polynomial:

$$\begin{array}{ccc} (\mathbb{R}^2, \mathcal{P}, \tau') & \xrightarrow[h']{} & (\mathbb{C}, \mathcal{O}_{\tau'}, \sigma) \\ g \downarrow & & \downarrow f \\ (\mathbb{R}^2, \mathcal{P}, \tau) & \xrightarrow[h]{} & (\mathbb{C}, \mathcal{O}_{\tau}, \sigma) \end{array} \tag{39.1}$$

Notice, however, that f does not have a dynamical meaning as $h'(\mathcal{P}) \neq h(\mathcal{P})$. Moreover, we have two independent scaling factors to normalize the maps h and h', and there are several useful ways to do so. For instance,

N1: Let $z_1 = 1$ while $z'_{p-1} = i$; then $f(z) = z^2 + 1$.

N2: Let $z_1 = z'_1 = 1$; then $f(z) = (z_2 - 1)\, z^2 + 1$.

39.2.3. *Ambiguity in the Moduli space.* The operator g^* does not descend to the moduli space $\mathcal{M}$: the Riemann surface $(\mathbb{C}, \mathcal{O}_{\tau'})$ in diagram (39.1) is not uniquely determined by $(\mathbb{C}, \mathcal{O}_{\tau})$. However, the ambiguity is bounded:

LEMMA 39.1. *For a given Riemann surface $(\mathbb{C}, \mathcal{O}_{\tau})$, there exists a bounded (in terms of $p = |\mathcal{O}|$) number of Riemann surfaces $(\hat{\mathbb{C}}, \mathcal{O}_{\tau'})$. Moreover, if $(\mathbb{C}, \mathcal{O}_{\tau})$ belongs to a compact subset $\mathcal{K}$ of $\mathcal{M}$ then $(\mathbb{C}, \mathcal{O}_{\tau'})$ belongs to a compact subset $\mathcal{K}' \Subset \mathcal{M}$ as well.*

PROOF. Let us use normalization $N1$ for h and h', so $f(z) = z^2 + 1$. Then the set $f^{-1}(\mathcal{O}_{\tau})$ is uniquely defined by $\mathcal{O}_{\tau}$:

$$z'_i = \pm\sqrt{z_{i+1} - 1}. \tag{39.2}$$

Since $\mathcal{O}_{\tau'} \subset f^{-1}(\mathcal{O}_{\tau})$, we have only finitely many (at most 2^{p-2}) options for $\mathcal{O}_{\tau'}$. The choice of $\pm$ signs in (39.2) is determined by the marking $z'_i = h'(0_i)$, and formulas (39.2) express the pullback operator g^* in the local coordinates:

$$g^* : (z_2, \ldots, z_{p-1}) \mapsto (z'_1, \ldots, z'_{p-2}).$$

Let $\mathcal{Z}_{\tau} = \mathcal{O}_{\tau} \cup \{\infty\}$. If $(\mathbb{C}, \mathcal{O}_{\tau}) \in \mathcal{K}$ then the points of $\mathcal{Z}_{\tau}$ are ε-separated in the spherical metric for some $\varepsilon = \varepsilon(\mathcal{K}) > 0$ (see Lemma 18.10). But then the points of $\mathcal{Z}_{\tau'}$ given by (39.2) are ε'-separated for some $\varepsilon' > 0$ depending only on ε, and the conclusion follows. $\square$

39.2.4. *Fixed points of g^*.*

PROPOSITION 39.2. *A branched covering $g : (\mathbb{R}^2, \mathcal{P}) \to (\mathbb{R}^2, \mathcal{P})$ is realizable if and only if the pullback operator $g^* : \mathcal{T} \to \mathcal{T}$ has a fixed point.*

PROOF. If g is realizable then by definition there is a superattracting quadratic polynomial f_c with the postcritical set $\mathcal{O}_c$ and homeomorphisms h and h' homotopic rel $\mathcal{P}$ such that the diagram is valid:

$$\begin{array}{ccc} (\mathbb{R}^2, \mathcal{P}) & \xrightarrow[h']{} & (\mathbb{C}, \mathcal{O}_c) \\ g \downarrow & & \downarrow f_c \\ (\mathbb{R}^2, \mathcal{P}) & \xrightarrow[h]{} & (\mathbb{C}, \mathcal{O}_c) \end{array}$$

Comparing it with diagram (39.1) we see that $[h'] = g^*[h]$. But $[h] = [h']$ by definition of a point in $\mathcal{T}$. So, $[h]$ is a fixed point of g^*.

Vice versa, assume that a homeomorphism h in diagram (39.1) represents a fixed point of f^*. Then $[h'] = [h]$, which means by definition that after postcomposing h' with a scaling $z \mapsto \lambda z$ we have $h' \simeq h$ rel $\mathcal{P}$. But then the quadratic polynomial $\lambda^{-1} f$ is Thurston equivalent to g.

□

39.2.5. *Co-differential as a push-forward operator.* Recall from §18.5 that the cotangent space[9] $\mathrm{T}^{\#}_{\tau}\mathcal{T}$ to the Teichmüller space $\mathcal{T}$ is isometric to the space $\mathcal{Q}^1(\mathbb{C} \smallsetminus \mathcal{O}_\tau)$ of integrable meromorphic quadratic differentials. So, the co-differential

$$Dg^*(\tau)^{\#} : \mathrm{T}^{\#}_{\tau'}\mathcal{T} \to \mathrm{T}^{\#}_{\tau}\mathcal{T}$$

can be viewed as an operator

$$Dg^*(\tau)^{\#} : \mathcal{Q}^1(\mathbb{C} \smallsetminus \mathcal{O}_{\tau'}) \to \mathcal{Q}^1(\mathbb{C} \smallsetminus \mathcal{O}_\tau).$$

On the other hand, the quadratic map f from diagram (39.1) induces the push-forward operator between the same spaces:

$$f_* : \mathcal{Q}^1(\mathbb{C} \smallsetminus \mathcal{O}_{\tau'}) \to \mathcal{Q}^1(\mathbb{C} \smallsetminus \mathcal{O}_\tau)$$

(see §2.11.3). It turns out that these two operators are the same:

LEMMA 39.3. *Up to the above isometries, the co-differential of $Dg^*(\tau)^{\#}$ is equal to the push-forward operator f_*.*

PROOF. Let us take a smooth vector field v on $\hat{\mathbb{C}}$ representing a tangent vector to $\mathcal{T}$ at τ. The differential $Dg^*(\tau)$ acts on v as the pullback f^*v. Since the $\bar\partial$-operator behaves naturally under holomorphic pullbacks, we have $\bar\partial(f^*v) = f^*(\bar\partial v)$. Using now duality between the pullback and push-forward (Lemma 2.110), we obtain for any quadratic differential $q \in \mathcal{Q}^1(\hat{\mathbb{C}} \smallsetminus \mathcal{O}_{\tau'})$:

$$< q, f^*v >= \int q \cdot \bar\partial(f^*v) = \int q \cdot f^*(\bar\partial v) = \int f_*q \cdot \bar\partial v =< f_*q, v > .$$

□

39.2.6. *Non-escaping in $\mathcal{M}$ creates the fixed point.* The previous discussion implies that the pullback operator g^* is *uniformly contracting* depending only on the location of the Riemann surface $(\mathbb{C}, \mathcal{O}_\tau)$ in the *moduli space*:

LEMMA 39.4. *For any compact subset $\mathcal{K}$ in $\mathcal{M}$ there exists $\rho = \rho(\mathcal{K}) < 1$ such that $\|Dg^{\#}(\tau)\| \le \rho$ for any $\tau \in \mathcal{T}$ such that $(\mathbb{C}, \mathcal{O}_\tau) \equiv \pi(\tau) \in \mathcal{K}$.*

PROOF. Let us first show that g^* is a strict infinitesimal contraction. Otherwise there is a quadratic differential $q \in \mathcal{Q}^1(\hat{\mathbb{C}} \smallsetminus \mathcal{O}_{\tau'})$ such that $\|f_*q\| = \|q\|$. Then by Lemma 2.109, $f^*(q_1) = 2q$, where $q_1 = f_*q$.

Let $N = \#$ poles $-\#$ zeros of q (and N_1 is the repsective number for q_1). As we know (Exercise 2.105), $N = N_1 = 4$.

On the other hand, if a is a pole (or zero) of q_1 which is not the critical value of f, then both preimages of a under f are poles or zeros (respectively) of q. Hence $N \ge 2N_1 - 2 > N_1$ – contradiction.

[9]We use a non-standard notation $\mathrm{T}^{\#}$ for the cotangent space to avoid confusion with other appearances of the $*$ in our discussion.

To show that the amount of contraction depends only on $\mathcal{K}$, we will use Lemma 39.1. Indeed, it shows that g^* descends to a correspondence of $\mathcal{M}$ of bounded degree. This correspondence is strictly contracting and hence uniformly contracting on compact subsets, impying the desired. □

COROLLARY 39.5. *Let $\tau \in \mathcal{T}$ and $\tau_n = (g^*)^n \tau$. If $\pi(\tau_n)$ is non-escaping in $\mathcal{M}$ (i.e., it stays in some compact subset $\mathcal{K} \subset \mathcal{M}$) then the orbit $\{\tau_n\}$ converges to a fixed point.*

PROOF. In this case, the above lemma implies that $\operatorname{dist}(\tau_n, \tau_{n+1}) \to 0$ exponentially fast. □

39.2.7. *Escaping creates an invariant multicurve.* We will now analyze the situation when the Riemann surfaces $\pi(\tau_n) = \mathbb{C} \setminus \mathcal{O}^n$ escape in the Moduli space. By Lemma 18.11 and discussion of §6.7, this creates a family of canonical annuli A^n_γ on $\mathbb{C} \setminus \mathcal{O}^n$. labeled by the components of a multicurve $\mathcal{G}^n$ on $\mathbb{R}^2 \setminus \mathcal{P}$. Let

$$\mathcal{G}^n(\mu) = \{\gamma \in \mathcal{G}^n : \bmod A^n_\gamma \geq \mu\}.$$

Let us fix some $K \geq 1$ such that the first two marked Riemann surfaces in our orbit, τ and τ', are K-qc equivalent. (The best such K is equal to $\exp \operatorname{dist}_{\mathrm{T}}(\tau_0, \tau_1)$.)

LEMMA 39.6. *We have $\mathcal{G}^{n+1}(\mu) \subset \mathcal{G}^n(\mu/K - 2)$.*

PROOF. Since the operator $g^* : \mathcal{T} \to \mathcal{T}$ is contracting in the Teichmüller metric, we have $\operatorname{dist}_{\mathrm{T}}(\tau_{n+1}, \tau_n) \leq \log K$ for all $n = 0, 1, \ldots$. Then by definition of the Teichmüller distance, there exists a K-qc map $\phi_n : (\mathbb{C}, \mathcal{O}^{n+1}) \to (\mathbb{C}, \mathcal{O}^n)$ such that $\phi_n \circ h_{n+1} \simeq h_n$.

For a curve $\gamma \in \mathcal{G}^{n+1}(\mu)$, let us consider the corresponding canonical annulus $A^{n+1}_\gamma \subset \mathbb{C} \setminus \mathcal{O}^{n+1}$ homotopic to $h_{n+1}(\gamma)$. Then the annulus $\phi_n(A^{n+1}_\gamma) \subset \mathbb{C} \setminus \mathcal{O}^n$ is homotopic to $h_n(\gamma)$ and $\bmod \phi_n(A^{n+1}_\gamma) \geq \mu/K$. Hence $\bmod A^n_\gamma \geq \mu/K - 2$ for the canonical annulus A^n_γ, so $\gamma \in \mathcal{G}^n(\mu/K - 2)$. □

For a Jordan curve γ in $\mathbb{R}^2 \setminus \mathcal{P}$, let us consider the full preimage $g^{-1}(\gamma)$ in $\mathbb{R}^2 \setminus g^{-1}(\mathcal{P})$. It consists of one or two components (depending on whether γ surrounds the critical value or not). Let us now consider each of these components as curves in a bigger surface, $\mathbb{R}^2 \setminus \mathcal{P}$ (recall that $\mathcal{P} \subset g^{-1}(\mathcal{P})$). If some component happens to be peripheral (in particular, trivial), we throw it away. We are left with $k \in [0, 2]$ non-peripheral Jordan curves γ'_i in $\mathbb{R}^2 \setminus \mathcal{P}$. Each of these curves is called a *pullback* of γ.

For a multicurve $\mathcal{G}$ in $\mathbb{R}^2 \setminus \mathcal{P}$, we let $g^*(\mathcal{G})$ be the multicurve comprising the pullbacks of all the components of $\mathcal{G}$.

LEMMA 39.7. *We have: $g^*(\mathcal{G}^n(\mu)) \subset \mathcal{G}^{n+1}(\mu/2 - 2)$.*

PROOF. Let $\gamma \in \mathcal{G}^n(\mu)$, $A \equiv A^n_\gamma \subset \mathbb{C} \setminus \mathcal{O}^n$ be the corresponding canonical annulus, and let $f_n : \mathbb{C} \setminus \mathcal{O}^{n+1} \to \mathbb{C} \setminus \mathcal{O}^n$ be the quadratic polynomial from the diagram (39.1). Then $f_n^{-1}(A)$ consists of one or two symmetric annuli $A' \subset \mathbb{C} \setminus \mathcal{O}^{n+1}$ depending on whether A surrounds the critical value $v = f(0)$ or not. Moreover, $\bmod A' = (1/2) \bmod A$ in the former case, and $\bmod A' = \bmod A$ in the latter. Hence

$$\bmod A^{n+1}_{\gamma'} \geq \mu/2 - 2,$$

where $\gamma' \in \mathcal{G}^{n+1}$ and $h_{n+1}(\gamma')$ is homotopic to A'. The conclusion follows. □

Putting Lemmas 39.6 and 39.7 together, we obtain:

COROLLARY 39.8. *We have:* $g^*(\mathcal{G}^n(\mu)) \subset \mathcal{G}^n(\mu/2K-4)$.

A multicurve $\mathcal{G}$ on $\mathbb{R}^2 \smallsetminus \mathcal{P}$ is called *invariant* if $g^*(\mathcal{G}) \subset \mathcal{G}$. Let

$$M_n = \max\nolimits_{\gamma\in\mathcal{G}^n} \operatorname{mod} A^n_\gamma. \tag{39.3}$$

LEMMA 39.9. *Let $L > 4K$. If $M_n > L^p$ for some n, then for some $\underline{\mu} < L^p$, the multicurve $\mathcal{G}^n(\underline{\mu})$ is invariant.*

PROOF. Let us order monotonically the moduli $\mu^n(\gamma) := \operatorname{mod} A^n_\gamma$, $\gamma \in \mathcal{G}^n$, of the canonical annuli on $\mathbb{C} \smallsetminus \mathcal{O}^n$ (which are bigger than $4L$):

$$\mu^n_1 \geq \mu^n_2 \geq \cdots \geq \mu^n_l > 4L, \quad l = l_n \leq p-2.$$

If $M_n/4L > L^{p-2}$ then $\mu^n_i > L\mu^n_{i+1}$ for some $i \in [1, l]$ (in case $i = l$ we let $\mu^n_{l+1} := 4L$). Let us take the biggest i like this; then $\mu^n_{i+1} \leq L^{p-2}$. Taking any $\underline{\mu} \in (4L, L^p)$ satisfying

$$L\mu^n_{i+1} < \underline{\mu} < \mu^n_i \tag{39.4}$$

we see that $\mathcal{G}^n(\underline{\mu}/L) = \mathcal{G}^n(\underline{\mu})$. By Corollary 39.8, we have:

$$g^*(\mathcal{G}^n(\underline{\mu})) \subset \mathcal{G}^n(\underline{\mu}/2K-4) \subset \mathcal{G}^n(\underline{\mu}/L) = \mathcal{G}^n(\underline{\mu}).$$

□

LEMMA 39.10. *If in* Lemma 39.9, $L > K^p$, *then the multicurves $\mathcal{G}_{n+j}(\underline{\mu})$ are all the same for $j = 0, 1, \ldots, p$.*

PROOF. Lemma 39.6 and inequality (39.4) show that for all $j = 0, 1, \ldots p-1$, the threshold $\underline{\mu}$ is still squeezed in between μ^{n+j}_i and μ^{n+j}_{i+1}, so it selects the same components of the canonical multicurve. □

39.2.8. *Transformation rules.* To simplify notation, we skip the label "n" and mark the objects of level $n+1$ with prime. For instance, $\mathcal{O}^n \equiv \mathcal{O}$, $\mathcal{O}^{n+1} \equiv \mathcal{O}'$, $\mu(\gamma) \equiv \mu^n(\gamma)$, $\mu'(\gamma) \equiv \mu^{n+1}(\gamma)$.

LEMMA 39.11. *Let δ' be a canonical Jordan curve on $\mathbb{C} \smallsetminus f^{-1}(\mathcal{O})$. Then there is a Jordan curve δ on $\mathbb{C} \smallsetminus \mathcal{O}$ such that $\delta' \simeq f^*(\delta)$. Moreover, $d = \deg(f : \delta' \to \delta)$ is equal to 2 or 1 depending on whether δ' surrounds the critical point or not.*

PROOF. Obviously, the canonical multicurve is invariant under the central symmetry $z \mapsto -z$. Hence $-\delta'$ is either essentially disjoint from δ', or is essentially equal to it.

Assume the former. Then by replacing δ' with a homotopic Jordan curve (for instance, with a closed hyperbolic geodesic in $\mathbb{C} \smallsetminus f^{-1}(\mathcal{O})$) we ensure that $-\delta'$ is actually disjoint from δ'. It follows that f maps δ' invectively onto some Jordan curve δ in $\mathbb{C} \smallsetminus \mathcal{O}$.

Let us assume now that $-\delta' \simeq \delta'$. Then we can replace δ' with a central symmetric homotopic curve. Restriction of f to such a curve is equivalent to taking its quotient mod $z \mapsto -z$. So, it is a double covering onto the image, and its image is a Jordan curve in $\mathbb{C} \smallsetminus \mathcal{O}$.

So, δ' is a lift $f^*\delta$ of some Jordan curve $\delta \in \mathbb{C} \smallsetminus \mathcal{O}$. It has degree 2 or 1 over δ depending on whether δ surrounds the ramification point $v = f(0)$ or not. In the former case, δ' surrounds the critical point 0, while in the latter it does not. □

LEMMA 39.12. *Under the circumstances of the previous lemma, we have:*

$$d^{-1}\mu(\delta) - 2 \le \mu'(\delta') \le d^{-1}\mu(\delta) + 2.$$

PROOF. Let us consider the maximal annulus $\mathbf{A} = \mathbf{A}_\delta$ in $\mathbb{C} \smallsetminus \mathcal{O}$ homotopic to δ. It lifts to an annulus $\mathbf{A}'$ in $\mathbb{C} \smallsetminus f^{-1}(\mathcal{O})$ homotopic to δ'. Moreover, the map $f : \mathbf{A}' \to \mathbf{A}$ is a double covering of degree d, so $\operatorname{mod} \mathbf{A}' = d^{-1} \operatorname{mod} \mathbf{A}$. Hence the maximal modulus in the homotopy class of δ' is at least that big, while the canonical modulus (if non-zero) is by two units smaller.

To go in the opposite direction, let us consider the canonical annulus $A' \equiv A_{\delta'}$ in $\mathbb{C} \smallsetminus f^{-1}(\mathcal{O})$. Since the the family of canonical annuli is invariant under the central reflection $z \mapsto -z$, the symmetric annulus $-A'$ is also canonical. Two cases can occur:

a) A' does not surround the critical point. Then by Lemma 39.11, the homotopy classes $[A']$ and $[-A']$ are essentially disjoint. By Lemma 6.39, $-A'$ is disjoint from A', and hence f univalently maps A' onto its image A.

b) A' surrounds the critical point. Then the homotopy classes $[A']$ and $[-A']$ coincide. By the uniqueness of the canonical annulus in a given homotopy class (Theorem 6.38), the actual annuli coincide, so A' is central symmetric. It follows that $f|A'$ is a double covering over its image A.

In either case we have

$$\mu(\delta') = \operatorname{mod} A' = d^{-1} \operatorname{mod} A \le d^{-1}\mu(\delta) + 2.$$

$\square$

One can avoid using a deep Theorem 6.38 by replacing it with the following

EXERCISE 39.13. *In Case b) above there exists a central symmetric annulus $B \subset A'$ such that* $\operatorname{mod} B \ge \operatorname{mod} A' - 4$.

EXERCISE 39.14. *Generalize the above two lemmas to arbitrary finite degree branched coverings $f : S \to T$ between Riemann surfaces.*

39.2.9. *Thurston matrix.* Let $\mathcal{G} \equiv \mathcal{G}^n$ be the canonical invariant multicurve on $\mathbb{R}^2 \smallsetminus \mathcal{P}$ of size l. Let us associate to it the following $l \times l$-matrix $T = (t_{\gamma\delta})_{\gamma,\delta \in \mathcal{G}^n}$ with entries in $\{0, 1, 1/2\}$. If $[\gamma]$ is not a pullback $f^*[\delta]$ we let $t_{\gamma\delta} = 0$. Otherwise we let

$$t_{\gamma\delta} = \frac{1}{\deg(f : \delta' \to \delta)}$$

where δ' is the pullback $f^*\delta$ homotopic to γ. Let $\mu = (\mu(\gamma))_{\gamma \in \mathcal{G}}$ be the vector of the canonical moduli.

LEMMA 39.15. *Under the above circumstances, we have:*

$$\mu' = T \cdot \mu + O(1),$$

where the constant depends only on p and K.

PROOF. Let us take a canonical annulus $A_\gamma \subset \mathbb{C} \smallsetminus \mathcal{O}'$, $\gamma \in \mathcal{G}$, and uniformize it by a flat cylinder $\mathcal{C} = S^1 \times (0, \mu(\gamma))$. Let us mark on $\mathcal{C}$ the points corresponding to $f^{-1}(\mathcal{O})$ (which necessarily do not belong to $\mathcal{O}'$). The genuinely horizontal circles through these points partition $\mathcal{C}$ into several (at most p) flat cylinders $\mathcal{C}_j$ such that

$$\mu'(\gamma) \equiv \operatorname{mod} \mathcal{C} = \sum_j \operatorname{mod} \mathcal{C}_j. \tag{39.5}$$

Each cylinder $\mathcal{C}_j$ represents some homotopy class of a Jordan curve γ_j on $\mathbb{R}^2 \smallsetminus g^{-1}(\mathcal{P})$. If $\operatorname{mod} \mathcal{C}_j \geq \underline{\mu} + 2$ then by Lemma 39.11, γ_j is homotopic to a pullback $g^*(\delta_j)$ of a canonical Jordan curve δ_j on $\mathbb{R}^2 \smallsetminus \mathcal{P}$. By Lemma 39.12

$$\operatorname{mod} \mathcal{C}_j \leq t_{\gamma\delta_j}\, \mu(\delta_j) + 2. \tag{39.6}$$

It follows that $\mu(\delta_j) \geq \underline{\mu}$ and thus $\delta_j \in \mathcal{G}(\underline{\mu})$. Putting together (39.6) with (39.5), we conclude:

$$\mu'(\gamma) \leq \sum_{\delta} t_{\gamma\delta}\, \mu(\delta) + (2 + \underline{\mu})p.$$

This provides us with the more delicate bound.

To obtain the opposite bound, let us consider the canonical annuli

$$A_j \subset \mathbb{C} \smallsetminus f^{-1}(\mathcal{O})$$

representing the cylinders $\mathcal{C}_j$ (when exist). Since they are homotopic in $\mathbb{C} \smallsetminus \mathcal{O}'$, their union is contained in some annulus $A \subset \mathbb{C} \smallsetminus \mathcal{O}'$. Since the A_j are pairwise disjoint, the Gröztsch Inequality implies:

$$\mu'(\gamma) \geq \operatorname{mod} A - 2 \geq \sum \operatorname{mod} A_j - 2 \geq \sum_j t_{\gamma\delta_j}(\mu(\delta_j) - 2)$$

(where the last estimate comes from Lemma 39.12. Thus,

$$\mu'(\gamma) \geq \sum_{\delta} t_{\gamma\delta}\, \mu(\delta) - 2p.$$

$\square$

39.2.10. *Structure of an invariant multicurve.* Below we will make use of the basic Perron-Frobenius Theory, see Appendix §19.19.

Let us consider an invariant multicurve $\mathcal{G}$. Any component $\gamma \in \mathcal{G}$ surrounds some set $\mathcal{P}_\gamma \subset \mathcal{P}$ of punctures $0_i \in \mathcal{P}$. Moreover, since γ is non-peripheral,

$$k(\gamma) := \#\mathcal{P}_\gamma \in [2, p-1].$$

Let $\delta = g^*(\gamma)$. Then g maps injectively $\mathcal{P}_\delta$ to $\mathcal{P}_\gamma$, and hence $k(\delta) \leq k(\gamma)$ (so, $k(\gamma)$ is a "Lyapunov function" for the pullback).

Assume γ is a recurrent vertex in the graph Γ_T of the Thurston matrix. Let us consider a simple loop $(\gamma \equiv \gamma_0, \gamma_1, \dots, \gamma_{q-1}, \gamma_q)$ in Γ_T. By definition, this means:

(i) $\gamma_{m+1} = f^*(\gamma_m)$, $m = 0, 1, \dots, q-1$;

(ii) γ_m and γ_n are not homotopic for any $0 \leq m < n \leq q-1$;

(iii) $(g^q)^*(\gamma) \simeq \gamma$ rel $\mathcal{P}$.

Let D_m be the Jordan disk bounded by the γ_m. Then the following properties hold:

a) All the disks D_m contain the same number k of punctures, and

$$g(D_{m+1} \cap \mathcal{P}) = D_m \cap \mathcal{P};$$

b) γ_{m+1} is the only pullback of $f^*(\gamma_m)$;

c) The Jordan disks $D_m \subset \mathbb{R}^2$ are (essentially) pairwise disjoint for $m = 0, 1, \dots, q-1$;

d) $\displaystyle\bigcup_{m=0}^{q-1} D_m \supset \mathcal{P}$.

Since $k(\gamma_{m+1}) \le k(\gamma_m)$, $m = 0, 1 \ldots, q-1$, and $k(\gamma_q) = k(\gamma)$, we conclude that all these numbers must be equal:

$$k(\gamma) = k(\gamma_1) = \cdots = k(\gamma_q) := k,$$

which proves a). Moreover, the other component, γ'_{m+1}, of $g^{-1}(\gamma_m)$ (if exists) does not surround any punctures. So it is trivial, and we obtain b).

Let us now consider two curves γ_m and γ_n, $0 \le m < n \le q-1$. Since they are components of a multicurve, they are essentially disjoint. Replacing them with homotopic ones (rel $\mathcal{P}$), we can assume that they are disjoint in the first place. Then the corresponding Jordan disks D_m and D_n in $\mathbb{R}^2$ are either nested or disjoint. If the former holds, say $D_m \subset D_n$, then $\gamma_m \simeq \gamma_n$. (Indeed, D_m and D_n contain the same number of punctures, and hence the annulus $D_n \smallsetminus D_m$ does not contain any punctures.) Contradiction with assumption (ii) proves property c).

To see d), let us now take any point $0_i \in \mathcal{P}$ and find the smallest m such that $g^m(0_i) \in D_0$. Then property a) implies that $0_i \in D_m$.

Let us relabel the curves γ_m so that γ surrounds 0. Then the map $g : D_0 \to D_{q-1}$ has degree 2, while the other maps $g : D_{m+1} \to D_m$ have degree 1. Letting $t_{mn} \equiv t_{\gamma_m \gamma_n}$, we conclude that

$$t_{0,q-1} = 1/2, \quad t_{m+1,m} = 1,$$

while all other entries t_{mn} with $0 \le m, n \le q-1$ vanish. This is a weighted permutation with spectral radius $(1/2)^{1/q} < 1$.

AS $k((f^*)^n(\gamma)) = k(\gamma)$ for some n. the curve γ is periodic: $(f^*)^p(\gamma) = \gamma$. Let n_γ be the smallest period of γ. Using a similar argument as above, we see that the pullbacks $\gamma_m := (f^*)^m(\gamma)$, $m = 0, 1, \ldots, n_\gamma - 1$ bound essentially disjoint Jordan disks D_m in $\mathbb{R}^2$. Moreover, the degree of $[\gamma_{m+1}] \to [\gamma_m]$ is equal to 2 or 1 depending on whether γ_m surrounds the critical point or not. It follows that the spectral radius of the block of the Thurston matrix T corresponding to this cycle of curves is equal to $(1/2)^{1/n_\gamma}$.

We see that the recurrent part of the graph Γ_T is decomposed into (non-interacting) cycles of curves with spectral radius less than 1. Consequently, we obtain

Lemma 39.16. *Let*

$$g : (\mathbb{R}^2, \mathcal{P}) \to (\mathbb{R}^2, \mathcal{P}),$$

be a critically periodic degree two branched covering, and let T be the Thurston matrix of some invariant multicurve. Then the recurrent part of T is the weighted permutation with all cycles having multiplier $1/2$. Hence the spectral radius of T is bounded by $(1/2)^{1/p}$.

39.2.11. *Improvement of the canonical moduli.* Finally, we can show that if some Riemann surface $S_n = \mathbb{C} \smallsetminus \mathcal{O}^n = \pi(\tau_n)$ escapes far away in the Moduli space then after several more iterates it will come closer. We measure how far S_n escapes in the moduli space by the maximal canonical modulus $M_n = \|\mu^n\|_\infty$ (39.3).

Lemma 39.17. *There exists $\underline{M}$ (depending on p and K) such that if $M_n \ge M$ for some Riemann surface S_n then*

$$M_{n+p} \le \frac{3}{4} M_n.$$

PROOF. Lemma 39.16 implies that for $\underline{M}$ sufficiently big (depending on p and K), we have:

$$\|T^p \cdot \mu\|_\infty \le \frac{1}{2} \|\mu_n\|_\infty = M_n.$$

By Lemma 39.9, the threshold $\underline{\mu}$ can be selected so that the multicurves $\mathcal{G}^{n+k}(\underline{\mu})$ are all the same for $k = 0, 1, \ldots, p-1$. Lemma 39.15 implies the claim. $\square$

Notes

Probably, the first rough image of the "Mandelbrot set" $\mathcal{M}$ was made by Brooks and Matelski [**BrMa**]. A finer image produced by Mandelbrot [**Man**] sparked a great interest in this object.

Elementary structure of $\mathcal{M}$ was analyzed by Levin [**Le1**] by means of the Montel theory of normal families. In this approach, the boundary of $\mathcal{M}$ appears as the set of "irregular points" where the sequence of polynomials $c \mapsto f_c^n(0)$ fails to be normal (similarly to the classical definition of the Julia set).

Theorem 34.1 on the connectivity of $\mathcal{M}$ was proved by Douady and Hubbard [**DH1**],[**DH2**, Exposé VIII], with an input from Sibony (see the acknowledgment in [**DH1**]), in the early 1980s. The elementary proof given in §34.3 reproduces the original argument. The *Multiplier Theorem* is also due to Douady and Hubbard [**DH1**],[**DH2**]. The unifying insight on these results from the Teichmüller theory viewpoint is due to Sullivan (preprint IHES, 1982); It was further advanced by McMullen and Sullivan in [**McS**].

The *Phase-Parameter Relation* was famously formulated by Douady as a principle: *You plow in the Dynamical Plane and then harvest in the Parameter Plane.* In the course of this book we will encounter many applications of this principle (though sometime "harvesting" is harder than "plowing").

Foundations of the combinatorial theory of the Mandelbrot set were developed by Douady and Hubbard in the Orsay Notes [**DH2**], followed by Milnor, and their schools (by the same participants as the combinatorial theory for Julia sets). In particular, see [**DH1**], [**DH2**, Exposé XIII] and [**M5**] for the Wake Theorem.

The *Fatou Conjecture* on Density of Hyperbolicity (33.7) is an interpretation of several remarks that Fatou made on page 73 of [**F2**]. First, Fatou observed that *hyperbolicity* is preserved under perturbations. Then he conjectured that any rational map can be approximated by a *stable* one (compare with Theorem 36.1). He also suggested that unstable maps form some kind of algebraic set: apparently, he did not give a close look at this issue. Here is the quotation: "Il est probable, mais je n'ai pas approfindi la question, que cette propriété appartient à tout les substitutions generales, c'est-à-dire à celles dont les coefficients ne vérifient aucune relation particulière. Je signale, dans ce même ordre d'idées, l'intérêt qu'il y aurait à rechercher les conditions nécessares et suffisantes pour que l'ensembe $\mathcal{F}$ varie d'une manière continue, tant au point de vue de la position de ses points qu'au point de vue de la connexion des domains dans lesquels il divise le plan, lorsqu'on fait varier les coefficients de $R(z)$." (Here $R(z)$ is the iterated rational function and $\mathcal{F}$ is its "Julia set"):

The general theory of hyperbolicity and structural stability developed in the 1960s by Smale, Anosov, and many other people (see [**KaH**]), greatly clarified how Fatou's remarks should be interpreted. Theorem 36.2 on the J-stability was independently proven by Mañé-Sad-Sullivan [**MSS**] and the author [**L7**]. Formally speaking, it proves the original Fatou Conjecture.

The reduction of the Fatou Conjecture (on the density of hyperbolicity) to the No Invariant Line Fields Conjecture appeared in [**MSS**]. The latter Conjecture had a precise analogue for Kleinian groups that had been proved by Sullivan [**S5**].

Theorem 37.31 on the entropy growth was posed as a Conjecture in the preprint version of the Milnor-Thurston paper[**MT**]; it became a theorem in the published version. It was one of the first deep applications of the complex methods (Rigidity Theorem 35.1 for superattracting maps) to real dynamics. For further developments in the higher degree case, see [**MTr, Rad, BS**] and van Strien's survey [**vS**].

Probably, Myrberg was the first who studied hyperbolic windows in the real quadratic family and alluded to the problem of their density [**Myr2**]. The Density Theorem was proven in [**L10, GS**]; it will be discussed in vol. III.

The Thurston Realization Theorem became accessible to the math community due to the work of Douady and Hubbard [**DH4**]. The "spider algorithm" adapting it to the quadratic case appeared in [**HSc1**]. Now the theory is available as part of a recent book [**H3**].

As the parameter space of higher degree polynomials is multi-dimensional, it is much more difficult to visualize its bifurcation locus and to develop even a basic theory for it. The problem is approached from a variety of angles: by either taking natural one-dimensional slices of the space (see [**BKM**]) or by modeling it by a relevant space of geodesic laminations (see [**BO2**]), of by applying to it the pluripotential theory of currents (see [**DeM, Du**]).

CHAPTER 6

Straightening, puzzle, and attractors

40. Straightening

40.1. Geometric adjustments. The notion of a quadratic-like map with the fixed domain is too rigid, so we allow *adjustments* of the domains which do not affect the essential dynamics of the map (see Exercise 28.2). An appropriate adjustment allows one to improve the geometry of the map:

Lemma 40.1. *Consider a quadratic-like map $f: U \to U'$ with*

$$\operatorname{mod} A \geq \mu > 0 \tag{40.1}$$

and $f(0) \in U$. Then there exist κ, σ and C depending only on μ such that f admits an adjustment $g: V \to V'$ with the following properties:

(i) *The new domains V and V' are bounded by real analytic κ-quasicircles with σ-bounded shape around the origin.*

(ii) $\operatorname{mod}(V' \smallsetminus V) \geq \mu/2 > 0$.

(iii) *g admits a decomposition*

$$g = h \circ f_0, \tag{40.2}$$

where $f_0(z) = z^2$ and h is a univalent function on $W := f_0(V)$ with distortion bounded C.

Proof. Let us uniformize the fundamental annulus A of f by a round annulus, $\phi: \mathbb{A}(1/r, r) \to A$, where $r \geq e^{\pi\mu} \equiv r_0$. Then $\gamma' := \phi(\mathbb{T})$ is the equator of A. Consider the disk V' bounded by γ', and let $V = f^{-1}V'$. Since $f(0) \in V'$, V is a conformal disk and the restriction $f: V \to V'$ is a quadratic-like adjustment of f (see Exercise 28.2).

Restrict ϕ to the annulus $\mathbb{A}(1/r_0, r_0)$. Take an arc $\alpha = [a, b]$ on $\mathbb{T}$ of length at most $\delta := (1 - 1/r_0)/2$. By the Koebe Distortion and 1/4 Theorems in the disk $\mathbb{D}_{2\delta}(a)$, we have:

$$|\phi(b) - \phi(a)| \geq \frac{|b-a|}{4} |\phi'(a)|; \quad l(\phi(\alpha)) \leq K\, |\phi'(a)|\, l(\alpha),$$

where l stands for the arc length, and K is an absolute constant. Hence $\gamma' := \partial V' = \phi(\mathbb{T})$ is a quasicircle with the dilatation depending only on $r_0 = r_0(\mu)$.

Applying the same argument to the uniformization of $f^{-1}A$, we conclude that its equator $\gamma := \partial V$ is a quasicircle with bounded dilatation as well.

Since γ and γ' are 0-symmetric κ-quasicircle, the shape of these curves around 0 is bounded by some constant $C(\kappa)$ (see Exercise 15.13). This proves (i).

Property (ii) is obvious since $\operatorname{mod}(V' \smallsetminus V) \geq \operatorname{mod} \mathbb{A}(1, r_0) \geq \mu/2$.

Since g is even, it admits decomposition (40.2). Moreover, h admits a univalent extension to the disk $\tilde{W} = f_0(U)$, and

$$\operatorname{mod}(\tilde{W} \smallsetminus W) = 2 \operatorname{mod}(U \smallsetminus V) \geq \mu/2.$$

The Koebe Distortion Theorem (in the invariant form of Theorem 4.16) completes the proof. □

In the connected case the above adjustment can be refined further. To make the statement, it is convenient to normalize f (conjugating it by an appropriate complex rescaling $z \mapsto \lambda z$) so that it has the following Taylor expansion at the origin:

$$f(z) = c + z^2 + \dots \tag{40.3}$$

(i.e., we have made the coefficient at z^2 equal to 1).

Lemma 40.2. *Under the circumstances of* Lemma 40.1*, assume that f has connected Julia set and is normalized by* (40.3)*. Then f can be adjusted so that it satisfies properties* (i) *and* (iii) *of* Lemma 40.1 *and also the following two properties:*

(ii′) $\underline{\mu} \leq \operatorname{mod}(V' \smallsetminus V) \leq 1$*, where $\underline{\mu} > 0$ depends only on μ.*

(iv) *For some constant $\rho \in (0, 1)$ depending only on μ,*

$$\rho \leq r_V \leq R_{V'} \leq 1/\rho,$$

where r_V and $R_{V'}$ are the inner and outer radia of the respective domains around 0.

Proof. (ii′) Let $U^n := f^{-n}U'$ and let $A^n := U^{n-1} \smallsetminus U^n = f^{-(n-1)}(A)$. Without loss of generality we can assume that f satisfies properties (i)–(iii) from Lemma 40.1. Then the annulus $B := A \cup A^1$ is obtained by gluing A and A^1 along a $\kappa-$quasicircle, implying that $\operatorname{mod} B \asymp \operatorname{mod} A$ with constants depending only on μ (see Exercise 15.20). Let $B^n := f^{-(n-1)}(B)$.

Since the Julia set is connected, $v \equiv f(0) \in U^n$ for all n, so the restrictions $f : U^n \to U^{n-1}$ are quadratic-like maps obtained by consecutive adjustments of $f : U \to U'$. Since

$$\operatorname{mod} B^n = \operatorname{mod} B/2^{n-1} \asymp \operatorname{mod} A/2^{n-1} = \operatorname{mod} A^n,$$

we can select n in such a way that

$$2\underline{\mu} \leq \operatorname{mod} A^n \leq \operatorname{mod} B^n \leq 1,$$

with some $\underline{\mu} > 0$ depending only on μ.

Let us now adjust $f|\,U^n$ once more as in Lemma 40.1. We obtain a quadratic-like map $g : V \to V'$ satisfying property (ii′).

(iv) Assume now that f is normalized by (40.3), so is g. Then in representation (40.2), $g = h \circ f_0$, the univalent map $h : (W, 0) \to (V', c)$ is also normalized: $h'(0) = 1$. Since $W = f_0(V)$,

$$0 < C^{-1} r_W \leq r_{V',v} \leq R_{V',v} \leq C R_W,$$

where $C = C(\mu)$ is the distortion bound from Lemma 40.1. Hence

$$C^{-1} r_V^2 \leq r_{V',v} \leq R_{V',v} \leq C R_V^2. \tag{40.4}$$

But since $V' \supset V \ni v$, we have: $R_{V',v} \geq R_V/2$, and by the right-hand side of (40.4), $R_V \geq 1/(2C)$. Since V has a σ-bounded shape around the origin, the inner radius r_V is also bounded away from 0: $r_V \geq 1/(2C\sigma)$.

On the other hand, if $r_V = L \gg 1$ then the left-hand side of (40.4) (and bounded shape of V) implies that the annulus $V' \setminus V$ contains the round annulus whose inner radius is of order L and the outer radius is of order L^2, so that $\operatorname{mod}(V' \setminus V) \geq \log L - O(1)$, contradicting property (ii′). □

Under the circumstance of Lemma 40.1, we say that the ql map g has *L-bounded geometry*, where $L \geq \max\{2/\mu, \kappa, \sigma, C\}$. Under the circumstances of Lemma 40.2, we replace $2/\mu$ with $1/\underline{\mu}$ and say that g has *(L, ρ)-bounded geometry*. Also, if a ql map g admits decomposition (40.2), we call it *quadratic up to a bounded distortion*.

40.2. Douady-Hubbard Straightening Theorem. If the reader attempted to extend the basic dynamical theory from quadratic polynomials to quadratic-like maps, quite likely he/she had a problem with the No Wandering Domains Theorem. The only known proof of this theorem crucially uses the fact that a polynomial of a given degree depends on finitely many parameters. The flexibility offered by the infinite dimensional space of quadratic-like maps looks at this moment like a big disadvantage. It turns out, however, that the theorem is still valid for quadratic-like maps, and actually there is no need to prove it independently (as well as to repeat any other pieces of the topological theory). In fact, quadratic-like maps do not exhibit any new features of topological dynamics, since all of them are topologically equivalent to polynomials (restricted to appropriate domains)!

To state the result precisely, we need a few definitions. Two quadratic-like maps f and g are called *topologically conjugate* if they become such after some adjustments of their domains. Thus there exist adjustments $f : U \to U'$ and $g : V \to V'$ and a homeomorphism $h : (U', U) \to (V', V)$ such that the following diagram is commutative:

$$\begin{array}{ccc} U & \xrightarrow[f]{} & U' \\ h\downarrow & & \downarrow h \\ V & \xrightarrow[g]{} & V' \end{array}$$

In case when one of the maps is a polynomial, we allow to take any quadratic-like restriction of it.

If the homeomorphism h in the above definition can be selected quasiconformal (respectively: conformal or affine) then the maps f and g are called quasiconformally (respectively: conformally or affinely) conjugate. Two quadratic-like maps are called *hybrid* equivalent if they are qc conjugate by a map h with $\bar{\partial} h = 0$ a.e. on the filled Julia set $\mathcal{K}(f)$.

Remark. The last condition implies that h is conformal on the $\operatorname{int} \mathcal{K}(f)$. On the Julia set $\mathcal{J}(f)$ it gives an extra restriction only if $\mathcal{J}(f)$ has positive area.

The equivalence classes of hybrid (respectively: qc, topological etc.) conjugate quadratic-like maps are called *hybrid* (respectively: qc, topological etc.) *classes*.

STRAIGHTENING THEOREM. *Any quadratic-like map g is hybrid conjugate to a quadratic polynomial f_c. If $\mathcal{J}(g)$ is connected then the corresponding polynomial f_c is unique.*

This polynomial f_c is called the *straightening* of g.

The Straightening Theorem immediately reduces the Topological Dynamics for ql maps to that for polynomials. In particular, we conclude:

COROLLARY 40.3. *If $g : U \to U'$ is a quadratic-like map, then:*

(i) *Repelling periodic points are dense in $\mathcal{J}(g)$.*

(ii) *Iterated preimages of any point $z \in U$, except the superattracting fixed point (if exists), accumulate on the whole Julia set $\mathcal{J}(g)$.*

(iii) *The Julia set $\mathcal{J}(g)$ is the smallest completely invariant compact set (except for a possible superattracting fixed point). The filled Julia set $\mathcal{K}(g)$ is the smallest completely invariant hull.*

(iv) *There are no wandering components of* $\operatorname{int}\mathcal{K}(g)$.

(v) *If all periodic points of g are repelling then $\mathcal{K}(g) = \mathcal{J}(g)$.*

REMARK 40.4. If $\mathcal{J}(g)$ is a Cantor set, then the straightening is not unique. Indeed, by Theorem 36.18, all quadratic polynomials f_c, $c \in \mathbb{C} \smallsetminus \mathcal{M}$, are qc equivalent. Since their filled Julia sets have zero area, they are actually hybrid equivalent. Hence all of them are "straightenings" of g.

Existence of the straightening will be proven in the next section, while uniqueness will be postponed until §41.3.

40.3. Construction of the straightening. The idea is to "mate" g near $\mathcal{K}(g)$ with $f_0 : z \mapsto z^2$ near ∞.

Without loss of generality, we can assume that g is conventional. Take some $r > 1$. Consider two closed disks: the disk $\overline{U}'$ endowed with the map $g : \overline{U} \to \overline{U}'$ and the disk $\hat{\mathbb{C}} \smallsetminus \mathbb{D}_r$ endowed with the map $f_0 : \hat{\mathbb{C}} \smallsetminus \mathbb{D}_r \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}_{r^2}$. Let us view them as two hemi-spheres, $S^2_0 \equiv \overline{U}'$ and $S^2_\infty \equiv \hat{\mathbb{C}} \smallsetminus \mathbb{D}_r$ (see Figure 40.1). Glue them together by an orientation preserving *equivariant* qc homeomorphism $T : A \to \mathbb{A}[r, r^2]$ between the closed fundamental annuli. Here "equivariance" means that h *respects the boundary dynamical relation* (compare §19.4):

$$T(gz) = f_0(Tz) \quad \text{for } z \in \partial U. \tag{40.5}$$

Such a map $T = T_g$ is called a *tubing* of g.

EXERCISE 40.5. *Construct a tubing T (using that g is conventional). Do it so that* $\operatorname{Dil} T$ *is bounded in terms of* $\operatorname{mod} A$ *and qc dilatation of the quasicircles ∂U, $\partial U'$.*

In this way we obtain an oriented *qc sphere*

$$S^2 = S^2_0 \sqcup_T S^2_\infty \equiv \overline{U}' \sqcup_T (\hat{\mathbb{C}} \smallsetminus \mathbb{D}_r)$$

with the atlas of two local charts given by the identical maps $\phi_0 : S^2_0 \to \overline{U}'$ and $\phi_\infty : S^2_\infty \to \hat{\mathbb{C}} \smallsetminus \mathbb{D}_r$. Moreover, these hemi-spheres are quasidisks in S^2. For insttucance, in the local chart ϕ_0 the curve $\gamma := \partial S^2_\infty$ becomes $\phi_0(\gamma) = \partial U$ which is a quasicircle since f is conventional.

Define now a map $F : S^2 \to S^2$ by letting

$$F(z) = \begin{cases} \phi_0^{-1} \circ g \circ \phi_0(z) & \text{for } z \in \phi_0^{-1}\overline{U} \\ \phi_\infty^{-1} \circ f_0 \circ \phi_\infty(z) & \text{for } z \in \overline{S}^2_\infty \end{cases}$$

(It is certainly quite a pedantic way of writing since the maps ϕ_0 and ϕ_∞ are in fact identical.) Since T is equivariant (40.5), these two formulas match on γ. Hence F is a continuous endomorphism of S^2. Moreover, it is a double branched covering

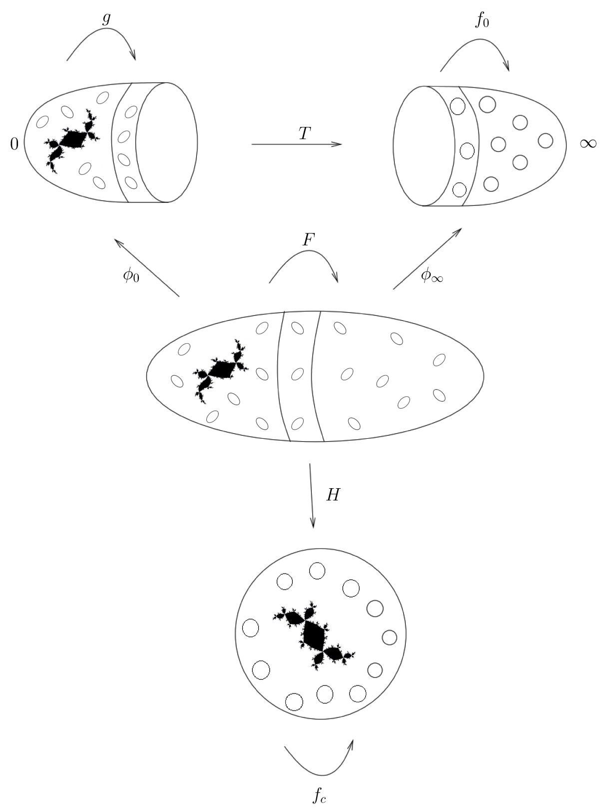

FIGURE 40.1. Straightening

of the sphere onto itself (with two simple branched points at "0"$\equiv \phi_0^{-1}(0)$ and "∞"$\equiv \phi_\infty^{-1}(\infty)$).

Since $F : S^2 \to S^2$ is holomorphic in the local charts ϕ_0 and ϕ_∞, it is quasiregular on $S^2 \smallsetminus \gamma$. Since γ is a quasicircle, it is removable (Lemma 16.3). Hence F is quasiregular on the whole sphere.

EXERCISE 40.6. *Let us adjust g so that ∂U is smooth. Then the gluing map T can be chosen so that S^2 is a smooth sphere and the map F is smooth.*

We will now construct an F-invariant conformal structure μ on S^2 (with a bounded dilatation with respect to the qc structure of the sphere S^2). Start in a neighborhood of ∞: $\mu | S^2_\infty = (\phi_\infty)^* \sigma$. Since σ is f_0-invariant, $\mu | S^2_\infty$ is F-invariant. Since ϕ_∞ is qc, $\mu | S^2_\infty$ has a bounded dilatation.

Next, look at this structure in the local chart $\phi_0 : S^2_0 \to U'$, and by means of Corollary 29.5 extend it canonically to an invariant structure on the whole sphere S^2 with the same dilatation. We will keep the same notation μ for the extension.

EXERCISE 40.7. *Work out details of this canonical extension.*

We obtain an F-invariant measurable conformal structure μ with bounded dilatation on the whole sphere S^2. By the Measurable Riemann Mapping Theorem, there exists a qc map $H : (S^2, \mu) \to \hat{\mathbb{C}}$ such that $H_*\mu = \sigma$ and normalized so that $H(0) = 0$, $H(\infty) = \infty$ and $H \circ \phi_\infty^{-1}(z) \sim z$ as $z \to \infty$. Then the map $f := H \circ F \circ H^{-1}$ is a quadratic polynomial (see §29.1.2) with the critical point at the origin and asymptotic to z^2 at ∞. Hence $f = f_c : z \mapsto z^2 + c$ for some c.

EXERCISE 40.8. *Show that* $\mathcal{K}(f) = H(\phi_0^{-1}\mathcal{K}(g))$.

The qc map $H \circ \phi_0^{-1}$ conjugates $g : U \to U'$ to a quadratic-like restriction of f. Moreover, restricting it to $\mathcal{K}(g)$, we see that

$$(H \circ \phi_0^{-1})_*\sigma = H_*\mu = \sigma,$$

so H is a hybrid conjugacy between g and the restriction of f. Thus, f is a straightening of g.

REMARK 40.9. The straightening construction of f_c was uniquely determined by the choice of the tubing $T : A \to \mathbb{A}[r, r^2]$. In fact, one can say more: *In the case of connected Julia set, the straightening is independent of the choice of tubing, while in the disconnected case, it depends only on the tubing position of the critical value* (see Proposition 41.12 below).

40.4. Addendum to the straightening construction. Here we will refine the straightening construction in several ways. In particular, we will extend the tubing to a bigger annulus, through a series of liftings (similarly to the extension of the Böttcher function carried in §23.5.2).

40.4.1. *Tubing equipotentials and rays.* The tubing $T : A \to \mathbb{A}[r, r^2]$ plays the role of the Böttcher coordinate for the (conventional) quadratic-like map g. In particular, we can use it to define equipotentials and rays for g as pullbacks by T of the round circles and radial intervals in $\mathbb{A}[r, r^2]$. In this way we obtain two foliations in the fundamental annulus A. There are natural radii/levels assigned to the equipotentials and external angles assigned to the rays. (For instance, the boundary equipotential ∂U has radius r and level $t = \log r$.)

By means of the dynamics, we can now extend these foliations to invariant (singular) foliations in $\overline{U}' \smallsetminus \mathcal{K}(g)$. If $\mathcal{K}(g)$ is connected then these foliations are in fact non-singular. In the disconnected case, they have simple "cross-singularities" at the critical point 0 and its iterated preimages. In this case, the figure-eight equipotential passing through 0 is called *critical.* We let $\Omega \equiv \Omega_g \subset \overline{U}'$ be the semi-open topological annulus bounded by this equipotential (which is excluded from Ω) and the external boundary $\partial U'$ (which is part of Ω). In the connected case, we let $\Omega \equiv \Omega_g = U' \smallsetminus \mathcal{K}(g)$. (Everything is similar to the polynomial case.)

40.4.2. *Equivariant extension of the tubing.* Similarly to the Böttcher coordinate, the tubing can be equivariantly extended to the domain Ω_g (compare §23.5.2). It is based on a simple lifting step:

LEMMA 40.10. *Let us consider a nest of two (non-compact) Riemann surfaces $\Omega \subset \Omega'$ with boundary. We assume that the boundaries $\Gamma' := \partial\Omega'$ and $\Gamma := \partial\Omega \Subset \operatorname{int}\Omega'$ are (disjoint) topological circles, and that $A := \Omega' \smallsetminus \operatorname{int}\Omega$ is a closed annulus bounded by Γ and Γ' (its "inner" and "outer" boundaries respectively). Let $g : \Omega \to \Omega'$ be a holomorphic double covering map such that $g(\Gamma) = \Gamma'$.*

Consider also another map $\tilde{g} : \tilde{\Omega} \to \tilde{\Omega}'$ with the same properties (all corresponding objects for $\tilde{g}$ are marked with "tilde"). Let $h : A \to \tilde{A}$ be an equivariant homeomorphism, i.e., $h(gz) = \tilde{g}(hz)$ for $z \in \Gamma$.

Assume A and $\tilde{A}$ do not contain the critical values of g and $\tilde{g}$ (respectively). Then $A^1 := g^{-1}(A)$ and $\tilde{A}^1 := \tilde{g}^{-1}(\tilde{A})$ are annuli attached to A and $\tilde{A}$ respectively, and h extends uniquely to an equivariant homeomorphism $H : A \cup A^1 \to \tilde{A} \cup \tilde{A}^1$.

Moreover, if Γ is a quasicircle (inside Ω') and h is K-qc then H is K-qc as well.

PROOF. Since A does not contain the critical values of g, A^1 is an annulus. Since $g(\Gamma) = \Gamma'$, A^1 is attached to A along Γ, so together they form an annulus $A \cup A^1$.

By the general lifting theory (see the Lifting Criterion from §1.6.2 and Exercise 1.100), h lifts to a homeomorphism

$$h^1 : A^1 \to \tilde{A}^1$$

in two ways determined by the choice of value of h_1 at one point. But since $h : A \to \tilde{A}$ is equivariant, $h|\,\Gamma$ is a lift of $h|\,\Gamma'$. Hence the lift h_1 can be selected so that it coincides with h on Γ, and we obtain a single equivariant homeomorphism

$$H : A \cup A^1 \to \tilde{A} \cup \tilde{A}^1.$$

Uniqueness of such an extension is obvious.

If h is K-qc then so is h_1 (since g is holomorphic). If Γ is a quasicircle then H is K-qc as well (by Lemma 16.3). $\square$

By iterating this lifting construction, we obtain:

COROLLARY 40.11. *Let $g : U \to U'$ and $\tilde{g} : \tilde{U} \to \tilde{U}'$ be two (conventional) quadratic-like maps with fundamental annuli A and $\tilde{A}$. Let $h : A \to \tilde{A}$ be an equivariant homeomorphism between the fundamental annuli. If $\mathcal{K}(g)$ and $\mathcal{K}(\tilde{g})$ are connected then h extends uniquely to an external conjugacy $U' \smallsetminus \mathcal{K}(f) \to \tilde{U}' \smallsetminus \mathcal{K}(\tilde{g})$. Moreover, if h is K-qc then so is the extension.*

COROLLARY 40.12. *Let $T : A \to \mathbb{A}[r, r^2]$ be a K-qc tubing for g as above. Then it extends to an equivariant K-qc map $\Omega_g \to \overline{\mathbb{D}}_{r^2} \smallsetminus \overline{\mathbb{D}}_R$, where $R = 1$ in the connected case and $R > 1$ in the Cantor case.*

This extension will be denoted T and called a *tubing* as well. Its equivariance means that $T(gz) = T(z)^2$ for $z \in \Omega_g \cap \overline{U}$.

Note that in the Cantor case, we have $g(0) \in \Omega_g$, so the point $T(g(0))$ is well defined. We call it the *tubing position of the critical value*.

40.4.3. *Böttcher coordinate for the straightening.* The map $B := \phi_\infty \circ H^{-1}$ in the Straightening construction (see Figure 40.1) is the Böttcher coordinate for f on $\Omega(r) := H(S^2_\infty)$. Indeed, $B|\,\Omega(r)$ is conformal (since both ϕ_∞ and H transfer the conformal structure $\mu|S^2_\infty$ to σ) and B conjugates f to $f_0 : z \mapsto z^2$.

Since $B(\partial\Omega(r)) = \mathbb{T}_r$, $\partial\Omega(r) = \mathcal{E}^r$ is the equipotential of radius r for f. Thus, *we have conjugated $g : U \to U'$ to $f : \Sigma(r) \to \Sigma(r^2)$, where $\Sigma(r)$ is the subpotential disk of radius r for f* (see §23.5.4).

In the Cantor case (with the extended tubing) shows that *the tubing position of the critical value for a ql map g coincides with Böttcher position of the critical value for its straightening $f \equiv f_c$*:

$$T_g(g(0)) = B_c(c). \tag{40.6}$$

40.4.4. *Dilatation.* Finally, let us dwell on an important issue of a bound on the dilatation of the qc homeomorphism that straightens g.

Lemma 40.13. *Let $g : U \to U'$ be a quadratic-like map with* $\operatorname{mod} A \geq \delta > 0$. *Then g is hybrid conjugate to a straightening f_c by a K-qc map whose dilatation K depends only on δ.*

Proof. Let us first adjust g according to Lemma 40.1 (keeping the same notations for the domains U and U').

Let us now follow the proof of the Straightening Theorem. Take a look at the conformal structure $\mu|\, S_0^2$ in the local chart ϕ_0, i.e., consider the conformal structure $\nu = (\phi_0)_*(\mu|\, S_0^2)$ on U'. On $U' \smallsetminus \mathcal{K}(g)$, it is obtained by pulling back (by the conformal g-dynamics) the structure $T^*(\sigma)$ from the fundamental annulus A. On $\mathcal{K}(g)$ it is equal to the standard structure σ. Hence the dilatation of ν is equal to the dilatation of the tubing T.

The qc map $H \circ \phi_0^{-1}$ conjugating $g : U \to U'$ to $f : \Sigma(r) \to \Sigma(r^2)$ transfers ν to σ. Hence its dilatation is also equal to $\operatorname{Dil} T$. But the latter is bounded in terms of δ (see Exercise 40.5). □

40.4.5. *Standard equipment of ql maps.* Due to the Straightening Theorem, we can equip ql maps with the standard ammunition of quadratic polynomials. Notice first that the $\alpha-$ and $\beta-$ fixed points are well defined as long as the Julia set $\mathcal{J}(g)$ is connected: namely, the $\alpha-$fixed point is specified by being either non-repelling or dividing repelling. (This charachterization is topologically invariant due to Exercise 21.1.) Moreover, as pointed out in §40.4.1, once we select a tubing, we obtain the external foliations of rays and equipotentials in $U' \smallsetminus \mathcal{K}(g)$. Under the straightening conjugacy they are mapped to the corresponding foliations for the polynomial f_c. Since all the conjugacies agree on the Julia set and homotopic rel $\mathcal{J}$ (see Corollary 41.9 below), the landing properties of the rays are independent of the particular choice of the tubing. In particular, the $\beta-$fixed point is always the landing point for the 0-ray (which gives an alternative way to specify this point).

40.5. Quadratic-like germs.

40.5.1. *Notion.* Let us now introduce a notion that will give us an appropriate flexibility in changing the domain of a ql map.

A ql map $g : W \to W'$ is called a *ql restriction* of a ql map $f : U \to U'$ if $W \subset U$ and $g = f|\,W$. For instance, $\tilde{f}$ can be obtained from f by an elementary adjustment of the domain (see §28.1.1).

Lemma 40.14. *If g is a ql restriction of f then $\mathcal{K}(g) = \mathcal{K}(f)$.*

Proof. Obviously, $\mathcal{K}(g) \subset \mathcal{K}(f)$. Since $\deg f = \deg g$, the set $\mathcal{K}(g)$ is completely invariant under f. The conclusion follows from Corollary 40.3 (iii). □

We say that two ql maps f and $\tilde{f}$ *represent the same quadratic-like germ* if there is a sequence of ql maps $f = f_0, f_1, \dots f_n = \tilde{f}$ such that any two consecutive ones, f_k and f_{k+1}, have a common quadratic-like restriction.

Lemma 40.14 implies that *any ql germ has a well defined filled Julia set.*

EXERCISE 40.15. *Let f and $\tilde{f}$ be two ql maps with connected Julia set. Then the following are equivalent:*

(i) *f and $\tilde{f}$ represent the same germ;*

(ii) *they have the same filled Julia set $\mathcal{K}$ and $f = \tilde{f}$ in a neighborhood of $\mathcal{K}$;*

(iii) *they have a common ql restriction.*

EXERCISE 40.16. *Let $f : U \to U'$ and $\tilde{f} : \tilde{U} \to \tilde{U}'$ be two ql maps coinciding near* 0 (*which is the critical point for both*). *Let W be the component of $U \cap \tilde{U}$ containing* 0. *If $W \ni f(0)$ then $f|W = \tilde{f}|W$ is a common ql restriction for these two maps.*

In particular, any quadratic polynomial f defines a ql germ represented by any ql restriction $f : f^{-1}(\mathbb{D}_R) \to \mathbb{D}_R$ with $R > |f(0)|$, or, more canonically, by any ql restriction $f : \Sigma(r) \to \Sigma(r^2)$ to a subpotential domain $\Sigma(r) \ni 0$.

40.5.2. *Renormalization of germs.* In §28.4 we introduced a notion of a renormalizable ql map and its pre-renormalization. Here we will revisit this notion preparing ground to viewing the renormalization $f \mapsto Rf$ as an analytic operator in some infinite-dimensional complex space (of quadratic-like germs considered up to rescaling).

Theorem 31.22 implies:

COROLLARY 40.17. *Let $f : U \to U'$ and $\tilde{f} : \tilde{U} \to \tilde{U}'$ be two ql maps with connected Julia set representing the same ql germ. If f is renormalizable with period p, then so is $\tilde{f}$, and the corresponding pre-renormalizations, g and $\tilde{g}$, represent the same ql germ.*

Thus, we can promote the pre-renormalization to the *renormalization* acting on ql germs. Moreover, it is natural to consider the ql germs *up to rescaling*, i.e., up to a conjugacy by $z \mapsto \lambda z$, $\lambda \in \mathbb{C}^*$.[1] For instance, one can normalize them so that the β-fixed point is placed at 1.

For any renormalization period p, this provides us with a well defined renormalization operator $R = R_p$.

The above discussion applies to quadratic polynomials by considering the corresponding quadratic-like germs.

Notes. As we already mentioned in Notes to Ch. IV, the notion of polynomial-like map was introduced by Douady and Hubbard in [**DH3**]. Basic theory of these maps, including the Straightening Theorem, was developed in the same paper. The proof of this theorem was historically the first application of the method of quasiconformal surgery (see §29.1.2).[2]

Interesection of domains of ql maps (§40.5.1) was discussed in [**McM1**, §5.4]. The notion of ql germ was introduced in [**L12**].

[1] Recall that 0 is assumed to be the critical point. Without this normalization, one should consider ql germs up to conjugacy by affine maps $z \mapsto \lambda z + a$.

[2] We distinguish here "qc surgery" from "qc deformation" that had been introduced ealier.

41. External structure

41.1. External map. Before passing to the uniqueness part of the Straightening Theorem, let us dwell on an important relation between quadratic-like and expanding circle maps.

41.1.1. *Holomorphic extensions of expanding circle maps.* Expanding circle maps were introduced in §19.13. The basic example is provided by the doubling map $f_0 : z \mapsto z^2$. More generally, in §25.3 we encountered Blaschke maps.

We let $\mathcal{E} \equiv \mathcal{E}_2$ be the class of degree two analytic expanding circle maps $g : \mathbb{T} \to \mathbb{T}$ considered up to conjugacy by a circle rotation. We can also view $\mathcal{E}$ as the class of analytic expanding circle maps g normalized so that $g(1) = 1$.

As we know from Exercise 19.74, any map $g \in \mathcal{E}$ admits a holomorphic extension $g : V \to V'$, where $V \Subset V'$ are two $\mathbb{T}$-symmetric open (for definiteness) annuli with smooth boundary. Let $A := (\overline{V}' \setminus V) \setminus \mathbb{D}$ be the *external fundamental annulus* for g.

Given another map $\tilde{g} : \tilde{V} \to \tilde{V}'$ as above, we will mark the corresponding objects with "tilde". This following statement is a complex version of Proposition 19.67:

PROPOSITION 41.1. *Any two expanding circle maps, $g : V \to V'$ and $\tilde{g} : \tilde{V} \to \tilde{V}'$, are conjugate by a $\mathbb{T}$-symmetric qc map $h : (V', V, \mathbb{T}) \to (\tilde{V}', \tilde{V}, \mathbb{T})$. In fact, any equivariant qc map $H : A \to \tilde{A}$ between the external fundamental annuli admits a unique extension to a qc conjugacy h as above. Moreover* $\mathrm{Dil}(h) = \mathrm{Dil}(H)$.

PROOF. Consider an equivariant qc map H as above with dilatation K. By Lemma 40.10 it can be uniquely lifted to an equivariant K-qc homeomorphism $h : V' \setminus \overline{\mathbb{D}} \to \tilde{V}' \setminus \overline{\mathbb{D}}$. By Corollary 13.9, h admits a continuous extension to the unit circle. Reflecting it to the interior of the disk and then exploiting the Gluing Lemma (see §13.3), we obtain the desired K-qc conjugacy $h : V' \to \tilde{V}'$. □

REMARK 41.2. a) Using the expanding property of g, one can justify without using Corollary 13.9 that $h \,|\, V \setminus \overline{\mathbb{D}}$ extends continuously to $\mathbb{T}$ (and in fact, that it is Hölder continuous): compare Theorem 25.4 and Lemma 41.5 below.

b) One can modify the above proof in the spirit of the Pullback Argument (see §38.5) as follows:

- Start with a $\mathbb{T}$-symmetric K_0-qc map $h_0 : (V', V) \to (\tilde{V}', \tilde{V})$ equivariant on ∂A (and hence on the $\mathbb{T}$-symmetric annulus) with $\mathrm{Dil}(h_0 \,|\, A) = K$;
- Lift it by the iterated g gluing the lifts to K_0-qc homeomorphisms $h_n : V \to \tilde{V}$ which are equivariant and K-qc on the $\overline{V} \setminus g^{-n}(V)$;
- Pass to a limit.

This turns h_0 to the desired K-qc conjugacy.

41.1.2. *Connected case.* To any quadratic-like map $f : U \to U'$ one can naturally associate an expanding circle map $g \in \mathcal{E}$ that captures the "external dynamics" of f. For this reason g is called the *external map* of f.

The construction is very simple if the Julia set $\mathcal{J}(f)$ is connected. In this case the domain $\mathbb{C} \setminus \mathcal{K}(f)$ is simply connected and can be conformally mapped onto the complement of the unit disk:

$$\psi : \mathbb{C} \setminus \mathcal{K}(f) \to \mathbb{C} \setminus \overline{\mathbb{D}}.$$

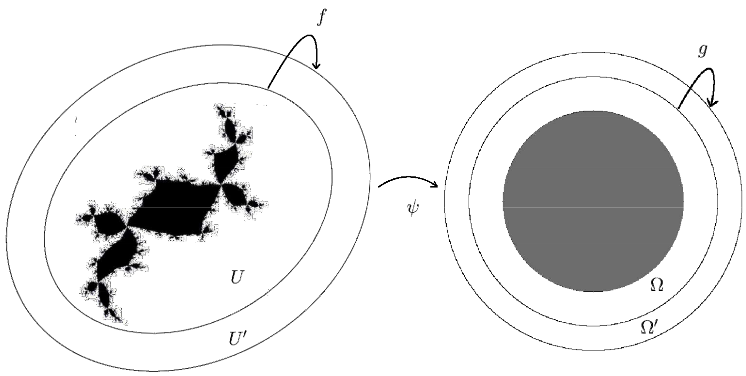

FIGURE 41.1. Construction of external map.

Let $\Omega = \psi(U \smallsetminus \mathcal{K}(f))$, $\Omega' = \psi(U' \smallsetminus \mathcal{K}(f))$. These are two conformal annuli with common inner boundary, the unit circle $\mathbb{T}$, and such that the outer boundary of Ω is contained in Ω'. Conjugating f by ψ we obtain a holomorphic double covering

$$g : \Omega \to \Omega', \quad g(z) = \psi \circ f \circ \psi^{-1}(z) \quad \text{for} \quad z \in \Omega.$$

By Lemma 5.7 and the Schwarz Reflection Principle, g can be extended to an expanding circle map.

Since the Riemann map ψ is defined up to post-composition with a rotation $\mathrm{R}_\theta : z \mapsto e(\theta)z$, $\theta \in \mathbb{R}/\mathbb{Z}$, the circle map g is defined up to conjugacy by R_θ, so it represents an element of the space $\mathcal{E}$. Acccording to our convention, we normalize g by putting its fixed point at $1 \in \mathbb{T}$.

Note also that if f is replaced by an affinely conjugate map $A^{-1} \circ f \circ A$, where $A : z \mapsto \lambda z$, $\lambda \in \mathbb{C}^*$, then the Riemann map ψ is replaced by $\psi \circ A$, and the external map g remains the same. Thus, to any quadratic-like map f (with connected Julia set) prescribed *up to affine conjugacy* corresponds an expanding circle map $g \in \mathcal{E}$ well-defined *up to conjugacy by rotation*.

41.1.3. *Genuine vs degenerate ql maps.* We will now apply the above construction to a slightly more general situation to give a criterion when a degenerate ql map can be restricted to a genuine one (quantitatively).

PROPOSITION 41.3. *Let $(f : U \to U')$ be a ql map, perhaps degenerate. Assume it has a completely invariant compact hull $\mathcal{K} \Subset U$. Then f restricts to a genuine ql map $W \to W'$ whose filled Julia set is $\mathcal{K}$. Moreover, if* $\operatorname{mod}(U' \smallsetminus \mathcal{K}) \geq \mu > 0$ *then the restriction can be selected so that* $\operatorname{mod}(W' \smallsetminus W) \geq \nu(\mu) > 0$.

PROOF. Note first that $0 \in \mathcal{K}$. Indeed, the map $f : U \smallsetminus \mathcal{K} \to U' \smallsetminus \mathcal{K}$ is a double branched covering between two annuli, so it must be unbranched by the Riemann-Hurwitz Formula.

Now, the construction of the external map can be applied to this situation ad verbuatim using the uniformlization $\psi : \mathbb{C} \smallsetminus \mathcal{K} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$. It provides as with a double covering $g : (V, \mathbb{T}) \to (V', \mathbb{T})$ between two $\mathbb{T}$-symmetric annuli $V \subset V'$.

Let $S' \subset V'$ be the hyperbolic neighborhood of $\mathbb{T}$ in V' of radius 1, and let $S := g^{-1}(S')$. Then S' is a $\mathbb{T}$-symmetric domain such that $g : S \to S'$ is a double covering. Since g strictly expands the infinitesimal hyperbolic metric of V', $g^{-1}(S) \Subset S'$.

Letting $W := \mathcal{K} \cup \psi^{-1}(S)$ and $W' := \mathcal{K} \cup \psi^{-1}(S')$, we obtain a desired ql restriction of f.

The last assertion follows from an observation that the fundamental annulus $S' \smallsetminus (S \cup \mathbb{D})$ depends continuously on g, and the latter varies within a compact family of maps (see Exercise 7.32). □

41.1.4. *General case.* In the case of a disconnected Julia set the construction is more subtle.

Take a fundamental annulus $A = \overline{U}' \smallsetminus U$ bounded by real analytic curves $E = \partial U'$ and $I = \partial U$. Then $f : I \to E$ is a real analytic double covering.

Let $\mu = \operatorname{mod} A$. Let us consider an abstract double covering $\xi_1 : A_1 \to A$ of an annulus A_1 of modulus $\mu/2$ over A. Let I_1 and E_1 be the "inner" and "outer" boundary of A_1, i.e., ξ_1 maps I_1 onto I and E_1 onto E. Then there is a real analytic diffeomorphism $\theta_1 : E_1 \to I$ such that $\xi_1 = f \circ \theta_1$. This allows us to stick the annulus A_1 to the domain $\mathbb{C} \smallsetminus U$ bounded by I. We obtain a Riemann surface $T_1 = (\mathbb{C} \smallsetminus U) \cup_{\theta_1} A_1$. Moreover, the maps f on I and ξ_1 on A_1 match to form an analytic double covering $f_1 : A_1 \to A$.

This map f_1 restricts to a real analytic double covering of the inner boundary of A_1 onto its outer boundary. This allows us to repeat this procedure: we can attach to the inner boundary of T_1 an annulus A_2 of modulus $\frac{1}{4}\mu$, and extend f_1 to the new annulus T_2. Proceeding in this way, we will construct a Riemann surface

$$T \equiv T^A(f) = \varinjlim T_n := (\mathbb{C} \smallsetminus U) \cup_{\theta_1} A_1 \cup_{\theta_2} A_2 \dots \tag{41.1}$$

and an analytic double covering $F : \bigcup_{n \geq 1} A_n \to \bigcup_{n \geq 0} A_n$ extending f.

The inner end of T can be represented by a puncture or by an ideal circle. But in the former case, after filling that puncture we would obtain a superattracting fixed point α (since the map F is a double covering near α). This would contradict the property that the trajectories of F are repelled from the inner end of T.

Thus, the inner end of T is a circle. Hence T can be uniformized by $\mathbb{C} \smallsetminus \mathbb{D}$ (with the inner ideal boundary uniformized by the unit circle $\mathbb{T}$). Now by the Reflection Principle, this conformal representation of F can be extended to an analytic expanding endomorphism $g \equiv g_A : \mathbb{T} \to \mathbb{T}$.

For a given choice of the fundamental annulus A, the map $g_A : V \to V'$ (which comes together with the domains (V, V')) is well-defined up to rotation. Indeed, by construction, for two such maps g_A and $\tilde{g}_A$ there is a conformal isomorphism $h : \mathbb{C} \smallsetminus \mathbb{D} \to \mathbb{C} \smallsetminus \mathbb{D}$ conjugating them on an outer neighborhood of the circle. Reflecting h to the unit disk, we conclude that h is a rotation conjugating g_A and $\tilde{g}_A$ near the circle.

Exercise 41.4. (i) *Show that in the connected case this construction leads to the same result as the construction of* §41.1.2.

(ii) *Show that the external map is equal to* $z \mapsto z^2$ *iff* f *is a restricted quadratic polynomial* f_c.

41.2. External automorphisms. Let us say that h is an *external automorphism* of an analytic expanding circle map $g : (V, \mathbb{T}) \to (V', \mathbb{T})$ if h is a homeomorphism between external neighborhoods, $h : \Omega \to \Omega'$, commuting with g.

LEMMA 41.5. *Let h be an orientation preserving external automorphism of an expanding circle map $g \in \mathcal{E}$. Then h admits a continuous extension to the circle $\mathbb{T}$ as* id.

PROOF. (Compare Proposition 19.58.)

Since by Proposition 41.1 all expanding cicle maps of class $\mathcal{E}$ are topologically conjugate near $\mathbb{T}$, we can assume without loss of generality that $g = f_0 : z \mapsto z^2$.

Let us begin with the real ray $\mathcal{R}^0 = \mathbb{R}_+ \smallsetminus \overline{\mathbb{D}}$ (or rather: its germ near $\mathbb{T}$). Since h is an automorphism, the image $\Gamma^0 := h(\mathcal{R}^0)$ (taken near $\mathbb{T}$) is a g-invariant (in the sense of germs) curve. By the standard argument (see Theorem 24.3), it must land at the g-fixed point $1 \in \mathbb{T}$. Thus, h continuously extends to 1 along this ray, and $h(1) = 1$.

Let us now consider the respective lifts $\mathcal{R}^{1/2}$, $\Gamma^{1/2}$ of $\mathcal{R}^0$, Γ^0 under f landing at $-1 \in \mathbb{T}$. Since h commutes with g, it maps $\mathcal{R}^{1/2}$ to $\Gamma^{1/2}$. Thus, h constinuously extends to -1 along $\mathcal{R}^{1/2}$ and $h(-1) = -1$.

Let us now consider dyadic rectangles Δ^1_0 and Δ^1_1 attached to $\mathbb{T}$ obtained by cutting an annulus $\mathbb{A}(1, r]$ (with $r > 1$ sufficiently close to 1) by the rays $\mathcal{R}^0$ and $\mathcal{R}^{1/2}$ (see §23.6.2). They are mapped by h to respective topological recatngles Π^1_0 and Π^1_1 obtained by cutting the topological annulus $h(\mathbb{A}(1, r])$ by the rays Γ^0 and $\Gamma^{1/2}$. Moreover, since h is orientation preserving, Π^1_0 is attached to the upper semicircle, while Π^1_1 is attached to the lower one, so h "preserves" the upper and lower external neighborhoods of the circle.

Let us now consider rays $\mathcal{R}^{1/4}$ and $\mathcal{R}^{3/4}$, which are the lifts of $\mathcal{R}^{1/2}$. Respectively, consider curves $\Gamma^{1/4}$ and $\Gamma^{3/4}$ which are lifts of the curve $\Gamma^{1/2}$. Since the automorphism h preserves the upper and lower external neighborhoods of the circle, it maps $\Gamma^{k/4}$ to the corresponding $\mathcal{R}^{k/4}$, $k \in \{1, 3\}$. It follows that h constinuously extends along the rays $\mathcal{R}^{k/4}$ to points $e(k/4) \in \mathbb{T}$ fixing both of them. Moreover, it maps the dyadic rectanges $\Delta^2_{i_0,i_1}$ to corresponding topological rectangles $\Pi^2_{i_0 i_1}$ attached to the same circle arcs.

Proceeding this way, we will conclude that h continuously extends to all dyadic points $e(\mathfrak{p}/2^n) \in \mathbb{T}$ mapping the dyadic rectangles $\Delta^n_{i_0 \dots i_{n-1}}$ to topolgical rectangles $\Pi^n_{i_0 \dots i_{n-1}}$ attached to the same dyadic arcs.

Finally, since the diameters of these rectangles go to 0 as $n \to 0$, the map h, extended as id to the circle, is continuous. □

LEMMA 41.6. *Under the circumstances of* Lemma 41.5*, the automorphism h moves points by a bounded hyperbolic distance:*

$$\rho_{\mathbb{C} \smallsetminus \overline{\mathbb{D}}}(h(z), z) \le R \quad \forall\, z \text{ near } \mathbb{T}.$$

PROOF. Since the hyperbolic distances on $\mathbb{C} \smallsetminus \overline{\mathbb{D}}$ and on $V' \smallsetminus \overline{\mathbb{D}}$ are comparable near $\mathbb{T}$ (see Exercise 7.13), we can work with the latter. By Corollary 7.11, the map $g : V \smallsetminus \overline{\mathbb{D}} \to V' \smallsetminus \overline{\mathbb{D}}$ expands this distance. It follows that for any smooth curve γ in $V' \smallsetminus \overline{\mathbb{D}}$, any lift γ_i^{-1} $(i = 0, 1)$ of it by g is hyperbolically shorter than γ.

By Lemma 41.5, any point $v \in V'$ near $\mathbb{T}$ can be connected to $\zeta := h(v) \in V'$ by a Euclideanly short curve γ. Then the lifts γ_i^{-1} are also Euclideanly short, while stay a definite Euclidean distance apart. Aplying Lemma 41.5 once again, we see that $h(\gamma_i^{-1}) \cap \gamma_{1-i}^{-1} = \emptyset$ for both $i = 0, 1$. On the other hand, by the equivariance of h, we have $h(v_i^{-1}) = \zeta_k^{-1}$ for some $k = k(i) \in \{0, 1\}$. (Here v_i^{-1} and ζ_i^{-1} are naturally labeled endpoints of γ_i^{-1}.) We conclude that $k(i) = i$, so γ_i is bounded by v_i^{-1} and $h(v_i^{-1})$. It follows that

$$\rho_{\text{hyp}}(h(v_i^{-1}), v_i^{-1}) \le \rho_{\text{hyp}}(v, h(v)).$$

Thus, there exists and external annulus $\Omega_\circ \subset V' \smallsetminus \overline{\mathbb{D}}$ such that for any point $z \in \Omega_\circ$ we have:

$$\rho_{\text{hyp}}(h(z), z) \le \rho_{\text{hyp}}(gz, h(gz)). \tag{41.2}$$

Take now an external fundamental annulus $S := g^{-n}(A)$ compactly contained in $\Omega_\circ$, and let $\Omega \subset \Omega_\circ$ be the external annulus bounded by $\mathbb{T}$ and $\partial^i S$. Since $S \Subset \Omega_\circ$, we have

$$\rho_{\text{hyp}}(z, h(z)) \le R \quad \forall\, z \in S.$$

At the same time, for any $z \in \Omega$ there is an $m \in \mathbb{Z}_+$ such that $z_m \equiv g^m z \in S$, while $z_k \in \Omega$ for $k = 0, \dots, m-1$. Then (41.2) applies to all iterates z_k, $k = 0, \dots, m-1$, yielding

$$\rho_{\text{hyp}}(z, h(z)) \le \rho_{\text{hyp}}(z_m, h(z_m)) \le R,$$

as asserted. □

Lemma 41.7. *Under the circumstance of the above lemmas, the automorphism h is homotopic to* id *near $\mathbb{T}$ through a family of maps commuting with g.*

Proof. Again, without loss of generality we can assume that $g = f_0$. Let us consider the universal covering $e^* : \overline{\mathbb{H}} \to \mathbb{C} \smallsetminus \mathbb{D}$, $e^*(z) = e(-z)$, and let us lift h to a homeomorphism $\hat{h} : \overline{\mathbb{H}} \to \overline{\mathbb{H}}$ equal to id on $\mathbb{R}$. Then $\hat{h}$ commutes with the translation $L : \zeta \mapsto \zeta + 1$ (a deck transformation for e^*) and the doubling $T : \zeta \mapsto 2\zeta$ (a lift of f_0).

Let us now consider a linear homotopy connecting $\hat{h}$ to id:

$$\hat{h}_t(z) = \hat{h}(z) + t(z - \hat{h}(z)), \quad 0 \le t \le 1.$$

It commutes with both L and T and hence descends to a homotopy h_t commuting with f_0. □

41.3. Uniqueness of the straightening.

41.3.1. *Connected case.* Let us first show that an "external automorphism" of a quadratic-like map admits a continuous extension to the Julia set by the identity.

Lemma 41.8. *Let $f : U \to U'$ be a quadratic-like map with connected Julia set. Let $W \subset U$ and $W' \subset U$ be two (open) annuli whose inner boundary is $\mathcal{J}(f)$. Let $h : W \to W'$ be an orientation preserving automorphism of f. Then h admits a continuous extension to a map $W \cup \mathcal{J}(f) \to W' \cup \mathcal{J}(f)$ which is the identity on the Julia set.*

Proof. Consider the Riemann mapping $\psi : \mathbb{C} \smallsetminus \mathcal{K}(f) \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ and the external circle map $g : V \to V'$, $g | V \smallsetminus \overline{\mathbb{D}} = \psi \circ f \circ \psi^{-1}$. Transfer the annuli W and W' to the g-plane. We obtain two annuli $\Omega = \psi(W)$ and $\Omega' = \psi(W')$ in $V \smallsetminus \overline{\mathbb{D}}$ attached to the unit circle $\mathbb{T}$. The homeomorphism $k : \Omega \to \Omega'$, $k = \psi \circ h \circ \psi^{-1}$, commutes with g.

By Lemma 41.6, the map k moves points near $\mathbb{T}$ by a bounded hyperbolic distance:

$$\rho_{\mathbb{C}\smallsetminus\overline{\mathbb{D}}}(k(z), z) \le R.$$

Since the Riemann mapping $\psi : \mathbb{C} \smallsetminus \mathcal{K}(f) \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ is a hyperbolic isometry, the same is true for h:

$$\rho_{\mathbb{C}\smallsetminus\mathcal{K}(f)}(h(z), z) \le R \quad \text{for } z \in W \text{ near } \mathcal{J}(f).$$

By Proposition 7.5, the Euclidean distance $|z - h(z)|$ goes to 0 as $z \to \mathcal{J}(f)$. It follows that the extension of h by the identity to the Julia set is continuous. □

Corollary 41.9. *Let f and $\tilde{f}$ be two quadratic-like maps, and let h and h' be two orientaton preserving homeomorphisms conjugating f to $\tilde{f}$ in some neighborhoods of the Julia sets. Then $h = h'$ on $\mathcal{J}(f)$. Moreover, h is homotopic to h' on an exterior neighborhood real $\mathcal{J}(f)$ through a family of conjugacies between f and $\tilde{f}$.*

Proof. For the last statement use Lemma 41.7. □

Let us now put together the above results:

Theorem 41.10. *Let us consider two quadratic-like maps $f : U \to U'$ and $\tilde{f} : \tilde{U} \to \tilde{U}'$ with connected Julia sets. Assume that they are topologically conjugate near their Julia sets by an orientation preserving homeomorphism $\psi : V \to \tilde{V}$. Assume also that we are given an equivariant homeomorphism $H : A \to \tilde{A}$ between the (closed) fundamental annuli of f and $\tilde{f}$.*

Then there exists a unique homeomorphism $h : U' \to \tilde{U}'$ conjugating f to $\tilde{f}$, coinciding with ψ on the Julia set $\mathcal{J}(f)$, and coinciding with H on A.

If H is qc, then $h|U \smallsetminus \mathcal{K}(f)$ is also qc with the same dilatation. If both H and ψ are qc, then h is qc, and

$$\mathrm{Dil}(h) = \max(\mathrm{Dil}\, H,\, \mathrm{Dil}(\psi|\, \mathcal{K}(f)).$$

In particular, if ψ is a hybrid equivalence, then $\mathrm{Dil}(h) = \mathrm{Dil}(H)$.

Proof. By the Lifting Construction of Corollary 40.11, H admits a unique equivariant extension to a homeomorphism $h : U \smallsetminus \mathcal{K}(f) \to \tilde{U} \smallsetminus \mathcal{K}(\tilde{f})$. This extension continuously matches with ψ on the filled Julia set. Indeed, $\psi^{-1} \circ h$ commutes with f on some external neighborhood of $\mathcal{K}(f)$. By Lemma 41.8, this map continuously extends to the filled Julia set as the identity. Hence h continuously extends to the filled Julia set as ψ.

If H is qc then $h|\, U \smallsetminus \mathcal{K}(f)$ is qc with the same dilatation (Corollary 40.11). All the rest follows from Bers' Gluing Lemma (see §13.3). □

Of course, we can always construct an equivariant qc map H between the fundamental annuli. Hence if two quadratic-like maps are topologically equivalent, then the conjugacy can be selected to be quasiconformal outside the filled Julia set. If they are hybrid equivalent, then the dilatation of the conjugacy is completely controlled by the dilatation of H, which is in turn controlled by the geometry of the fundamental annuli (see Lemma 40.13). In the polynomial case we can do even better:

COROLLARY 41.11. *Consider two quadratic polynomials $f : z \mapsto z^2 + c$ and $\tilde{f} : z \mapsto z^2 + \tilde{c}$ with connected Julia sets. If they are topologically conjugate near their filled Julia sets by an orientation preserving map h_0, then there exists a unique global conjugacy $h : \mathbb{C} \to \mathbb{C}$ that coincides with h_0 on $\mathcal{K}(f)$ and is conformal on the basin of ∞. If h_0 is qc then so is h, and $\operatorname{Dil} h = \operatorname{Dil}(h_0 | \mathcal{K}(f))$. If h_0 is hybrid then $h = \mathrm{id}$ and $f = \tilde{f}$.*

PROOF. By Theorem 23.25, the Riemann-Böttcher map $B_f : \mathcal{D}_f(\infty) \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ conjugates f to $z \mapsto z^2$, and similarly for $\tilde{f}$. Hence the Böttcher conjugacy

$$H = B_{\tilde{f}}^{-1} \circ B_f : \mathcal{D}_f(\infty) \to \mathcal{D}_{\tilde{f}}(\infty) \tag{41.3}$$

conformally conjugates f to $\tilde{f}$ on their basins of ∞. By the previous theorem, this conjugacy matches with the topological conjugacy on the filled Julia set giving us a desired global conjugacy h.

The qc part of the statement follows from the same theorem. Moreover, if f and $\tilde{f}$ are hybrid equivalent, then $\operatorname{Dil}(h) = 0$ a.e. on $\mathbb{C}$. By Weyl's Lemma, h is conformal and hence affine. As $h(0) = 0$ and $h(z) \sim z$ near ∞, we conclude that $h = \mathrm{id}$.

The uniqueness of h follows from the fact that id is the only conformal automorphism $\mathbb{C} \smallsetminus \overline{\mathbb{D}} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ commuting with $z \mapsto z^2$ (and hence (41.3) is the only conformal isomorphism $\mathcal{D}_f(\infty) \to \mathcal{D}_{\tilde{f}}(\infty)$ conjugating f to $\tilde{f}$). □

The last statement of the above Corollary gives the uniqueness part of the Straightening Theorem in the connected case.

41.3.2. *Disconnected case.*

PROPOSITION 41.12. *For a ql map g with disconnected Julia set, the tubing position of the critical value, $T_g(g(0))$, determines the straightening $f_c : z \mapsto z^2 + c$.*

PROOF. By (40.6), the tubing position of the critical value for g is equal to the Böttcher position $B_c(c)$ of the critical value for f_c. But by Theorem 34.1, the latter is equal to the Riemann position $\Psi_{\mathcal{M}}(c)$ of the parameter c. (Recall that $\Psi_{\mathcal{M}} : \mathbb{C} \smallsetminus \mathcal{M} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$ is the Riemann mapping for the complement of the Mandelbrot set.) As $\Psi_{\mathcal{M}}(c)$ determines c, the conclusion follows. □

41.4. Mating of f_c with $g \in \mathcal{E}$. We have associated to any quadratic-like map f with connected Julia set its straightening $f_c : z \mapsto z^2 + c$ and its external map $g : \mathbb{T} \to \mathbb{T}$. Recall that quadratic-like maps are considered up to affine conjugacy, while expanding circle maps are considered up to rotation. We normalize these maps so that

$$f(z) = c + z^2 + \text{h.o.t.}$$

near the origin, while g fixes 1. Now we will reverse the above construction producing a *mating* between f_c and g:

PROPOSITION 41.13. *Given a parameter $c \in \mathcal{M}$ and an expanding circle map $g : (V, \mathbb{T}) \to (V', \mathbb{T})$, there exists a unique quadratic-like map f (up to affine conjugacy) such that f_c and g are the straightening and the external map of f, respectively.*

PROOF. The proof is similar to the proof of the Straightening Theorem, so we will just sketch it.

Existence. Let us consider a (conventional) quadratic-like restriction $f_c : U \to U'$ of our quadratic polynomial (e.g., we can select U as a disk bounded by by some equipotential of f_c). Take some equivariant diffeomorphism $h_0 : \overline{U'} \smallsetminus U \to \overline{V'} \smallsetminus V$ and extend it by Lemma 40.10 to an equivariant qc map $h : \overline{U'} \smallsetminus \mathcal{K}(f) \to \overline{V'} \smallsetminus \overline{\mathbb{D}}$. Now glue two hemi-spheres $S^2_0 := U'$ and $S^2_\infty := \hat{\mathbb{C}} \smallsetminus \overline{\mathbb{D}}$ by means of h to obtain a qc sphere S^2. Define a map

$$F : U \sqcup_h (V \smallsetminus \overline{\mathbb{D}}) \to U' \sqcup_h (V' \smallsetminus \overline{\mathbb{D}})$$

as f_c on $U \subset S^2_0$ and as g on $V \smallsetminus \overline{\mathbb{D}} \subset S^2_\infty$. It is a well defined quasiregular double branched covering. Moreover, it preserves the conformal structure μ which is standard on $\mathcal{K}(f) \subset S^2_0$ and on S^2_∞. By means of the Measurable Riemann Mapping Theorem, F can be turned into the desired quadratic-like map f.

Uniqueness. Assume that two quadratic-like maps $f : U \to U'$ and $\tilde{f} : \tilde{U} \to \tilde{U}'$ have the same straightenings and the same normalized external maps. Then they are hybrid conjugate by a qc map $h : U' \to \tilde{U}'$ near their filled Julia sets, and are conformally conjugate by a map $\phi : \mathbb{C} \smallsetminus \mathcal{K}(f) \to \mathbb{C} \smallsetminus \mathcal{K}(\tilde{f})$. By Theorem 41.10, these two conjugacies match on the Julia set and glue together into a global conformal (and hence affine) map $\mathbb{C} \to \mathbb{C}$. $\square$

Notes. External maps were constructed by Douady and Hubbard [**DH3**]. Mating of quadratic polynomials with external maps (§41.4) appeared in the same paper. Note, however, that the discussion in [**DH3**] was carried on the level of *conformal* rather than *affine* classes. The affine adjustment, more suitable for the renormalization theory, appeared in [**L12**].

Proposition 41.3 on ql restrictions is due to McMullen ([**McM2**, Prop. 4.10]).

42. Quadratic-like families

42.1. Definitions. Let $\Lambda \subset \mathbb{C}$ be a domain in the complex plane. A *DH quadratic-like family*[3] $\mathbf{g}$ over Λ is a family of quadratic-like maps $g_\lambda : U_\lambda \to U'_\lambda$ depending on $\lambda \in \Lambda$ such that:

- The *tube* $\mathbb{U} = \{(\lambda, z) : \ \lambda \in \Lambda, \ z \in U_\lambda\}$ is a domain in $\mathbb{C}^2$;
- $g_\lambda(z)$ is holomorphic in two variables on $\mathbb{U}$.

As usual, we assume that the critical point of each f_λ is located at the origin, and that U_λ and U'_λ are 0-symmetric quasidisks.

We will now make several additional assumptions. The first of them is minor. We say that $\mathbf{g}$ *extends beyond* $\mathbb{U}$ if there exists a domain $\Lambda' \Supset \Lambda$ and a quadratic-like family $G_\lambda : V_\lambda \to V'_\lambda$ over Λ' such that for $\lambda \in \Lambda$, g_λ is an adjustment (see §§28.1.1, 40.1) of G_λ.

We call a quadratic-like family $\mathbf{g} : U_\lambda \to U'_\lambda$ over Λ *proper* if

- $\mathbf{g}$ admits an extension beyond $\mathbb{U}$;
- For $\lambda \in \partial\Lambda$, $g_\lambda(0) \in \partial U'_\lambda$.

Obviously $g_\lambda(0) \neq 0$ for $\lambda \in \partial\Lambda$, so assuming that Λ *is a Jordan disk*, we have a well defined *winding number* of the curve $\lambda \mapsto g_\lambda(0), \ \lambda \in \partial\Lambda$, around 0. We call it the winding number of $\mathbf{g}$ and denote $w(\mathbf{g})$. A proper family $\mathbf{g}$ is called *unfolded* if $w(\mathbf{g}) = 1$. By the Argument Principle, any proper unfolded quadratic-like family

[3] Here "DH" stands for "Douady-Hubbard". Later on (see §47.3), "generalized ql families" will appear. Until then, there is no ambiguity, and "ql family" will stand for "DH ql family".

has a unique parameter value $\lambda_\circ$ such that $f_\circ \equiv f_{\lambda_\circ}$ has a superattracting fixed point, i.e., $f_\circ(0) = 0$. We will select $\lambda_\circ$ as the *base point* in Λ.

Finally, we want the fundamental annulus $A_\lambda = \overline{U'_\lambda} \smallsetminus U_\lambda$ of g_λ to move holomorphically with λ. So, assume that there is an equivariant holomorphic motion $h_\lambda : A_\circ \to A_\lambda$, i.e., such that

$$h_\lambda(g_\circ z) = g_\lambda(h_\lambda(z)) \quad \text{for} \quad z \in \partial U_\circ.$$

Moreover, we will make a technical

Assumption H: The motion of any compact subset $Q \subset \overline{U'_\circ} \smallsetminus \overline{U}_\circ$ extends to a slightly bigger disk $\Lambda_Q \Supset \Lambda$.

Remark 42.1. Note that the motion of $\partial U_\circ$ cannot be extended beyond Λ since for $\lambda \in \partial\Lambda$ the boundary curve ∂U_λ *pinches* at the critical point 0 (becoming a figure-eight curve).

Denote this holomorphic motion by $\mathbf{h}$. We say that the quadratic-like family $\mathbf{g}$ is *equipped* with the holomorphic motion $\mathbf{h}$. Sometimes we will use notation $(\mathbf{g}, \mathbf{h})$ for an equipped quadratic-like family.

For equipped families, there is a *natural choice of tubing* (see §40.3) continuously depending on λ. Namely, select any tubing $T_\circ : A_\circ \to \mathbb{A}[r, r^2]$ for the base point, and then let

$$T_\lambda = T_\circ \circ h_\lambda^{-1}. \tag{42.1}$$

These are tubings since the holomorphic motion h_λ is equivariant.

The *Mandelbrot set* of the quadratic-like family is defined as

$$\mathcal{M}(\mathbf{g}) = \{\lambda \in \Lambda : \ \mathcal{J}(g_\lambda) \text{ is connected}\}.$$

If $\mathbf{g}$ is proper, then $\mathcal{M}(\mathbf{g})$ is compactly contained in Λ.

Let us finish with a few terminological and notational remarks. Let $\pi : \mathbb{C}^2 \to \mathbb{C}$ stand for the projection onto the first coordinate. We call a set $\mathbb{U} \subset \mathbb{C}^2$ a *tube over* $\Lambda = \pi(\mathbb{U}) \subset \mathbb{C}$ if it is a fiber bundle over Λ whose *fibers* $U_\lambda := \mathbb{U} \cap \pi^{-1}(\lambda)$ are Jordan disks (either open or closed). For $X \subset \Lambda$, we let $\mathbb{U}|X = \mathbb{U} \cap \pi^{-1}X$.

42.2. Restricted quadratic family. In this section we will show that the quadratic family $(f_c)_{c\in\mathbb{C}}$ can be naturally restricted to a proper unfolded equipped quadratic-like family.

Fix some $r > 1$. Restrict the parameter plane $\mathbb{C}$ to the subpotential disk $\Sigma_{\text{par}} \equiv \Sigma_{\text{par}}(r^2)$ bounded by the parameter equipotential of radius r^2 (see §34.2). According to formula (34.1), this parameter domain is specified by the property that

$$c = f_c(0) \in \Sigma_c(r^2) \equiv \Sigma'_c\,.$$

(Recall that $\Sigma_c(\rho)$ stands for the dynamical subpotential disk of radius ρ, see §23.5.4). Hence for $c \in \Sigma_{\text{par}}$, f_c restricts to a quadratic-like map $f_c : \Sigma_c \to \Sigma'_c$, where $\Sigma_c \equiv \Sigma_c(r)$. These quadratic-like maps obviously form a quadratic-like family over Σ_{par}, which we will call a *restricted quadratic family*.

The restricted quadratic family is proper. The first property of the definition is obvious. The main property, $f_c(0) \in \partial\Sigma'_c$ for $c \in \partial\Sigma_{\text{par}}$, follows from formula (34.1). The winding number of this family is equal to 1. Indeed, when the *parameter c* runs once along the boundary $\partial\Sigma_{\text{par}}$, the *critical value* $c = f_c(0)$ runs once around $0 \in \Sigma_{\text{par}}$.

The restricted quadratic family is equipped with the Böttcher motion (see §34.4) of the fundamental annulus. Select 0 as the base point in $\Sigma_{\rm par}$ and let

$$B_c^{-1} : \mathbb{A}[r, r^2] \to \overline{\Sigma_c'} \smallsetminus \Sigma_c \tag{42.2}$$

(note that $\mathbb{A}[r, r^2] = \overline{\Sigma_0'} \smallsetminus \Sigma_0$). Since the Böttcher function $B_c(z)$ is holomorphic it two variables (see §23.6.3) $\{B_c^{-1}\}_{c\in\Sigma_{\rm par}}$ is a biholomorphic motion.

Finally, note that for any slightly smaller annulus $\mathbb{A}[\rho, r^2]$, $\rho > r$, the Böttcher motion (42.2) extends to a slightly bigger subpotential domain, $\Sigma_{\rm par}(\rho^2) \Supset \Sigma_{\rm par}$.

Thus, the restricted quadratic family satisfies all the properties required for an equipped proper unfolded quadratic-like family.

42.3. Straightening of quadratic-like families. The Mandelbrot set $\mathcal{M}(\mathbf{g})$ of any quadratic-like family $\mathbf{g}$ can be canonically mapped to the genuine Mandelbrot set $\mathcal{M}$. Namely, by the Straightening Theorem, for any $\lambda \in \mathcal{M}(\mathbf{g})$ there is a unique quadratic polynomial $f_{c(\lambda)} : z \mapsto z^2 + c(\lambda)$, $c(\lambda) \in \mathcal{M}$, which is hybrid equivalent to g_λ. The map $\chi : \lambda \mapsto c(\lambda)$ is called the *straightening* of $\mathcal{M}(\mathbf{g})$.

We know that the straightening is not canonically defined outside the Mandelbrot set but rather depends on the choice of the tubing. But for equipped families there is a natural choice given by (42.1). With this choice, the straightening χ admits an extension to the whole parameter domain Λ, which will still be denoted by χ.

We can now formulate a fundamental result of the theory of quadratic-like families:

THEOREM 42.2. *Let* $\mathbf{g}$ *be a proper unfolded equipped quadratic-like family over* Λ. *Endow it with the natural tubing* (42.1). *Then the corresponding straightening* χ *is a homeomorphism from* Λ *onto* $\Sigma_{\rm par}$ *mapping* $\mathcal{M}(\mathbf{g})$ *onto* $\mathcal{M}$.

The proof of this theorem will be split into several pieces each of which is interesting in its own right.

42.4. The critical value moves transversally to h. We say that a holomorphic curve $\Gamma \subset \mathbb{C}^2$ is a *global transversal* to a holomorphic motion $\mathbf{h}$ if it transversally intersects each leaf of $\mathbf{h}$ at a single point.

LEMMA 42.3. *Under the assumptions of* Theorem 42.2, *the graph of the function* $\lambda \mapsto g_\lambda(0)$, $\lambda \in \Lambda$, *is a global transversal to the holomorphic motion* $\mathbf{h}$ *on* $\mathbb{U}' \smallsetminus \mathbb{U}$.

We will also express it by saying that *the critical value moves transversally to* $\mathbf{h}$. The moral of this lemma is that *in the complex setting the transversality can be achieved for purely topological reasons.*

PROOF. Take a point $z \in A_\circ = \overline{U_\circ'} \smallsetminus U_\circ$ and consider its orbit

$$\psi_z : \lambda \mapsto h_\lambda(z)$$

under the motion $\mathbf{h}$. By Assumption H of §42.1, for $z \in \overline{U_\circ'} \smallsetminus \overline{U}_\circ$ the function ψ_z admits a holomorphic extension to a slightly bigger parameter domain $\Lambda_z \Supset \Lambda$. For $z \in \partial U_\circ$, equivariance equation

$$f_\lambda(\psi_z(\lambda)) = \psi_{f_\circ z}(\lambda)$$

implies that ψ_z admits an extension to the domain $\Lambda_{f_\circ z} \Supset \Lambda$ (note that $f_\circ z \in \partial U_\circ'$) as a multiply valued holomorphic function with only algebraic singularities. Such a function is continuous up to the boundary of Λ.

Thus, for any $z \in A_\circ$, the function ψ_z admits a continuous extension to $\overline{\Lambda}$. Moreover,

$$\psi_z(\lambda) \in U'_\lambda \text{ for any } z \in U'_\circ \setminus U_\circ \text{ and } \lambda \in \overline{\Lambda}. \tag{42.3}$$

For $z \in U_\circ \setminus \overline{U}_\circ$, this follows immediately from Assumption H. To see it for $z \in \partial U_\circ$, let us take any intermediate Jordan disk, $U_\circ \Subset W_\circ \Subset U'_\circ$, and let W_λ be the Jordan disk bounded by $h_\lambda(\partial W_\circ)$, $\lambda \in \overline{\Lambda}$. Then we have:

- $W_\lambda \Subset U'_\lambda$ for any $\lambda \in \overline{\Lambda}$ (by Assumption H);
- $\psi_z(\lambda) \in W_\lambda$ for any $\lambda \in \Lambda$, and by continuity, $\psi_z(\lambda) \in \overline{W}_\lambda$ for $\lambda \in \partial\Lambda$,

implying (42.3).

Fix now some $z \in U'_\circ \setminus U_\circ$ and let $\psi \equiv \psi_z$. Since the tube $\mathbb{V} \equiv \mathbb{U}'|\,\partial\Lambda$ is homeomorphic to the solid torus $\partial\Lambda \times \mathbb{D}$ over $\partial\Lambda$, the curve $\lambda \mapsto \psi(\lambda)$, $\lambda \in \partial\Lambda$ (which is a "parallel" of the torus $\mathbb{V}$) is homotopic to the "core" $\lambda \mapsto 0$ of this torus, i.e., these two curves can be joined by a continuous family of curves

$$\psi^t : \partial\Lambda \to \mathbb{V}, \quad 0 \le t \le 1.$$

Consider now the curve

$$\phi : \lambda \mapsto g_\lambda(0), \quad \lambda \in \partial\Lambda. \tag{42.4}$$

Since $\mathbf{g}$ is proper, $\phi(\lambda) \in \partial V_\lambda$. Hence $\phi(\lambda) - \psi^t(\lambda) \neq 0$ for any $t \in [0,1]$, $\lambda \in \partial\Lambda$. It follows that the curves $\lambda \mapsto \phi(\lambda) - \psi(\lambda)$ and $\lambda \mapsto \phi(\lambda)$, $\lambda \in \partial\Lambda$, have the same winding number around 0. But the latter number is equal to 1, since $\mathbf{g}$ is unfolded. Hence the former number is equal to 1 as well. By the classical Argument Principle, the graphs of the functions ϕ and ψ have a single transverse intersection, as asserted. □

42.5. External uniformization. In this section we will construct a dynamical (quasiconformal) uniformization of $\Lambda \setminus \mathcal{M}(\mathbf{g})$ which generalizes the uniformization of $\mathbb{C} \setminus \mathcal{M}$ described in §§34.1, 34.2). This construction provides us with an illustration of the *Phase-Parameter Relation* between the parameter and dynamical planes by means of holomorphic motions (compare §17.5). See Figure 42.5.

Let us consider a set $P^0 = \{\lambda \in \Lambda : \ g_\lambda(0) \in U'_\lambda \setminus U_\lambda\}$ (i.e., the set of parameters for which the critical point escapes under the first iterate through the half-closed fundamental annulus $A^0_\lambda := U'_\lambda \setminus U_\lambda$). Since $\mathbf{g}$ is proper, all points in Λ sufficiently close to $\partial\Lambda$ belong to P^0. We will show that P^0 is an annulus naturally homeomorphic to the base fundamental annulus $A^0_\circ$.

To this end consider the graph of the function $\phi : \lambda \mapsto g_\lambda(0)$,

$$\Gamma = \{(\lambda, z) \in \mathbb{C}^2 : \ \lambda \in \Lambda, \ z = g_\lambda(0)\}.$$

By Lemma 42.3, this graph is a global transversal to the holomorphic motion $\mathbf{h}$ of $A^0_\circ$. Hence there is a well defined holonomy $\gamma^0 : A^0_\circ \to \Gamma$ along the leaves of $\mathbf{g}$, and it maps $A^0_\circ$ homeomorphically onto a topological annulus $B^0 \subset \Gamma$ (compare §17.4.2). Obviously, $\pi(B^0) = P^0$. Altogether, we have a homeomorphism $\pi \circ \gamma^0$ from $A^0_\circ$ onto P^0. It follows, in particular that P^0 is a topological annulus, whose inner boundary is a Jordan curve $\pi \circ \gamma^0(\partial U_\circ)$ in Λ and the outer boundary is $\partial\Lambda$.

Let us consider the domain $\Lambda^1 = \Lambda \setminus P^0$. The restriction of our quadratic-like family to this parameter domain is not proper any more. To restore this property, we have to restrict the dynamical domains as well. Let $U^1_\lambda = g^{-1}_\lambda U_\lambda$. For any $\lambda \in \Lambda^1$,

$g_\lambda(0) \in U_\lambda$; hence U^1_λ is a topological disk and $g_\lambda : U^1_\lambda \to U_\lambda$ is a quadratic-like map. This gives us a quadratic-like family over Λ^1.

It is proper since by construction $g_\lambda(0) \in \partial U_\lambda$ for $\lambda \in \partial\Lambda^1$. It has winding number 1 since the function $\phi : \lambda \mapsto g_\lambda(0)$ does not have zeros in the annulus $\overline{P}^0$. It follows that the boundary curves $\phi : \partial\Lambda \to \mathbb{C}^*$ and $\phi : \partial\Lambda^1 \to \mathbb{C}^*$ are homotopic (after parameterizing $\partial\Lambda$ and $\partial\Lambda^1$ by the standard circle) and hence they have the same winding number around 0.

Let us now equip this family with a holomorphic motion $h^1_\lambda : A^1_\circ \to A^1_\lambda$ of the fundamental annulus $A^1_\lambda := U_\lambda \smallsetminus U^1_\lambda$. This motion is obtained by lifting the motion h_λ by means of the double coverings $g_\lambda : A^1_\lambda \to A^0_\lambda$ (see Lemma 17.10):

$$\begin{array}{ccc} A^1_\circ & \xrightarrow[h^1_\lambda]{} & A^1_\lambda \\ g_\circ \downarrow & & \downarrow g_\lambda \\ A_\circ & \xrightarrow[h_\lambda]{} & A_\lambda \end{array}$$

By the First λ-lemma, the original holomorphic motion $\mathbf{h}$ matches with $\mathbf{h}'$ on the common boundary[4] $\partial^i A^0_\lambda = \partial^o A^1_\lambda$, so that together they provide a single holomorphic motion of the union $A^0_\lambda \cup A^1_\lambda$ over Λ^1.

Let $P^1 = \{\lambda \in \Lambda^1 : \; g_\lambda(0) \in A^1_\lambda\}$. Applying the above result to the restricted quadratic-like family, we obtain a homeomorphism $\pi \circ \gamma^1 : A^1_\circ \to P^1$, where $\gamma^1 : A^1_\circ \to \Gamma$ is the holonomy along $\mathbf{h}^1$. Since γ^1 matches with γ^0 on the common boundary of the annuli, they give us a homeomorphism of the union of the dynamical annuli onto the union of parameter annuli, $A^0 \cup A^1 \to P^0 \cup P^1$.

Proceeding in the same way, we construct:

• A nest of parameter annuli P^n attached one to the next and the corresponding parameter domains $\Lambda^n = \Lambda^{n-1} \smallsetminus P^{n-1}$ (where $\Lambda^0 \equiv \Lambda$). Moreover,

$$\bigcup P^n = \Lambda \smallsetminus \mathcal{M}(\mathbf{g}).$$

• A sequence of proper unfolded quadratic-like families

$$g_{n,\lambda} \equiv g_\lambda : U^n_\lambda \to U^{n-1}_\lambda \text{ over } \Lambda^n,$$

where $U^n_\lambda = g^{-n}_\lambda U_\lambda$ (thus $U^0_\lambda \equiv U_\lambda,\ U^{-1}_\lambda \equiv U'_\lambda$).

• A sequence of holomorphic motions $h_{n,\lambda}$ of the fundamental annulus $A^n_\lambda := U^{n-1}_\lambda \smallsetminus U^n_\lambda$ over Λ^n that equip $g_{n,\lambda}$; moreover, the $h_{n,\lambda}$ are obtained by lifting the $h_{n-1,\lambda}$ by means of the coverings $g_\lambda : A^n_\lambda \to A^{n-1}_\lambda$. These holomorphic motions match on the common boundaries of the fundamental annuli.

Let $\gamma_n : A^n_\circ \to \Gamma$ be the holonomy along $\mathbf{h}_n$. Since the holomorphic motions match on the common boundaries, these holonomies also match, and determine a continuous injection

$$\gamma : U'_\circ \smallsetminus \mathcal{K}(f_\circ) \to \Gamma.$$

Composing it with the projection π, we obtain a homeomorphism

$$\pi \circ \gamma : U'_\circ \smallsetminus \mathcal{K}(f_\circ) \to \Lambda \smallsetminus \mathcal{M}(\mathbf{g}) \tag{42.5}$$

[4]Recall from §1.7.12 that $\partial^i A$ and $\partial^o A$ mean the inner and the outer boundary components of A.

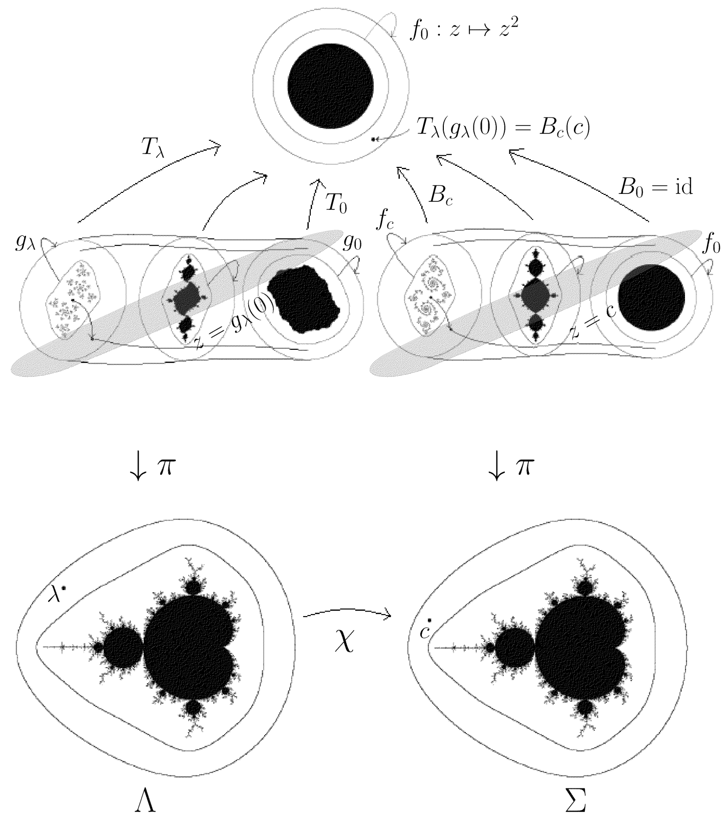

FIGURE 42.1. External straightening.

between the dynamical and parameter annuli. Note that the inverse map is equal to $\gamma^{-1} \circ \Phi$, where

$$\Phi : \Lambda \to \Gamma, \quad \Phi(\lambda) = (\lambda, g_\lambda(0)).$$

This is the *Phase-Parameter Relation* we alluded earlier.

Composing the above homeomorphism with the tubing (42.1), we obtain a "uniformization" $\psi_{\mathcal{M}(\mathbf{g})} \equiv \psi_{\mathbf{g}}$ of $\Lambda \smallsetminus \mathcal{M}(\mathbf{g})$ by a round annulus:

$$\psi_{\mathbf{g}} = T_{\circ} \circ \gamma^{-1} \circ \Phi : \ \Lambda \smallsetminus M(\mathbf{g}) \to \mathbb{A}(1, r^2), \quad \psi_{\mathbf{g}}(\lambda) = T_\lambda(g_\lambda(0)). \tag{42.6}$$

We see that this uniformization is given by *the tubing position of the critical value of* g_λ (compare §40.4.3).

COROLLARY 42.4. *The Mandelbrot set* $\mathcal{M}(\mathbf{g})$ *is a hull.*

In particular, $\mathcal{M}(\mathbf{g})$ is connected. Note that this argument can be applied to the quadratic family restricted to a high level subpotential domain $\Sigma_{\mathrm{par}}(r^2)$, $r >> 1$ (see §42.2) providing us with yet another proof of connectivity of the actual Mandelbrot set $\mathcal{M}$.

The above uniformization of $\Lambda \smallsetminus \mathcal{M}(\mathbf{g})$ is generally not conformal. However, in the case of a restricted quadratic family (see §42.2), it is a restriction of the

Riemann map $\Psi_{\mathcal{M}} : \mathbb{C} \smallsetminus \mathcal{M} \to \mathbb{C} \smallsetminus \overline{\mathbb{D}}$. Indeed, in this case, the tubing T_λ turns into the Böttcher maps B_c (see (42.2)), the critical value $g_\lambda(0)$ turns into c, and formula (42.6) turns into formula (34.1) for the Riemann map $\Psi_{\mathcal{M}}$.

42.6. External straightening. We are now ready to prove that the straightening is a homeomorphism outside the Mandelbrot sets.

LEMMA 42.5. *Under the assumptions of* Theorem 42.2*, the straightening*

$$\chi : \Lambda \smallsetminus \mathcal{M}(\mathbf{g}) \to \Sigma_{\mathrm{par}} \smallsetminus \mathcal{M}$$

is a homeomorphism.

PROOF. Let us consider the uniformizations $\psi_{\mathbf{g}} : \Lambda \smallsetminus \mathcal{M}(\mathbf{g}) \to \mathbb{A}(1, r^2)$ and $\Psi_{\mathcal{M}} : \Sigma_{\mathrm{par}} \smallsetminus \mathcal{M} \to \mathbb{A}(1, r^2)$ constructed above. Then

$$\chi = {\Psi_{\mathcal{M}}}^{-1} \circ \psi_{\mathbf{g}}. \tag{42.7}$$

Indeed, let $\lambda \in \Lambda \smallsetminus \mathcal{M}(\mathbf{g})$ and $c = \chi(\lambda) \in \Sigma_{\mathrm{par}} \smallsetminus \mathcal{M}$. Putting together (34.1), (42.6) and (40.6), we obtain:

$$\psi_{\mathbf{g}}(\lambda) = T_\lambda(g_\lambda(0)) = B_c(c) = \Psi_{\mathcal{M}}(c),$$

which is exactly (42.7). Since $\psi_{\mathbf{g}}$ and $\Psi_{\mathcal{M}}$ are both homeomorphisms, χ is a homeomorphism as well. $\square$

42.7. Quasiconformality. We willl show next that the external straightening from Lemma 42.5 can be selected to be quasiconformal (quantitatively).

The Phase-Parameter Relation (Lemma 17.14) implies:

LEMMA 42.6. *Under the assumptions of* Theorem 42.2*, suppose that the tubing $T_\circ : A_\circ \to \mathbb{A}[r, r^2]$ and the holomorphic motion* $\mathbf{h}$ *are K-qc. Then the uniformization $\Phi_{\mathbf{g}} : \Lambda \smallsetminus \mathcal{M}(\mathbf{g}) \to \mathbb{A}(1, r^2)$ (42.6) is K-qc as well.*

In fact, we can make the dilatation depend only on $\operatorname{mod} A_\circ$ and $\operatorname{mod}(\Lambda \smallsetminus \Lambda')$, after an appropriate adjustment of the family $\mathbf{g}$:

LEMMA 42.7. *Let us consider a quadratic-like family* $\mathbf{g}$ *over Λ satisfying the assumptions of* Theorem 42.2. *This family can be adjusted to a ql family $\tilde{\mathbf{g}}$ over a disk $\tilde{\Lambda} \subset \Lambda$ (equipped in the same way as* $\mathbf{g}$*) in such a way that the dilatation of the straightening $\tilde{\chi} : \tilde{\Lambda} \smallsetminus \mathcal{M}(\tilde{\mathbf{g}}) \to \Sigma \smallsetminus \mathcal{M}$ will depend only on* $\operatorname{mod} A_\circ$ *and* $\operatorname{mod}(\Lambda \smallsetminus \Lambda^1)$.

PROOF. Let us first restrict our family to the family $g_\lambda : U^1_\lambda \to U_\lambda$ over Λ^1, equipped with the restricted holomorphic motion, $h_\lambda : A_\circ \to A_\lambda$ as described in §42.5. By the Second λ-Lemma, $\operatorname{Dil}(\mathbf{h}|\,\Lambda^1) \le K$, where K depends only on $\operatorname{mod}(\Lambda \smallsetminus \Lambda^1)$.

Next, we can adjust the quadratic-like map $g_\circ : U^1_\circ \to U_\circ$ to a ql map $\tilde{g}_\circ : \tilde{U}^1_\circ \to \tilde{U}_\circ$ with L-bounded geometry, where L depends only on $\operatorname{mod}(A^1_\circ) \ge (1/2) \operatorname{mod} A_\circ$ (as described in Lemma 40.1). This leads to a natural adjustment $(\tilde{\mathbf{g}}, \tilde{\mathbf{h}})$ over $\tilde{\Lambda}$ of the ql family $(\mathbf{g}, \mathbf{h})$ over Λ^1.

The tubing $\tilde{T}_\circ : \tilde{A}_\circ \to \mathbb{A}[r, r^2]$ for the adjusted map $\tilde{g}_\circ$ can be selected with dilatation depending only on L, and hence only on $\operatorname{mod} A_\circ$.

By Lemma 42.6, dilatation of the straightening of $(\tilde{\mathbf{g}}, \tilde{\mathbf{h}})$ depends only on the dilatation of $\tilde{B}_\circ$ and $\tilde{\mathbf{h}}$, so ultimately it depends only on $\operatorname{mod} A_\circ$ and $\operatorname{mod}(\Lambda \smallsetminus \Lambda^1)$. $\square$

REMARK 42.8. If the holomorphic motion $\mathbf{h}$ and the tubing $T_\circ$ are assumed to be smooth/(real analytic) on $U'_\circ \smallsetminus \mathcal{K}_\circ$ (which can actually be done in general and which is obviously the case in our main examples of parapuzzle ql families producing primitive M-copies: see §47.7), then the external straightening

$$\chi : \Lambda \smallsetminus \mathcal{M}(\mathbf{g}) \to \Sigma_{\text{par}} \smallsetminus \mathcal{M}$$

has the same quality.

42.8. Miracle of continuity. We will now show that the straightening is continuous on the boundary of $\mathcal{M}(\mathbf{g})$:

LEMMA 42.9. *Under the assumptions of* Theorem 42.2*, the straightening χ is continuous at any point $\lambda \in \partial\mathcal{M}(\mathbf{g})$ and moreover $\chi(\lambda) \in \partial\mathcal{M}$.*

PROOF. First we will show that $\chi|\,\partial\mathcal{M}(\mathbf{g})$ is a continuous extension of $\chi|\,\Lambda \smallsetminus \mathcal{M}(\mathbf{g})$. Let $\lambda_n \in \Lambda \smallsetminus \mathcal{M}(\mathbf{g})$ be a sequence of parameter values converging to some $\lambda \in \partial\mathcal{M}(\mathbf{g})$. Let $c_n = \chi(\lambda_n)$ and $c = \chi(\lambda) \in \mathcal{M}$. We should show that $c_n \to c$. Let $g_\lambda : U \to U'$, $f_c : \Sigma \to \Sigma'$.

By Lemma 42.5, the map $\chi : \Lambda \smallsetminus \operatorname{int}\mathcal{M}(\mathbf{g}) \to \Sigma_{\text{par}} \smallsetminus \operatorname{int}\mathcal{M}$ is proper, and hence any limit point d of $\{c_n\} \subset \Sigma_{\text{par}} \smallsetminus \mathcal{M}$ belongs to $\partial\mathcal{M}$. We assert that $g_\lambda : U \to U'$ is qc conjugate to $f_d : V \to V'$. Indeed, the $g_{\lambda_n} : U_n \to U'_n$ are hybrid equivalent to the $f_{c_n} : \Sigma_n \to \Sigma'_n$ by means of some qc maps $\psi_n : U'_n \to \Sigma'_n$. By the straightening construction (see the proof of Lemma 40.13), the dilatation of ψ_n is equal to the dilatation of the tubing $T_{\lambda_n} = T_\circ \circ h_\lambda^{-1}$, which is locally bounded by the Second λ-lemma. By Exercise 13.14, the sequence ψ_n is precompact in the topology of uniform convergence on compact subsets of U'. Take a limit map $\psi : U' \to \Sigma'$. Since $g_{\lambda_n} \to g_\lambda$ uniformly on compact subsets of U and $f_{c_n} \to f_d$ (along a subsequence) uniformly on compact subsets of Σ, the map ψ conjugates g_λ to f_d, as was asserted.

But g_λ is also hybrid equivalent to f_c. Thus, f_c and f_d are qc conjugate in some neighborhoods of their filled Julia sets. By Corollary 41.11, they are qc conjugate on the whole complex plane. Since $d \in \partial\mathcal{M}$, Proposition 36.18 implies the desired: $c = d$ (and, in particular, $c \in \partial\mathcal{M}$).

The above argument implies that χ continuously maps $\Lambda \smallsetminus \operatorname{int}\mathcal{M}(\mathbf{g})$ into $\Sigma_{\text{par}} \smallsetminus \operatorname{int}\mathcal{M}$. We still need to show that χ is continuous at any point $\lambda \in \partial\mathcal{M}(\mathbf{g})$ even if it is approached from the interior of $\mathcal{M}(\mathbf{g})$. The argument is similar to the above except one detail. So, let now $\{\lambda_n\}$ be any sequence in Λ converging to λ. Let c_n, c and d be as above. Then the above argument shows that f_c is qc equivalent to f_d. But now we already know that $c \in \partial\mathcal{M}$ (though this time we do not know this for d). Applying Proposition 36.18 once again, we conclude that $c = d$. $\square$

"Only by miracle can one ensure the continuity of straightening in degree 2" said Adrien Douady [**D2**]. As we have seen, a reason behind this miracle is quasiconformal rigidity of the quadratic maps f_c with $c \in \partial\mathcal{M}$ (Proposition 36.18). Another reason is the λ-lemma. All these reasons are valid only for one-parameter families. There are no miracles in the polynomial families with more (essential) parameters.

42.9. Hyperbolic components. As in the case of the genuine Mandelbrot set, a component Q of $\operatorname{int}\mathcal{M}(\mathbf{g})$ is called hyperbolic if it contains a hyperbolic parameter value.

EXERCISE 42.10. *Show that:*

(i) *All parameter values in a hyperbolic component of* $\operatorname{int}\mathcal{M}(\mathbf{g})$ *are hyperbolic;*

(ii) *Neutral parameter values belong to* $\partial\mathcal{M}(\mathbf{g})$.

LEMMA 42.11. *If Q is a hyperbolic component of* $\operatorname{int}\mathcal{M}(\mathbf{g})$ *then there exists a hyperbolic component Δ of* $\operatorname{int}\mathcal{M}$ *such that $\chi : Q \to \Delta$ is a proper holomorphic map.*

PROOF. Obviously the straightening of a hyperbolic map is hyperbolic. Hence $\chi(Q)$ is contained in some hyperbolic component Δ of $\operatorname{int}\mathcal{M}$. Moreover, since the hybrid conjugacy is conformal on the interior of the filled Julia set, it preserves the multiplies of attracting cycles. Hence

$$\rho_Q(\lambda) = \rho_\Delta(c) \text{ for } \lambda \in Q,\ c = \chi(\lambda),$$

where ρ_Q and ρ_Δ are the multiplier functions on the domains Q and Δ respectively. By the Implicit Function Theorem, both these functions are holomorphic. Moreover, by the Multiplier Theorem, ρ_Δ is a conformal isomorphism onto $\mathbb{D}$. Hence $\chi = \rho_\Delta^{-1} \circ \rho_Q$ is holomorphic as well.

By Lemma 42.9, the map $\chi : Q \to \Delta$ is continuous up to the boundary and $\chi(\partial Q) \subset \partial\Delta$. Hence it is proper. □

42.10. Queer components. As in the quadratic case, a non-hyperbolic component Q of $\operatorname{int}\mathcal{M}(\mathbf{g})$ is called queer. In this section we will prove, using the dynamical uniformization of queer components (§36.7.4), that the straightening χ is holomorphic on Q. Let us begin with an extension of Corollary 36.6 to quadratic-like families:

LEMMA 42.12. *Let Q be a queer component of $\mathcal{M}(\mathbf{g})$. Take a base point $\lambda_\circ \in Q$. Then there is a holomorphic motion $H_\lambda : U'_\circ \to U'_\lambda$ conjugating $g_\circ$ to g_λ.*

PROOF. Since $\mathcal{M}(\mathbf{g})$ is equipped, there is an equivariant holomorphic motion $h_\lambda : A_\circ \to A_\lambda$. Let $A^n_\lambda = g_\lambda^{-n} A_\lambda$. Since the critical point is non-escaping under the iterates of g_λ, the A^n_λ are annuli and the maps $g^n_\lambda : A^n_\lambda \to A_\lambda$ are double coverings. By Lemma 17.10, $\mathbf{h}$ can be consecutively lifted to holomorphic motions $h_{n,\lambda} : A^n_\circ \to A^n_\lambda$. By the First λ-lemma, they automatically match on the common boundaries of the annuli, so that we obtain an equivariant holomorphic motion $H_\lambda : U'_\circ \smallsetminus \mathcal{K}(g_\circ) \to U'_\lambda \smallsetminus \mathcal{K}(g_\lambda)$. Since the sets $\mathcal{K}(g_\lambda)$ are nowhere dense (see Corollary 40.3(v) and Exercise 42.10(ii)), the First λ-lemma implies that the H_λ extend to an equivariant holomorphic motion $U'_\circ \to U'_\lambda$. □

LEMMA 42.13. *The straightening χ is holomorphic on any queer component Q of* $\operatorname{int}\mathcal{M}(\mathbf{g})$.

PROOF. Select a base point $\lambda_\circ \in Q$, and let $\phi : U' \to \Sigma'$ denote the hybrid conjugacy between $g_\circ : U \to U'$ and its straightening $f_\circ \equiv f_{c_\circ} : \Sigma \to \Sigma'$. Let H_λ be the holomorphic motion constructed in the previous lemma. By Lemma 17.15, the Beltrami differential

$$\mu_\lambda(z) = \begin{cases} \frac{\bar\partial H_\lambda(z)}{\partial H_\lambda(z)}, & z \in \mathcal{J}(g_\circ) = \mathcal{K}(g_\circ), \\ 0, & z \in \mathbb{C} \smallsetminus \mathcal{J}(g_\circ), \end{cases} \tag{42.8}$$

depends holomorphically on $\lambda \in Q$. Push the structures $\mu_\lambda \,|\, \mathcal{J}(g_\circ)$ forward to the $f_\circ$-plane by the straightening conjugacy ϕ, i.e., let ν_λ be the $f_\circ$-invariant Beltrami differential equal to $\phi_*(\mu_\lambda)$ on $\mathcal{J}(f_\circ)$ and vanishing on $\mathbb{C} \setminus \mathcal{J}(f_\circ)$. Since ϕ is conformal a.e. on the Julia set,

$$\nu_\lambda = \left(\mu_\lambda \, \phi' / \overline{\phi'}\right) \circ \phi^{-1},$$

which is obviously holomorphic in $\lambda \in Q$. Let $h_\lambda : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ be the solution of the Beltrami equation for ν_λ tangent to the identity at ∞. Then the maps $f_\lambda := h_\lambda \circ f_\circ \circ h_\lambda^{-1}$ are quadratic polynomials $z \mapsto z^2 + c(\lambda)$, and by Corollary 29.3, $c(\lambda)$ is holomorphic in λ. Finally, note that f_λ is the straightening of g_λ by means of the hybrid conjugacy $h_\lambda \circ \phi \circ H_\lambda^{-1}$. □

Remark 42.14. At this stage we do not yet know that $\chi \,|\, Q$ is non-constant, but if it is, then it is easy to show that χ *properly* maps Q onto some *queer* component of $\mathcal{M}$.

42.11. Discreteness of the fibers.

Lemma 42.15. *For any $c \in \mathcal{M}$, the fiber $\chi^{-1}(c)$ is finite.*

Proof. Since $\mathcal{M}(\mathbf{g})$ is compact, it is enough to show that the fibers are discrete. Assume that there exists some $c \in \mathcal{M}$ with an infinite fiber $\chi^{-1}(c)$. Then this fiber contains a sequence of distinct parameter values $\lambda_n \in \chi^{-1}(c)$ converging to some point $\lambda_\infty \in \chi^{-1}(c)$. Let $g \equiv g_\infty : U \to U'$.

Without loss of generality, we can assume that $\lambda_\infty \in \partial\mathcal{M}(\mathbf{g})$. [Otherwise, consider the component U of $\operatorname{int} \mathcal{M}(\mathbf{g})$ containing λ_∞. Since χ is holomorphic on U and continuous on $\overline{U}$, we conclude that $\chi | \overline{U} \equiv$ const. But then we can replace λ_∞ by any boundary point of U.]

Let us select λ_∞ as the base point in Λ. Since the quadratic-like family $g_\lambda : U_\lambda \to U'_\lambda$ is equipped, there exists an equivariant holomorphic motion $h_\lambda : A \to A_\lambda$ of the closed fundamental annulus $A_\lambda = \overline{U}'_\lambda \setminus U_\lambda$ over Λ (where $A \equiv \overline{U}' \setminus U$). Extend it by the Third λ-lemma[5] to a holomorphic motion $h_\lambda : \mathbb{C} \setminus U \to \mathbb{C} \setminus U_\lambda$ over a neighborhood $\Lambda' \subset \Lambda$ of λ_∞ (keeping the same notation for the extension). We will now construct a holomorphic family of hybrid deformations G_λ of g over Λ' naturally generated by this holomorphic motion.

To this end let us first pull back the standard conformal structure to $\mathbb{C} \setminus U$, $\mu_\lambda = h_\lambda^*(\sigma)$. Then extend μ_λ to a g-invariant conformal structure on $\mathbb{C} \setminus \mathcal{K}(g)$ by pulling it back by iterates of g. Finally, extend it to $\mathcal{K}(g)$ as the standard structure. This gives us a holomorphic family of g-invariant conformal structures on $\mathbb{C}$. We will keep the same notation μ_λ for these structures. Solving the Beltrami equations, we obtain a holomorphic family of qc maps $H_\lambda : \mathbb{C} \to \mathbb{C}$ such that $\mu_\lambda = (H_\lambda)^*(\sigma)$, and in particular, $\bar\partial H_\lambda(z) = 0$ a.e. on $\mathcal{K}(g)$. Conjugating g by these maps, we obtain a desired hybrid deformation $G_\lambda = H_\lambda \circ g \circ H_\lambda^{-1}$, $\lambda \in \Lambda'$.

On the other hand, for maps $g_n \equiv g_{\lambda_n}$, we can construct the Beltrami differentials $\mu_n \equiv \mu_{\lambda_n}$ in a different way. Namely, since the map g_n is hybrid equivalent to g, the equivariant map $h_n \equiv h_{\lambda_n}$ uniquely extends from $\mathbb{C} \setminus U$ to a hybrid conjugacy (Theorem 41.10). Let us keep the same notation h_n for this conjugacy.

The above two constructions naturally agree: $(h_n)^*\sigma = \mu_n$. Indeed, it is true on $\mathbb{C} \setminus U$ by definition. It is then true on $U \setminus \mathcal{K}(g)$, since both Beltrami differentials

[5] If we assume that the holomorphic motion $\mathbf{h}$ is smooth (see Remark 42.8) then we can use the Elementary λ-Lemma (Lemma 17.1) instead.

are g-invariant. Finally, it is true on the filled Julia set $\mathcal{K}(g)$ since h_n is conformal a.e. on it.

Thus, the qc maps $H_n \equiv H_{\lambda_n} : \mathbb{C} \to \mathbb{C}$ and $h_n : \mathbb{C} \to \mathbb{C}$ satisfy the same Beltrami equation. They also coincide at two points, e.g., at the critical point and at the β-fixed point of g (in fact, by Corollary 41.9 they coincide on the whole Julia set of g). By uniqueness of the solution of the Beltrami equation, $H_n = h_n$. Hence

$$G_{\lambda_n}(z) = g_{\lambda_n}(z). \tag{42.9}$$

Take an $\varepsilon > 0$ such that both functions $G_\lambda(z)$ and $g_\lambda(z)$ are well-defined in the bidisk $\{(\lambda, z) \in \mathbb{C}^2 : \ |\lambda - \lambda_\circ| < \varepsilon, \ z \in V \equiv g^{-1}U\}$. For any $z \in V$, consider two holomorphic functions of λ:

$$\Phi_z(\lambda) = G_\lambda(z) \quad \text{and} \quad \phi_z(\lambda) = g_\lambda(z), \quad |\lambda - \lambda_\infty| < \varepsilon.$$

By (42.9), they are equal at points λ_n converging to λ_∞. Hence they are identically equal.

Thus for $|\lambda| < \varepsilon$, two quadratic-like maps, G_λ and g_λ, coincide on V. But it is impossible since the Julia set of G_λ is always connected, while the Julia set of g_λ is disconnected for some λ arbitrary close to λ_∞ (recall that we assume that $\lambda_\infty \in \partial\mathcal{M}(\mathbf{g})$). □

Corollary 42.16. $\chi(\operatorname{int}\mathcal{M}(\mathbf{g})) \subset \operatorname{int}\mathcal{M}$.

Remark. Of course, it is not obvious only for queer components.

Proof. Take a component Q of $\operatorname{int}\mathcal{M}(\mathbf{g})$. We have proven that $\chi\,|\,Q$ is a *non-constant* holomorphic function. Hence the image $\chi(Q)$ is open. Since it is obviously contained in $\mathcal{M}$, it must be contained in $\operatorname{int}\mathcal{M}$. □

42.12. Bijectivity. What is left to show is that the map $\chi : \mathcal{M}(\mathbf{g}) \to \mathcal{M}$ is bijective. By §42.5, the winding number of the curve $\chi : \partial\Lambda \to \mathbb{C}$ around any point $c \in \Sigma_{\text{par}}$ is equal to 1. By the Topological Argument Principle (§3.3),

$$\sum_{a \in \chi^{-1}c} \operatorname{ind}_a(\chi) = w_c(\chi \circ \partial\Lambda) = 1, \quad c \in \Sigma_{\text{par}}. \tag{42.10}$$

It immediately follows that the map $\chi : \Lambda \to \Sigma_{\text{par}}$ is *surjective* (for otherwise the sum in the left-hand side would vanish for some $c \in \Sigma_{\text{par}}$).

Let us show that χ is injective on the interior of $\mathcal{M}(\mathbf{g})$. Indeed, if $a_0 \in \operatorname{int}\mathcal{M}(\mathbf{g})$, then by Corollary 42.16 $c = \chi(a_0) \in \operatorname{int}\mathcal{M}$, and by Lemma 42.9, $\chi^{-1}(c) \subset \operatorname{int}\mathcal{M}(\mathbf{g})$. But since $\chi|\operatorname{int}\mathcal{M}(\mathbf{g})$ is holomorphic (see §§42.9 and 42.10), we have $\operatorname{ind}_a(\chi) > 0$ for any $a \in \operatorname{int}\mathcal{M}$. It follows that the sum in the left-hand side of (42.10) actually contains only one term, so that c has only one preimage, a_0.

Finally, assume that there is a point $c \in \partial\mathcal{M}$ with more than one preimage. By the Topological Argument Principle, χ has a non-zero index at one of those preimages, say, a_1. Take another preimage a_2. Both a_1 and a_2 belong to $\partial\mathcal{M}(\mathbf{g})$.

Take a point $a_2' \notin \partial\mathcal{M}(\mathbf{g})$ near a_2, and let $c' = \chi(a_2')$. By Exercise 3.10, χ is locally surjective near a_1, so that c' has a preimage a_1' over there. This contradicts injectivity of χ on $\Lambda \smallsetminus \partial\mathcal{M}(\mathbf{g})$.[6]

This completes the proof of Theorem 42.2.

[6]At this point, one can use only injectivity on $\Lambda \smallsetminus \mathcal{M}(\mathbf{g})$.

43. QL families over complex renormalization windows

43.1. Canonical homeomorphism. Let us go back to the quadratic family $f_c : z \mapsto z^2 + c$. Take some superattracting parameter $c_\circ$ of period $p > 1$. It is the center of the renormalization window $\mathcal{V} \equiv \mathcal{V}_\circ$ described in §37.11.1. Recall that it is equal to the truncated parabolic wake $\mathcal{W}^{\mathrm{par}} \equiv \mathcal{W}^{\mathrm{par}}_\circ$ centered at $c_\circ$ (and rooted at the corresponding root $\mathfrak{r}$). Moreover, $\partial\mathcal{V}$ crosses $\mathcal{M}$ at $\mathfrak{r}$.[7]

For any polynomial f_c, $c \in \mathcal{V}$, we have constructed a quadratic-like map $g_c = f_c^p : V_c \to V_c'$ around the critical point. If the Julia set $\mathcal{J}(g_c)$ of this map is connected then f_c is renormalizable with combinatorics given by the Hubbard tree $\mathcal{T} \equiv \mathcal{T}_\circ$ of the superattracting parameter $c_\circ$ (see §§28.4.5, 37.11.2). Let

$$M \equiv M_\circ = \{c \in \mathcal{V} : \ f_c \textit{ is renormalizable with combinatorics } \mathcal{T}\} \cup \{\mathfrak{r}\}.$$

Notice that in the primitive case, the root map $f_\mathfrak{r}$ is still renormalizable with combinatorics $\mathcal{T}$ (see Exercise 28.24).

Theorem 43.1. *The set M is canonically ambiently homeomorphic to the Mandelbrot set $\mathcal{M}$. Moreover,*

(i) *This homeomorphism is conformal on* $\operatorname{int} M$.

(ii) *In the primitive case, it admits a homeomorphic extension to a neighborhood of M which is externally qc.*

(iii) *In the satellite case, it admits a homeomorphic extension to a neighborhood of M which is externally qc outside an arbitrary small neighborhood of the root $\mathfrak{r}$.*

Remark 43.2. In fact, as will be shown in vol. III, the above extensions are qc on the *whole* neighborhoods as above (not only "externally"). Moreover, externally, they can be made smooth (compare Remark 42.8).

The canonical homeomorphism $\chi : M \to \mathcal{M}$ is given by the straightening of the renormalization, i.e. for $c \in M \smallsetminus \{\mathfrak{r}\}$, $\chi(c) \in \mathcal{M}$ is defined so that the quadratic polynomial $f_{\chi(c)} : z \mapsto z^2 + \chi(c)$ is hybrid equivalent to the renormalization g_c. At the root, we let by definition: $\chi(\mathfrak{r}) = 1/4$.

Theorem 42.2 is designed to imply this result. However, it does not do it since the quadratic-like family (g_c) over $\mathcal{V}$ *is not full*: it misses the root $\mathfrak{r}$. This problem can be fixed for the primitive renormalization, as the map $f_\mathfrak{r}$ is also renormalizable (with the same combinatorics) in this case. However, in the satellite case, it is not fixable (see §28.4.4).

Another issue is the existence of a global equipment of a family with holomorphically moving fundamental annulus. For primitive copies, we will deal with it in §47. In general, as we will see below, any family can be locally equipped, which is sufficient for most purposes.

In this section we will give a proof of Theorem 43.1 that will produce for us all little M-copies. For primitive copies, an alternative proof will be given in §47 where we will construct the corresponding fully equipped quadratic-like families to which Theorem 42.2 can be applied.

In what follows we will notationally distinguish the parameter plane for the quadratic family f_c from the parameter plane of its renormalization g_λ (though

[7]In §37.11.1, the window and associated objects are labeled by the root rather than the center: hopefully, it will not cause a confusion.

formally, the latter is part of the former). In particular, a base point $\lambda_\circ \in \mathcal{V}$ will not be assumed equal to $c_\circ$, the center of M.

43.2. Straightening of general ql families. Before passing to a proof of Theorem 43.1, let us summarize some features of Theorem 42.2 that are valid for more general ql families, without assumptions that they are equipped, proper, and unfolded.

LEMMA 43.3. *Let* $\mathbf{g} = (g_\lambda)$ *be a quadratic-like family over a domain* $\Lambda \subset \mathbb{C}$. *Then:*

(i) *The canonical straightening* $\chi : \mathcal{M}(\mathbf{g}) \to \mathcal{M}$ *is continuous;*

(ii) *It is holomorphic on* $\operatorname{int} \mathcal{M}(\mathbf{g})$*;*

(iii) χ *admits a continuous extension to a neighborhood of any point* $\lambda_\circ \in \partial\mathcal{M}(\mathbf{g})$, *and this extension is locally quasiregular on the complement of* $\mathcal{M}(\mathbf{g})$*;*

(iv) χ *maps* $\partial\mathcal{M}(\mathbf{g})$ *to* $\partial\mathcal{M}$*;*

(v) *If* $\mathcal{M}(\mathbf{g}) \neq \Lambda$ *then for any* $c \in \mathcal{M}$, *the fiber* $\chi^{-1}(c)$ *is discrete;*

(vi) *If a component* Q *of* $\operatorname{int} \mathcal{M}(\mathbf{g})$ *is compactly contained in* Λ, *then* χ *properly maps it onto some component of* $\operatorname{int} \mathcal{M}$.

For the proof, we need to locally equip our family (see §42.1):

LEMMA 43.4. *Let* $\mathbf{g}$ *be a ql family over a domain* Λ, *and let* $\lambda_\circ \in \Lambda$. *Then* $\mathbf{g}$ *can be equipped with an equivariant smooth holomorphic motion of some fundamental annulus,*

$$h_\lambda : A_\circ \to A_\lambda, \tag{43.1}$$

over some neighborhood $\Lambda_\circ \ni \lambda_\circ$.

PROOF. Make the outer boundary of $A_\circ$ stay still, lift it to an equivariant motion of the inner boundary, and interpolate it inside $A_\circ$ by means of the Elementary λ–Lemma. □

PROOF. Let us start with the external part of (iii). Let $\lambda_\circ \in \partial\mathcal{M}(\mathbf{g})$. Let us locally equip this family by Lemma 43.4. Let us also select a tubing $T_\circ : A_\circ \to \mathbb{A}[r, r^2]$ for $g_\circ$ (see §40.3). This determines an extension of the straightening to $\Lambda_\circ$. By (34.1) and (40.6), the Riemann position of $\chi(\lambda)$ is equal to the tubing position of the critical value outside $\mathcal{M}(\mathbf{g})$:

$$\Psi_{\mathcal{M}}(\chi(\lambda)) = T_\circ(h_\lambda^{-1}(g_\lambda(0))),$$

where $h_\lambda^{-1} : \Sigma_\lambda \to \Sigma_\circ$ is the lift of the inverse of (43.1) to the exterior of the figure-eight centered at 0 (see §40.4.2). Since for μ near λ, the composition $\mu \mapsto h_\lambda^{-1} \circ h_\mu$ is a holomorphic motion centered at λ, the map $\chi(\lambda)$ is locally quasiregular in $\Lambda_\circ \smallsetminus \mathcal{M}(\mathbf{g})$ (see Lemma 17.9). All the more, it is continuous.

The rest is proved along the lines of Theorem 42.2: For (ii) see §§42.9 and 42.10; for (i) (and completion of (iii)) see §42.8; it implies (iv); for (v) see Lemma 42.15; and (vi) follows. □

In fact, as the above proof shows, if the family happens to be globally equipped then the straightening extends to the whole parameter domain:

PROPOSITION 43.5. *Let* $\mathbf{g}$ *be an equipped quadratic-like family over a domain* $\Lambda \subset \mathbb{C}$. *Then the canonical straightening* $\chi : \mathcal{M}(\mathbf{g}) \to \mathcal{M}$ *admits a continuous extension to* Λ *satisfying properties* (i)–(vi) *of* Lemma 43.4.

43.3. Proof of Theorem 43.1. Proposition 43.3 (applied to the renormalized family $\mathbf{g}$ over Λ) already provides us with some interesting information about the canonical map $\chi : M \to \mathcal{M}$. In fact, some of it can be obtained in a more direct way by making use of the explicit uniformizations of components of $\operatorname{int} \mathcal{M}$ described by the Multiplier and Queer Theorems:[8]

LEMMA 43.6. (i) *Any hyperbolic component Q of $\operatorname{int} M$ is conformally mapped under the straightening χ onto some hyperbolic component Δ of $\mathcal{M}$.*

(ii) *Any queer component Q of $\operatorname{int} M$ is either conformally mapped under the straightening χ onto some queer component Δ of $\mathcal{M}$, or else collapses to some point of $\partial\mathcal{M}$.*

PROOF. (i) Since the property of having an attracting cycle is preserved under the renormalization and the straightening, hyperbolicity of Q implies that the image $\chi(Q)$ is contained in some hyperbolic component Δ of $\mathcal{M}$. Moreover, by the Multiplier Theorem, the components Q and Δ are conformally mapped onto $\mathbb{D}$ by the multipliers functions ρ_Q and ρ_Δ. Since χ preserves the multiplier of the attracting cycle, we have: $\chi = \rho_\Delta^{-1} \circ \rho_Q$, implying the conclusion.

(ii) In the queer case, fix a base point $\lambda_\circ \in Q$ and let $c_\circ := \chi(\lambda_\circ) \in \Delta$. Let $\mu_\circ$ be an invariant Beltrami differential on the Julia set $\mathcal{J}_\circ$ of $g_{\lambda_\circ}$. By the ergodicity (Lemma 36.17) it is either supported on the grand orbit

$$\operatorname{Orb} J_\circ = \bigcup_{n=0}^{\infty} g_{\lambda_\circ}^{-n}(J_\circ)$$

of the little Julia set $J_\circ$, or on its complement. Moreover, in the latter case, there are no invariant line fields on $\operatorname{Orb} J_\circ$, implying that the whole component Q collapses to a single point. By continuity, this point must belong to $\partial\mathcal{M}$ (see Lemma 43.3(iii), (iv)).

In the former case, the restriction $\nu_\circ := \mu_\circ |\, J_\circ$ is an invariant Beltrami differential on $J_\circ$. By the Queer Theorem, the family of $g_{\lambda_\circ}$-invariant Beltrami differentials $t\mu_\circ$ $(t \in \mathbb{D})$ on $\mathcal{J}_\circ$ naturally parameterizes the queer component Q. Restricting it to the little Julia set $J_\circ$ and straightening the renormalization, we obtain an $f_{c_\circ}$−invariant family of Beltrami differentials $th_*(\nu_\circ)$ $(t \in \mathbb{D})$ parameterizing Δ. (We use that the straightening conjugacy h is conformal on $J_\circ$.) Putting these pieces together, we come to the desired conclusion. $\square$

What is left is to show that the straightening $\chi : M \to \mathcal{M}$ is injective, surjective, continuous at the root (in the satellite case), and admits an external qc extension.

43.3.1. *Injectivity.* Assume that for two points $c, \tilde{c} \in M$, we have $\chi(c) = \chi(\tilde{c})$. Let $f = f_c : z \mapsto z^2 + c$ and $\tilde{f} = f_{\tilde{c}} : z \mapsto z^2 + \tilde{c}$. Assume first that both points c and $\tilde{c}$ are different from $\mathfrak{r}$, so both are renormalizable (with some period p). Then by definition of the straightening χ, their renormalizations $g : U \to U'$ and $\tilde{g} : \tilde{U} \to \tilde{U}'$ are hybrid equivalent. Using the Pullback Argument (and a bit of measurable dynamics), we will show that the maps f and $\tilde{f}$ are hybrid equivalent as well. Since they are quadratic polynomials, it will follow that $c = \tilde{c}$.

[8]Though clarifying, this lemma is not needed for the proof of Theorem 43.1. In fact, the latter is independent of the Multiplier Theorem and supersedes it.

FIGURE 43.1. An illustration of a partial cell decomposition for period 12 renormalizable map (the period 4 elephant eye tuned by the rabbit).

Let $\mathbf{K} = \bigcup K_i$ be the union of the filled little Julia sets of f (corresponding to the renormalization under consideration). Let $\mathfrak{R}$ be the configuration of the f-rays landing the $\beta-$ and $\beta'-$fixed points of these little Julia sets. Let $\mathcal{E}$ be some f-equipotential, and let $\mathcal{E}' = f^p(\mathcal{E})$. Let us consider a configuration $\mathcal{C}$ of these two equipotentials and the rays truncated at the level of $\mathcal{E}'$. This configuration is marked with the Böttcher coordinate.

Moreover, the configuration $\mathcal{C}$ produces a degenerate renormalization $g = (f^p : W \to W')$ around the little Julia sets K, so that the corresponding renormalization $U \to U'$ is obtained by the thickening of the domains W, W' (see §28.4.3). In fact, it produces a degenerate renormalization $g = (f^p : W_i \to W_i')$ around each little Julia set K_i,

As usual, we mark with "tilde" the corresponding objects for $\tilde{f}$.

Since the maps f and $\tilde{f}$ are renormalizable with the same combinatorics, there exists a homeomorphism

$$h : (\mathbb{C}, \mathcal{C}) \to (\mathbb{C}, \tilde{\mathcal{C}}) \tag{43.2}$$

respecting the marking and coinciding with the Böttcher conjugacy outside $\mathcal{E}$. Moreover, by Exercise 24.19, h can be selected to be quasiconformal.

REMARK 43.7. This also follows from the Fourth λ-lemma since by the Wake Theorem our configuration moves holomorphically over the corresponding renormalization window.

Map (43.2) can be modified so that its restriction to the domain W' is a hybrid conjugacy between the renormalizations g and $\tilde{g}$ (keeping all other properties of it). To see it, consider the configuration $\Gamma \subset \partial W'$ of four ray segments landing at the $\beta-$ and $\beta'-$points of K that are used to construct the renormalization, and the corresponding configuration $\tilde{\Gamma}$ for $\tilde{g}$. Modify h in the fundamental annulus $A = \overline{U}' \smallsetminus U$ so that it maps the ray segments of $\Gamma \cap A$ to the corresponding segments $\tilde{\Gamma} \cap \tilde{A}$, respecting the boundary marking. Lifting h to the preimages of U and passing to a limit (like in the Pullback Argument), we obtain a hybrid conjugacy that maps $\partial W'$ to $\partial \tilde{W}'$ respecting the boundary marking.

Similarly, we can modify h near every little Julia set K_i, to turn it to a hybrid conjugacy on each of them. Then we obtain a map that conjugates $f \,|\, \mathbf{K}$ to $\tilde{f} | \tilde{\mathbf{K}}$. In particularly, it conjugates f and $\tilde{f}$ on their postcritical sets.

Let us now consider the combinatorial Hubbard tree $\mathcal{T}$ of f corresponding to the configuration $\mathcal{C}$ (see §37.11.2). It has marked points $(0_n)_{n=0}^{p-1}$ and is endowed with a piecewise linear map F cyclically permuting these points. (Recall that F is linear on each component of $\mathcal{T} \smallsetminus (\mathbf{0} \cup \mathbf{b})$, where $\mathbf{b}$ is the set of branched points.) Since h maps $\mathcal{C}$ to $\tilde{\mathcal{C}}$, it respects the corresponding cell decomposition, and hence induces a map $H : \mathcal{T} \to \tilde{\mathcal{T}}$ respecting the marking. Hence it conjugates F to $\tilde{F}$ on the set of marked points. As these maps are piecewise linear in between, H conjugates F to $\tilde{F}$.

Thus the $(F,\tilde{F})-$lift of H is equal to H. It follows that the $(f,\tilde{f})-$lift of h is a map h_1 respecting our cell decomposition. By the Alexander Trick, h_1 is homotopic to h rel $\partial\mathcal{C}$.

As $h : K_i \to \tilde{K}_i$ is a conjugacy on the little Julia sets (and they are contained in our cells), h_1 coincides with h on them. It follows that h_1 is homotopic to h rel $\mathbf{K} \supset \overline{\mathcal{P}}$. Thus, we obtain a qc Thurston equivalence between f and $\tilde{f}$.

By means of the Pullback Argument (see §38.5), h can be transformed into a qc-conjugacy between f and $\tilde{f}$ that coincides with the Böttcher conjugacy outside $\mathcal{K}(f)$ and with the hybrid conjugacy on $\mathbf{K}$. Then $\bar{\partial}h = 0$ a.e. on $\operatorname{Orb} K = \bigcup_{n=0}^{\infty} f^{-n}(\mathbf{K})$. By Lemma 28.33, this set has full measure in $\mathcal{K}(f)$. Hence $\bar{\partial}h = 0$ a.e. on the whole complex plane. By Weyl's Lemma, h is affine.

43.3.2. *Surjectivity.* For any $b \in \mathcal{M}$, we want to find $c \in M$. such that $\chi(c) = b$. Let us start with a superattracting b with the Hubbard tree $\mathcal{T}_b$. Let us tune the Hubbard tree $\mathcal{T}_\circ$ with $\mathcal{T}_b$ (see Exercise 37.42). It is easy to see that we obtain an admissible Hubbard tree, which can be realized by a desired superattracting parameter value c (see §37.9).

Moreover, by Lemma 43.6(i)[9] χ maps conformally the hyperbolic component Q_c centered at c onto the hyperbolic component Δ_b, so the latter is contained in $\operatorname{Im}\chi$. Since χ is continuous and superattracting parameters accumulate onto the whole boundary $\partial\mathcal{M}$ (see Exercise 33.6), the latter is also contained in $\operatorname{Im}\chi$.

What is left, are possible queer components. But if one of them were missing in $\operatorname{Im}\chi$ then $\operatorname{Im}\chi$ would not be full. On the other hand, M is full, and hence has vanishing first Alexander cohomology (Theorem 1.23). As the latter property is invariant under homeomorphisms, $H^1_{\mathrm{A}}(\operatorname{Im}\chi) = 0$ as well, and hence $\operatorname{Im}\chi$ is full.

43.3.3. *Continuity at the root.*

LEMMA 43.8. *The map $\chi : M \to \mathcal{M}$ is continuous at the roof $\mathfrak{r}$ of M.*

PROOF. Recall that $\Delta_0 \ni 0$ stands for the main hyperbolic component of $\mathcal{M}$ (bounded by the main cardioid). Let $Q_0 \ni c_\circ$ be the main hyperbolic component of M. By Lemma 43.6(i), the map $\chi : Q_0 \to \Delta_0$ is univalent, and hence continuous up to the boundary.

Let us now take a sequence of points $c_n \in M \smallsetminus \overline{Q}_0$, $c_n \to \mathfrak{r}$. Then c_n belongs to a limb $\mathcal{L}_n$ attached to ∂Q_0. By Lemma 37.22, $\operatorname{diam}\mathcal{L}_n \to 0$; hence the root points $\mathfrak{r}_n$ of the limbs converge to $\mathfrak{r}$. By continuity of $\chi|\,\overline{Q}_0$, we have $\chi(\mathfrak{r}_n) \to 1/4$. Applying Lemma 37.22 once again, we conclude that $\operatorname{diam}\chi(\mathcal{L}_n) \to 0$, and hence $\chi(c_n) \to 1/4$ as well. □

43.3.4. *Homeomorphic extension.* Let us now show that χ admits a homeomorphic extension to a neighborhood of M (compare with §42.5). It is convenient to use renormalization $g_\lambda : W_\lambda \to W'_\lambda$ near the critical value as depicted on Figure 28.3. Let us consider the *fundamental tile* $T_\lambda := \operatorname{cl}(W'_\lambda \smallsetminus W_\lambda)$. Then the complement of the valuable little Julia set, $\overline{W}'_\lambda \smallsetminus K_{v,\lambda}$, is tessellated with its pullbacks $T^n_{i,\lambda}$ under the iterates g^n_λ.

The boundary of T_λ moves holomorphically over the corresponding truncated parameter wake $\mathcal{W}'_{\mathrm{par}}$. By the Fourth $\lambda-$Lemma, this motion can be extended to

[9] One can use Proposition 43.3 (vi) instead.

a holomorphic motion of the whole tile T_λ over $\mathcal{W}_{\rm par}$. By the Phase-Parameter Relation, we obtain a *fundamental parameter tile* $T_{\rm par} \subset W'_{\rm par}$ (attached to the Chebyshev parameter of M) naturally homeomorphic to the tile $T_\circ$ at the base parameter.

Next, the above holomorphic motion can be lifted to a holomorphic motion of the tiles of level one, $T^1_{i,\lambda}$, over $W'_{\rm par} \setminus T_{\rm par}$. By the Phase Parameter Relation, we obtain two parameter tiles $T^1_{i,\rm par}$ (attached to appropriate dyadic tips of M) naturally homeomorphic to the corresponding dynamical tiles $T^1_{i,\circ}$ at the base parameter. Moreover, this homeomorphism matches with the previous one on the common boundary of the parameter tiles (since the motions of the tiles match on their common boundary).

Proceeding this way, we obtain a tessellation of

$$\Lambda^{\rm out} := (\mathcal{W}_{\rm par} \setminus M) \cup \{\text{dydic tips of } M\}$$

by parameter tiles $T^n_{i,\rm par}$. Moreover, there are natural homeomorphisms from these tiles to the corresponding dynamical tiles $T^n_{i,\circ}$ that altogether form a homeomorphism from $\Lambda^{\rm out}$ onto the dynamical escaping set at the base parameter,

$$D^{\rm out}_\circ := (W'_\circ \setminus K_{v,\circ}) \cup \{\text{dydic tips of } K_{v,\circ}\}.$$

Repeating the same procedure for the actual quadratic family (f_c), we obtain a natural homeomorphism from $\Lambda^{\rm out}$ onto the corresponding truncated wake of $\mathbb{C} \setminus \mathcal{M}$ (via the natural homeomorphism of the corresponding dynamical sets at the base parameters).

Now the miracle of continuity ensures that this homeomorphism matches with the canonical straightening $\chi : M \to \mathcal{M}$. This produces a homeomorphic extension of χ to the truncated wake $\mathcal{W}_{\rm par}$. Finally, extend it arbitrarily to a small outer neighborhood of the root.

Let us now show that the straightening homeomorphism χ is externally quasiconformal in a slighly smaller wake-like region $\tilde{\mathcal{W}}_{\rm par} \setminus M \subset \mathcal{W}_{\rm par} \setminus M$ around M truncated near the root. To this end, let us slightly adjust the base dynamical wake $W'_\circ$ by replacing the rays landing at the root by nearby curves ("adjusted rays") invariant under $g_\circ$ (also landing at the root) and by replacing the outer equipotential with a higher level equipotential (compare with Lemma 42.6). Then the base tile $T_\circ$ get adjusted accordingly and the tiles T_λ get adjusted via the holomorphic motion over the adjusted parameter wake $\tilde{\mathcal{W}}_{\rm par}$. (We will mark the adjusted objects with the tilde.)

By the Phase-Parameter Relation, we obtain a truncated parameter tile $\tilde{T}_{\rm par} \subset T_{\rm par}$ naturally homeomorhic to $\tilde{T}_\circ$. Moreover, as the motion (T_λ) exends to the original wake $\mathcal{W}_{\rm par}$, it has a bounded dilatation a truncated wake over $\tilde{\mathcal{W}}^{\rm tr}_{\rm par} \Supset T_{\rm par}$ obtained by removing from $\tilde{\mathcal{W}}_{\rm par}$ a neighborhhood of the root. (by the Second $\lambda-$Lemma). Then all the lifts of this motion to the adjusted tiles $\tilde{T}^n_{i,\lambda}$ over $\tilde{\mathcal{W}}^{\rm tr}_{\rm par}$ have the same uniformly bounded dilatation. It follows that the straightening homeomorphisms have uniformly bounded dilatation on the truncated parameter tiles $\tilde{T}^{\rm tr}_{\rm par}$. As these tiles cover $\tilde{\mathcal{W}}^{\rm tr}_{\rm par} \setminus M$ (indeed, they cover $\tilde{\mathcal{W}}_{\rm par} \setminus M$ truncated by two adjusted parameter rays landing near the root), the conclusion follows (using the Little Gluing Lemma).

This completes the proof of Theorem 43.1, except for an external qc extension to a neigborhood of the root in the primitive case. It can be done by a carefully matching a local extension with the previously constructed one. However, we will neglect this issue for the moment. An alternative approach to the primitive case via puzzle given in §47 will automatically take care of it.

43.4. Tuning and renormalization combinatorics (once again). We are now in a position to complete our discussion of the renormalization combinatorics. Its real counterpart appeared in §37.11.3 (Exercise 37.51(ii)).

Recall from §37.11.2 that for a quadratic polynomials, we have four way of recording the renormalization combinatorics: by means of a periodic ray portrait Θ, a superattracting parameter c, a Hubbard tree $\mathcal{T}$, and a little M-copy M (and we know by now that this copy is homeomorphic indeed to the big Mandelbrot set $\mathcal{M}$).

Remark 43.9. We allow ourselves the trivial combinatorics represented by $\Theta = \emptyset$ (no rays in the portrait), or $c = 0$, or $\mathcal{T} = \{*\}$, (the tree degenerates to the singleton), or $M = \mathcal{M}$ (the "little" M is equal to the whole Mandelbrot set). We will usually identify it as $*$.

We can now extend all these descriptions to the quadratic-like case. Renormalizable ql maps (and their renormalizations) are defined in the same way as in the polynomial case (see §28.4). As this notion is invariant under topological conjugacies, there is a natural relation between renormalizations of a ql map g and its straightening $f \equiv f_c == \chi(g)$. Which allows us to define the *renormalization combinatorics* of g in terms of f.

Assume that g is several times renormalizable (maybe, infinitely many). Let (g_n) be the list of its consequtive renormalizations, $n = 0, 1, \dots,$ with "absolute" combinatorics $\tau^{[n]}$ (which can be represented by a ray portrait $\Theta^{[n]}$, by a superattracting parameter $c^{[n]} \in \mathcal{M}$, by a Hubbard tree $\mathcal{T}^{[n]}$, or by a little M-copy $M^{[n]}$.). Then g_{n+1} is the renormalization of g_n with some *relative combinatorics* τ_n which can be represented by a ray portrait Θ_n, by a superattracting parameter $c_n \in \mathcal{M}$, by a Hubbard tree $\mathcal{T}_n$, or by a little M-copy M_n. (Notice that we use the lower indices to distinguish the relative combinatorics from the absolute ones.) The string of these relative combinatorics,

$$\bar{\tau} := (\{*\} = \tau_0, \tau_1, \dots),$$

represents the *full renormalization combinatorics* for g. Though it contains the same amount of information as the string of absolute combinators, the relative data often provides a better grasp of the renormalization structure in question. In particular, *any* string of prime Hubbard trees is realizable as a renormalization combinatorics:

Proposition 43.10. (i) *For any finite string of abstract prime Hubbard trees,*

$$\bar{\tau} = (\mathcal{T}_0, \mathcal{T}_1, \dots, \mathcal{T}_N), \quad N \in \mathbb{N},$$

there exists a little M-copy $M(\bar{\tau})$ comprising all the parameters $c \in \mathcal{M}$ that are renormalizable with renormalizatoin combinatorics $\bar{\tau}$.

(ii) *Any infinite string of abstract prime Hubbard trees,*

$$\bar{\tau} = (\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots,),$$

determine a combinatorial class $\mathcal{C}(\bar{\tau}) \subset \mathcal{M}$ of parameters which are infinitely renormalizable with combinatorics τ. Moreover, $\mathcal{C}(\bar{\tau})$ is either a hull or a singleton.

PROOF. (i) By Exercise 37.42, there is a superattracting parameter c corresponding to the Hubbard tree with combinatorics $\bar{\tau}$. The little M-copy centered at this parameter is desired.

(ii) (Compare Corollary 9.4.) Let us consider truncated renormalization combinatorics $\bar{\tau}^{[n]} := (\mathcal{T}_\circ, \dots, \mathcal{T}_n)$. Then $\mathcal{C}(\bar{\tau})$ is the intersection of the nest of the corresponding little copies $M^{[n]} = \mathcal{C}(\bar{\tau}^{[n]})$, each of which is a subhull in the previous one. □

Let us consider a little Mandelbrot copy $M_\circ$ centered at a superattracting parameter $c_\circ$, and let $\chi_\circ : M_\circ \to \mathcal{M}$ be the corresponding straightening homeomorphism. Take some $c \in M_\circ$ and let $d = \chi(c)$. Thus, the map f_c is renormalizable with combinatorics specified by $c_\circ$ (or by the whole copy $M_\circ$ for that matter) and f_d is the straigtening of its renormlization Rf_c. Under these circumstnces, f_c is called the *tuning* of $f_\circ \equiv f_{c_\circ}$ by f_d. Accordingly, the parameter c is called the tuning of $c_\circ$ by d, and one writes

$$c = c_\circ * d.$$

Let us describe what happens under tuning with the combinatorial laminations. Let p be the period of 0 under $f_\circ$. Let us consider the lamination $\mathcal{L}_\circ \equiv \mathcal{L}_{c_\circ}$ of $f_\circ$, and let Q_0 be the central gap of this lamination. For book keeping reasons, it is convenient to denote the doubling map associated with $f_\circ$ by $T_\circ : \theta \mapsto 2\theta \mod 1$. It induces the return map $\hat{T}_\circ^p : Q_0 \to Q_0$. The quotient $\overline{Q}_0/\mathcal{L}_\circ$ is naturally identified with the disk $\mathbb{D}$ in such a way that the above return map induces the squaring map $f_0 : z \mapsto z^2$ of $\bar{\mathbb{D}}$ (corresponding, via the Riemann uniformization, to the return of the hyperbolic map $f_\circ$ to the immediate basin $\mathcal{D}_0$ of 0, see Corollary 25.9). Let $\pi : \overline{Q}_0 \to \overline{\mathbb{D}}$ be the corresponding projection. Let us consider some lamination $\mathcal{L}$ on $\mathbb{D}$ and pull it back by π to the gap Q_0. Then spread it around by the iterated doubling map $\hat{T}_\circ$ to all other gaps of $\mathcal{L}_\circ$. We obtain the lamination $\mathcal{L}_\circ * \mathcal{L}$ called the *tuning* of $\mathcal{L}_\circ$ by $\mathcal{L}$. Intuitively, this procedure means pinching of the basin $\mathcal{D}_0$ along the lamination $\mathcal{L}$ and pinching all other components of $\operatorname{int} \mathcal{K}_\circ$ along the pullbacks of $\mathcal{L}$.

In what follows (in this section), $\mathcal{L}_c \equiv \mathcal{L}_{c,\text{puz}}$ stands for the combinatorial lamination for f_c.

PROPOSITION 43.11. (i) *For $c = c_\circ * d$, we have $\mathcal{L}_c = \mathcal{L}_\circ * \mathcal{L}_d$.*

(ii) *If f_c is n times renormalizable with renormalization combinatorics $(c_\circ, \dots, c_{n-1})$ and the last renormalization straightening f_d (i.e., $c = c_\circ * c_1 * \dots * c_{n-1} * d$) then*

$$\mathcal{L}_c = \mathcal{L}_{c_\circ} * \mathcal{L}_{c_1} * \dots * \mathcal{L}_{c_{n-1}} * \mathcal{L}_d.$$

(iii) *If f_c is infinitely renormalizable with renormalization combinatorics $(c_\circ, c_1, c_2, \dots)$ (i.e., $c = c_\circ * c_1 * c_2 \dots$) then*

$$\mathcal{L}_c = \mathcal{L}_{c_\circ} * \mathcal{L}_{c_1} * \mathcal{L}_{c_2} \dots := \operatorname{cl} \bigcup_{n=1}^{\infty} \mathcal{L}_\circ * \mathcal{L}_1 * \dots * \mathcal{L}_{n-1}.$$

Let us do some preparation. Let $g = f_c^p : W \to W'$ be the pre-renormalization of f_c with combinatorics $c_\circ$. It straightens to f_d, so $f_d = h \circ g \circ h^{-1}$, where $h : W' \to \mathbb{C}$ is the straightening conjugacy, and $d = \chi_\circ(c)$. The α-fixed point α_g of g generates a cycle $\boldsymbol{\alpha}_g$ of period p for f_c. On the other hand, under the straightening conjugacy, α_g becomes the α-fixed point $\alpha_d = h(\alpha_g)$ for f_d.

LEMMA 43.12. *Under the above circumstances, assume that the cycle $\boldsymbol{\alpha}_c$ is either repelling or parabolic with non-zero rotation number. Then:*

(i) *The number of accesses to α_g from $\mathbb{C} \smallsetminus \mathcal{K}(f_c)$ and to α_d from $\mathbb{C} \smallsetminus \mathcal{K}(f_d)$ is the same (with a natural relation betweeen them).*

(ii) *The combinatorial rotation numbers of $\boldsymbol{\alpha}_g$ and α_d are equal.*

PROOF. Under the straightening, the number of accesses and the rotation number are preserved, so they are the same for α_d viewed from $\mathbb{C} \smallsetminus \mathcal{K}(f_d)$ and for α_g viewed from $\mathbb{C} \smallsetminus \mathcal{K}(g)$. But Lemma 28.36 tells us that the latter can be viewed from $\mathbb{C} \smallsetminus \mathcal{K}(f_c)$ as well. □

Proof of Proposition 43.11. Let us start with the main hyperbolic component $\Delta_\circ$ of $M_\circ$. By Corollary 35.11, the lamination $\mathcal{L}_c$ stays the same over it: $\mathcal{L}_c = \mathcal{L}_\circ$ for $c \in \Delta_\circ$. But $\chi_\circ(\Delta_\circ) = \Delta_0$, where the latter is the main hyperbolic component of $\mathcal{M}$. As $\mathcal{L}_d$ is trivial for $d \in \Delta_0$, we are done.

Let us now go to the main cardioid $\partial\Delta_\circ$ of $M_\circ$. Let $\theta \mapsto c(\theta)$ be its parametrization by the rotation number $\theta \in \mathbb{R}/\mathbb{Z}$ of the cycle $\boldsymbol{\alpha}_c$, while $\theta \mapsto d(\theta)$ be the parametrization of the main cardioid $\partial\Delta_0$ by the rotation number of the fixed point α_d. Under the straightening $\chi : \partial\Delta_\circ \to \partial\Delta_0$, the rotatoin number is preserved: $\chi(c(\theta)) = d(\theta)$.

For $\theta = 0$ (i.e., at the root $\mathfrak{r}_\circ$ of $\Delta_\circ$) we have by Theorem 35.27: $\mathcal{L}_{c(0)} = \mathcal{L}_\circ$. The corresponding parameter $d(0)$ is the root $1/4$ if of the main cardioid. The corresponding Julia set (cauliflower) is a Jordan curve (Theorem 26.1), so its lamination is trivial. The conclusion follows.

For an irrational θ, the Wake Decomposition Theorem, 37.15, implies that $\mathcal{L}_{c(\theta)} = \mathcal{L}_\circ$, while $\mathcal{L}_d$ is trivial, and we are done again.

For any other $c \in M_\circ$ (including non-root parabolics $c(\mathfrak{p}/\mathfrak{q}) \in \partial\Delta_\circ \smallsetminus \{\mathfrak{r}_\circ\}$), let $\mathfrak{R}_c = \{\mathcal{R}_c^{\theta_k}\}$ be the configuration of rays for f_c landing at the $g-$fixed point $\alpha_g \in \mathcal{K}(g)$, and let $\Theta_c := \{\theta_k\} \subset \mathbb{T}$ be the corresponding set of angles. This is a tuned rotation cycle in the sense defined in §28.4.10. By Lemma 43.12, the image configuration $h(\mathfrak{R}_c)$ represents all accesses to α_d (cyclically permuted by f_d with rotation number $\mathfrak{p}/\mathfrak{q}$). By the Lindelöf Theorem, this configuration is represented by a cycle $\mathfrak{R}_d = \{\mathcal{R}_d^{\omega_i}\}$ of external rays in $\mathbb{C} \smallsetminus \mathcal{K}(f_d)$, where $\{\omega_i\} = \Theta_{\mathfrak{p}/\mathfrak{q}}$ is the rotation set on $\mathbb{T}$ with rotation number $\mathfrak{p}/\mathfrak{q}$. By definition, the pullback of $\Theta_{\mathfrak{p}/\mathfrak{q}}$ by π to $\mathbb{T}$ (where π is defined before Proposition 43.11) is the tuning of $\mathcal{L}_\circ$ by $\Theta_{\mathfrak{p}/\mathfrak{q}}$. On the other hand, it is exactly Θ_c.

We have shown that $\mathcal{L}_c$ contains $\mathcal{L}_\circ * \Lambda_{\mathfrak{p}/\mathfrak{q}}$, where $\Lambda_{\mathfrak{p}/\mathfrak{q}}$ is the ideal polygon spanned by $\Theta_{\mathfrak{p}/\mathfrak{q}}$. It follows that $\mathcal{L}_c$ contains $\mathcal{L}_\circ * \mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$, where $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}}$ is the completely invariant lamination generated by $\Lambda_{\mathfrak{p}/\mathfrak{q}}$ (which coincides with the puzzle lamination $\mathcal{L}_d^{[0]}$ of zero level for f_d, see §32.1.5). So, $\mathcal{L}_c \succ \mathcal{L}_\circ * \mathcal{L}_d^{[0]}$.

If our f_d is non-renormalizable and has both fixed point repelling (so, c is not a parabolic point on $\partial\Delta_\circ$), then by Proposition 32.9 (i), the lamination $\mathcal{L}_d^{[0]}$ coincides with the whole combinatorial lamination $\mathcal{L}_d$ and the latter is maximal among clean laminations (since polygonal), implying the desired result:

$$\mathcal{L}_c = \mathcal{L}_\circ * \mathcal{L}_d^{[0]}. \tag{43.3}$$

Assume f_d is exactly once renormalizable, maybe in the degenerate sense (so c can be a parabolic point on $\partial\Delta_\circ$). This renormalization corresponds to some little copy $M^{[1]} \subset M_\circ$ containing c which is centered at some superatracting parameter

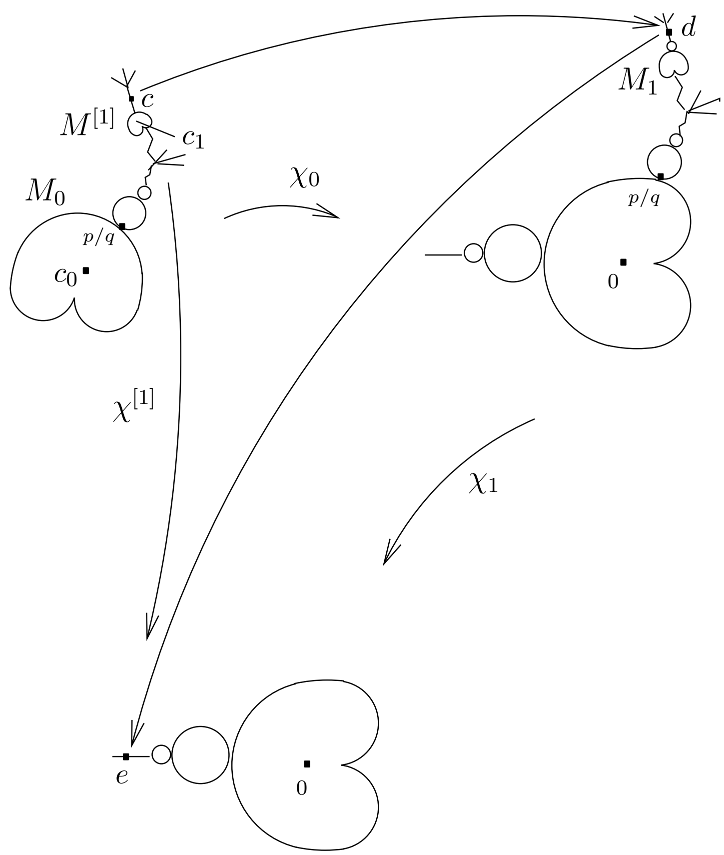

FIGURE 43.2. Composition of two tunings.

c_1. Let $\chi^{[1]} : M^1 \to \mathcal{M}$ be the straightening of this copy (see Figure 43.2). Then $c = c_1 * e$ for $e = \chi^{[1]}(c)$. Applying (43.3) to this situation (replacing $c_\circ$ with c_1 and d with e), we conclude that $\mathcal{L}_c$ coincides with the lamination $\mathcal{L}_1 * \mathcal{L}_e^{[0]}$, where $\mathcal{L}_1$ is the lamination for f_{c_1} while $\mathcal{L}_e^{[0]}$ is the combinatorial lamination for f_e. Moreover, $\mathcal{L}_1 = \mathcal{L}_\circ * \mathcal{L}_{d_1}$, where $d_1 = \chi_0(c_1)$ is the center of the little copy $M_1 := \chi_\circ(M^{[1]})$. Hence

$$\mathcal{L}_c = (\mathcal{L}_\circ * \mathcal{L}_{d_1}) * \mathcal{L}_e^{[0]} = \mathcal{L}_\circ * (\mathcal{L}_{d_1} * \mathcal{L}_e^{[0]}) = \mathcal{L}_\circ * \mathcal{L}_d^{[0]},$$

where the last equality follows from the above discussion applied to the straightening

$$\chi_1 = \chi^{[1]} \circ \chi_\circ^{-1} : (M^{[1]}, d_1, d) \to (\mathcal{M}, 0, e).$$

If f_d is twice renormalizable, we proceed inductively.

In this way, we prove the assertion for all periodically repelling maps f_b which are at most finitely renormalizable. In the infinitely renomalizable case, we conclude that

$$\mathcal{L}_c \succ \mathcal{L}_\circ * \mathcal{L}_1 * \mathcal{L}_2 * \dots$$

But the latter is the lamination corresponding to the whole puzze $\mathcal{Y}$. By Theorem 32.9, it coincides with the combinatorial lamination $\mathcal{L}_c$. Applying this result to the lamination $\mathcal{L}_1 * \mathcal{L}_2 * \dots$, we conclude that it coincides with $\mathcal{L}_d$, concluding the proof. □

COROLLARY 43.13. *Under the above circumstances, for any (pre-)periodic point $z \in \mathcal{K}(g)$, except the β−fixed point (for g) and its iterated preimages, the number of accesses to z from $\mathbb{C} \smallsetminus \mathcal{K}(g)$ and from $\mathbb{C} \smallsetminus \mathcal{K}(f_c)$ is the same.*

To formulate a parameter counterpart of this statement (and of Lemma 28.36). let us introduce a *dyadic wake* of a little copy M as the prepreriodic wake of $\mathcal{M}$ rooted at a dyadic tip of M. The most prominent one is the *Chebyshev wake* rooted at the Chebyshev point of M. The number of such wakes at each dyadic tip is equal to $\mathfrak{q} - 1$, where $\mathfrak{p}/\mathfrak{q}$ is the rotation number of the parabolic cycle at the root of M.

EXERCISE 43.14. (i) *Any little copy M is obtained from $\mathcal{M}$ by chopping off the body sector attached at the root $\mathfrak{r}$ of M, together with all the dyadic wakes of M.*

(ii) *For ant other parabolic or Misiurewicz point $z \in M$, the number of accesses to z from $\mathbb{C} \smallsetminus \mathcal{M}$ and from $\mathbb{C} \smallsetminus M$ is the same (in a natural way).*

Let us now relate different combinatorial notions in the infinitely renormalizable case:

COROLLARY 43.15. *Two infinitely renormalizable quadratic polynomials are combinatorially equivalent if and only if they have the same renormalization combinatorics.*

PROOF. As we know from Proposition 25.61, the Hubbard tree $\mathcal{T}_f$ of a hyperbolic map f determines the lamination $\mathcal{L}_f$ (which can be thus denoted $\mathcal{L}_{\mathcal{T}}$), and the other way around. Hence the renormlization combinatorics $\bar\tau = (\mathcal{T}_\circ, \mathcal{T}_1, \mathcal{T}_2, \dots)$ of an infinitely renormalizable map f determines its combinatorial laminatoin $\mathcal{L}_f = \mathcal{L}_{\mathcal{T}_\circ} * \mathcal{L}_{\mathcal{T}_1} * \mathcal{L}_{\mathcal{T}_2} \dots$, and the other way around. □

REMARK 43.16. As written, this corollary looks like a purely combinatorial relation between laminations and combinatorial Hubbard trees. And indeed, it could be viewed as such. However, the Thurston theory guarantees that any admissible Hubbard tree is realized by a unique superttracting parameter (see also Rigidity Theorem 35.1 for the uniqueness part). Then the Douady-Hubbard theory of little copies guarantees that any renormalization combinatorics $\bar\tau$ is realized by some infinitely renormalizable map.

43.5. Renormalization structure of $\mathcal{M}$.

43.5.1. *Complex case.* It is tautological to say that $\mathcal{M}$ is decomposed into subsets of renormalizable and non-renormalizable parameters. It becomes much less so if we are able to describe the structure of these two subsets. Already at this stage, we posses quite a deep information:

THEOREM 43.17. *The Mandelbrot set $\mathcal{M}$ admits the following decomposition:*

$$\mathcal{M} = \Delta_0 \cup \mathcal{N}_0 \cup \bigcup \mathcal{M}_{\mathfrak{p}/\mathfrak{q}} \cup \bigcup \mathcal{M}_{\mathcal{T}} \cup \mathfrak{Y}_0, \tag{43.4}$$

where

- *Δ_0 is the main hyperbolic component of* int $\mathcal{M}$ *bounded by*
- *the main cardioid $\mathcal{N}_0$;*

• $\mathcal{M}_{\mathfrak{p}/\mathfrak{q}}$ *is the satellite M-copy attached to the main cardioid at the parabolic points with rotation numbers $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$ (corresponding to the immediate renormalization with this rotation number);*

• $\mathcal{M}_{\mathcal{T}}$ *is the maximal primitive M-copy corresponding to a prime Hubbard tree $\mathcal{T}$;*

• $\mathfrak{Y}_0$ *is the set of non-renormalizable parameters with both fixed points repelling.*

This decomposition is disjoint except for the bifurcation points $\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}} \in \mathcal{N}_0$ on the main cardioid where the corresponding satellite M-copies are attached.

Proof. Tautologically, we have the following decomposition:

$$\mathcal{M} = \Delta_0 \sqcup \mathcal{N}_0 \sqcup \{\text{immediately renormalizable parameters}\}$$

$$\sqcup \{\text{non} - \text{immediately renormalizable parameters}\} \sqcup \mathfrak{Y}_0 \tag{43.5}$$

Theorem 43.1 implies that the set of immediately renormalizable parameters is decomposed into the disjoint union of unrooted satellite copies $\mathcal{M}^*_{\mathfrak{p}/\mathfrak{q}}$ (whose closures are the satellite copies

$$\mathcal{M}_{\mathfrak{p}/\mathfrak{q}} = \operatorname{cl} \mathcal{M}^*_{\mathfrak{p}/\mathfrak{q}} = \mathcal{M}^*_{\mathfrak{q}/\mathfrak{p}} \cup \{\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}}\},$$

obtained by adding the corresponding roots.)

Theorem 31.17 asserts that non-immediately renormalizable parameters are primitively renormalizable. Theorem 43.1 implies that the set of such parameters is decomposed into the disjoint union of the maximal primitive M-copies according to the combinatorics of the first renormalization, which is encoded by a prime Hubbard tree.

Altogether, this provides us with the desired decomposition. □

Applying this decomposition to the M-copies themselves, we obtain a decomposition of $\mathcal{M}$ according to the second renormalization type, etc. At the limit, we obtain the following decomposition according to the infinite renormalization type:

Corollary 43.18. *The Mandelbrot set $\mathcal{M}$ admits the following decomposition:*

$$\mathcal{M} = \mathfrak{H} \sqcup \mathfrak{N} \sqcup \mathfrak{Y} \sqcup \mathfrak{I},$$

where

• $\mathfrak{H}$ *is the set of all hyperbolic parameters (i.e., the union of all hyperbolic components of* $\operatorname{int} \mathcal{M}$*);*

• $\mathfrak{N}$ *is the set of all neutral parameters (i.e., the union of the boundaries of all hyperbolic components);*

• $\mathfrak{Y}$ *is the set of all Yoccoz parameters (i.e., periodically repelling at most finitely renormalizable parameters);*

• $\mathfrak{I}$ *is set of infinitely renormalizable parameters: each of these parameters belongs to an infinite nest of little M-copies.*

Remark 43.19. Except for the last assertion, this decomposition is actually tautological.

43.5.2. *Real case.* Taking the real slice of the above decomposition, we obtain:

THEOREM 43.20. *The real Mandelbrot set $\mathcal{M}_{\mathbb{R}}$ admits the following decomposition:*

$$\mathcal{M}_{\mathbb{R}} = \{c_{\text{caul}}\} \cup \Delta_0^{\mathbb{R}} \cup \{c_{\text{Myr}}\} \cup \bigcup M_{\text{doub}}^{\mathbb{R}} \cup \bigcup M_{\text{Kn}}^{\mathbb{R}} \cup \mathfrak{Y}_0^{\mathbb{R}},$$

where

- $c_{\text{caul}} = 1/4$ *is the cauliflower parameter;*
- $\Delta_0^{\mathbb{R}} = (-3/4, 1/4)$ *is the main hyperbolic window of* $\mathcal{M}_{\mathbb{R}}$*;*
- $c_{\text{Myr}} = -3/4$ *is the Myrberg doubling bifurcation parameter;*
- $M_{\text{doub}}^{\mathbb{R}} = [c_{\mathfrak{t}}, -3/4]$ *is the* (*closed*) *doubling renormalization window, where* $c_{\mathfrak{t}}$ *is its tip* (*solving the equation* $f_c^2(c) = -f_c(c)$)*;*
- $M_{\text{Kn}}^{\mathbb{R}}$ *is the maximal primitive renormalization interval corresponding to a prime superattracting kneading sequence* Kn*.*
- $\mathfrak{Y}_0^{\mathbb{R}}$ *is the set of real non-renormalizable parameters with both fixed points repelling.*

This decomposition is disjoint except that the Myrberg parameter c_{Myr} *belongs to the closed doubling renormalization interval* $M_{\text{doub}}^{\mathbb{R}}$*.*

Several remarks are due:

– We single out the doubling renormalization as it is the only real satellite renormalization type (for the first renormalization).

– We single out the Myrberg doubling bifurcation parameter since it is the only parameter which is almost renormalizable, but not renormalizable (according to our definitions).

– The definitions are designed so that the renormalization intervals are closed, even in the doubling case.

The real slice of Corollary 43.18 assumes the following form:

COROLLARY 43.21. *The real Mandelbrot set $\mathcal{M}_{\mathbb{R}}$ admits the following decomposition:*

$$\mathcal{M}_{\mathbb{R}} = \mathfrak{H}_{\mathbb{R}} \sqcup \mathfrak{N}_{\text{sn}}^{\mathbb{R}} \sqcup \mathfrak{N}_{\text{Myr}}^{\mathbb{R}} \sqcup \mathfrak{Y}_{\mathbb{R}} \sqcup \mathfrak{I}_{\mathbb{R}},$$

where

- $\mathfrak{H}_{\mathbb{R}}$ *is the set of all real hyperbolic parameters* (*i.e., the union of all hyperbolic windows of* $\mathcal{M}_{\mathbb{R}}$)*;*
- $\mathfrak{N}_{\text{sn}}^{\mathbb{R}}$ *is the set of all saddle-node parameters* (*where the saddle-node bifurcation occur*)*;*
- $\mathfrak{N}_{\text{Myr}}^{\mathbb{R}}$ *is the set of all Myrberg parameters* (*where the doubling bifurcations occur*)*;*
- $\mathfrak{Y}_{\mathbb{R}}$ *is the set of all real Yoccoz parameters* (*i.e., real periodically repelling at most finitely renormalizable parameters*)*;*
- $\mathfrak{I}_{\mathbb{R}}$ *is set of all real infinitely renormalizable parameters: each of these parameters belongs to an infinite nest of renormalization intervals.*

43.5.3. *Preperiodic-superattracting bifurcation.* Let us first consider the Chebyshev map $f_\circ = f_{-2}$. Since the interval $\mathcal{I}_+ := [0,2]$ is mapped homeomorphically onto the whole interval $\mathcal{I} = [-2,2]$, we obtain a backward orbit of 0,

$$\cdots \mapsto x_{-2} \mapsto x_{-1} \mapsto 0, \quad x_{-n} \in (0,2),$$

converging to the fixed point $\beta = 2$. Moreover, these points robust under perturbation: for c near 2, we have the corresponding backward orbit

$$\cdots \mapsto x_{-2}(c) \mapsto x_{-1}(c) \mapsto 0$$

depending holomorphically on c. Solving equation $c^2 + c = x_{-n}(c)$, we obtain a sequence of superattracting parameters c_n (of period $n+2$) converging to $c_\circ = -2$ at rate 4^{-n}.

LEMMA 43.22. *The above superattracting parameters c_n are the centers of little M-copies M_n of size $\asymp 16^{-n}$. Moreover, the characteristic rays $\mathcal{R}_n^\pm$ through the roots $\mathfrak{r}_n$ of M_n converge to the ray $(-\infty, -2]$.*

More generally, we have:

LEMMA 43.23. *Let $c_\circ$ be a preperiodic parameter with multiplier ρ. Then there is a sequence c_n of superattracting parameters converging to $c_\circ$ at rate $\asymp \rho^{-n}$ which are the centers of little M-copies M_n of size $\asymp \rho^{-2n}$. Moreover, the characteristic rays $\mathcal{R}_n^\pm$ through the roots $\mathfrak{r}_n$ of M_n converge (subsequentially) to the valuable parameter rays $\mathcal{R}^{\eta_i}_{\mathrm{par}}$ landing at $c_\circ$ (described in* Theorem 37.35*).*

COROLLARY 43.24. *Any preperiodic parameter $c_\circ$ is well branched with $\mathcal{R}^{\mathrm{par}}_{\eta_i}$ being the only rays landing at $c_\circ$.*

Notes. The Straightening Theorem for ql families, justifying existence of little M-copies, is one of the most fundamental results in Holomorphic Dynamics. It was proven by Douady and Hubbard in the mid 1990s [**DH3**].

The higher degree case was studied in [**IK**]. No miracles of continuity occur in this case, see [**DH3, I**]. In the *anti-holomorphic* world, the continuity breaks down even in the unimodal case, see [**IMu**]. The corresponding image (the bifurcation locus for the family $\bar{z}^2 + c$) is known under the name of *Tricorn*: see Figure 43.3 [**M7, HSc2**].

The idea of *tuning* was articulated in [**D2**].

44. Complex *a priori bounds*: first results

In the renormalization contexts that we discuss *complex a priori bounds* usually mean lower bounds on the moduli of certain annuli that control the geometry of the Julia sets. In this section we will consider two quite different situations when such bounds are valid: for Yoccoz maps (Theorem 44.9) and for real Feigenbaum maps (Theorem 44.11).

44.1. Principal moduli. Let us introduce the *principal moduli*:

$$\mu_n := \operatorname{mod} A^n \equiv \operatorname{mod}(V^{n-1} \smallsetminus V^n)$$

Below we will give *a priori bounds* on the μ_n that will provide us with a strong geometric control of the whole puzzle. In particular, they control the *distortion* of the generalized renormalizations. To this end, we will introduce *asymmetric moduli* and prove that they do not decrease under the generalized renormalization. The precise formulation (Theorem 44.9) is given at the end of the section.

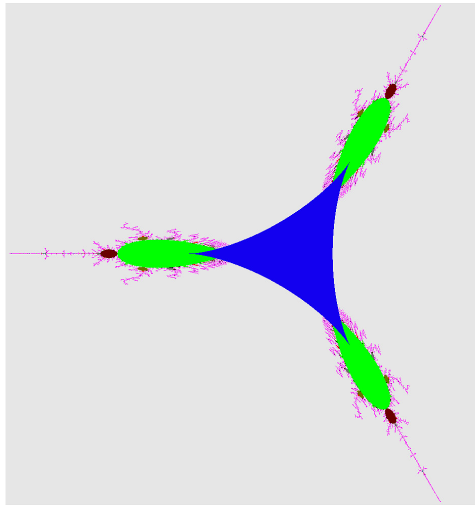

FIGURE 43.3. Tricorn.

44.2. Fibonacci case. Let us first consider the Fibonacci case which is combinatorially easy but already reveals the main geometric idea. So, we assume that the generalized renormalizations have a form (see Figure 31.8.1):

$$g_n : V_0^n \cup V_1^n \to V^{n-1}, \quad g_n(0) \in V_1^n,\ g_n^2(0) \in V_0^n.$$

Let $S_{n-1} := V^{n-1} \smallsetminus (V_0^n \cup V_1^n)$, and let

$$\operatorname{mod}[\partial V_0^n] := \sup \operatorname{mod} R_0^n, \quad \operatorname{mod}[\partial V_1^n] := \sup \operatorname{mod} R_1^n,$$

where the sup are taken over all annuli $R_0^n \subset S^{n-1}$ and $R_1^n \subset S^{n-1}$ homotopic to ∂V_0^n and ∂V_1^n respectively (compare §6.7). Then the *asymmetric modulus* is defined as

$$\sigma_n := \operatorname{mod}[\partial V_0^n] + \frac{1}{2} \operatorname{mod}[\partial V_1^n].$$

PROPOSITION 44.1. *The asymmetric moduli do not decrease under the generalized renormalization: $\sigma_{n+1} \geq \sigma_n$.*

PROOF. Let us take an annulus R_0^n as above and lift it to an annulus $R_1^{n+1} \subset S^n$ homotopic to ∂V_1^{n+1}, so that $g_n : R_1^{n+1} \to R_0^n$ is a conformal isomorphism. Then

$$\operatorname{mod}[\partial V_1^{n+1}] \geq \operatorname{mod} R_1^{n+1} = \operatorname{mod} R_0^n.$$

Similarly:
– lifting an annulus $R_1^n \subset S^{n-1}$ homotopic to ∂V_1^n we obtain an annulus $R_0^{n+1} \subset S^n$ homotopic to ∂V_0^{n+1} that double covers R_1^n,
– lifting the annulus $V_1^n \smallsetminus g_n(V_0^{n+1})$, we obtain a disjoint annulus $T_0^{n+1} \subset S^n$ in the same homotopy class that double covers $V_1^n \smallsetminus g_n(V_0^{n+1})$.

By the Grötzsch Inequality,

$$\operatorname{mod}[\partial V_0^{n+1}] \geq \operatorname{mod} R_0^{n+1} + \operatorname{mod} T_0^{n+1} = \frac{1}{2}(\operatorname{mod} R_1^n + \operatorname{mod}(V_1^n \smallsetminus g_n(V_0^{n+1}))$$

$$= \frac{1}{2}(\operatorname{mod} R_1^n + \mu_n) \geq \frac{1}{2}(\operatorname{mod} R_1^n + \operatorname{mod} R_0^n) \tag{44.1}$$

where the bottom line equality holds since g_n conformally maps $V_1^n \setminus g_n(V_0^{n+1})$ onto A^n. Taking the asymmetric combination of the above moduli, we obtain

$$\operatorname{mod}[\partial V_0^{n+1}] + \frac{1}{2}\operatorname{mod}[\partial V_1^{n+1}] \geq \operatorname{mod} R_0^n + \frac{1}{2}\operatorname{mod} R_1^n. \tag{44.2}$$

Taking the supremum over all annuli R_0^n, R_1^n in question, we obtain the desired estimate. □

COROLLARY 44.2. *If f has the Fibonacci combinatorics on all principal levels, then on all these levels we have: $\mu_n \geq \frac{1}{4}\mu_1$.*

PROOF. To pass from the asymmetric to the principal moduli, use (44.1):

$$\mu_{n+1} \geq \operatorname{mod}[\partial V_0^{n+1}] \geq \frac{1}{2}(\operatorname{mod} R_1^n + \operatorname{mod} R_0^n) \geq \frac{1}{2}\sigma_n \geq \frac{1}{2}\sigma_2 \quad (n \geq 2).$$

But $\sigma_2 \geq \frac{1}{2}\mu_1$, implying $\mu_n \geq \frac{1}{4}\mu_1$ for $n \geq 3$. Finally, note that (44.1) also yields $\mu_{n+1} \geq \frac{1}{2}\mu_n$, so $\mu_2 \geq \frac{1}{2}\mu_1$. □

44.3. General case: first estimates. Let L_n be the first landing map to V^n. By (31.14), for each V_i^n we have a decomposition:

$$g_n | V_i^n = L_{n-1} \circ f | V_i^n, \tag{44.3}$$

where L_{n-1} conformally maps $f(V_i^n)$ onto V^{n-1}.

Lemma 31.14, Exercise 6.13(i), and the Koebe Distortion Theorem imply:

LEMMA 44.3. *If $\mu_{n-1} \geq \underline{\mu}$ the distortion of all branches $L_{n-1} | f(V_i^n)$ are bounded by $C(\underline{\mu})\exp(-\mu_{n-1})$).*

Let us fix a level $n > 0$, denote $V^{n-1} = \Delta$, $V_i = V_i^n$, $g = g_n$, $A = A^n = \Delta \setminus V_0$, $\mu = \mu_n$, and mark the objects of the next level $n+1$ with prime. Thus $\Delta' \equiv V \equiv V_0$, and $g' : \cup V_i' \to \Delta'$. (We restore the index n whenever we need it).

LEMMA 44.4. *Let $D' \subset \Delta'$ be a puzzle piece such that $g^k D' \subset V_{i(k)}$, $k = 1, \dots, l$, with $i(k) \neq 0$ for $0 < k < l$, while $g^{l+1}(D') = \Delta'$. Then*

$$\operatorname{mod}(\Delta' \setminus D') \geq \frac{1}{2}\sum_{k=1}^{l+1} \operatorname{mod}(\Delta \setminus V_{i(k)}).$$

PROOF. Let us consider the following nest of topological disks:

$$\Delta' \equiv W_1 \supset \dots \supset W_{l+1} \supset W_{l+2} = D',$$

where W_{k+1} is defined inductively as the pullback of $V_{i(k)}$ under $g^k : W_k \to \Delta$, $k = 1, \dots l$ (compare with the Telescope from §3.6). Since $\deg(g^k : W_k \to \Delta) \leq 2$,

$$\operatorname{mod}(W_k \setminus W_{k+1}) \geq \frac{1}{2}\operatorname{mod}(\Delta \setminus V_{i(k)}), \quad 1 \leq k \leq l.$$

But by the Grötzsch inequality

$$\operatorname{mod}(\Delta' \setminus D') \geq \sum_{k=1}^{l+1} \operatorname{mod}(W_k \setminus W_{k+1}),$$

and the desired estimate follows. □

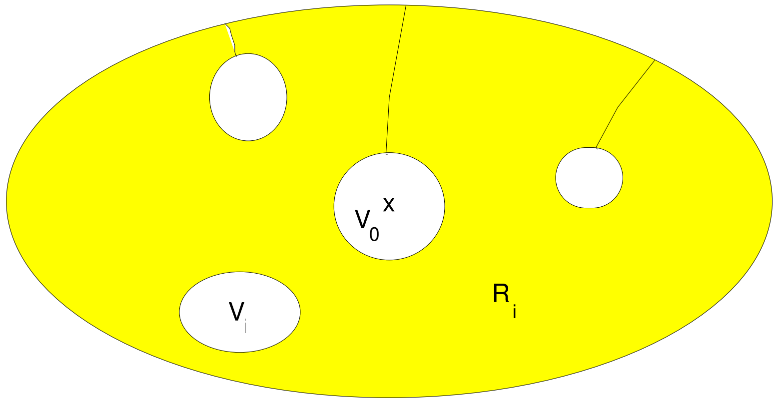

FIGURE 44.1. Annulus R_i.

COROLLARY 44.5. *Given a puzzle piece V'_j, we have*

$$\operatorname{mod}(\Delta' \smallsetminus V'_j) \geq \frac{1}{2}\mu.$$

Moreover, if the return to level n is non-central, that is $g(0) \in V_i$ with $i \neq 0$, then

$$\operatorname{mod}(\Delta' \smallsetminus V'_j) \geq \frac{1}{2}(\mu + \operatorname{mod}(\Delta \smallsetminus V_i)).$$

So, a definite principal modulus on some level produces a definite space around all the puzzle pieces of the next level.

44.4. Isles and asymmetric moduli. Let $\mathcal{V}^n$ stand for the family of all pieces V_i^n of level n. Let $\{V_i\}_{i\in\mathcal{I}} \subset \mathcal{V}^n$ be a finite family of disjoint puzzle pieces consisting of at least two pieces (that is $|\mathcal{I}| \geq 2$) and containing a critical puzzle piece V_0. Let us call such a family *admissible*. We will freely identify the label set $\mathcal{I}$ with the family itself.

Given a puzzle piece D, let $\mathcal{I}|D$ denote the family of puzzle pieces of $\mathcal{I}$ contained in D. Let D be a puzzle piece containing at least two pieces of family $\mathcal{I}$. For $V_i \subset D$ let

$$R_i \equiv R_i(\mathcal{I}|D) \subset D \smallsetminus \bigcup_{j\in\mathcal{I}|D} V_j$$

be an annulus of maximal modulus enclosing V_i but not enclosing other pieces of the family $\mathcal{I}$. Such an annulus exists by Theorem 6.38 (see Figure 44.1). We will briefly call it the *maximal annulus* enclosing V_i in D (rel the family $\mathcal{I}$).

Let us define *the asymmetric modulus of the family $\mathcal{I}$ in D* as

$$\sigma(\mathcal{I}|D) = \sum_{i\in\mathcal{I}} \frac{1}{2^{1-\delta_{i0}}} \operatorname{mod} R_i(\mathcal{I}|D),$$

where δ_{ji} is the Kronecker symbol. So the critical modulus is supplied with weight 1, while the off-critical moduli are supplied with weights $1/2$ (if D is off-critical then all the weights are actually $1/2$).

For $D = V^{n-1}$, let $\sigma_n(\mathcal{I}) \equiv \sigma(\mathcal{I}|V^{n-1})$. The asymmetric modulus of level n is defined as follows:

$$\sigma_n = \min_{\mathcal{I}} \sigma_n(\mathcal{I}),$$

where $\mathcal{I}$ runs over all admissible subfamilies of $\mathcal{V}^n$.

The principal moduli μ_n and the asymmetric moduli σ_n are the main geometric parameters of the renormalized maps g_n. Again, in what follows the label n will be suppressed as long as the level is not changed.

Let $\{V_i'\}_{i\in\mathcal{I}'}$ be an admissible subfamily of $\mathcal{V}'$. Let us organize the pieces of this family in *isles* in the following way. A puzzle piece $D'\subset\Delta'$ is called an *island* (for family I') if

- D' contains at least two puzzle pieces of family $\mathcal{I}'$;
- There is a $t\geq 1$ such that $g^k D'\subset V_{i(k)}$, $k=1,\dots t-1$, with $i(k)\neq 0$, while $g^t D=\Delta$.

Given an island D', let $\phi_{D'}=g^t : D'\to\Delta$. This map is either a double covering or a biholomorphic isomorphism depending on whether D' is critical or not. In the former case, $D'\supset V_0'$ (for otherwise $D'\subset V_0'$ contradicting the first part of the definition of isles).

We call a puzzle piece $V_j'\subset D'$ $\phi_{D'}$*-precritical* if $\phi_{D'}(V_j')=V_0$. There are at most two precritical pieces in any D'. If there are actually two of them, then they are off-critical and symmetric with respect to the critical point 0. In this case D' must also contain the critical puzzle piece V_0'.

Let $\mathcal{D}'=\mathcal{D}(\mathcal{I}')$ be the family of isles associated with $\mathcal{I}'$. Let us consider the asymmetric moduli $\sigma(\mathcal{I}'|D')$ as a functional on this family. This functional is clearly monotone:

$$\sigma(\mathcal{I}'|D')\geq\sigma(\mathcal{I}'|D_1')\quad\text{if}\quad D'\supset D_1', \tag{44.4}$$

and superadditive:

$$\sigma(\mathcal{I}'|D')\geq\sigma(\mathcal{I}'|D_1')+\sigma(\mathcal{I}'|D_2'),$$

provided the D_i' are disjoint sub-isles in D'.

Let us call an island D' *innermost* if it does not contain any other isles of the family $\mathcal{D}(\mathcal{I}')$. As this family is finite, innermost isles exist.

44.5. Non-decreasing of the asymmetric moduli.

LEMMA 44.6. *Let $\mathcal{I}'$ be an admissible family of puzzle pieces. Let D' be an innermost island associated to the family $\mathcal{I}'$, and let $\mathcal{J}'=\mathcal{I}'|D$. For $j\in\mathcal{J}'$, let us define $i(j)$ by the property $\phi_{D'}(V_j')\subset V_{i(j)}$, and let $\mathcal{I}=\{i(j): j\in\mathcal{J}'\}\cup\{0\}$. Then $\{V_i\}_{i\in\mathcal{I}}$ is an admissible family of puzzle pieces, and*

$$\sigma(\mathcal{I}'|D')\geq\frac{1}{2}\left((|\mathcal{J}'|-s)\mu+s\bmod R_0+\sum_{j\in\mathcal{J}',\, i(j)\neq 0}\bmod R_{i(j)}\right), \tag{44.5}$$

where $s=\#\{j: i(j)=0\}$ is the number of $\phi_{D'}$-precritical pieces, and R_i are the maximal annuli enclosing V_i in Δ rel $\mathcal{I}$.

PROOF. Let $\phi\equiv\phi_{D'}$. Let us show first that the family $\mathcal{I}$ is admissible. This family is finite since $\mathcal{J}'\subset\mathcal{I}'$ is finite. The critical puzzle piece belongs to $\mathcal{I}$ by definition. So the only property to check is that $|\mathcal{I}|\geq 2$. But otherwise $\mathcal{J}'$ would consist of two precritical puzzle pieces. But then D' would be critical, and thus should have also contained the critical piece V_0', which is a contradiction.

Let us observe next that

$$\operatorname{mod}(V_{i(j)}\smallsetminus\phi(V_j'))\geq\mu\quad\text{if}\quad i(j)\neq 0. \tag{44.6}$$

Indeed, in this case $g^m(\phi(V'_j)) = V_0$ for some $m > 0$. Let $W \subset V_{i(j)}$ be the pull-back of Δ under g^m. Then the annulus $W \smallsetminus \phi(V'_j)$ is univalently mapped by g^m onto the annulus $\Delta \smallsetminus V_0$. Hence $\operatorname{mod}(W \smallsetminus \phi(V'_j)) = \operatorname{mod}(\Delta \smallsetminus V_0) = \mu$, and (44.6) follows.

Given an $i \in I$, let us consider a topological disk $Q_i = R_i \cup V_i \subset \Delta$ (the "filled annulus R_i"). By the Grötzsch inequality and (44.6),

$$\operatorname{mod}(Q_{i(j)} \smallsetminus \phi(V'_j)) \geq \operatorname{mod} R_{i(j)} + (1 - \delta_{0,i(j)})\mu. \tag{44.7}$$

For $j \in J'$, let us consider an annulus $B'_j \subset D'$, the component of $\phi^{-1}(R_{i(j)})$ enclosing V'_j. This annulus does not enclose any other pieces $V'_k \in \mathcal{J}'$, $k \neq j$. Indeed, otherwise the inner component of $\mathbb{C} \smallsetminus B'_j$ would be an island contained in D', despite the assumption that D' is innermost.

Let us now consider a topological disk P'_j obtained by filling the annulus B'_j. Then

$$\operatorname{mod} R'_j \geq \operatorname{mod}(P'_j \smallsetminus V'_j), \tag{44.8}$$

where $R'_j \subset D'$ is the maximal annulus enclosing V'_j rel $\mathcal{J}'$. Moreover $\phi : P'_j \to Q_{i(j)}$ is univalent or double covering depending on whether $j \neq 0$ or $j = 0$. Hence

$$\operatorname{mod}(P'_j \smallsetminus V'_j) \geq \frac{1}{2^{\delta_{j0}}} \operatorname{mod}(Q_{i(j)} \smallsetminus \phi(V'_j)). \tag{44.9}$$

Inequalities (44.7)-(44.9) yield

$$\operatorname{mod} R'_j \geq \frac{1}{2^{\delta_{j0}}}(\operatorname{mod} R_{i(j)} + (1 - \delta_{0,i(j)})\mu). \tag{44.10}$$

Summing up estimates (44.10) over $\mathcal{J}'$ with weights $1/2^{1-\delta_{j0}}$, we obtain the desired inequality. $\square$

Corollary 44.7. *For any island D' of the family $\mathcal{D}'$ the following estimates hold:*

$$\sigma(\mathcal{I}'|D') \geq \frac{1}{2}\mu \quad \textit{and} \quad \sigma(\mathcal{I}'|D') \geq \sigma(\mathcal{I}) \geq \sigma.$$

Hence $\sigma' \geq \sigma$.

Proof. By monotonicity (44.4), it is enough to check the case of an innermost island D'. Let us use the notations of the previous lemma. Since the family $\mathcal{I}$ is admissible, it contains an off-critical piece. Hence $|\mathcal{J}'|$ is always strictly greater than the number s of precritical pieces in D', and (44.5) implies the first of the above inequality.

Furthermore, as $\mu \geq \operatorname{mod}(R_0)$ and $|\mathcal{J}'| \geq 2$, the right-hand side in (44.5) is bounded from below by

$$\frac{1}{2}\left(|\mathcal{J}'| \operatorname{mod} R_0 + \sum_{i \in \mathcal{I}, i \neq 0} \operatorname{mod} R_i\right) \geq \sigma(\mathcal{I}).$$

(Note that $\sigma(\mathcal{I})$ makes sense since $\mathcal{I}$ is admissible). Finally $\sigma(\mathcal{I}) \geq \sigma$, and the second inequality follows.

Taking the infimum over all admissible families, we obtain the last conclusion. $\square$

COROLLARY 44.8. *If the return under consideration is non-central* (*i.e.*, $g(0) \in V_i$ *with* $i \neq 0$) *then*

$$\mu' \geq \operatorname{mod} R'_0 \geq \frac{1}{2}\sigma.$$

PROOF. Estimate (44.10) for $j = 0$ gives

$$\operatorname{mod} R'_0 \geq \frac{1}{2}(\mu + \operatorname{mod} R_i) \geq \frac{1}{2}\sigma.$$

□

Let (m_k) be the sequence of non-central levels of the Principal Nest. Let us define a function $d(n)$ as follows: $d(1) = 0$, while for $n \geq 2$, let

$$d(n) = n - 2 - \max\{m_k : \ m_k \leq n - 2\} \in \mathbb{Z}_+.$$

Thus, $d(n) = 0$ iff the level $n - 2$ is non-central; otherwise, $d(n)$ measures the distance from $n - 2$ to the closest preceding non-central level.

THEOREM 44.9. *Let* $g : \bigcup V_i^1 \to V^0$ *be a generalized quadratic-like map with*

$$\mu_1 \equiv \operatorname{mod}(V^0 \smallsetminus V^1) \geq \underline{\nu} > 0.$$

Then:

(i) *The asymmetric moduli* σ_n *grow monotonically and hence stay away from 0 on all levels:* $\sigma_n \geq \underline{\sigma} > 0$.

(ii) *The principal moduli* μ_n *are bounded from below by* $\underline{\mu}\, 2^{-d(n)}$, *with* $\underline{\mu}$ *depending only on* $\underline{\nu}$. *In particular,* $\mu_n \geq \underline{\mu}$ *if the level* $n-2$ *is non-central* (*i.e., when* $d(n) = 0$).

(iii) *The distortion of all branches of the first landing maps*

$$L_n : \bigsqcup_i Q_i^n \to V^n$$

are bounded on all levels by some $K(\underline{\nu})$. *In particular, the distortion of the diffeomorphisms* L_{n-1} *from* (31.14) *on* $f(V_i^n)$ *are bounded by* $K(\underline{\nu})$.

PROOF. The first assertion follows from Corollary 44.7. Together with Corollary 44.8 it implies the second one. Together with Lemma 31.20 and the Koebe Distortion Theorem, it implies the last one. □

COROLLARY 44.10. *Let* f *be a non-renormalizable quadratic polynomial with both fixed points repelling. If all the levels of the Principal Nest are non-central, then all the principal moduli stay away from zero:* $\mu_n \geq \underline{\mu} > 0$ (*with* $\underline{\mu}$ *depending only on the lower bound* $\underline{\nu} > 0$ *on the first principal modulus*).

44.6. Complex bounds for real Feigenbaum maps.

44.6.1. *Complex bounds for infinitely renormalizable maps.* Let us now introduce a most desired geometric quality of infinitely renormalizable maps; if satisfied it gives a great control of the geometry of its Julia set, as well as the geometry of the Mandelbrot set near the corresponding parameter value.

Let f be an infinitely renormalizable ql germ, and let $R^n f$, $n = 0, 1, 2, \dots,$ be all its renormalizations Let us say that f has (*complex*) *a priori bounds* if there is an $\underline{\mu} > 0$ and a choice of ql representatives $R^n f : U^{[n]} \to V^{[n]}$ such that

$$\operatorname{mod}(V^{[n]} \smallsetminus U^{[n]}) \geq \underline{\mu} \quad \text{for all } n \in \mathbb{N}.$$

We will now provide a supply of maps that possess *complex a priori* bounds. Note that a *real* polynomial f_c means that c is real; the map itself can still be viewed on the complex plane. (A similar convention applies to more general "real" maps.)

THEOREM 44.11. *Complex a priori bounds are valid for any real Feigenbaum polynomial.*

First, let us do some geometric preparation.

44.6.2. *Towers of Epstein class.* Let $g \equiv g_0 : U \to V$ be a ql map (perhaps, degenerate) of class $\mathfrak{Q}'$. Assume it is anti-renormalizable, so it is the prerenormalization $g = Rg_{-1}$ of a map $g_{-1} : U^{[-1]} \to V^{[-1]}$ of class $\mathfrak{Q}'$. (So, if the renormalization period is p_{-1} then $g = g_{-1}^{p_{-1}} \,|\, U^{[-1]}$.) In turn, if g_{-1} is anti-renormalizable then we have an anti-renormalization $g_{-2} : U^{[-2]} \to V^{[-2]}$ defined on a bigger domain; so $g_{-1} = Rg_{-2} = g_{-2}^{p_{-2}} \,|\, U^{[-1]}$. Proceeding this way as long as possible, we obtain a nest of anti-renormalizations

$$g_n : U^{[n]} \to V^{[n]}, \quad g_n = Rg_{n-1} = g_{n-1}^{p_{n-1}} \,|\, U^{[n]}, \quad n \in -\mathbb{N},$$

called a *McMullen (ql) tower* (it can be finite or infinite.) Topology on the space of towers is defined naturally (note that an infinite tower can be obrained as a limit of a sequence of finite towers of growing size).

We say that a McMullen tower

$$(g_n : U^{[n]} \to V^{[n]}), \quad n \in -\mathbb{N}$$

has a *bounded combinatorics* if the renormalization periods p_n are bounded by some $\bar{p}$. It has *a priori* bounds if $\mathrm{mod}(V^n \smallsetminus U^n) \geq \underline{\mu} > 0$ on all levels $n \in -\mathbb{N}$. We let $\mathfrak{T}_{\mathfrak{Q}}(\bar{p}, \underline{\mu})$ be the space of towers (finite or infinite) combinatorics bounded by $\bar{p}$ and geometry bounded by $\underline{\mu}$.

Towers of Epstein class, as well as their combinatorial and geometric bounds, are defined naturally (where a gemetric bound $\underline{\nu}$ means that $g_n \in \mathfrak{E}(\underline{\mu})$). We let $\mathfrak{T}_{\mathfrak{E}}(\bar{p}, \underline{\nu})$ be space of towers of Epstein class (finite or infinite) with the corresponding bounds.

LEMMA 44.12. *For any $\bar{p}$ and $\underline{\nu} > 0$, the space $\mathfrak{T}_{\mathfrak{E}}(\bar{p}, \underline{\nu})$ of towers (g_n) of Epstein class with the corresponding combinatorial and geometric bounds is compact.*

PROOF. It follows from compactness of $\mathfrak{E}(\underline{\nu})$ by means of the diagonal procedure. $\square$

44.6.3. *Domain of analyticity for a tower of Epstein class.* The map g^p is naturally defined on the disk $\Omega^{-1} := g^{-p}(V') \supset U \equiv \Omega^0$, providing us with an analytic extension of f to a bigger domain.

If the map f is n times anti-renormalizable (not necessarily with the same combinatorics), we can repeat this procedure, obtaining an increasing nest of domain

$$\Omega^0 \subset \Omega^{-1} \subset \cdots \subset \Omega^{-n}$$

to which f is consecutively analytically extended. If f is infinitely anti-renormalizable, then f ultimately extends to the domain $\Omega^{-\infty} := \bigcup \Omega^{-n}$.

PROPOSITION 44.13. *Let $(g_{-n})_{n=0}^{\infty}$ be a McMullen tower with a bounded combinatorics and a priori bounds, $f \equiv g_0$. Then:*

(i) *The domain $\Omega^{-\infty}$ contains the real line.*

(ii) *$f : \Omega^{-\infty} \to V$ is a branched covering of infinite degree onto its image.*

(iii) *The slits $[\pm b, \pm\infty)$ of $\mathbb{C}_L$ lift to proper curves $\Gamma_\pm \subset \partial\Omega^{-\infty}$ intersecting the real line only at points $\pm a\partial H$.*

44.6.4. *Proof of Complex Beau Bounds.* We are ready to prove that any tower $(g_n)_{n=0}^{-\infty}$ of Epstein class $\mathfrak{T}_{\mathfrak{E}}(\bar{p}, \underline{\nu})$ restricts to a McMullen ql tower of class $\mathfrak{T}_{\mathfrak{Q}}(\bar{p}, \underline{\mu})$ with $\underline{\mu}$ depending only on $\underline{\nu}$. The idea is that in a small wedge around $\mathbb{R}$, real *a priori* bounds control the pullbacks by the tower maps, while outside this wedge the pullbacks are decomposed into big scales Lipschitz maps (Lemma 7.3) followed by the square root map. This yields a strong contraction in big scales.

LEMMA 44.14. *Let $g : (U, \mathcal{I}) \to (\mathbb{C}(L), \mathcal{I})$ be a map of Epstein class $\mathfrak{E}(\underline{\nu})$, where $g : \mathcal{I} \to \mathcal{I}$ is a proper unimodal map. Let us select an intemediate interval $\mathbb{I}_\rho$ so that*

$$(1+\varepsilon)\,|\mathcal{I}| < |\mathbb{I}_\rho| < 2\,|\mathcal{I}| \quad \text{and} \quad \mathrm{mod}_{\mathbb{R}}(L : \mathbb{I}_\rho) \geq \varepsilon > 0,$$

where ε depends only on $\underline{\nu}$. Then there exist $\lambda \in (0,1)$ and $\theta \in (0, \pi/2)$ depending only on ε such that

$$g^{-1}(\mathbb{D}_\rho) \subset \mathbb{D}_{\rho/\lambda} \cup (\mathbb{D}_{\lambda\rho} \smallsetminus \mathbb{W}_\theta). \tag{44.11}$$

PROOF. Let us represent g as $\psi \circ f_0$, where $f_0(z) = z^2$ and ψ is the inverse to a univalent embedding $\phi : \mathbb{C}_L \to \mathbb{C}_L$. Let

$$\Delta := g^{-1}(\mathbb{D}_\rho), \quad W := \phi(\mathbb{D}_\rho) = f_0(\Delta).$$

By the Symmetric Schwarz Lemma, W is contained in the round disk $\mathbb{D}(J)$ based upon the interval $J = \phi(\mathbb{I}_\rho)$.

Furthermore, by the Koebe Distortion Theorem, ϕ has a bounded distortion on $\mathbb{I}_\rho$. Since the critical value $v = g(0)$ is contained in $\mathcal{I} \Subset \mathbb{I}_\rho$, it divides $\mathbb{I}_\rho$ into comparable parts (with the constant depending on ε only). Hence $\phi(v) = f_0(0) = 0$ divides J into comparable parts as well.

It follows that the domain $f_0^{-1}(\mathbb{D}(J))$ is an oval with bounded geometry based upon the interval $(g\,|\,\mathbb{R})^{-1}(\mathbb{I}_\rho) \Subset \mathbb{I}_\rho$. Moreover, since the boundary of $\partial\mathcal{I}$ is repulsive and the Epstein class $\mathfrak{E}(\underline{\nu})$ is compact, there is a definite space in $\mathbb{I}_\rho$ around $(g\,|\,\mathbb{R})^{-1}(\mathbb{I}_\rho)$. Altogether, these imply that $f_0^{-1}(\mathbb{D}(J))$ is contained in a domain specified in the right-hand side of (44.11).

As $g^{-1}(\mathbb{D}_\rho) = f_0^{-1}(W) \subset f_0^{-1}(\mathbb{D}(J))$, the conclusion follows. $\square$

LEMMA 44.15. *For any $\underline{\nu} > 0$ there exist $\rho > 0$ and $\underline{\mu} > 0$ with the following property. For any infinite tower $(g_n)_{n=0}^{-\infty}$ of Epstein class $\mathfrak{T}_{\mathfrak{E}}(\bar{p}, \underline{\nu})$, the base map $g \equiv g_0(U, \mathcal{I}) \to (\mathbb{C}(L), \mathcal{I})$ admits a ql restriction with range $\mathbb{D}_\rho(L)$ and $\mathrm{mod}\, g \geq \underline{\mu}$.*

PROOF. Let us normalize the tower so that $\mathcal{I} = \mathbb{I}$. Select a high level $n < 0$ (to be specified below), and consider the corresponding tower map

$$g_n : (U^{[n]}, H^{[n]}) \to (\mathbb{C}(L^{[n]}), L^{[n]}),$$

where $H^{[n]}$ is the real slice of $U^{[n]}$ and $g_n : H^{[n]} \to L^{[n]}$ is a proper unimodal map. Moreover, according to the real *a priori* bounds, the nest of intervals

$$\mathcal{I}^{[n]} \Subset H^{[n]} \Subset L^{[n]} \tag{44.12}$$

is a configuration with a definite space in between two cosecutive intervals. Hence we can select $|\mathcal{I}^{[n]}| < \rho < |L^{[n]}|$ so that the nest $\mathcal{I}^{[n]} \Subset \mathbb{I}_\rho \Subset L^{[n]}$ has the same property and the interval $\mathbb{I}_\rho$ is comparable with $\mathcal{I}^{[n]}$ (with all constants depending only on $\underline{\nu}$).

Let $\Omega^{[n]} := g_n^{-1}(\mathbb{D}_\rho)$, so the map $g_n : \Omega^{[n]} \to \mathbb{D}_\rho$ is a double branched covering . Let $\mathbb{W}_\theta$ stand for the union of two $\mathbb{R}$-symmetric wedges of size 2θ:

$$\mathbb{W}_\theta = \{z : \ |\arg z| < \theta \text{ or } |\arg z - \pi| < \theta\}.$$

Let $\rho := |L^{[n]}|$. By Lemma 44.14, there exist $\lambda \in (0,1)$ and $\theta \in (0, \pi/2)$ depending only on the geometry of the nest (44.12) such that

$$\Omega^{[n]} \subset \mathbb{D}_{\rho/\lambda} \cup (\mathbb{D}_{\lambda\rho} \smallsetminus \mathbb{W}_\theta). \tag{44.13}$$

To simplify notation, let $V := \mathbb{D}_\rho$ and let $W := g^{-1}(V)$. The map $g : W \to V$ is a double branched covering. We claim that $W \Subset V$ (provided n is sufficiently big), which makes this map quadratic-like. First, let us show that for n sufficiently big, we have:

$$W \subset \mathbb{D}_{\rho/\lambda}. \tag{44.14}$$

So, taking any point $z \in V$ and letting $g^{-1}(z) = \{\pm w\}$, we need to check that $|w| < \rho/\lambda$. Otherwise, let us coinsider the orbit $(w_k \equiv g_n^k(w))_{k=0}^p$, where $g_n^p = g$, and take the last moment $k \in [0, p-1]$ within it for which $w_k \notin \mathbb{D}_{\rho/\lambda}$. Then $w_{k+1} \in \mathbb{D}_\rho$ (even $w_{k=1} \in \mathbb{D}_{\rho/\lambda}$ unless $k = p-1$), and (44.13) implies that $w_k \in \mathbb{D}_{\lambda\rho} \smallsetminus \mathbb{W}_\theta$.

Let us now represent g_n^k as $\psi \circ f_0$, where $f_0(z) = z^2$, ψ is univalent, and $\psi^{-1} : \mathbb{C}(L^{[n]}) \to \mathbb{C}(L^{[n]})$ is a map of class $\mathfrak{U}$. By the Lipschitz control of ψ^{-1} (Lemma 7.3), we have:

$$|\psi^{-1}(w_k)| \le C|w_k| \le C\rho,$$

with some constant $C > 0$ depending only on θ. Applying the square root map f_0^{-1}, we concude that for ρ sufficiently big, $|w| < \sqrt{C\rho} < \rho/\lambda$, as was desired.

We see that W is compactly contained in $\mathbb{D}_\rho$. Let us show that ∂W does not touch the slits L^c either. Otherwise ∂W would contain an $\mathbb{R}-$symmetric open arc Γ in $\mathbb{C} \smallsetminus L^c$ whose both ends land at some point $a \in L^c$. Since Γ is in the domain of g, it is in the domain in any g_n, $n \le 0$ of the tower. But for n sufficiently big, the point a is also in the domain of g_n. Hence the closed arc $\hat{\Gamma} := \Gamma \cup \{a\}$ is in the domain of g_n. Since $\operatorname{Dom} g_n$ is simply connected, the whole Jordan disk D boundsd by $\hat{\Gamma}$ is inteh domain of g_n. In this way we a holomorphic function on a domain D whose boundary is mapped to $\mathbb{R}$, which is impossible for an open map. □

45. Geometry of Julia sets

In this section, we will exlore further the problem of local connectivity of Julia sets:

JLC Problem. *Give a combinatorial criterion for local connectivity of the Julia set $\mathcal{J}(f_c)$ of a quadratic polynomial.*

Above, we have already seen several "tame" (local connected) creatures (hyperbolic, parabolic, and postcritically non-recurrent maps), as well some "wild" (non-local connected) creatures (Cremer maps). In this section we will enrich the tame zoo with Yoccoz and real Feigenbaum maps. On the other hand, the wild zoo will be enlarged with some infinitely renormalizable maps of unbounded satellite type.

We will also show that the Julia sets of Yoccoz maps all have zero area.

45.1. Statements. Let us say that a quadratic polynomial f with connected Julia set belongs to the *Yoccoz class* $\mathfrak{Y}$ if it is periodically repelling and at most finitely renormalizable.

THEOREM 45.1. *For any quadratic polynomial f of Yoccoz class, the Julia set $\mathcal{J}(f)$ is locally connected.*

We let $\mathfrak{Y}[0]$ be the class of non-renormalizable quadratic polynomials f_c with both fixed points repelling. Note that if f_c has a non-repelling cycle of period $p > 1$ then it is renormalizable since c belongs to a renormalization window of period p (see §37.11) or is the root of a primitive hyperbolic component. Hence $\mathfrak{Y}[0] \subset \mathfrak{Y}$.

We let $\mathfrak{Y}[n]$ be the class of quadratic polynomials which are exactly n times renormalizable with both fixed points of the last renormalization $R^n f$ repelling (in other words, $R^n f$ straightens to a polynomial of class $\mathfrak{Y}[0]$). For the same reason as above, $\mathfrak{Y}[n] \subset \mathfrak{Y}$, and by definitions,

$$\mathfrak{Y} = \bigcup_{n=0}^{\infty} \mathfrak{Y}[n].$$

Here is the main particular case of the above theorem:

COROLLARY 45.2. *For $f \in \mathfrak{Y}[0]$, the Julia set $\mathcal{J}(f)$ is locally connected.*

We will complement the above result with the measure-theoretic one:

THEOREM 45.3. *For any quadratic polynomial f of Yoccoz class,*

$$\operatorname{area} \mathcal{J}(f) = 0.$$

In what follows, the proof is written for $f \in \mathfrak{Y}[0]$. The argument for $f \in \mathfrak{Y}[n]$ is the same except one should use the puzzle associated with the nth renormalization level (introduced in §31.9).

We will proceed with a general discussion of the JLC probelm for infinitely renormalizable maps, and will complete this section with the following result:

THEOREM 45.4. *The Julia set of a real Feigenbaum map is locally connected.*

45.2. Shrinking of puzzle pieces. Corollary 45.2 follows from the following result (due to Corollary 9.9):

THEOREM 45.5. *For $f \in \mathfrak{Y}[0]$, we have:*

$$\max_i \operatorname{diam} Y_i^{(n)} \to 0 \quad \text{as } n \to \infty,$$

which is the complex counterpart of Theorem 31.25.

EXERCISE 45.6. *Show that it is sufficient to prove that for any $z \in \mathcal{J} \setminus \partial\mathcal{Y}$,*

$$\operatorname{diam} Y^{(n)}(z) \to 0 \quad \text{as } n \to \infty. \tag{45.1}$$

Recall from §9.1.3 that a point $z \in \mathcal{J} \setminus \partial\mathcal{Y}$ satisfying (45.1) is called *rigid*, and note that the set of rigid points is backward invariant. Let us start proving rigidity with points whose orbits avoid some central puzzle piece:

LEMMA 45.7. *For any central puzzle piece $Y^{(l)} \equiv Y_0^{(l)}$, the set*

$$Q^{(l)} := \{z \in \mathcal{J} :\ f^n z \notin \operatorname{int} Y^{(l)},\ n = 0, 1, 2, \dots\}$$

is hyperbolic.

REMARK 45.8. Compare this assertion with Proposition 30.33.

PROOF. Let us consider all non-critical puzzle pieces $Y_i \equiv Y_i^{(l)}$, $i \neq 0$, of depth l. They can be slightly *thickened* to open disks $D_i \Supset Y_i$ so that f is univalent on each D_i and

$$f(Y_i) \supset Y_j \implies f(D_i) \Supset D_j, \quad i, j \neq 0$$

(compare §28.4.3). By the Schwarz Lemma, if $z \in Y_i$ and $f(z) \in Y_j$ then $\|Df(z)\|_{\text{hyp}} \geq \lambda_{ij} > 1$ where the norm is calculated from the hyperbolic metric on D_i to that on D_j.

Let $\lambda := \min \lambda_{ij} > 1$. Let us consider a point $z \in Q$ such that

$$f^k z \in Y_{j(k)}, \quad k = 0, 1, \dots, n, \tag{45.2}$$

for some $j(k) \neq 0$. Then $\|Df^n(z)\|_{\text{hyp}} \geq \lambda^n$, where the norm is calculated from the hyperbolic metric of $D_{j(0)}$ to that of $D_{j(n)}$. Since the hyperbolic metrics on the D_i restricted to Y_i are comparable with the Euclidean ones (and there are only finitely many of the D_i), we obtain that

$$|Df^n(z)| \geq a\lambda^n,$$

with some constant $a > 0$ independent of z, as asserted. □

COROLLARY 45.9. *If $f^m z \in Q^{(l)} \setminus \partial\mathcal{Y}$ for some $l, m \in \mathbb{N}$, then z is rigid.*

Here is another noteworthy consequence:

COROLLARY 45.10. *Let $L \equiv L_{(l)}$ be the landing map to $Y \equiv Y^{(l)}$, and let D_n be the set of points z whose first landing time is equal to n (so $Lz = f^n z$). Then*

$$\operatorname{area} D_n \leq C\rho^n, \quad \text{where } \rho = \rho(l) < 1.$$

Let us now proceed with a proof of Theorem 45.5.

Case 0: $f^n 0 = \alpha$ for some $n \in \mathbb{Z}_+$. It is a Misiurewicz case covered in §27.1.5.

Otherwise the critical puzzle pieces $Y^{(n)}$ (whose interiors contain 0) are well defined for all $n \in \mathbb{N}$.

Case 1: f is combinatorially non-recurrent, i.e., there is a critical puzzle piece $Y^{(l)}$ and $m \in \mathbb{Z}_+$ such that $f^k(0) \notin Y^{(l)}$ for all $k = m, m+1, \dots$. Then $f^m(0)$ belongs to the hyperbolic set $Q^{(l)}$ from Lemma 45.7.[10] By Corollary 45.9, 0 is rigid, so the puzzle pieces $Y^{(n)}$ shrink to 0.

It follows that there is a puzzle piece $Y^{(l')} \subset Y^{(l)}$ that does not contain points $f^k(0)$, $k = 1, \dots, m-1$, and hence does not contain any points of $\operatorname{orb} f(0)$. By adjusting notation, we can let $Y^{(l)}$ be such a piece.

Furthermore, taking further preimages of 0, we conclude that any precritical point $z \in \mathcal{J}$ is rigid.

Take now an arbitrary non-precritical point $z \in \mathcal{J}$. If $\operatorname{orb} z$ does not enter some critical puzzle piece $Y^{(l)}$, rigidity follows again from Corollary 45.9.

[10]Note that in the non-recurrent case, Lemma 45.7 follows from Corollary 27.23.

Otherwise, $\operatorname{orb} z$ lands infinitely many times in the interior of any puzzle piece $Y^{(n)}$. Let us consider the landing moments $m_k \to \infty$ of $\operatorname{orb} z$ in the interior of some puzzle piece $Y^{(n)} \Subset Y^{(l)}$. The pieces $Z_k := Y^{n+m_k}(z)$ are the pullbacks of $Y^{(n)}$ under these landing maps. Since the orbit of $f(0)$ does not enter $\operatorname{int} Y^{(l)}$, the inverse branches $f^{-m_k} : Y^{(n)} \to Z_k$ extend univalently to $Y^{(l)}$. By the Shrinking Lemma, $\operatorname{diam} Z_k \to 0$ as $k \to \infty$, as desired.

Case 2 (main): f is combinatorially recurrent. In this case, the Principal nest $V^0 \supset V^1 \supset \dots$ is infinite. Let $V^\infty := \bigcap V^n$.

Since f is non-renormalizable, there are infinitely many non-central levels n_k in the nest. By Theorem 44.9 (i), $\operatorname{mod}(V^{n_k+1} \smallsetminus V^{n_k+2}) \geq \underline{\mu} > 0$. By the Gröztsch Inequality, $\operatorname{mod}(V^0 \smallsetminus V^\infty) = \infty$, so V^∞ is a single point. Hence $\operatorname{diam} V^n \to 0$, so the critical puzzle pieces $Y^{(n)}$ shrink. Thus, 0 is rigid, and so are all its iterated preimages.

Let us now prove rigidity at any other point $z \in \mathcal{J} \smallsetminus \partial\mathcal{Y}$. If $\operatorname{orb} z$ does not accumulate on 0 then z belongs to one of the hyperbolic sets $Q^{(l)}$, and the conclusion follows from Corollary 45.9. Otherwise, $\operatorname{orb} z$ lands in all puzzle pieces V^n. Let us consider the corresponding branches of the first landing maps,

$$L_n = f^{m_n} : W^n \to V^n, \quad W^n \ni z,\ m_n \to \infty.$$

By Theorem 44.9 (ii), they have a bounded distortion, which allows us to apply the refined Shrinking Lemma (Exercise 21.35). It implies that the puzzle pieces W^n shrink to z, completing the proof of Theorem 45.5. □

Given an $\varepsilon > 0$, let

$$\mathcal{J}_\varepsilon \equiv \mathcal{J}_\varepsilon(f) := \{z \in \mathcal{J} :\ f^n z \notin \mathbb{D}_\varepsilon,\ n = 0, 1, \dots\}.$$

It is a compact invariant subset of $\mathcal{J}(f)$.

COROLLARY 45.11. (i) *For any map f of Yoccoz class, the sets $\mathcal{J}_\varepsilon$ are hyperbolic.*

(ii) $\operatorname{area} \mathcal{J}_\varepsilon(f) = 0$.

(iii) *Let $D_n := \{z :\ f^k z \notin \mathbb{D}_\varepsilon,\ k = 0, \dots, n-1\}$. Then*

$$\operatorname{area} D_n \leq C\rho^n, \quad \text{where } \rho = \rho(\varepsilon) < 1.$$

PROOF. (i) Since the puzzle pieces $Y^{(n)}$ shrink to 0, any set $\mathcal{J}_\varepsilon$ is contained in some set $Q^{(n)}$, which is hyperbolic by Lemma 45.7.

(ii) The assertion follows from Exercise 25.24.

□

We let

$$\mathcal{J}_0 \equiv \mathcal{J}_0(f) := \bigcup_{\varepsilon > 0} \mathcal{J}_\varepsilon. \tag{45.3}$$

By the above Corollary, $\operatorname{area} \mathcal{J}_0(f) = 0$. Note that

$$\mathcal{J} \smallsetminus \mathcal{J}_0 = \{z \in \mathcal{J} :\ \operatorname{cl}(\operatorname{orb} z) \ni 0\}.$$

45.3. Persistent recurrence and minimality. Let P be a critical puzzle piece. In §31.2.2 we introduced a notion of the first kid of P. More generally, a *kid* of P is a critical puzzle piece Q which is a degree two pullback of P. In other words, there is a double branched covering $f^m : Q \to P$. All kids of P form a nest $Q_1 \supset Q_2 \supset \dots$, where Q_n is called the *nth kid of* P.

Let f be a map of Yoccoz class, and let $\mathcal{Y}$ be the puzzle for f corresponding to the deepest renormalization level.

Since the puzzle pieces shrink, the *combinatorial recurrence* of the critical point is equivalent to the *topological recurrence*, so we do not need to distinguish these properties anymore.

A recurrent critical point is called *reluctantly recurrent* if there is a critical puzzle piece $Y^{(l)}$ with infinitely many kids. Otherwise f is called *persistently recurrent*.

Exercise 45.12. *Show that f is reluctantly recurrent if and only if there is an $\varepsilon > 0$ and a sequence of moment $n_k \to \infty$ such that $f^{n_k}(0) \in \mathbb{D}_\varepsilon$ and the pullback of $\mathbb{D}_\varepsilon$ to 0 is univalent.*

Recall from §23.3 the notion of a *regular* backward orbit

$$\hat{z} = (z \equiv z_0, z_{-1}, z_{-2}, \dots) \in \widehat{\mathbb{C}}. \tag{45.4}$$

A compact invariant set K is called *completely irregular* if all backward orbits $\hat{z} \in \hat{K}$ (i.e., ones staying in K) are irregular. Equivalently, *no point $z \in K$ has a neighborhood U that can be univalently pulled back along some backward orbit $\hat{z} \in \hat{K}$.*

Proposition 45.13. *Let f be a recurrent map of Yoccoz class. A map is persistently recurrent if and only if the dynamics on the postcritical set $\omega(0)$ is completely irregular.*

Proof. Assume f is reluctantly recurrent. Let $Y^{(l)}$ be a critical puzzle piece with infinitely many kids $Y^{(l+n_k)}$. Then the pullbacks of $Y^{(l)}$ along the orbits $(0_m)_{m=1}^{n_k}$ are univalent. Taking a limit of the corresponding backward orbits, we obtain an infinite backward orbit $\hat{z} \in \hat{\omega}(0)$ along which the pullback of $Y^{(l)}$ is univalent, so it is regular.

Vice versa, assume there is a regular backward orbit $\hat{z} \in \hat{\omega}(0)$ as (45.4). Then there is a neighborhood $U \ni z$ that can be univalently pulled back along $\hat{z}$; let $U_{-k} \ni z_{-k}$, $k = 0, 1, \dots$, be this pullback. Since all puzzle pieces shrink, there exists a piece $Y^{(l)}(z) \subset U$, so we can assume that $U = Y^{(l)}(z)$ in the first place.

Let us consider some U_{-k}, $k \in \mathbb{N}$. Since $z_{-k} \in \omega(0)$, there is a postcritical point $0_j \equiv f^j(0) \in U_{-k}$. For the smallest $j = j_k$ like this, the pullback of U_{-k} along $(0_n)_{n=0}^{j}$ is a double branched covering (Corollary 31.6 (ii)).

Let us now consider a postcritical point $0_{j+k} \in U$. Since 0 is recurrent, $\operatorname{orb} 0_{j+k}$ lands in the critical puzzle piece $Y^{(l)}$; let $m \in \mathbb{N}$ be the first landing moment. By Corollary 31.6 (i)), the map $f^m : Y^{(l+m)}(0_{j+k}) \to Y^{(l)}$ is univalent. Moreover,

$$Y^{(l+m)}(0_{j+k}) \subset Y^{(l)}(0_{j+k}) = Y^{(l)}(z) \equiv U.$$

Hence the pullback of $Y^{(l+m)}(0_{j+k})$ along the orbit $(0_n)_{n=0}^{j+k}$ is a double branched covering, implying that $Y^{(l+mj+k)}$ is a kid of $Y^{(l)}$.

Since k is arbitrary big, $Y^{(l)}$ has infinitely many kids, so f is reluctantly recurrent. $\square$

COROLLARY 45.14. *If f is persistently recurrent then the dynamics on $\omega(0)$ is minimal.*

PROOF. Otherwise, there is a critical puzzle piece $Y \equiv Y^{(l)}$ such that the first landing times $n(z)$ for points $z \in \omega(0)$ to $\operatorname{int} Y$ are unbounded (see Exercise 19.5). The pullbacks of Y along the corresponding backward orbits are univalent. Taking a limit of these backward orbits, we obtain an infinite univalent pullback of Y along $\omega(0)$. $\square$

Let us finish with a fun exercise:

EXERCISE 45.15. *If $\mathcal{Y}$ corresponds to the deepest renormalization level then any critical puzzle piece has at least two kids.*

45.4. Area of Julia sets. In this section, we will prove Theorem 45.3.

Let $\mathcal{Y}$ be the puzzle associated with the deepest renormalization of f (see §31.9). Let us consider two cases.

Case (i): *f is reluctantly recurrent* , so there is a critical puzzle piece P with infinitely many kids Q_k. Let $f^{m_k} : Q_k \to P$ be the corresponding branched double coverings.

Since P is a quasidisk, for any point $\zeta \in \mathcal{J} \cap \operatorname{int} P$ there a round disk $D(\zeta) \equiv \mathbb{D}_r((\zeta) \subset P$ centered at ζ such that

$$r \equiv r(\zeta) \asymp \operatorname{dist}(D(\zeta), \partial P) \asymp \operatorname{dist}(\zeta, \mathcal{J} \cap \partial P). \tag{45.5}$$

Since $\mathcal{J} \cap \partial P$ consists of finitely many pre-α-fixed points, there is an $l = l(\zeta)$ such that f^l maps $D(\zeta)$ with bounded distortion onto an oval $O(\zeta)$ of definite size. It follows that each $D(\zeta)$ contains a gap in the Julia set of definite size.

Let us now take any point $z \in \mathcal{J} \smallsetminus \mathcal{J}_0$ (see (45.3), so $\operatorname{cl}(\operatorname{orb} z) \ni 0$. Then there exist n_k such that $f^{n_k} z \in \operatorname{int} Q_k$. Let $\zeta_k := f^{n_k+m_k} z \in \operatorname{int} P$, and let S_k be the pullback of P to z by $f^{n_k+m_k}$. As it has degree two, the corresponding pullback of $D(\zeta_k)$ has degree at most two. Since $D(\zeta_k)$ contains a definite gap in the Julia set, so does S_k. Since $\operatorname{diam} S_k \to 0$, the Julia set is porous at z. As this happens for all $z \in \mathcal{J} \smallsetminus \mathcal{J}_0$, we obtain $\operatorname{area}(\mathcal{J} \smallsetminus \mathcal{J}_0) = 0$ by the Lebesgue Theorem. On the other hand, $\operatorname{area} \mathcal{J}_0 = 0$ by Corollary 45.11, and the conclusion follows.

Case (ii): *Persistent recurrence.*

Let us consider a non-central puzzle piece V^{n-1} of the principal nest and the corresponding generalized renormalization

$$g_n : \bigcup V_i^n \to V^{n-1}.$$

Recall that R_0^n stands for the maximal annulus in $V^{n-1} \smallsetminus \bigcup V_i^n$ surrounding V^n but not the lateral pieces. By Corollary (44.8), the moduli $\operatorname{mod} R_0^n$ are bounded away from 0.

Hence R_0^n contains a definite gap in $\mathcal{J}$. Spreading it around (for all V^n as above), we conclude again that the set $\mathcal{J} \smallsetminus \mathcal{J}_0$ is porous.

45.5. Infinitely renormalizable maps. Part of the above analysis carries through in the infinitely renormalizable case as well:

EXERCISE 45.16. *Let f be a periodically repelling quadratic polynomial. Let us consider the puzzle $\mathcal{Y}^{[m]}$, $m \in \mathbb{N}$, of any existing renormalization level m. Then* *for any admissible l, the set $Q \equiv Q^{(l)}$ from* Lemma 45.7 *is hyperbolic, and for any $z \in Q$,*

$$\operatorname{diam} Y^{(n)}(z) \to 0 \text{ as } n \to \infty.$$

Thus, all points of Q are rigid. In particular, all periodic points of f are rigid.

For an infinitely renormalizable quadratic polynomial, recall from §28.4.8 the postcritical impression $\mathcal{O} \equiv \mathcal{O}_f$, accompanied with the little Julia sets $K_i^{[n]} = J_i^{[n]}$ (28.8), and in particular, the canonical Julia nest $K^{[0]} \supset K^{[1]} = K^{[0]} \supset \dots$ (see also §31.10). For $z \in \mathcal{O}$, we let

$$K^{[0]}(z) = J^{[n]}(z) \supset K^{[1]}(z) = J^{[1]}(z) \supset \dots$$

be the nest of little Julia sets containing z.

More generally, take any point z in the realm of attraction $\mathcal{R}(\mathcal{O}$, i.e., such that $\omega(z) \subset \mathcal{O}$.

Exercise 45.17. *Show that:*

(i) *For $z \in \mathcal{R}(\mathcal{O})$, we have: $\omega(z) = \mathcal{O}$.*

(ii) *$z \in \mathcal{R}(\mathcal{O})$ iff the orbit of z eventually lands in every little Julia set $\mathcal{K}^{[n]}$.*

Let l_n be the first landing moment in the nth little Julia set, i.e., the first time when $f^{l_n} z \in \mathcal{K}^{[n]}$. Taking the corresponding pullback of $K^{[n]}$ under f^{l_n}, we obtain a Julia set $K^{[n]}(z)$ around z.

Lemma 45.18. *For an infinitely renormalizable quadratic polynomial f, the Julia set $\mathcal{K}(f) = \mathcal{J}(f)$ is (perfectly) rigid at $z \in \mathcal{R}(\mathcal{O})$ iff the little Julia sets $K^{[n]}(z)$ shrink:* $\operatorname{diam} K^{[n]}(z) \to 0$ *as $n \to \infty$. In particular, the Julia set is (perfectly) rigid at the critical point iff the canonical Julia nest shrinks.*

Proof. As we know (see §28.4.6 and Theorem 31.17), little Julia sets $K^{[n]}(z)$ are intersections of nests of perfect puzzle pieces. Hence, if the nest

$$K^{[0]}(z) \supset K^{[1]}(z) \supset \dots$$

shrinks to z, then there is a nest of perfect puzzle pieces shrinking to z, implying that $\mathcal{K}$ is perfectly rigid at z.

Vice versa, assume $\mathcal{K}$ is rigid at some $z \in \operatorname{Orb} \mathcal{O}$. Then the puzzle pieces around z (in the general sense of §9.1.1) shrink. Since $\mathcal{K}_{\text{lc}} = \mathcal{K}_{\text{puz}}$ (Theorem 32.10)), the Yoccoz puzzle pieces $Y^{(n)}(z)$ around z shrink as well. As

$$\bigcap_n Y^{(n)}(z) = \bigcap_n K^{[n]}(z),$$

the conclusion follows. □

Proposition 45.19. *An infinitely renormalizable quadratic polynomial f has a locally connected Julia set iff for any $z \in \mathcal{R}(\mathcal{O})$, the little Julia sets $K^{[n]}(z)$ shrink:*

$$\operatorname{diam} K^{[n]}(z) \to 0 \quad \text{as } n \to \infty. \tag{45.6}$$

Proof. The only property to be checked (in view Lemma 45.18) is that that under shrinking assumption (45.6), $\mathcal{J}$ is locally connected at any point $z \in \mathcal{J} \smallsetminus \mathcal{R}(\mathcal{O})$. For such a point, there exists a sequence of moments $l_k \to \infty$ such that $\operatorname{dist}(f^{l_k} z, \mathcal{O}) \geq \delta > 0$. Take an accumulation point $\zeta \in \mathcal{J} \smallsetminus \mathcal{O}$ for $(f^{l_k} z)$.

Since under the shrinking assumption, the puzzle pieces around points of $\mathcal{O}$ shrink, ζ belongs to the interior of some puzzle piece $P \Subset \mathbb{C}$. Pulling P back by the f^{l_k}, we obtain a nest of puzzle pieces P^k containing z in their interiors. By the Shrinking Lemma, they shrink to z, and the conclusion follows. $\square$

COROLLARY 45.20. *For an infinitely renormalizable quadratic polynomial f with locally connected Julia set, the postcritical impression $\mathcal{O}$ is a Cantor set coinciding with $\omega(0)$ and the dynamics on this set is conjugate to the adding machine $\tau_{\mathbf{q}}$ corresponding to the sequence $\mathbf{q} = (q_n)$ of relative renormalization periods.*

It was shown in Proposition 27.25 that Cremer Julia sets are wild, i.e., not locally connected. We can now construct quite different examples of wild Julia sets:

EXAMPLE 45.21. *There exists a rate $\varepsilon_n \to 0$ such that for any sequence of non-vanishing rotation numbers $\mathfrak{p}_n/\mathfrak{q}_n \in (-\varepsilon_n, \varepsilon_n)$, $n \in \mathbb{Z}_+$, the following holds: If f is an infinitely renormalizable quadratic polynomial of satellite type $(\mathfrak{p}_n/\mathfrak{q}_n)_{n\in\mathbb{Z}_+}$ then the Julia set $\mathcal{J}(f)$ is not locally connected at 0.*

PROOF. Let β_n, $n \in \mathbb{N}$, be the β-fixed point of the little Julia set $J^{[n]}$ of the n-fold pre-renormalization f_n (which is equal to the α-fixed point of f_{n+1}). Given any rate $\delta_n \to 0$, we can select a sufficiently fast rate (ε_n) in such a way that $|\beta_{n+1} - \beta_n| < \delta_n$. Hence $\underline{\beta} := \inf |\beta_n| > 0$. But the little Julia set $J^{[n]}$ is $0-$symmetric and hence contains all three points $\beta_n, 0 - \beta_n$, so $\operatorname{diam} J^{[n]} \geq \underline{\beta} > 0$. Application of Lemma 45.18 concludes the proof. $\square$

45.6. JLC for real Feigenbaum maps. Let us now prove Theorem 45.4.

Let f be a real Feigenbaum map, let $(J^{[n]} = K^{[n]})$ be the canonical next of its little Julia sets, and let $(\mathcal{I}^{[n]} = K^{[n]} \cap \mathbb{R}$ be the corresponding nest of their real slices.

A priori bounds and Lemma 40.13 imply that the renormalizations $R^n f$ are uniformly $K-$qc conjugate to quadragtic polynomials f_{c_n}, $c_n \in [-2, 1/4]$. For the latter, the Julia sets $\mathcal{J}(f_{c_n})$ and their real slices $\mathcal{I}(f_{c_n})$ have size of order 1. Therefore, the sizes of the Julia sets $J^{[n]}$ are comparable with the sizes of the intervals $\mathcal{I}^{[n]}$: $\operatorname{diam} J^{[n]} \asymp \operatorname{diam} \mathcal{I}^{[n]}$.

By Corollary 30.46 (i), the intervals $\mathcal{I}^{[n]}$ shrink. Hence the little Julia sets $J^{[n]}$ shrink as well, implying the local connectivity of $\mathcal{J}(f)$ at the critical point (Lemma 45.18).

The same argument shows that $\mathcal{J}(f)$ is locally connected at every point $z \in \mathcal{O}$. To complete a proof, we need to show (in view of Proposition 45.19) that the little Julia sets $K^{[n]}(z)$ shrink for any point $z \in \mathcal{R}(\mathcal{O})$.

Let us show that there is a definite gap (in the relative scale) between the little Julia sets and the rest of the postcritical set, i.e., there exists a $\delta > 0$ such that

$$\operatorname{dist}(\mathcal{O} \smallsetminus J^{[n]}) \geq \delta \operatorname{diam} J^{[n]}.$$

Indeed, let us consider pre-renormalizations $f_n : U^{[n]} \to V^{[n]}$ with a bounded geometry described in §40.1, and let $U_k^{[n]} := f_n^{-k}(U^{[n]})$. Then

$$\frac{\operatorname{diam} U_k^{[n]}}{\operatorname{diam} K^{[n]}} \to 1 \text{ as } k \to \infty \quad \text{(uniformly in } n).$$

On the other hand, since $\mathcal{O}$ has a bounded geometry (Theorem 30.49), the intervals $\mathcal{I}^{[n]}$ stay a definite relative distance from the rest of the postcritical set, i.e., there exists an $\varepsilon > 0$ such that

$$\operatorname{dist}(\mathcal{I}^{[n]}, \mathcal{O} \smallsetminus \mathcal{I}^{[n]}) \geq \varepsilon \, |\mathcal{I}^{[n]}|.$$

It follows that for k sufficiently big (independently of n), the real traces of the domains $U_k^{[n]}$ stay definite relative distance away from $\mathcal{O} \smallsetminus I^{[n]}$. But for any such a k, the domains $U_k^{[n]}$ have a bounded geometry (depending on k but uniform in n). Hence the whole domains stay a definite relative distance away from $\mathcal{O} \smallsetminus \mathcal{I}^{[n]}$.

By the Koebe Distortion Theorem, the pullbacks of the $U_k^{[n]}$ to z by f^{l_n} are disks $W_k^{[n]} \ni z$ with a bounded geometry as well. Together with the topological exactness of f, this implies that they shrink as $n \to \infty$ (compare with the Shrinking Lemma). All the more, the little Julia sets $K^{[n]}(z)$ shrink.

Notes. The Puzzle techniques and its applications to the JLC and MLC Problems (Theorems 45.1 and 47.17) were developed around 1990 by Yoccoz. This theory is closely related to an earlier work by Branner and Hubbard [**BH**] concerning the dynamics of cubic polynomials with one escaping critical points. Yoccoz never published his results: their first accounts appeared in the expositions by Hubbard [**H2**] and Milnor [**M4**]. The above approach based on the Principal Nest and Generalized Renormalization and leading to the bounds of Theorems 44.9 was developed by the author in [**L10**].

Theorem 45.3 on the area of $\mathcal{J}$ is due to Lyubich [**L8**] and Shishikura [**Sh1**].

JLC for real Feigenbaum maps (Theorem 45.4) was proven by J. Hu and Y. Jiang [**HJ, J**]. See also McMullen [**McM2**, §8.1].

46. Measurable Dynamics of real maps

In this section we will study Measurable Dynamics of real unimodal maps: ergodicity, description of attractors, and the problem of existence of an absolutely continuous invariant measure (*acim*).

Throughout this chapter, we assume that $f : \mathcal{I} \to \mathcal{I}$ belongs to the class $\mathfrak{Q}_{\mathbb{R}}$ of real ql maps. Recall from Theorem 30.52 that such a map has a unique *topological attractor* $\mathcal{A}^{\mathrm{t}} \equiv \mathcal{A}^{\mathrm{t}}_f$, which is either an attracting or parabolic cycle, or the cycle of an exact periodic interval, or a Feigenbaum attractor.

As we know from Theorem 30.10, in the hyperbolic or parabolic cases, the dynamics is *regular*, i.e., almost all orbits converge to the corresponding hyperbolic or parabolic cycle. Below we will explore what happens in the irregular cases.

46.1. Following the critical point. Let us start with the real version of Theorem 22.2 (which admits a similar proof):

LEMMA 46.1. *Let $f : \mathcal{I} \to \mathcal{I}$ be a real ql map. Assume it has a closed nowhere dense invariant set $K \subset \mathcal{I}$ of positive length. Then $\omega(x) \subset \omega(0)$ for almost all $x \in K$.*

Let us now introduce a real version of the sets $\mathcal{J}_\varepsilon$:

$$\mathcal{I}_\varepsilon \equiv \mathcal{I}_\varepsilon(f) := \{z \in \mathcal{I} : |f^n z| \geq \varepsilon, \ n = 0, 1, \dots\}.$$

Lemma 45.11 (together with Exercise 30.11) implies:

LEMMA 46.2. *For any real ql map $f : \mathcal{I} \to \mathcal{I}$ of Yoccoz class and for any $\varepsilon > 0$, we have:*

(i) *The set* $\mathcal{I}_\varepsilon$ *is hyperbolic;*
(ii) $\operatorname{length} \mathcal{I}_\varepsilon = 0$;
(iv) *Let* $D_n^{\mathbb{R}} := \{x \in \mathcal{I} : \ f^k x \notin \mathbb{D}_\varepsilon, \ k = 0, \dots, n-1\}$. *Then*

$$\operatorname{length} D_n^{\mathbb{R}} \le C\rho^n, \quad \text{where } \rho = \rho(\varepsilon) < 1.$$

Consequently, $\omega(x) \ni 0$ *for a.e.* $x \in \mathcal{I}$.

46.2. *A priori* bounds. *Scaling factors* for the Real Principal Nest (see §31.11) are defined as

$$\tau_n \equiv \tau_n(f) := \frac{|I^n|}{|I^{n-1}|} < 1.$$

By restricting *a priori* bounds of Theorem 44.9 to the real line we obtain:

THEOREM 46.3. *Let* g *be a real symmetric generalized quadratic-like map with* $\operatorname{mod} g \ge \underline{\nu} > 0$. *Then:*
(i) *The scaling factors* τ_n *are bounded by some* $\bar{\tau} < 1$ *depending only on* $\underline{\nu}$ *and* $d(n)$. *In particular,* $\tau_n \le \bar{\tau}(\underline{\nu}) < 1$ *if the level* $n-2$ *is non-central (i.e., when* $d(n) = 0$*).*
(ii) *The distortion of all branches of the first landing maps* (31.32)

$$L_n : \bigsqcup_i J_i^n \to I^n$$

are bounded on all levels by some $K(\underline{\nu})$. *In particular, the distortion of the diffeomorphisms* h_n *from* (31.34) *are bounded by* $K(\underline{\nu})$.

46.3. Density Lemma.

LEMMA 46.4. *Let* $f : \mathcal{I} \to \mathcal{I}$ *be a non-regular map. Then there exists a nest of* 0*-symmetric intervals* $I^0 \supset I^1 \supset \dots$ *such that for any invariant measurable set* $X \subset \mathcal{I}$ *of positive length, we have:* $\operatorname{dens}(X|\, I^n) \to 1$ *as* $n \to \infty$.

PROOF. By Proposition 30.48 and Lemma 46.2, $\omega(x) \ni 0$ for a.e. $x \in \mathcal{I}$, so we can assume without loss of generality that $\omega(x) \ni 0$ for a.e. $x \in X$.

Let us consider two cases.

Case a) *f is at most finitely renormalizable.* Then f has an exact periodic interval $\mathcal{T} \ni 0$ of some period p (see Theorem 30.37). Let $I^0 \supset I^1 \supset \dots$ be the Real Principal Nest for the return map $f^p|\,\mathcal{T}$. By Theorem 31.25, it shrinks to 0.

Let $x \in X$ be a Lebesgue density point for X. Since $\omega(x) \ni 0$, x belongs to the domain of each first landing map $L_n : \bigsqcup_i J_i^n \to I^n$. Let J^n be the interval of this domain containing x.

Using exactness once again, we conclude that $\operatorname{diam} J^n \to 0$. Hence

$$\operatorname{dens}(X|\, J^n) \to 1.$$

By Theorem 46.3 (ii), the landing branches $L_n : J^n \to I^n$ have a bounded distortion. It follows that $\operatorname{dens}(X|\, I^n) \to 1$.

Case b) *f is infinitely renormalizable.* In this case, use nest (30.8)

$$\mathcal{I}^0 \supset \mathcal{I}^1 \supset \dots \ni 0$$

of periodic intervals instead of the Principal Nest. By Corollary 30.46 , it shrinks to 0. By Corollary 30.42, the first landing maps L_n to the $\mathcal{I}^n$ have bounded distortion. This is all needed to apply the argument of Case a) to this situation. $\square$

Lemma 46.5. *Let $f : \mathcal{I} \to \mathcal{I}$ be a map of real Yoccoz class . Then* 0 *is the Lebesgue density point for any invariant measurable set $X \subset \mathcal{I}$ of positive length.*

Proof. Lemma 46.4 implies that $\mathrm{dens}(X \mid T^n) \to$ as as $|T^n| \to 0$ along any shrinking sequence of intervals $T^n \ni 0$ commensurable with the I^n (i.e. such that $C^{-1}|I^n| \le |T^n| \le C|I^n|$ with some *a priori* selected constant $C > 1$). So, we only need to fill-in intermediate scales in between two (incommensurable) intevals $I^n \subset I^{n-1}$.

Let us go back to the proof of Lemma 46.4, Case a). By Lemma 31.14, the interval $J^n \ni x$ can be enlarged to an interval $\tilde{J}^n$ mapped by L_n diffeomorphically onto I^{n-1}. Let us consider any intermediate interval $T \supset I^n$, contained in $(1/2)\, I^{n-1}$, and let $S \subset \tilde{J}^n$ be its pullback under $L_n : \tilde{J}^n \to I^{n-1}$ By the Koebe Distortion Theorem, $L_n : S \to T$ has a bounded distortion. Hence X has high density in T, implying the desired. □

The *internal side* of the crititical value $v = f(0)$ is the component of $\mathbb{R} \smallsetminus \{v\}$ containing $f(\mathcal{I}) \smallsetminus \{v\}$. We let $\mathrm{dens}^i(X \mid v)$ be the density of a set X on the internal side of v.

Corollary 46.6. *For any invariant measurable set $X \subset \mathcal{I}$ of positive length, the critical value v is a one-sided density point for X, on the internal side:*

$$\mathrm{dens}^i(X \mid v) = 1.$$

46.4. Ergodicity.

Theorem 46.7. *Any non-regular map $f : \mathcal{I} \to \mathcal{I}$ is ergodic with respect to the Lebesgue measure.*

Proof. If f is not ergodic then $\mathcal{I}$ can be decomposed into two disjoint invariant measurable subsets X_1 and X_2 of positive measure. Hence

$$\mathrm{dens}(X_1 \mid I) + \mathrm{dens}(X_2 \mid I) = 1$$

for any interval $I \subset \mathcal{I}$, On the other hand, by the Density Lemma, there exists a nest of intervals I^n such that both densities, $\mathrm{dens}(X_1 \mid I^n)$ and $\mathrm{dens}(X_2 \mid I^n)$, go to 1 as $n \to \infty$. Contradiction. □

46.5. Measure-theoretic attractor.

46.5.1. *Classification.*

Theorem 46.8. *Any unimodal map $f : \mathcal{I} \to \mathcal{I}$ of class $\mathfrak{Q}_{\mathbb{R}}$ has a unique measure-theoretic attractor $\mathcal{A}^{\mathfrak{m}} \equiv \mathcal{A}^{\mathfrak{m}}_f$ attracting Lebesgue almost all orbits:*

$$\omega(x) = \mathcal{A}^{\mathfrak{m}} \text{ for a.e. } x \in \mathcal{I}.$$

Moreover, $\mathcal{A}^{\mathfrak{m}}$ coincides with the topological attractor $\mathcal{A}^{\mathfrak{t}}$, except for one theoretical possibility: when $\mathcal{A}^{\mathfrak{t}}$ is a cycle of an exact interval then $\mathcal{A}^{\mathfrak{m}}$ may be a Cantor subset of $\mathcal{A}^{\mathfrak{t}}$.

In the latter case, $\mathcal{A}^{\mathfrak{m}}$ is called a *wild* attractor.

Proof. Let us consider one by one the cases of Theorem 30.52 describing topological attractors.

If f is *hyperbolic* or *parabolic* then $\omega(x) = \mathcal{A}^{\mathfrak{t}}$ for a.e. $x \in \mathcal{I}$ by Theorem 30.10.

If f is *infinitely renormalizable* then the assertion follows from Theorem 30.48.

Assume finally that f is of *real Yoccoz class.* Then $\mathcal{A}^{\mathfrak{t}}$ is a transitive cycle of intervals. By Corollary 30.34, almost any orbit eventually lands in $\mathcal{A}^{\mathfrak{t}}$, so we only need to study the measurable dynamics on $\mathcal{A}^{\mathfrak{t}}$.

By Lemma 46.2, $\omega(x) \ni 0$ for a.e. $x \in \mathcal{A}^{\mathfrak{t}}$. Hence $\omega(x) \supset \omega(0)$ for a.e. $x \in \mathcal{A}^{\mathfrak{t}}$.

If $\omega(x) = \mathcal{A}^{\mathfrak{t}}$ for a.e. $x \in \mathcal{A}^{\mathfrak{t}}$ then $\mathcal{A}^{\mathfrak{t}}$ is the unique measure-theoretic attractor. So, assume $\omega(x) \neq \mathcal{A}^{\mathfrak{t}}$ for a subset of $\mathcal{A}^{\mathfrak{t}}$ of positive measure. Since

$$T := \{x \in \mathcal{A}^{\mathfrak{t}} : \ \omega(x) = \mathcal{A}^{\mathfrak{t}}\}$$

is a completely invariant (under $f \,|\, \mathcal{A}^{\mathfrak{t}}$) measurable subset $\mathcal{A}^{\mathfrak{t}}$, it has zero measure by ergodicity (Theorem 46.7).

Let us consider a countable base of open intervals J_k in $\mathcal{A}^{\mathfrak{t}}$, and let

$$X_k := \{x \in \mathcal{A}^{\mathfrak{t}} : \ f^n(x) \notin J_k, \ n = 0, 1, \dots\}, \quad X_\infty := \bigcup X_k.$$

Note X_∞ is the complement of T, so it has the full measure.

Obviously, each X_k is compact and invariant. Since $f|\,\mathcal{A}^{\mathfrak{t}}$ is topologically transitive, X_k is nowhere dense. By Lemma 46.1, $\omega(x) \subset \omega(0)$ for a.e. $x \in X_k$. Consequently, $\omega(x) = \omega(0)$ for a.e. $x \in X_k$, and hence $\omega(x) = \omega(0)$ for a.e. $x \in X_\infty$. The conclusion follows. $\square$

46.5.2. *Persistent recurrence of wild attractors.* Let us start with some preparation.

Lemma 46.9. *If f has a wild attractor $\mathcal{A}^{\mathfrak{m}}$ then there exists an invariant measurable set X of positive measure such that* $\operatorname{dens}(X \,|\, I) < 1$ *for any interval I contained in the transitive cycle of intervals $\mathcal{A}^{\mathfrak{t}}$.*

Proof. By definition, the realm of attraction $\mathcal{R} \equiv \mathcal{R}(\mathcal{A}^{\mathfrak{m}})$ has full measure and $\omega(x) = \mathcal{A}^{\mathfrak{m}}$ for a.e. $x \in \mathcal{R}$.

Since $\mathcal{A}^{\mathfrak{m}}$ is nowhere dense, there exists a closed interval $M \subset \mathcal{A}^{\mathfrak{t}} \smallsetminus \mathcal{A}^{\mathfrak{m}}$. Then for any $x \in \mathcal{R}(\mathcal{A}^{\mathfrak{m}})$, there exists an $n(x)$ such that $f^n \notin M$ for $n \geq n(x)$. Since $n(x)$ assume only countable set of values, there is a value $m \in \mathbb{N}$ such that

$$Z_m := \{x \in \mathcal{R} : \ n(x) = m\}$$

has positive measure. Hence $l(f^m(Z_m)) > 0$, and moreover, this set is contained in

$$X := \{x \in \mathcal{R} : \ f^n x \notin M, \ n = 0, 1, 2, \dots\} \subset \mathcal{I} \smallsetminus M.$$

We conclude that X has positive measure, too (and of course, it is invariant).

Since $\mathcal{A}^{\mathfrak{t}}$ is the cycle of a topologically exact interval, the orbit of any interval $I \subset \mathcal{A}^{\mathfrak{t}}$ covers the whole $\mathcal{A}^{\mathfrak{t}}$. Hence, if $\operatorname{dens}(X \,|\, I) = 1$ then X would have full measure in $\mathcal{A}^{\mathfrak{t}}$, which would contradict $X \cap M = \emptyset$. $\square$

Lemma 46.10. *The α-fixed point does not belong to the wild attractor $\mathcal{A}^{\mathfrak{m}}$.*

Proof. Assume $\alpha \in \mathcal{A}^{\mathfrak{m}}$, so $\omega(0) \ni 0$. Since $\mathcal{A}^{\mathfrak{m}}$ is nowhere dense, there is an open interval $\tilde{T} \subset \mathcal{I}$ around α containing two gaps G_i $(i = 1, 2)$ in $\mathcal{A}^{\mathfrak{m}}$ such that $\partial\mathcal{I} \subset \partial G_1 \cup \partial G_2$. Let a_i be the midpoint of G_i and let $T := (a_1, a_2)_\#$.

Let X be an invaraint set from Lemma 46.9. Then $\operatorname{dens}(X \,|\, G_i) < 1$ for both $i = 1, 2$.

Let $\tilde{T}^n = f^{-n}(\tilde{T})$ where the f^{-n} mean the branches fixing α. Let $T^n \subset \tilde{T}^n$ and G_i^n be corresponding pullbacks of T and G_i. Then there exists an $\varepsilon > 0$ such that

$$|G_i^n| \geq \varepsilon |T^n| \text{ and } \operatorname{dens}(X \,|\, G_i) < 1 - \varepsilon, \quad i = 1, 2, \ n \in \mathbb{N}.$$

Since $\omega(0) \ni \alpha$, there is a sequence of moments $l_n \to \infty$ such that $f^{l_n}(0) \in \tilde{T}^n$, and hence $f^{l_n}(0) \in T^n$. Let us select the smallest l_n like that (the first landing moments to the corresponding intervals). Then each $\tilde{T}^n$ can be univalently pulled back along the orbit $\{f^m(v)\}_{m=0}^{l_n-1}$. Let $\tilde{J}^n \ni v$ be these pullbacks, and let J^n, H_i^n be the corresponding pullbacks of the T^n and G_i^n.

By the Koebe Distortion Theorem, $|H_i^n| \asymp |J^n|$ and $\mathrm{dens}(X \mid H_i^n) \leq \rho < 1$. Hence $\underline{\mathrm{dens}}(X \mid v)$ is bounded away from 1 on both sides of v, contradicting the Density Lemma (Corollary 46.6). □

THEOREM 46.11. *If f has a wild attractor $\mathcal{A}^{\mathfrak{m}}$ then f is persistently recurrent.*

PROOF. If f is reluctantly recurrent then there is a real puzzle piece $\tilde{P} \ni 0$ with infinitely many kids Q_k. Let $f^{n_k} : Q_k \to \tilde{P}$ be the corresponding proper unimodal maps. As the boundary points of $\tilde{P}$ are α-prefixed, $\partial\tilde{P} \cap \mathcal{A}^{\mathfrak{m}} = \emptyset$ (by Lemma 46.10). So there are two disjoint intervals G_i $(i = 1, 2)$ in $\tilde{P} \smallsetminus \mathcal{A}^{\mathfrak{m}}$ such that $\partial\tilde{P} \subset \partial G_1 \cup \partial G_2$. Let a_i be the midpoint of G_i and let $P := (a_1, a_2)_{\#}$.

The rest of the proof is similar to that of Lemma 46.10. Pulling $\tilde{P}$ back along the orbit $\{f^m v\}_{m=0}^{n_k-1}$, we obtain intervals $\tilde{J}^k \ni v$ mapped by f^{n_k-1} univalently onto $\tilde{P}$. Let J^k and H_i^k, be the corresponding pullbacks of P and G_i, $i = 1, 2$. By the Koebe Distortion Theorem, $f^{n_k-1} : J^k \to P$ are diffeomorphisms with bouned distortion.

Let X be an invaraint set from Lemma 46.9. As $\mathrm{dens}(X \mid G_i) < 1$ for $i = 1, 2$, it follows that the $\mathrm{dens}^{\pm}(X \mid J_k)$ are bounded away from 1 on both sides of v. Hence $\underline{\mathrm{dens}}(X \mid v) < 1$ on both sides of v, contradicitng Corollary 46.6. □

COROLLARY 46.12. *The dynamics on the wild attractor $\mathcal{A}^{\mathfrak{m}}$ is minimal. In particular, $\mathcal{A}^{\mathfrak{m}}$ does not contain periodic points.*

46.6. Stochastic maps. A real ql map $f : \mathcal{I} \to \mathcal{I}$ that has an absolutely continuous invariant measure (acim) μ is called *stochastic*.

46.6.1. *An acim is physical.*

PROPOSITION 46.13. *A stochastic map f has a unique acim μ. Moreover, this is a physical measure that governs behavior of Lebesgue almost all orbits:*

$$\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k x} \to \mu \tag{46.1}$$

for Leb a.e. $x \in \mathcal{I}$.

PROOF. An acim is ergodic and unique by Theorem 46.7. By the Ergodic Theorem, there is an invariant set X_0 of positive length such that (46.1) holds for all $x \in X_0$. Then this holds for all $x \in X := \mathrm{Orb}_-(X)$. Since X is completely invariant, it has full Leb measure (once again, by Theorem 46.7), and we are done. □

46.6.2. *Entropy.* As in the Kneading theory (§32.4) let us consider the tiling $\mathcal{P}$ of the interval $\mathcal{I}$ into its halves $\bar{\mathcal{I}}_- = \{x \in \mathcal{I} : \ x \leq 0\}$ and $\bar{\mathcal{I}}_+ = \{x \in \mathcal{I} : \ x \geq 0\}$. If f is not superattracting then $\mu\{0\} = 0$ for any invariant measure μ, so $\mathcal{P}$ can be considered as a partition for (f, μ).

LEMMA 46.14. *Assume $f : \mathcal{I} \to \mathcal{I}$ is neither hyperbolic nor parabolic. Then the partition $\mathcal{P}$ is a generator for any invariant measure μ.*

Proof. Indeed, the partitions $\mathcal{P}^n$ shrink by Corollary 30.51 from the No Wandering Intervals Theorem, $\square$

Remark 46.15. In fact, one does not need to use the No Wandering Intervals Theorem, since by the Poincaré recurrence, wandering intervals do not carry any invariant measure.

Corollary 46.16. *Under the circumstances of* Lemma 46.14,

$$h_\mu(f) = h_\mu(f, \mathcal{P}).$$

Theorem 46.17. *An acim μ of a stochastic map $f : \mathcal{I} \to \mathcal{I}$ has positive entropy:*

$$h_\mu(f) > 0.$$

Proof. Assume $h_\mu(f) = 0$. Then by Theorem 46.42, f is one-to-one μ-a.e. Since μ is absolutely continuous, there is an invariant measurable set $X \subset \mathcal{I}$ of positive length such that $f : X \to X$ is a bijection (and in particular, X is disjoint from the symmetric set $X' \equiv -X$).

On the other hand, by the Density Lemma, $\operatorname{dens}(X| I^n) \to 1$ for some shrinking nest of 0-symmetric intervals, forcing $X \cap I^n$ to overlap with $X' \cap I^n$. $\square$

46.6.3. *Lyapunov exponent.* Let μ be a (probability) invariant measure for a map $f : \mathcal{I} \to \mathcal{I}$. The *Lyapunov* (or *characteristic*) *exponent* of μ is defined as

$$\chi_\mu \equiv \chi_\mu(f) := \int \log |Df| \, d\mu$$

(which could be equal to $-\infty$). By the Ergodic Theorem, if μ is ergodic then

$$\chi_\mu = \lim_{n\to\infty} \frac{1}{n} \log |Df^n(x)| \quad \text{for } \mu - \text{a.e. } x.$$

Pesin Formula. For an acim μ of a stochastic map $f : \mathcal{I} \to \mathcal{I}$, we have

$$h_\mu(f) = \chi_\mu(f).$$

Proof. By Corollary 46.39 (combined with Lemma 46.14),

$$h_\mu(f) = \int \log \operatorname{Jac}_\mu f \, d\mu.$$

Let ρ, be the density of μ with respect to the Lebesgue measure m. Then

$$\operatorname{Jac}_\mu f = \frac{\rho \circ f}{\rho} |Df|.$$

Hence

$$\log \operatorname{Jac}_\mu f = \log |Df| + \log \rho \circ f - \log \rho. \tag{46.2}$$

If we knew that $\log \rho$ is integrable, we would immediately conclude that

$$\int \log \operatorname{Jac}_\mu f \, d\mu = \int \log |Df| \, d\mu \tag{46.3}$$

(since the measure μ is f-invariant), implying the desired by the Jacobian Formula (Corollary 46.39). In general, (46.3) can be derived from the Ergodic Theorem as follows. Taking Birkhoff averages of (46.2), we obtain:

$$\frac{1}{n} \log \operatorname{Jac}_\mu f^n = \frac{1}{n} \log |Df^n| + \frac{1}{n} \log \rho \circ f^n - \frac{1}{n} \log \rho.$$

The last two terms go to 0 in measure (since μ is f-invariant), so

$$\frac{1}{n}\log \mathrm{Jac}_\mu f^n - \frac{1}{n}\log|Df^n| \to 0 \quad \text{in measure.}$$

On the other hand, by the Ergodic Theorem, this expression converges a.e. to the difference between the integrals in (46.3). It follows that this difference must vanish. $\square$

Corollary 46.18. *An acim μ of a stochastic map $f : \mathcal{I} \to \mathcal{I}$ has positive Lyapunov exponent: $\chi_\mu(f) > 0$.*

This shows that μ-almost all orbits are *exponentially unstable*, which is a basic feature of chaotic dynamics.

46.6.4. *Pesin unstable manifolds.* Let $\hat{f} : \hat{\mathcal{I}} \to \hat{\mathcal{I}}$ be the natural extension of f. Any f-invariant measure μ lifts to a $\hat{f}$-invariant measure $\hat{\mu}$. Let us consider a backward orbit

$$\hat{x} = (x_0, x_{-1}, x_{-2}, \dots) \in \hat{\mathcal{I}}$$

Assume we have a topological disk $W_0 \ni x_0$ that can be univalently pulled back along $\hat{x}$, and let

$$\hat{W} := (W_0, W_{-1}, W_{-2}, \dots)$$

be the corresponding pullback. It is called a *local leaf* of $\hat{\mathcal{I}}$ at $\hat{x}$.

Theorem 46.19. *Let μ be an invariant measure with positive characteristic exponent $\chi \equiv \chi_\mu > 0$. Then for $\hat{\mu}$-a.e. $\hat{x} \in \hat{\mathcal{I}}$, there exists a local leaf $\hat{W}^u(\hat{x})$ such that for any $\varepsilon > 0$ and any $\hat{y} \equiv (y_0, y_{-1}, y_{-2}, \dots) \in \hat{W}^u(\hat{x})$ we have*

$$C^{-1}\exp(-n(\chi+\varepsilon)) \le |y_{-n} - x_{-n}| \le C\exp(-n(\chi-\varepsilon)),$$
$$n = 0, 1, 2, \dots, \quad \text{with } C = C(\varepsilon, \hat{y}).$$

Under these circumstances, the leaves $\hat{W}^u(\hat{x})$ are called *Pesin local unstable manifolds.*

46.6.5. *Subhyperbolic maps.* As we know, real postcritically preperiodic maps are stochastic (Theorem 27.15). Let us now prove a more general result:

Theorem 46.20. *A real subhyperbolic map $f : \mathcal{I} \to \mathcal{I}$ is stochastic.*

Proof. By definition, a subhyperbolic map is not regular. By Corollary 30.46, it is not infinitely renormalizable either. Hence, by the Structural Theorem for real ql maps (30.52), f has a periodic exact interval $\mathcal{T}'$ of some period p.

Let α' be the α-fixed for $f^p : \mathcal{T}' \to \mathcal{T}'$. By Proposition 30.22, its iterated preimages accumulate on 0. Let a_{-n} be the closest to 0 preimage of order $\le n$. Then the nest of nice intervals $J^n := [-\alpha_{-n}, \alpha_{-n}]_\# \ni 0$ shrinks to 0, so we can select a nice interval $J \equiv J^n$ which is disjoint from the post-valuable set $\overline{\mathcal{P}}$.

Let T be the first return map to J. Let us show that it is a good expanding Bernoulli map, as defined in Appendix 1 (§46.8.2).

As we know from §31.9, this map is Bernoulli: its domain $\mathrm{Dom}\, T$ consists of intervals Q_i with disjoint interiors mapped diffeomorphically onto J . (Note that the central interval is absent since the critical point does not return to J.)

By Lemma 46.2, $\mathrm{Dom}\, T$ has full measure in J, securing property (G1) of good Bernoulli maps.

As the post-valuable set $\overline{\mathcal{P}}$ is disjoint from J, the iterated inverse branches $T_i^{-n} : J \to Q_i^n$ extend to a bigger interval $\hat{J} \Supset J$. By the Koebe Distortion Theorem, they have a uniformly bounded distortion (securing property FG2) of a good Bernoulli map). Since f^p is exact on $\mathcal{T}'$,

$$\sup_i \operatorname{diam} Q_i^n \to 0 \quad \text{as } n \to \infty.$$

By the bounded distortion property, $\operatorname{diam} Q_i^n \asymp |DT_i^{-n}(x)|$ for any $x \in J$, so

$$\sup_i \max_{x\in J} |DT_i^{-n}(x)| \to 0 \quad \text{as } n \to \infty.$$

Thus, T is expanding.

By Theorem 46.25, T has an acim μ_J. By Lemma 46.24 (from Appendix 46.8), we can spread μ_J around to obtain an acim μ for f itself (finite or infinite). To justify its finiteness, we need to show that the return time $n(x)$ to J is integrable with respect to μ_J. It follows from the exponential decay of the tails:

$$\exists\, \sigma \in (0,1) \text{ such that } m\{x \in J:\ n(x) \geq n\} \leq C\sigma^n, \quad \forall\, n \in \mathbb{N},$$

which is a consequence of Lemma 46.2 □

46.7. Pathological examples.

46.7.1. *Non-renormalizable but non-stochastic map.*

EXAMPLE 46.21. *There exists a non-stochastic topologically exact map $f \in \mathfrak{Q}_{\mathbb{R}}$.*

46.7.2. *A map without physical measures.*

EXAMPLE 46.22. *There exists a topologically exact map $f \in \mathfrak{Q}_{\mathbb{R}}$ that does not have any physical measure.*

46.7.3. *A weird physical measure.*

EXAMPLE 46.23. *There exists a topologically exact map $f \in \mathfrak{Q}_{\mathbb{R}}$ whose physical measure is supported at the (repelling) α-fixed point.*

46.8. Appendix 1: Preliminaries on acim.

46.8.1. *Spreading around.*

LEMMA 46.24. *Let m be a quasi-invariant measure for a map $f : Y \to Y$, and let $X \subset Y$ be a measurable subset of positive measure . Let $T : X \to X$ be the first return map to X, $Tx = f^{n(x)}x$. Assume that T has an acim $d\nu = \rho\, dm$. Then f has an acim $d\mu = q\, dm$ (maybe, infinite) whose restriction to X is equal to $d\nu$. Moreover, the total mass of μ is equal to $\displaystyle\int_X n(x)\, d\nu$.*

PROOF. For $y \in Y$, let

$$\Phi_y := \{x \in \operatorname{Dom} T:\ \exists k = k(x,y) < n(x) \in \mathbb{N} \text{ such that } f^k x = y\},$$

and let

$$q(y)\, dm(y) := \sum_{x\in\Phi_y} f_*^{k(x)}(\rho(x)\, dm(x)), \quad \text{i.e., } q(y) = \sum_{x\in\Phi_y} \frac{\rho(x)}{\operatorname{Jac}_m f^{k(x)}(x)}$$

(with understanding that $q(y) = 0$ if $\Phi_y = \emptyset$). □

46.8.2. *Stochasticity for expanding Bernoulli maps.* Let us adapt the notion of Bernoulli map from §19.11.3 and §31.7.1 to the one-dimensional setting. Let J be an interval and let $Q_i \subset J$ be a family of disjoint subintervals (which can be countable). Let us consider a C^2-map $g : (\mathrm{Dom}_g \equiv \bigcup Q_i) \to J$. Such a map is called *(unbranched) Bernoulli* if each branch $g : Q_i \to J$ is a diffeomorphism. It is called expanding if there exists an $n \in \mathbb{Z}_+$ and $\lambda > 1$ such that

$$|Dg^n(x)| \geq \lambda \quad \text{for all } x \in \mathrm{Dom}\, f^n.$$

An expanding Bernoulli map $f : \bigcup Q_i \to J$ is called *good* if

(G1) $m(J \smallsetminus \bigcup Q_i) = 0$ (where m is the Leb measure on J);

(G2) The branches $g : Q_i \to J$ have a uniformly bounded distortion.

THEOREM 46.25. *A good expanding Bernoulli map is stochastic.*

PROOF. Let us adapt the proof of Theorem 19.76 to this more general setting. Expanding property and uniformly bounded distortion assumption (G2) imply a uniform distortion bound for all branches of the iterates g^n (as in Lemma 19.68). Hence the densities ρ_n of the push-forward measures $m_n := (g^n)_*(m)$ have uniformly bounded distortion as well.

By Property (G1), the total mass of the push-forward measures remains the same: $m_n(J) = m(g^{-n}(J)) = m(J)$, so

$$\int \rho_n \, dm = m(J).$$

It follows that the densities ρ_n are uniformly bounded from above and below. Now Proposition 19.18 implies stochasticity. □

46.9. Appendix 2: Rokhlin Theory of measurable partitions.

46.9.1. *Measurable partitions.* Let (X, μ) be a Lebesgue space (all measure spaces in what follows are assumed to be Lebesgue). A simple example of a measurable partition of X is a finite or countable partition into measurable subsets. A *trivial* particular case is the partition ν whose only piece is the whole space X.

A more interesting example is the partition of the square $\mathbb{I} \times \mathbb{I}$ by the vertical fibers $\{x\} \times \mathbb{I}$. Though every piece of this partition has zero measure, the partition as a whole is a non-trivial object.

In general, a *measurable partition* $\mathcal{P}$ of X can be defined as a partition into fibers of some measurable map $\pi : X \to \bar{X}$. A particular case is the *singleton partition* ε into single points. A typical example of a non-measurable partition is provided by orbits of an ergodic transformation of an infinite space.

The space $\bar{X}$ is naturally identifies with the quotient $X/\mathcal{P}$ mod the equivalence relation corresponding to the partition $\mathcal{P}$. We let $\bar{\mu} := \pi_*(\mu)$ be the corresponding quotient measure.

Two measurable partitions, $\mathcal{P}$ of X and $\tilde{\mathcal{P}}$ of $\tilde{X}$, are viewed to be *equivalent* (or *equal* mod 0) if there exists a measure-preserving bijection $h : Y \to \tilde{Y}$ between full-subsets that carries $\mathcal{P} \,|\, Y$ to $\tilde{\mathcal{P}} \,|\, \tilde{Y}$. In what follows, all partitions are considered mod 0.

A subset $Z \subset X$ is called $\mathcal{P}$-*saturated* if it is the union of some pieces of $\mathcal{P}$. (In other words, each piece $P \in \mathcal{P}$ is either contained in Z or disjoint from Z.) The family of all measurable $\mathcal{P}$-saturated $(\bmod 0)$[11] subsets forms a σ-algebra $\mathfrak{S}(\mathcal{P})$ called the *envelope* of $\mathcal{P}$.

A family of saturated sets $X_i \in \mathfrak{S}(\mathcal{P})$ is called $\mathcal{P}$-*separating* if for any two pieces $P_1, P_2 \in \mathcal{P}$ there is a set X_i containing P_1 but disjoint from P_2.

Proposition 46.26. *A partition $\mathcal{P}$ is measurable iff its envelope $\mathfrak{S}(\mathcal{P})$ contains a countable $\mathcal{P}$-separating family of sets X_i. Moreover, the correspondence $\mathcal{P} \mapsto \mathfrak{S}(\mathcal{P})$ is as a bijection between measurable partitions and σ-subalgebras of the original σ-algebra $\mathfrak{S} = \mathfrak{S}(\varepsilon)$.*

Idea of the proof. To reconstruct a partition from a σ-subalgebra $\mathfrak{S}' \subset \mathfrak{S}$, take at most countable family of sets $X_i \in \mathfrak{S}$, $i = 0, 1, \dots$, generating $\mathfrak{S}$, let $X_i^0 := X_i$, $X_i^1 := X \smallsetminus X_i$, and define pieces of $\mathcal{P}$ as all possible intersections $X_0^{j_0} \cap X_1^{j_1} \cap \dots$ with , $j_i \in \{0, 1\}$. QED

If $\mathfrak{S}(\mathcal{P})$ consists of nul-sets and full-sets only, then $\mathcal{P}$ is called *ergodic*.

Let us say that a measurable partition is *standard* if it is isomorphic to the vertical partition of the product $Y \times F$ of two Lebesgue spaces (where the measure is not assumed to be the product measure). The measure-theoretic category is so flexible that any partition can be reduced to a trivial one:

Theorem 46.27. *Let $\mathcal{P}$ be a measurable partition of a Lebesgue space X. Then X can be decomposed into at most countably many disjoint measurable saturated subsets, $X = \bigsqcup X_i$, such that the restriction of the partition $\mathcal{P}$ to each X_i is standard.*

46.9.2. *Conditional measures.* Conditional measures are well defined on subsets of positive measure, but not so on nul-sets. However, if we have a *measurable partition* into nul-sets, conditional measures can still be defined for *almost all* pieces.

The model case is the vertical foliation of the square $\mathbb{I}^2$ with the 2D Leb measure dm (area). The classical Fubini Theorem allows us to disintegrate any test function on the box $\mathbb{I}^2$ along the vertical fibers $\mathbb{I}_t := \{t\} \times \mathbb{I}$:

$$\int_{\mathbb{I}^2} \phi \, dm = \int_{\mathbb{I}} \left(\int_{\mathbb{I}_t} \phi(t, y) dy \right) \, dt, \quad \text{where } \mathbb{I}_t := t \times \mathbb{I},$$

and in particular, for any measurable set $X \subset \mathbb{I}^2$, we have

$$\text{area}(Z) = \int_{\bar{Z}} \text{length}(Z \cap \mathbb{I}_t) \, dt,$$

where $\bar{Z}$ is the horizontal projection of Z. In this sense, $\text{length}(Z \cap \mathbb{I}_t)$ can be interpreted as the *conditional probability* of the event Z assuming the "event" $t \in \bar{Z} \in \mathbb{I}$ happened.

In a general setting, let us consider a measurable partition $\mathcal{P}$ of a Lebesgue space (X, μ). For a point $x \in X$ and the corresponding point $\bar{x} \in \bar{X} \equiv X/\mathcal{P}$, we let $P_x \equiv P_{\bar{x}} \ni x$ be the corresponding piece of $\mathcal{P}$.[12] Assume we have a family of probability measures $\mu_{\bar{x}}$ supported on the pieces $P_{\bar{x}}$, $\bar{x} \in \bar{X}$, with the following properties:

[11] According to our convention, Z is $\mathcal{P}$-saturated $(\bmod 0)$ if there exists a $\mathcal{P}$-saturated set Z' such that the symmetric difference $Z \oplus Z'$ is a nul-set.

[12] Formally speaking, $\bar{x} \equiv P_{\bar{x}}$, but we prefer to think of $\bar{x}$ as the "parameter" for the corresponding piece.

• Any test function $\phi \in L^1(X, \mu)$ is integrable on a.e. pieces $(P_{\bar{x}}, \mu_{\bar{x}})$;
• The function

$$\bar{x} \mapsto \int_{P_{\bar{x}}} \phi \, d\mu_{\bar{x}} \tag{46.4}$$

is integrable with respect to $\bar{\mu}$; and
•

$$\int_X \phi \, d\mu = \int_{\bar{X}} \left(\int_{P_{\bar{x}}} \phi \, d\mu_{\bar{x}} \right) \, d\bar{\mu}. \tag{46.5}$$

The latter property is called *disintegration* of ϕ over $\mathcal{P}$.

In particular, for any measurable subset $Z \subset X$, we have:

$$\mu(Z) = \int_{\bar{Z}} \mu_{\bar{x}}(Z \cap P_{\bar{x}}) \, d\bar{\mu}, \tag{46.6}$$

where the $\mu_{\bar{x}}(Z \cap P_{\bar{x}})$ are interpreted as *conditional probabilities* of the event Z.

Theorem 46.28. *For any measurable partition $\mathcal{P}$ on a Lebesgue space (X, μ), there exists a unique* $(\operatorname{mod} 0)$ *family of conditional measures $\mu_{\bar{x}}$.*

Note that we can let $\nu_x := \nu_{\bar{x}}$ and interpret (46.4) as a function on X constant on the pieces of $\mathcal{P}$. This function is called the *conditional expectation of ϕ with respect to $\mathcal{P}$* and is denoted $\mathrm{E}(\phi \,|\, \mathcal{P})$. Then by definition of the push-forward measure $d\bar{\mu}$, formulas (46.5) and (46.6) can be written without a reference to the quotient:

$$\int_X \phi \, d\mu = \int_X \mathrm{E}(\phi \,|\, \mathcal{P}) \, d\mu; \qquad \mu(Z) = \int_Z \mu_x(Z \cap P_x) \, d\mu(x).$$

Note that the function $\mathrm{E}(\phi \,|\, \mathcal{P})$ is measurable with respect to the sigma-algebra $\mathfrak{S}(\mathcal{P})$. (We aslo say that it is *measurable with respect to the partition $\mathcal{P}$*.) Moreover, the map $\phi \mapsto \mathrm{E}(\phi \,|\, \mathcal{P})$ is the projection $L^1(X, \mathfrak{S}, \mu) \to L^1(X, \mathfrak{S}(\mathcal{P}), \mu)$ with norm 1.

Exercise 46.29. *Show that the family of conditional measures is invariant under measure-theoretic isomorphisms* mod 0*. (Part of the problem is to formulate the statement rigorously.)*

46.9.3. *Martingales.* Partition are partially ordered: $\mathcal{P} \succ \mathcal{Q}$ if any piece of $\mathcal{P}$ is contained in some piece of $\mathcal{Q}$ (i.e., $\mathcal{P}$ is *finer* than $\mathcal{Q}$). An increasing sequence of measurable partitions,

$$\mathcal{P}^0 \prec \mathcal{P}^1 \prec \mathcal{P}^2 \prec \dots$$

is called a *filtration*. Let $\mathcal{P}^\infty = \bigvee \mathcal{P}^n$: elements of $\mathcal{P}^\infty$ are the intersections of various nests $P^1 \supset P^2 \supset \dots$ with $P^n \in \mathcal{P}^n$.

A sequence of random variables $\phi_n \in L^1$ is *subordinated* to a filtration $(\mathcal{P}^n)$ if ϕ_n is measurable with respect to $\mathcal{P}^n$. Such a seuqence is called a *martingale* if

$$\mathrm{E}(\phi_n \,|\, \mathcal{P}^m) = \phi_m \quad \text{for any } m \leq n. \tag{46.7}$$

For instance, given a function $\phi \in L^1(X)$, the sequence of projections

$$\phi_n := \mathrm{E}(\phi \,|\, \mathcal{P}^n) \tag{46.8}$$

is a martingale.

Doob Theorem. *Let* (ϕ_n) *be a martingale subordinated to a filtation* $(\mathcal{P}^n)$. *Assume it is uniformly integrable, i.e.,*

$$\int |\phi_n|\, d\mu \leq C.$$

Then $\phi_n \to \phi$ *a.e., where* $\phi \in L^(X, \mathcal{P}^\infty, \mu)$. *Moreover, the* ϕ_n *are the projections of* ϕ *to the partitions* $\mathcal{P}^n$, as in (46.8).

46.9.4. *Notes.* Theory of meaurable partitions of the Lebesgue space was developed in the classical paper by Rokhlin [**Ro1**]. In the Western literature, the equivalent, but less intuitive, language of σ-subalgebras is usually used.

For the theory of martingales and the Doob Theorem, see [**Doob**] or e.g., [**Bi, KoS**]

46.10. Appendix 3: Elements of Entropy Theory.

46.10.1. *Definition.* According to Shannon's insight, the amount of information hidden in an event of probability p is equal to $-\log p$. (So, if a low probability event happens, a big amount of information is revealed.) If we have a statistical system that can exist in d states whose probability distribution is $\mathbf{p} := (p_1, \dots, p_d) \in \mathbf{\Delta}^{d-1}$, then the average amount of information hidden in this system is equal to

$$H(\mathbf{p}) := -\sum_{i=1}^{d} p_i \log p_i \equiv \sum_{i=1}^{d} \eta(p_i), \quad \text{where } \eta(p) = -p \log p.$$

This quantity is called the *Shannon entropy* of the system. Note that $H : \mathbf{\Delta}^{d-1} \to \mathbb{R}_+$ is a concave continuous function on the probabilistic simplex.

A homogeneous system contains the maximum amount of information:

Exercise 46.30. *The entropy function* $H : \Delta^{d-1} \to \mathbb{R}_+$ *has a unique point of maximum* $p_\circ = (1/d, \dots, 1/d)$, *and*

$$\max_{p \in \mathbf{\Delta}^{d-1}} H(\mathbf{p}) = H(p_\circ) = \log d.$$

(The minima of $H(\mathbf{p})$ *are attained at the vertices of* $\mathbf{\Delta}^{d-1}$*, where* H *vanishes.)*

Let us now consider a probability space (X, μ). For a finite measurable partition $\mathcal{P} = \bigsqcup_{k=1}^{d} P_k$ of X, the measures $(\mu(P_k))_{k=1}^d \in \mathbf{\Delta}^{d-1}$ form a probability vector. So, we can consider its Shannon entropy:

$$H(\mathcal{P}) \equiv H_\mu(\mathcal{P}) := -\sum \mu(P_k) \log \mu(P_k) \equiv \sum_{P \in \mathcal{P}} \eta\left(\mu(P)\right).$$

Given another finite measurable partition $\mathcal{Q} = \{Q_j\}$ of X, let $\mathcal{P} \vee \mathcal{Q}$ be the partition comprising all pairwise intersections $P_k \cap Q_j$.[13]

Lemma 46.31. *For any two finite partitions, we have:*

$$H(\mathcal{P} \vee \mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}).$$

The equality is attained iff the partitions are independent.

[13] Of course, measure zero intersections can be thrown away, but they can also be kept as the matter of convenience of writing, as their presence would not affect any calculation.

Proof. We have:

$$H(\mathcal{P}\vee\mathcal{Q}) = -\sum_{i,j}\mu(P_i\cap Q_j)\log\mu(P_i\cap Q_j) = -\sum_{i,j}\mu(P_i\cap Q_j)\log(\mu(P_i)\,\mu(Q_j\,|\,P_i))$$
$$= -\sum_{i,j}\mu(P_i\cap Q_j)\log(\mu(P_i)) - \sum_{i,j}\mu(P_i)\,\mu(Q_j\,|\,P_i)\log\mu(Q_j\,|\,P_i))$$
$$= H(\mathcal{P}) - \sum_j\sum_i\mu(P_i)\,\mu(Q_j\,|\,P_i)\log\mu(Q_j\,|\,P_i)).$$

Applying Jensen's Inequality to the concave function $\eta(\mu) = -\mu\log\mu$, we bound the last sum by the value of η at $\sum_i \mu(P_i)\,\mu(Q_j\,|\,P_i) = \mu(Q_j)$. This bounds the double sum by $\sum_j \eta(\mu(Q_j)) = H(\mathcal{Q})$.

To have an equality in the Jensen's Inequality, all the conditional measures $\mu(Q_j\,|\,P_i)$ must coincide for any given j, and hence be equal to $\mu(Q_j)$. This is equivalent to the independency of $\mathcal{P}$ and $\mathcal{Q}$. □

Let us now consider the one-sided shift $\sigma:\Sigma\to\Sigma$ in d symbols, where $\Sigma\equiv\Sigma_d^+$ (see §19.10). Let μ be an invariant measure for σ. Truncating sequences at length n, we obtain a finite probability space comprising d^n cylinders $\Sigma_{\bar{j}}^n$ (where some of them may have zero measure). The Shannon entropy of this probability space is equal to

$$H^n \equiv H^n_\mu := -\sum_{\bar{j}}\mu(\Sigma^n_{\bar{j}})\log\mu(\Sigma^n_{\bar{j}}).$$

If strings of symbols are interpreted as encoded messages then H^n is the average amount of information contained in a message of length n. Thus, $\dfrac{H^n}{n}$ is the average amount of information *per symbol* contained in a message of length n.

Lemma 46.32. *Under the above circumstances, there exists a limit*

$$h_\mu(\sigma) := \lim_{n\to\infty}\frac{H^n_\mu}{n}.$$

This limit is called the *entropy of σ with respect to μ.*

Proof. By Lemma 46.31 and the shift invariance of the entropy, the sequence H^n is subadditive: $H^{n+m}\le H^n + H^m$. The conclusion follows from Fekete's Lemma. □

Let us now consider a general transformation $f:X\to X$ preserving a probability measure μ. Given a partition $\mathcal{P}$ of X into d measurable pieces P_i, we obtain a coding $\pi: X\to\Sigma\equiv\Sigma_d^+$ of f (see §19.11). Pushing μ forward to Σ we obtain a shift invariant measure $\nu := \pi_*(\mu)$ on Σ. Then the *entropy of f with respect to the partition $\mathcal{P}$* is defined as

$$h_\mu(f,\mathcal{P}) := h_\nu(\sigma).$$

Finally, the *entropy of f with respect to μ* is defined as

$$h_\mu(f) := \sup_{\mathcal{P}} h_\mu(f,\mathcal{P}).$$

Sometimes $h_\mu(f)$ is also called a *measure-theoretic* entropy[14] of f, to contrast it to the *topological* entropy (see §48.4).

Note in conclusion that cylinders $\Sigma^n_{\bar{j}}$ of the partition $\mathcal{P}$ correspond to sets

$$P_{j_0} \cap f^{-1}(P_{j_2}) \cap \cdots \cap f^{-(n-1}(P_{j_{n-1}})$$

that form a partition

$$\mathcal{P}^n := \mathcal{P} \vee f^{-1}(\mathcal{P}) \vee \cdots \vee f^{-(n-1)}(\mathcal{P}).$$

Then we can introduce $H^n(f,\mathcal{P}) \equiv H^n_\mu(f,\mathcal{P}) := H_\mu(\mathcal{P}^n)$ and define the entropy directly in terms of f, without a reference to the corresponding shift:

$$h_\mu(f,\mathcal{P}) := \lim_{n\to\infty} \frac{H^n(f,\mathcal{P})}{n}, \quad h_\mu(f) = \sup_{\mathcal{P}} h_\mu(f,\mathcal{P}).$$

(as it is commonly done).

46.10.2. *Conditional entropy.* Given a finite measurable partition $\mathcal{P}$ and a measurable set $Q \subset X$, we can define the conditional entropy $H_\mu(\mathcal{P}|\, Q)$ as the entropy of the slice of $\mathcal{P}$ by Q with respect to the conditional measure on Q. If we have a second partition $\mathcal{Q}$, then we can average the conditional entropy of $\mathcal{P}$ over all the pieces of $\mathcal{Q}$ to obtain the conditional entropy:

(46.9)

$$H(\mathcal{P} \,|\, \mathcal{Q}) \equiv H_\mu(\mathcal{P}|\, \mathcal{Q}) := -\sum H_\mu(\mathcal{P}|\, Q_i)\, \mu(Q_i) = -\sum \log \mu(P_j|\, Q_i)\, \mu(Q_i)$$

$$= -\sum (\log \mu(P_j \cap Q_i) - \log \mu(Q_i))\, \mu(P_j \cap Q_i) = H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{Q}).$$

By Lemma 46.31, for any two finite partitions $\mathcal{P}$ and $\mathcal{Q}$, we have:

$$H(\mathcal{P} \,|\, \mathcal{Q}) \le H(\mathcal{P}),$$

and the equality is attained iff the partitions are independent. The intuitive meaning of this statement is that the information hidden in $\mathcal{P}$ drops after some $\mathcal{Q}$-event happens (However, if $\mathcal{Q}$ is independent of $\mathcal{P}$, then $\mathcal{Q}$ does not reveal anything about $\mathcal{P}$.) More generally, we have:

EXERCISE 46.33. *Let us consider three finite measurable partitions: $\mathcal{Q}_1 \prec \mathcal{Q}_2$, and $\mathcal{P}$. Then $H(\mathcal{P} \,|\, \mathcal{Q}_2) \le H(\mathcal{P} \,|\, \mathcal{Q}_1)$.*

Using the general notion of conditional measures, definition (46.9) can be extended to an arbitrary measurable partition $\mathcal{Q}$:

$$H_\mu(\mathcal{P}|\, \mathcal{Q}) := \int_{\bar{X}} H_\mu(\mathcal{P}|\; Q(\bar{x}))\, d\bar{\mu}(\bar{x}) = -\int_X \log \mu(P(x)|\, Q(x))\, d\mu(x),$$

where $\bar{X} = X/\mathcal{Q}$ and $\bar{\mu}$ is the push-forward of μ to the quotient space. It can also be written in the following forms:

$$H_\mu(\mathcal{P}|\, \mathcal{Q}) = -\int_X \log \mu(P(x)|\, Q(x))\, d\mu(x) = \sum_{P\in\mathcal{P}} \eta\,(\mu(P \,|\, Q(x))\, d\mu(x). \tag{46.10}$$

In particular, let us consider a partition

$$\mathcal{P}^\infty := \bigvee_{n=0}^{\infty} \mathcal{P}^n,$$

[14] In the Russian literature, it is called *metric entropy.*

called the *tail* partition. It is measurable and invariant: $f^{-1}(\mathcal{P}^\infty) \prec \mathcal{P}^\infty$. Note also that $P \vee f^{-1}(\mathcal{P}) = \mathcal{P}$. If $\mathcal{P}$ is a generator then $\mathcal{P}^\infty = \varepsilon$ is the partition into points, while $f^{-1}(\mathcal{P}_\infty)$ is the partition into the fibers of f.

PROPOSITION 46.34. *Let $f : (X, \mu) \to (X\mu)$ be a measure-preserving transformation. For any finite measurable partition $\mathcal{P}$, we have*

$$h_\mu(f, \mathcal{P}) = H_\mu(P|\, f^{-1}(\mathcal{P}^\infty)) = H_\mu(\mathcal{P}^\infty|\, f^{-1}(\mathcal{P}^\infty)).$$

PROOF. As $\mathcal{P}^n = \mathcal{P} \vee f^{-1}(\mathcal{P}^{n-1})$, we have (skipping "$\mu$" in the notation):

$$\begin{aligned} H(\mathcal{P}^n) &= H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-1})) + H(f^{-1}(\mathcal{P}^{n-1})) = H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-1})) + H(\mathcal{P}^{n-1}) \\ &= H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-1})) + H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-2})) + H(\mathcal{P}^{n-2}) \\ &= H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-1})) + H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^{n-2})) + \dots + H(\mathcal{P}\,|\,f^{-1}(\mathcal{P})) + H(\mathcal{P}). \end{aligned}$$

Since the conditional entropies $H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^k))$ decrease (see Exercise 46.33), they converge to a limit, implying that

$$h_\mu(f) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{P}^n) = \lim_{k\to\infty} H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^k)).$$

Finally, let us show that the last limit is equal to $H(\mathcal{P}\,|\,f^{-1}(\mathcal{P}^\infty))$. Let $\mathbb{Q}^k := f^{-1}(\mathcal{P}^k)$, $k \in \mathbb{Z}_+ \cup \{\infty\}$. By the Doob Theorem, for any $P \in \mathcal{P}$, we have

$$\mu(P\,|\,\mathcal{Q}^k) \to \mu(P\,|\,\mathcal{Q}^\infty) \text{ a.e.} \quad \text{as } k \to \infty.$$

As the function $\eta(\mu) = -\mu \log \mu$ is continuous,

$$\eta\left(\mu(P\,|\,\mathcal{Q}^k)\right) \to \eta\left(\mu(P\,|\,\mathcal{Q}^\infty)\right) \text{ a.e.} \quad \text{as } k \to \infty.$$

as well. Moreover, as the above functions are bounded, their convergence survives integration. This yields the desired by (46.10). □

This formula assumes a nice interpretation in case of the shift σ partitioned into rank-one cylinders Σ^1_i, particularly, if to view the string of symbols $(i_0 i_1 \dots)$ as coming from the "past", so that i_0 is the last symbol we have received. Then it is equal to the *average amount of information received when the last symbol arrives, assuming the whole infinite pre-history* $(i_1 i_2 \dots)$ *is known.*

46.10.3. *Generators.* Recall from §46.9.1 that ε stands for the partition into singletons. A finte partition $\mathcal{P}$ is called a *generator* if $\mathcal{P}^\infty = \varepsilon \mod 0$. In this case, the coding with respect to $\mathcal{P}$ induces a measure-theoretic isomorphism between (f, μ) and the one-sided shift (σ_d, ν) endowed with some invariant measure ν. (Of course, for the shift itself the partition $\mathcal{P}^0$ into the rank 1 cylinders Σ^1_i, $i \in \{1, \dots, d\}$, is the generator for any measure.)

Let $p(x)$ be the conditional probability of x in the fiber $F_x := f^{-1}(fx)$. It is a positive measurable function on X. Notice that

$$H_\mu(f^{-1}\varepsilon|\,\varepsilon) = \int \left(-\sum_{F_x} p(y) \log p(y) \right) d\mu(x) - \int \log p(x)\, d\mu(x)$$

Proposition 46.34 immediately implies:

COROLLARY 46.35. *If an invariant measure μ has a generator $\mathcal{P}$ then*

$$h_\mu(f, \mathcal{P}) = H_\mu(f^{-1}\varepsilon|\,\varepsilon) = -\int \log p(x)\, d\mu(x).$$

We also see that the entropy $h_\mu(f,\mathcal{P})$ is the same for all generators $\mathcal{P}$. In fact, it gives the whole entropy of f:

LEMMA 46.36. *If a partition $\mathcal{P}$ is generating then $h_\mu(f,\mathcal{P}) = h_\mu(f)$.*

46.10.4. *Radon-Nikodym Jacobian.* Let us first consider a *quasi-isomorphism* $f : (X,\mu) \to (Y,\nu)$ between measure spaces, i.e. an invertible transformation such that $f_*\mu \sim \nu$. The *Jacobian* of f is defined as the following Radon-Nikodym density

$$\operatorname{Jac} f \equiv \operatorname{Jac}_{\mu,\nu} f := \frac{d(f^*\nu)}{d\mu}.$$

(If $(X,\mu) = (Y,\nu)$ then we abbreviate the notation to $\operatorname{Jac}_\mu f$.) It satisfies the property that for any measurable subset $Z \subset X$ we have:

$$\nu(fZ) = \int_Z \operatorname{Jac} f \, d\mu. \tag{46.11}$$

EXERCISE 46.37. (i) *Check the standard Chain Rules (almost everywhere):*

$$\operatorname{Jac}(f\circ g)\,(x) = \operatorname{Jac} f(gx)\,\cdot \operatorname{Jac} g(x), \quad \operatorname{Jac} f^{-1}(fx) = 1/\operatorname{Jac} f(x).$$

(ii) *f is an isomorphism iff* $\operatorname{Jac} f \equiv 1$ *a.e.*

Let us now consider a *quasi-endomorphism* $f : (X,\mu) \to (Y,\nu)$, i.e., a measurable transformation wiith at most countable preimages such that $f_*\mu \sim \nu$. For such a transformation, the space X can be decomposed into a disjoint union of measurable subsets X_i such that the restrictions $f : X_i \to Y_i\,(\equiv f(X_i))$ are quasi-isomorphisms. Then we let

$$f^*\nu := \sum (f\,|\,X_i)^*(\nu\,|\,Y_i), \quad \operatorname{Jac} f := \frac{d(f^*\nu)}{d\mu}.$$

EXERCISE 46.38. (i) *Check that this definition is independent of the choice of a partition and so provides us with a well-defined positive (a.e.) measurable function* $\operatorname{Jac} f \equiv \operatorname{Jac}_{\mu,\nu} f(x)$ *on X.*

(ii) *It is determined by property* (46.11) *for any subset $Z \subset X$ for which the restriction $f : Z \to f(Z)$ is an isomorphism .*

(iii) *The pullback $f^*\nu$ can also be defined by integrating over $d\nu$ the homogeneous measures $\sum_{x\in f^{-1}y} \delta_x$ on the fibers of f, i.e., by letting for any measurable $Z \subset X$,*

$$(f^*\nu)(Z) = \int_Y |Z \cap f^{-1}y|\, d\nu(y).$$

(iv) *f is an endomorphism iff*

$$\sum_{x\in f^{-1}y} \frac{1}{\operatorname{Jac} f(x)} \equiv 1 \quad \textit{for a.e. } y \in Y.$$

(v) *If f is an endomorphism then* $\operatorname{Jac} f(x) = 1/p(x)$ *a.e. (Recall that $p(x)$ is the conditional measure of x in the fiber F_x.)*

Putting the last property together with formula (46.35), (and Lemma 46.36) we come up with a nice interpretation of entropy as the *exponential rate of measure expansion.*

COROLLARY 46.39. (i) *If an invariant measure* μ *has a generator* $\mathcal{P}$ *then*

$$h_\mu(f) = \int \log \mathrm{Jac}_\mu f \, d\mu.$$

(ii) *If additionally* μ *is ergodic then*

$$\frac{1}{n} \log \mathrm{Jac}\, f^n(x) \to h_\mu(f) \quad \text{for a.e. } x.$$

PROOF. (ii) By the Chain Rule and the Ergodic Theorem, we have:

$$\frac{1}{n} \log \mathrm{Jac}\, f^n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \log \mathrm{Jac}\, f(f^k x) \to \int \log \mathrm{Jac}\, f \, d\mu \quad \text{for a.e. } x.$$

□

46.10.5. *Shannon-McMillan-Breiman Formula.* Let us first state it in the case of shift. Recall that for an infinite string $\bar{j} \in \Sigma \equiv \Sigma_d^+$, we let $\bar{j}^n := (j_0 \dots j_{n-1})$.

THEOREM 46.40. *Let* μ *be an ergodic invariant measure for the shift* $\sigma : \Sigma \to \Sigma$. *Then for a.e.* $\bar{j} \in \Sigma$, *we have:*

$$- \lim_{n\to\infty} \frac{1}{n} \log \mu(\Sigma^n_{\bar{j}_n}) = h_\mu(\sigma).$$

If $\bar{j}$ is interpreted as an encoded infinite message, then the quantity on the left becomes the *asymptotic amount of information per symbol contained in this message*. Thus, *for a typical message, this amount is equal to the entropy of the transmission channel*.

For general maps, the Shannon-McMillan-Breiman Formula assumes the following form:

COROLLARY 46.41. *Let* $f : X \to X$ *be a transformation preserving an ergodic measure* μ, *and let* $\mathcal{P}$ *be a finite measurable partition. Then for a.e., we have:*

$$- \lim_{n\to\infty} \frac{1}{n} \log \mu(P^n(x)) = h_\mu(f, \mathcal{P}),$$

where $P^n(x)$ *stands for the piece of* $\mathcal{P}^n$ *containing* $x \in X$.

So, the exponential rate of decay of the measures $\mu(P^n(x))$ is equal (a.e.) to the entropy.

Proof of the SMB Theorem. Let $\mathcal{P}$ be the standard generator for the shift. By Corollary 46.39 and the Egorov Theorem, we have

$$\frac{1}{n} \log \mathrm{Jac}\, f^n(x) \to h \equiv h_\mu(f)$$

uniformly on a subset $Z \subset X$ of measure $> 1 - \delta$, so

$$C^{-1} e^{((h-\varepsilon)n)} \le \mathrm{Jac}\, f^n(x) \le C e^{(h+\varepsilon)n}.$$

Take a density point $x \in Z$ and let $P^n \equiv P^n(x)$. Since $f^n \mid P^n(x)$ is injective, we have

$$\mu(P^n) \le 2\mu(P^n) \cap Z) \le 2 \int_{f^n(P^n)} (\mathrm{Jac}\, f^n(x))^{-1} \, d\mu(y) \le 2C e^{(-h+\varepsilon)n},$$

so eventually

$$-\frac{1}{n} \log \mu(P^n) \ge h - 2\varepsilon$$

As the integral go to h, the upper bound follows as well. QED

46.10.6. *Zero entropy.*

PROPOSITION 46.42. *A measure preserving endomorphism $f : (X, \mu) \to (X, \mu)$ with zero entropy is an isomorphism* mod 0*.*

PROOF. In case when μ has a (one-sided) generator, this follows immediately from Corollary 46.35. □

46.10.7. *Examples.* Let us start with the Bernoulli measures $\mu_{\mathbf{p}}$, $\mathbf{p} \in \boldsymbol{\Delta}$ for the shift $\sigma \equiv \sigma_d : \Sigma_d \to \Sigma_d$ with d symbols (see §19.10.3). It is easy to classify these these dynamical systems:

PROPOSITION 46.43. *Two (one-sided!) Bernoulli shifts, $\sigma : (\Sigma_d^+, \mu_{\mathbf{p}} \to \Sigma_d^+, \mu_{\mathbf{p}})$ and $\sigma : (\Sigma_{\tilde{d}}^+, \mu_{\tilde{\mathbf{p}}} \to \Sigma_{\tilde{d}}^+, \mu_{\tilde{\mathbf{p}}})$, are isomorphic (*mod 0*) iff $d = \tilde{d}$ and the initial probability distributions coincide up to a permutation, i.e., $\{p_1, \dots, p_d\} = \{\tilde{p}_1, \dots, \tilde{p}_d\}$.*

PROOF. The conditional distributions on $\sigma^{-1}(\varepsilon)$ for $\mu_{\mathbf{p}}$ are equal to

$$\mu_{\mathbf{p}}(i \,|\, i_1 i_2 \dots) = p_i, \quad i = 1, \dots, d.$$

Since the conditonal distributions are preserved isomorphisms (mod 0), the conclusion follows. □

REMARK 46.44. The situation with two-sided Bernoulli shifts is much more delicate: see the Notes below.

PROPOSITION 46.45. *We have: $h_{\mu_{\mathbf{p}}}(\sigma) = H(\mathbf{p})$. In particular, $h_\mu(\sigma) = \log d$ for the balanced measure.*

PROOF. Let us consider the partition $\mathcal{P}$ of $\Sigma \equiv \Sigma_d$ into the cylinders of rank 1. Since it is generating, $h_{\mu_{\mathbf{p}}}(\sigma) = h_{\mu_{\mathbf{p}}}(\sigma, \mathcal{P})$. Since the partitions $\sigma^{-k}(\mathcal{P})$ are independent,

$$H(\mathcal{P}^n) = \sum_{k=0}^{n-1} H(\sigma^{-k}(\mathcal{P})) = n\, H(\mathcal{P}).$$

The conclusion follows. □

COROLLARY 46.46. *Let us consider two probability distributions*[15] $\mathbf{p}$ *and* $\tilde{\mathbf{p}}$ *with different Shannon entropies: $H(\mathbf{p}) \neq H(\tilde{\mathbf{p}})$. Then the corresponding Bernoulli measures $\mu_{\mathbf{p}}$ and $\mu_{\tilde{\mathbf{p}}}$ are not equivariantly isomorphic* mod 0*.*

We also see that the balanced measure maximizes entropy among all Bernoulli measures (with the same number of states). In fact, it remains true among all measures:

THEOREM 46.47. *For any integer $d \geq 2$, the balanced measure μ is the unique measure of maximal entropy for the Bernoulli shift $\sigma : \Sigma_d^+ \to \Sigma_d^+$.*

PROOF. It is easy to see that μ is a mesure of maximal entropy, as for any invariant measure ν we have:

$$H_\nu(\sigma, \mathcal{P}^n) \leq \#\{\text{cylinders of rank } n\} = d^n.$$

[15] not necessarily with the same number of states

To prove uniqueness, let us use formula (46.35):

$$h_\nu(\sigma) = H_\nu(\sigma^{-1}\varepsilon \,|\, \varepsilon) = \int_{\bar{X}} H(\mathbf{p}(\bar{x}))\, d\bar{\nu}(\bar{x}),$$

where $\bar{X} = \Sigma/\sigma^{-1}(\mathcal{P})$, $\bar{\nu}$ is the quotient measure, and $\mathbf{p}(\bar{x})$ are the conditional probability distributions on the fibers of the natural projection $X \to \bar{X}$.

Each fiber of this projection contains d points, so $H(\mathbf{p}(\bar{x})) \leq \log d$ everywhere. If $\nu \neq \mu$ then $\mathbf{p}(\bar{x})$ is not the uniform distribution over some set of positive $\bar{\nu}$-measure $\bar{Z} \subset \bar{X}$. Hence $H(\mathbf{p}(\bar{x})) < \log d$ for $\bar{x} \in \bar{Z}$. It follows that $H_\nu(\sigma^{-1}\varepsilon \,|\, \varepsilon) < \log d$ as well. □

Similarly, we have:

Exercise 46.48. *Let $\sigma_A : \Sigma_A^+ \to \Sigma_A^+$ be an irreducible (one-sided) Markov shift, and let μ_A be its invariant balanced measure* (see §19.14.5) *. Then:*

(i) $h_{\mu_A}(\sigma_A) = \log r(A)$*, where $r(A)$ is the spectral radius of A;*

(ii) μ_A *is a measure of maximal entropy.*

As in the Bernoulli case, μ_A is the *unique* measure of maximal entropy for σ_A, but it is less obvious, and we omit this discussion for now.

46.10.8. *Notes.* The Dynamical Entropy Theory was founded by Kolmogorov [**K**] and Sinai [**Si**] in the late 1950s. It is less known that some version of this notion had appeared earlier in the unpublished Master Thesis by Arov (see [**Ar**]). Demonstration that the Bernoulli measures for *two-sided*(!) shifts are not all equivalent (which was the long-standing problem of the time) was the first spectacular success of the new theory. (Much later (around 1970) it was proven by Ornstein that the entropy fully classfies these shifts.) Rokhlin's survey [**Ro2**] from 1967 still remains invaluable source in this field. For other accounts, see e.g., [**Bi, KSF, MaE**].

The measure of maximal entropy for Markov maps was identified by Parry [**Par**].

47. Parapuzzle and its Principal Nest

In this section we will further refine the Renormalization Structure of the Mandelbrot set $\mathcal{M}$ (see §43.5) by constructing a *Generalized Renormalization Hierarchy* of parapuzzle tilings of the last two subsets of decomposition (43.4). The corresponding nests of parapuzzle pieces either shrink to primitive M-copies or to combinatorial classes of non-renormalizable parameters, which will give us a systematic way of constructing these objects. An important outcome is a construction of *proper unfolded ql families* that produce *primitive* M-copies.

47.1. Satellite dyadic tips, wakes, and decorations.

47.1.1. *Description.* Let us take a look at the Renormalization Decomposition of Theorem 43.17. It begins with the main hyperbolic component Δ_0 bounded by the main cardioid $\mathcal{N}_0$. Attached to the main cardioid at non-zero rational points $\mathfrak{r}_{\mathfrak{p}/\mathfrak{q}}$, $\mathfrak{p}/\mathfrak{q} \in (\mathbb{Q}/\mathbb{Z})^*$, are wakes $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$. For $c \in \mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$, the map f_c has $\mathfrak{q}$ rays landing at α_c that are permuted with combinatorial rotation number $\mathfrak{p}/\mathfrak{q}$ (see Corollary 37.10). Let $\mathfrak{R}_c \equiv \mathfrak{R}(\alpha_c)$ be this configuration of rays, and let $\mathfrak{R}'_c \equiv \mathfrak{R}(\alpha'_c) = f_c^{-1}(\mathfrak{R}_c)$ be the symmetric configuration of rays landing at $\alpha'_c = -\alpha_c$.

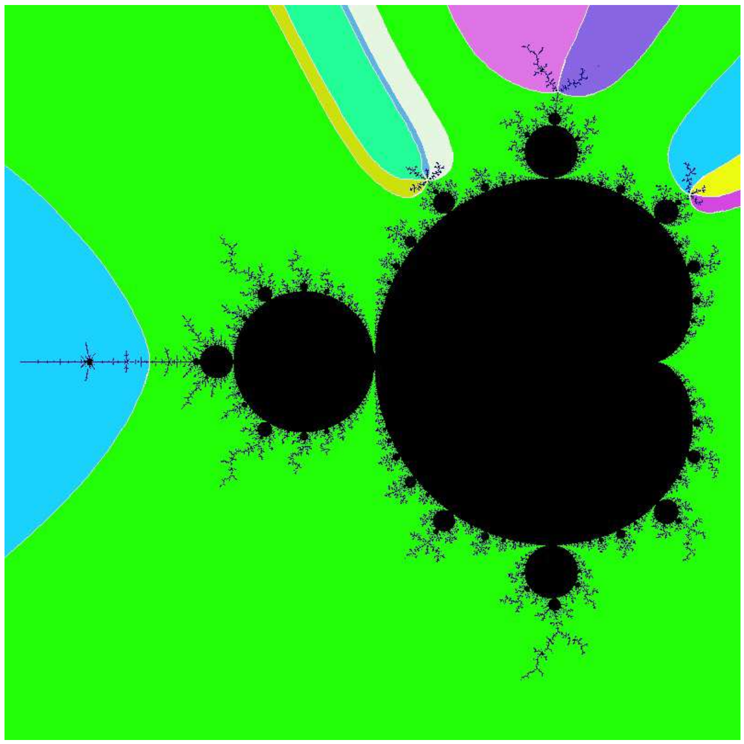

FIGURE 47.1. Chebyshev wakes attached to the tips of satellite M-copies with rotation numbers 1/2, 2/5, 1/3 and 1/4.

Inside the parabolic limb $\mathcal{L}_{\mathfrak{p}/\mathfrak{q}} = (\mathcal{W}^{\mathrm{par}}_{\mathfrak{p}/\mathfrak{q}} \cap \mathcal{M}) \cup \mathfrak{r}_{\mathfrak{p}/\mathfrak{q}}$ we see the satellite copy $\mathcal{M}_{\mathfrak{p}/\mathfrak{q}}$ of the Mandelbrot set $\mathcal{M}$. Each of these copies has a *Chebyshev tip* $\mathfrak{t} \equiv \mathfrak{t}^{\mathfrak{p}/\mathfrak{q}}_{\text{Ч}}$ for which $f^{\mathfrak{q}}_{\mathfrak{t}}(0) = \alpha_{\mathfrak{t}}$. Under the straightening $\chi : \mathcal{M}_{\mathfrak{p}/\mathfrak{q}} \to \mathcal{M}$, it corresponds to the Chebyshev parameter $c_{\text{Ч}} = -2$ of $\mathcal{M}$. Pulling back the rays landing at $\alpha_{\mathfrak{t}}$, we obtain $\mathfrak{q}$ rays landing at the critical value $v_{\mathfrak{t}} = f_{\mathfrak{t}}(0)$. By the corresponding parameter result (Theorem 37.35), each Chebyshev tip $\mathfrak{t}$ is the landing point of $\mathfrak{q}$ parameter rays. They divide the parameter plane into $\mathfrak{q}$ sectors, one *main sector* containing 0, and $\mathfrak{q}-1$ *satellite Chebyshev wakes* $\mathcal{V}^{\text{Ч}}_i$, $i = 1, \dots \mathfrak{q}-1$ (compare §43.4). The wake $\mathcal{V}^{\text{Ч}}_i$ is dynamically specified by the property that

$$F_c(0) = f^{\mathfrak{q}}_c(0) \in S'_{i;c},$$

where $S'_{i;c} = -S_{i;c}$, $i = 1, \dots, \mathfrak{q}-1$, are the lateral sectors attached to α'_c (see §24.4.3) and F_c is the *accelerated map* (24.2) from §24.4.3.

More generally, for each $\mathbf{n}\in\mathbb{Z}_+$, the satellite M-copy $\mathcal{M}_{\mathfrak{p}/\mathfrak{q}}$ has $2^{\mathbf{n}-1}$ *dyadic tips* $\mathfrak{t}\equiv\mathfrak{t}_j^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$, $j=1,\dots,2^{\mathbf{n}-1}-1$, for which

$$F_{\mathfrak{t}}^m(0)=f_{\mathfrak{t}}^{\mathfrak{q}m}(0)\in\Pi_{\mathfrak{t}},\ m=0,\dots,\mathbf{n}-1,\quad F_{\mathfrak{t}}^{\mathbf{n}}(0)=f_{\mathfrak{t}}^{\mathfrak{q}\mathbf{n}}(0)=\alpha'_{\mathfrak{t}},$$

where Π_c is the central complementary strip to the ray configuration $\mathfrak{R}_c\cup\mathfrak{R}'_c$. Under the straightening $\chi:\mathcal{M}_{\mathfrak{p}/\mathfrak{q}}\to\mathcal{M}$, they correspond to the dyadic tips $\mathfrak{t}_j^{\mathbf{n}}$ of $\mathcal{M}$ introduced in §37.10. There are $\mathfrak{q}-1$ *satellite dyadic wakes* $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$, $i=1,\dots,\mathfrak{q}-1$, attached to each of these tips in the same way as for the Chebyshev tips. The corresponding *satellite decorations* are

$$\mathcal{T}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}:=(\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}\cap\mathcal{M})\cup\{\mathfrak{t}_j^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}\}.$$

EXERCISE 47.1. *Looking at the picture of $\mathcal{M}$, how to recognize rotation numbers $\mathfrak{p}/\mathfrak{q}$ of the bifurcation points on the main cardioid?*

Let $\mathfrak{R}^1\equiv\mathfrak{R}'$, and let us define $\mathfrak{R}^{\mathbf{n}}$ inductively as the preimage of the ray configuration $\mathfrak{R}^{\mathbf{n}-1}$ by the double covering $F:\Pi\to S_0$. Let $A^{\mathbf{n}}$ be the component of $\mathbb{C}\smallsetminus\bigcup_{m\le\mathbf{n}}\mathfrak{R}^m$ containing 0. (As usual, sets depending on a parameter are marked with the corresponding subscript, e.g., $A_c^{\mathbf{n}}$.)

LEMMA 47.2. (i) *The ray configuration $\mathfrak{R}_c^{\mathbf{n}}$ moves holomorphically* (*under the Böttcher motion*) *over the domain*

$$\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}-1}=\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}\smallsetminus\bigcup_{m=1}^{\mathbf{n}-1}\bigcup_{i,j}\operatorname{cl}\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,m,}$$

(*the parabolic wake $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$ with the closures of all the satellite dyadic wakes of level $\le\mathbf{n}-1$ removed*). *In particular, it moves holomorphically over any dyadic wake $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$, $i=1,\dots,\mathfrak{q}-1$, $j=0,\dots,2^{\mathbf{n}-1}-1$.*

(ii) *For $c\in\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$, the map $F_c:A_c^{\mathbf{n}}\to A_c^{\mathbf{n}-1}$ is a double branched covering.*

PROOF. By the Wake Theorem (Corollary 37.10), the ray configuration $\mathfrak{R}_c$ moves holomorphically over the parabolic wake $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}$. Since $f_c(0)\notin\mathfrak{R}_c$ for $c\in\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}^{\mathrm{par}}$, the configuration $\mathfrak{R}'_c=f_c^{-1}(\mathfrak{R}_c)$ moves holomorphically as well. Moreover, the critical orbit $f_c^n(0)$, $n=0,1\dots,\mathfrak{q}-1$, does not hit the latter (since these $f_c^n(0)$ stay outside the critical sector $S_0(\alpha_c)$) (see §24.5.1) while the boundary of the Chebyshev wakes $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},1}$ is specified by the property that $F_c(0)\equiv f_c^{\mathfrak{q}}(0)\in\mathfrak{R}'_c$. By Lemma 17.10, the configuration $\mathfrak{R}_c^2=f_c^{-\mathfrak{q}}(\mathfrak{R}'_c)$ moves holomorphically outside this boundary, and hence inside $\Lambda^{\mathfrak{p}/\mathfrak{q},1}$ (compare Exercise 37.37). Note also that $F_c:A_c^2\to A_c^1\equiv\Pi_c$ is a double covering map for $c\in\Lambda^{\mathfrak{p}/\mathfrak{q},1}$.

Proceeding inductively, assume that the configuration $\mathfrak{R}_c^{\mathbf{n}}$ moves holomorphically over the parameter domain

$$\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}-1}=\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}\smallsetminus\bigcup_{m=1}^{\mathbf{n}-1}\bigcup_{i,j}\operatorname{cl}\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,m},$$

and that for $c\in\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}-1}$, the map F_c is a double covering $A_c^n\to A_c^{\mathbf{n}-1}$. Then the critical orbit $f_c^n(0)$, $n=0,1\dots,\mathfrak{q}-1$, does not hit $\mathfrak{R}_c^{\mathbf{n}-1}$ (for the same reason as above), while the boundary of the dyadic wakes $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$ is specified by the property that $F_c(0)\in\mathfrak{R}_c^{\mathbf{n}}$. Applying Lemma 17.10 once again, we conclude that the configuration $\mathfrak{R}_c^{\mathbf{n}+1} = F_c^{-1}(\mathfrak{R}_c^{\mathbf{n}})$ moves holomorphically outside this boundary, and in particular, it moves holomorphically over $\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$. Also, for $c \in \Lambda^{\mathfrak{p}/\mathfrak{q},\mathbf{n}}$, the map $F_c : A_c^{n+1} \to A_c^{\mathbf{n}}$ is obtained from $F_c : A_c^n \to A_c^{\mathbf{n}-1}$ by restricting the latter to $A_c^{\mathbf{n}+1} = F_c^{-1}(A_c^{\mathbf{n}})$, which yields a double covering. $\square$

PROBLEM 47.3. *Show that the straightening homeomorphism $\chi : \mathcal{M}^{\mathfrak{p}/\mathfrak{q}} \to \mathcal{M}$ admits a continuous extension to a map $\mathbb{C} \to \mathbb{C}$ that collapses each satellite wake (attached to the satellite dyadic tip $\mathfrak{t}_j^{\mathfrak{p}/\mathfrak{q},\,n}$) to the ray $\mathcal{R}_{\mathrm{par}}^{(2j-1)/2^{\mathbf{n}}}$ landing at the corresponding dyadic tip $\mathfrak{t}_j^{\mathbf{n}} = \chi(\mathfrak{t}_j^{\mathfrak{p}/\mathfrak{q},\,n})$.* (It can be called a *"sectorial Devil Staircase"*.)

Let $c_\circ \equiv c_\circ(\mathfrak{p}/\mathfrak{q}) \in \mathcal{M}_{\mathfrak{p}/\mathfrak{q}}$ be the center of the satellite M-copy. It is contained in all the domains $\Lambda^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}-1}$, so we can center the Böttcher motion $h_c : \mathfrak{R}_\circ^{\mathbf{n}} \to \mathfrak{R}_c^{\mathbf{n}}$ over this domain at $c_\circ$.

LEMMA 47.4. *The point $h_c^{-1}(F_c^{\mathbf{n}+1}(0))$ runs once over the boundary of the sector $S_{0;\,\circ}$ as c runs over the boundary of any satellite dyadic wake $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$.*

PROOF. Let us fix some wake $\mathcal{V} \equiv \mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$ rooted at a tip $\mathfrak{t}$. Let $\mathcal{R}_{\mathrm{par}}^{\pm}$ be the parameter rays that bound $\mathcal{V}$. At the tip, the corresponding dynamical rays $\mathcal{R}_{\mathfrak{t}}^{\pm}$ (with the same angles as $\mathcal{R}_{\mathrm{par}}^{\pm}$) land at the critical value $v_{\mathfrak{t}} = f_{\mathfrak{t}}(0)\,(= \mathfrak{t})$ bounding a sector $S(v_{\mathfrak{t}})$ (see Theorem 37.35). Moreover, under $f_{\mathfrak{t}}^{\mathfrak{q}\mathbf{n}}$, this sector is univalently mapped onto the lateral sector $S_{i;\,\mathfrak{t}}$. For $c \in \partial\mathcal{V}$ the critical value $v_c = f_c(0)\,(= c)$ belongs to one of the dynamical rays $\mathcal{R}_c^{\pm}$.

Let us consider the corresponding dynamical rays $\mathcal{R}_\circ^{\pm}$ for the base map $f_\circ$. They land at some preimage $w_\circ$ of $\alpha_\circ$ of order $\mathfrak{q}\mathbf{n}$ bounding a sector $S(w_\circ)$ (where $w_\circ$ is not the critical value anymore, but it corresponds to $v_{\mathfrak{t}}$ under the Böttcher motion, $v_{\mathfrak{t}} = h_{\mathfrak{t}}(w_\circ)$.) Moreover, under $F_\circ^{\mathbf{n}} \circ f_\circ^{\mathfrak{q}-1}$, the sector $S(w_\circ)$ is univalently mapped onto the sector $S_{0;\,\circ}$.

By the Phase-Parameter Relation (Lemma 34.9), the point $h_c^{-1}(v_c)$ runs once over the boundary of the sector $S(w_\circ)$ as c runs over the boundary of the dyadic wake $\mathcal{V}$. Since the Böttcher motion is equivariant, the point

$$h_c^{-1}(F_c^{\mathbf{n}+1}(0)) = h_c^{-1}(F_c^{\mathbf{n}} \circ f_c^{\mathfrak{q}-1}(v_c)) = F_\circ^{\mathbf{n}} \circ f_\circ^{\mathfrak{q}-1}(h_c^{-1}(v_c)), \quad c \in \partial\mathcal{V},$$

runs once around the boundary of $S_{0;\,\circ}$. $\square$

47.2. Motion of the initial tiling.

47.2.1. *Motion of the truncated satellite wakes.* Let select an equipotential level $t > 0$ that will serve as the initial equipotential for the Yoccoz puzzle (so the puzzle pieces $Y_{i;c}^0 \equiv Y_{i;c}^{(0)}$ of zero depth tile the subpotential domain $\Sigma_c(t)$). Then puzzle pieces $Y_{i,c}^{(n)}$ of depth n tile the subpotential domain $\Sigma_c^{(n)} := \Sigma_c(t/2^n)$. By the Phase-Parameter Relation, these domains move holomorphically over the parameter subpotential domain $\Sigma_{\mathrm{par}}(t/2^{n-1})$. Moreover, letting $\Sigma_{\mathrm{par}}^{(n)} := \Sigma_{\mathrm{par}}(t/2^n)$, we have: $v_c \in \partial\Sigma_c^{(n)}$ for $c \in \partial\Sigma_{\mathrm{par}}^{(n)}$ (where of course, $v_c = c$ for the quadratic family).

For a map f_c with escaping time $\mathbf{n}$, recall the initial tiling combining (31.22) and (31.25):

$$Y_c^0 \underset{\mathrm{ess}}{=} V_c^0 \cup (Q_{1;\,c} \cup Q_{2;\,c}) \cup \bigcup_{m=1}^{\mathbf{n}+1} \bigcup_{ij} Z_{ij;\,c}^m \tag{47.1}$$

endowed with the accelerated Markov map (31.26)

$$G_c : V_c^0 \cup (Q_{1;\,c} \cup Q_{2;\,c}) \cup \bigcup_{m=1}^{\mathbf{n}+1} \bigcup_{ij} Z_{ij;\,c}^m \to Y_c^0. \tag{47.2}$$

We let $\gamma(c) := G_c(0)$ be the corresponding *accelerated critical value.*

Let us define the *truncated parabolic wake of depth* n as $\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}^{(n)} = \mathcal{W}^{\mathfrak{p}/\mathfrak{q}} \cap \Sigma_{\mathrm{par}}^{(n)}$. The truncated parabolic wake of *level* $m \le \mathbf{n}$ is defined as

$$\mathcal{W}_{\mathfrak{p}/\mathfrak{q}}^m = \mathcal{W}_{\mathfrak{p}/\mathfrak{q}}^{(\mathfrak{q}m)} \quad (\text{note that } \mathfrak{q}m = \operatorname{depth} Y_c^m + \mathfrak{q} - 1).$$

It is dynamically specified by the property that $F_c(0) \equiv f_c^{\mathfrak{q}}(0) \in Y_c^m$.

Similarly, take a satellite dyadic wake $\mathcal{V} \equiv \mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$ and truncate it by the equipotential $\mathcal{E}_{\mathrm{par}}^{(l)}$ of depth $l \equiv l_{ij}^{\mathbf{n}} := \operatorname{depth} Z_{ij;\,c}^{\mathbf{n}} + \mathfrak{q} - 1$ (which is the same for all $c \in \mathcal{V}$). The truncated wake $\mathcal{V}_{\mathrm{tr}} \equiv \mathcal{V}_{ij;\,\mathrm{tr}}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$ is dynamically specified by the property that $F_c(0) \in Z_{ij;\,c}^{\mathbf{n}}$, see (31.23).

Exercise 47.5. *Check that the dyadic labeling of the pieces* $Z_{ij;\,c}^{\mathbf{n}}$ *by* $j \in \{0,1\}^{\mathbf{n}-1}$ (see Remark 31.19) *matches with the labeling of the corresponding wakes* $\mathcal{V}_{ij}^{\mathfrak{p}/\mathfrak{q},\mathbf{n}}$.

For a parameter set $\mathcal{V} \subset \mathbb{C}$, we let

$$\mathbf{Y}^0|\,\mathcal{V} := \bigcup_{c\in\mathcal{V}} Y_c^0.$$

Lemma 47.6. *The boundary of the initial tiling* (47.1) *moves holomorphically over any truncated satellite wake* $\mathcal{V}_{tr} \equiv \mathcal{V}_{ij;\,tr}^{\mathfrak{p}/\mathfrak{q},\,\mathbf{n}}$. *Moreover, the accelerated critical value* $\boldsymbol{\gamma}(c) := (c,\, \gamma(c))$ *is a proper map* $\boldsymbol{\gamma} : \mathcal{V}_{tr} \to \mathbf{Y}^0|\,\mathcal{V}_{tr}$ *with winding number one.*

Proof. Let us first consider pieces $Z_{ij;\,c}^m$ of level $m \le \mathbf{n}$, together with the puzzle piece $Y^{\mathbf{n}}$. By Lemma 47.2, the ray part of their boundary moves holomorphically over the wake $\mathcal{V}_{\mathrm{tr}}$. Also, the equipotentials of level $\ge \mathbf{n}$ move holomorphically over the subpotential domain $\Sigma_{\mathrm{par}}^{(\mathbf{n})}$. Hence the equipotential boundary of these puzzle pieces moves holomorphically over $\mathcal{V}_{\mathrm{tr}}$ as well.

All other pieces of the initial tiling are pullbacks of the pieces $Z_{j;\,c}^{\mathbf{n}}$ and $Y_c^{\mathbf{n}}$ by the map $F_c|\,Y_c^{\mathbf{n}}$. For $c \in \mathcal{V}_{\mathrm{tr}}$, the boundary of the latter pieces is not crossed by the critical value $\gamma(c) = F_c(0)$. By Lemma 17.10, all the pieces in question move holomorphically over $\mathcal{V}_{\mathrm{tr}}$.

Furthermore, by the Phase-Parameter Relation, $v_c \in \mathcal{E}_c^{(l)}$ as $c \in \mathcal{E}_{\mathrm{par}}^{(l)}$. Hence $f_c^{\mathfrak{q}-1}(v_c) \in \mathcal{E}_c^{(d)}$, where $d = \operatorname{depth} Z_{ij;\,c}^{\mathbf{n}}$, and $F_c^{\mathbf{n}}(f_c^{\mathfrak{q}-1}(v_c)) \in \mathcal{E}^{(0)}$. Thus,

$$\gamma(c) = G_c(0) = F_c^{\mathbf{n}+1}(0) = F_c^{\mathbf{n}}(f_c^{\mathfrak{q}-1}(v_c)) \in \mathcal{E}^{(0)}.$$

It follows that if c belongs to the equipotential piece of $\partial\mathcal{V}_{\mathrm{tr}}$ the $\gamma(c)$ belongs to the equipotential piece of ∂Y_c^0. On the other hand, if c belongs to the radial part of $\partial\mathcal{V}_{\mathrm{tr}}$ then $\gamma(c)$ belongs to the radial part of Y_c^0, by Lemma 47.4. The conclusion follows. $\square$

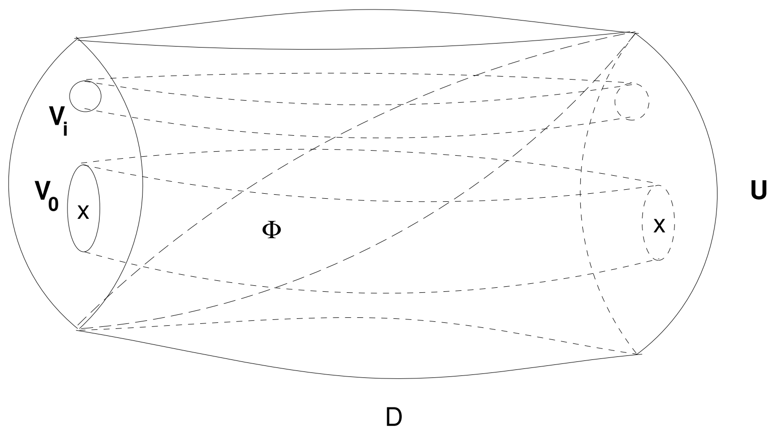

FIGURE 47.2. Generalized quadratic-like family.

47.3. Generalized quadratic-like families. We let $\pi : \mathbb{C}^2 \to \mathbb{C}$ be the projection to the first coordinate.

In what follows, we will refine the structure of the satellite dyadic limbs according to the combinatorics of the Principal Nest. To this end, we need a notion of generalized ql family.

Let us consider a topological disc $\Lambda \subset \mathbb{C}$ with a base point $\lambda_\circ \in D$, and a family of topological bidisks $\mathbb{V}_i \subset \mathbb{U} \subset \mathbb{C}^2$ over Λ (*tubes*), such that the $\mathbb{V}_i$ are pairwise disjoint. We assume that $V_{0,\lambda} \ni 0$. Let

$$\mathbf{g} : \bigcup \mathbb{V}_i \to \mathbb{U} \tag{47.3}$$

be a fiberwise map that admits a holomorphic extension to some neighborhoods of the $\mathbb{V}_i$ and whose fiber restrictions

$$\mathbf{g}(\lambda, \cdot) \equiv g_\lambda : \bigcup_i V_{i,\lambda} \to U_\lambda, \quad \lambda \in \Lambda,$$

are generalized quadratic-like maps with the critical point at $0 \in V_\lambda \equiv V_{0,\lambda}$ (see §31.3.1). We will assume that the discs U_λ and $V_{i,\lambda}$ are bounded by piecewise smooth quasicircles.

Let us also assume that there is a holomorphic motion $\mathbf{h}$ over $(\Lambda, \lambda_\circ)$,

$$h_\lambda : (\bar{U}_\circ, \bigcup_i \partial V_{i,\circ}) \to (\bar{U}_\lambda, \bigcup_i \partial V_{i,\lambda}), \tag{47.4}$$

which respects the boundary dynamics:

$$h_\lambda \circ g_\circ(z) = g_\lambda \circ h_\lambda(z) \quad \text{for} \quad z \in \bigcup \partial V_{i,\circ}. \tag{47.5}$$

A *generalized quadratic-like family* $(\mathbf{g}, \mathbf{h})$ *over* Λ is a map (47.3) together with a holomorphic motion (47.4) satisfying (47.5). We will sometimes reduce the notation to $\mathbf{g}$. In case when the domain of $\mathbf{g}$ consists of only one tube $\mathbb{V}_0$, we obtain a *DH quadratic-like family* in the sense of §42.1.

REMARK 47.7. It would be more consistent to call just $\mathbf{g}$ a holomorphic family, while to call the pair $(\mathbf{g}, \mathbf{h})$, say, an *equipped* holomorphic family. However, in this section we will assume that the families are equipped.

Let us now consider the *valuable section*

$$\gamma(\lambda) \equiv \gamma_{\mathbf{g}}(\lambda) = g_\lambda(0), \quad \boldsymbol{\gamma}(\lambda) \equiv \boldsymbol{\gamma}_{\mathbf{g}}(\lambda) = \mathbf{g}(\lambda, 0) \equiv (\lambda, \gamma(\lambda)).$$

Let us say that $\mathbf{g}$ is a *proper* (or *full*) holomorphic family if the fibration $\pi : \mathbb{U} \to \Lambda$ admits an extension to the boundary $\bar{\Lambda}$, $\bar{\mathbb{V}}_i \subset \mathbb{U}$, and $\boldsymbol{\gamma} : \Lambda \to \mathbb{U}$ is a proper section. Note that the fibration $\pi : V_0 \to \Lambda$ cannot be extended to $\bar{\Lambda}$, as the domains $V_{\lambda,0}$ pinch to figure-eights as λ approaches $\partial\Lambda$.

Given a proper holomorphic family $\mathbf{g}$ of generalized quadratic-like maps, let us define its *winding number* $w(\mathbf{g})$ as the winding number of the critical value $\gamma(\lambda)$ about the critical point 0. By the Argument Principle, it is equal to the winding number of the critical value about any section $\bar{\Lambda} \to \mathbb{U}$.

Let $\operatorname{mod} \mathbf{g} = \inf_{\lambda \in D} \operatorname{mod}(U_\lambda \smallsetminus V_{0,\lambda})$.

We also face the situation when $\mathbf{g}$ does not map every tube $\mathbb{V}_i$ onto the whole tube $\mathbb{U}$ but still satisfies the following *Markov property*: $\mathbf{g}(\mathbb{V}_i)$ either contains $\mathbb{V}_j$ or disjoint from it (and all the rest properties listed above are still valid, see §47.2.1). Then we call $\mathbf{g}$ a *holomorphic family of Markov maps.*

47.4. First generalized quadratic-like family. In what follows, the parameter will be denoted λ instead of c. Fix a truncated satellite dyadic wake $\mathcal{V}_{\text{tr}}$. The first generalized quadratic-like map $g_{1,\lambda} : \bigcup V^1_{i,\lambda} \to V^0_\lambda$ is defined as the first return map to V^0_λ (see §31.5). The itinerary of the critical point via the pieces of the initial tiling (47.1) determines the parameter tiling $\mathcal{D}^1$ of $\mathcal{V}_{\text{tr}}$ by the corresponding puzzle pieces. Let $\Delta^1(\lambda)$ stand for such parapuzzle piece containing λ.

More precisely, for any $\lambda \in \mathcal{V}_{\text{tr}}$, let us consider the first landing map

$$L_\lambda : \bigcup X_{\bar{i},\lambda} \to V^0_\lambda$$

(see §31.2.1). The puzzle piece $X_{\bar{i};\,\lambda}$ is specified by its itinerary $\bar{i} = (i_0, \dots, i_{s-1})$ under iterates of the accelerated Markov map G_λ (47.2) through non-central pieces P_i of the initial tiling[16] until the first landing in V^0_λ:

$$X_{\bar{i};\,\lambda} = \{z : G^k_\lambda(z) \in P_{i_k;\,\lambda},\ k = 0, \dots, s-1,\ \ G^s_\lambda(z) \in V^0_\lambda\}. \tag{47.6}$$

These tiles are organized in foliated tubes $\mathbf{X}_{\bar{i}}$ (endowed with holomorphically moving boundary). Moreover, the first landing map induces a foliated diffeomorphism $\mathbf{L} : \mathbf{X}_{\bar{i}} \to \mathbb{V}^0$ fibered over id.

Let $\bar{i}_\lambda$ stand for the itinerary of the accelerated critical value $\gamma(\lambda) = G_\lambda(0)$ through the initial tiling, so that, $G_\lambda(0) \in X_{\bar{i}_\lambda,\lambda}$. For a base point $\lambda_\circ \in \mathcal{V}_{\text{tr}}$ (which can be selected arbitrarily) , we let $\mathbf{X}_\circ \equiv \mathbf{X}_{\bar{i}_\circ}$. Then the parapuzzle piece of the tiling $\mathcal{D}^1$ centered at $\lambda_\circ$ is defined as follows:

$$\Delta^1(\lambda_\circ) = \boldsymbol{\gamma}^{-1}(\mathbf{X}_\circ) = \{\lambda \in \mathcal{V}_{\text{tr}} : G_\lambda(0) \in X_{\bar{i}_\circ,\lambda}\}.$$

Let $\mathbb{V}^1_j$ denote the components of $\mathbf{G}^{-1}(\mathbf{X}_{\bar{i}} | \Delta^1(\lambda_\circ))$ contained in $\mathbb{V}^0$, where $\mathbb{V}^1_0 \equiv \mathbb{V}^1$ is the critical component (i.e., the one containing $\mathbf{0}$). The first return map

$$\mathbf{g}_1 = \mathbf{L} \circ \mathbf{G} : \bigcup \mathbb{V}^1_j \to \mathbb{V}^0 \equiv \mathbb{U}^1$$

is the desired first generalized renormalization of $\mathbf{f}$.

By means of the first landing map $\mathbf{L}$, the holomorphic motion $\mathbf{h}$ over $\mathcal{V}$ can be lifted to the boundary of the tubes $\mathbf{X}_{\bar{i}}$. By the λ-lemma, this lift and the motion $\mathbf{h}$ of the boundary of the initial tiling (47.1) admit a common extension $\mathbf{H}_0$ over $\mathcal{V}_{\text{tr}}$.

[16] Here we switch to a homogeneous notation P_i for the pieces of the initial tiling.

Since the critical value $\boldsymbol{\gamma}(\lambda)$ lands at the tube $\mathbf{X}_\circ$ as λ ranges over $\Delta^1(\lambda_\circ)$, $\mathbf{H}_0$ can be lifted to a holomorphic motion of the annulus $V^0_\lambda \smallsetminus V^1_\lambda$ over $\Delta^1(\lambda_\circ)$. Let us extend this motion to V^1_λ by the λ-lemma. This provides us with a motion $\mathbf{h}_1$ which equips the generalized quadratic-like family $\mathbf{g}_1$.

Since the winding number of $\boldsymbol{\gamma}$ around $\mathbf{Y}^0_j$ over $\mathcal{V}_{\mathrm{tr}}$ is equal to 1 (by Lemma 47.6), the function $\boldsymbol{\gamma} : \Delta^1(\lambda_\circ) \to \mathbf{X}_\circ$ is proper with winding number 1. Since the first landing map $\mathbf{L}$ is a fiberwise diffeomorphism of every tube $\mathbf{X}_{\bar{i}}$ onto $\mathbb{V}_0$, it induces a homeomorphism between the marked tori $\delta\mathbb{L}_{\bar{j}} \to \delta\mathbb{V}_0$. Hence the function $\gamma_1(\lambda) = (\lambda, L_\lambda \circ \gamma(\lambda))$, $\Delta^1(\lambda) \to \mathbb{V}_0$, is also proper with winding number one. Thus, we obtain:

Lemma 47.8. *The first generalized renormalization*

$$(\mathbf{g}_1 : \bigcup \mathbb{V}^1_j \to \mathbb{V}^0 \equiv \mathbb{U}^1, \mathbf{h}_1)$$

is a proper family with winding number one over $\Delta^1(\lambda_\circ)$.

47.4.1. *Appendix: Extra parameter space.* Together with the tubes (47.6) let us also consider bigger tubes $\mathbb{W}_{\bar{i}}$ over $\mathcal{V}_{\mathrm{tr}}$ defined as follows. Let $P_{i_r,\lambda} = X^k_{j,\lambda}$ be the first "X-piece" in the itinerary $\{P_{i_m}\}^s_{m=0}$. Then

$$W_{\bar{i},\lambda} = \{z : G^m_\lambda z \in P_{i_m,\lambda},\ m = 0, \dots, r-1,\ \ G^r_\lambda(z) \in \Omega^k_{j,\lambda}\}, \tag{47.7}$$

where the domains $\Omega^k_{j,\lambda}$ are defined at the end of §47.2.1. Then

$$G^r_\lambda : W_{\bar{i},\lambda} \to \Omega^k_{j,\lambda}, \quad G^{s-r}_\lambda : \Omega^k_{j,\lambda} \to Y^{(1+(t-1)p)}_\lambda, \tag{47.8}$$

and both maps are univalent isomorphisms. Thus $\mathbf{G}^s : \mathbb{W}_{\bar{i}} \to \mathbb{Y}^{(1+(\mathbf{n}-1)p)}$ is a fiberwise conformal bundle diffeomorphism fibered over id.

Hence the holomorphic motion of $\mathbb{Y}^{(1+(\mathbf{n}-1)p)}$ (see Lemma 47.6) can be lifted to holomorphic motions of the $\mathbb{W}_{\bar{i}}$. Let $\mathbb{W}_\circ \equiv \mathbb{W}_{\bar{i}_\circ}$, where $\bar{i}_\circ$ is the itinerary of the critical value $\gamma(\lambda_\circ) = G_\circ(0)$ through the initial tiling. Let us introduce the following parameter domains in $\mathcal{V}_{\mathrm{tr}}$:

$$\Lambda^1(\lambda_\circ) = \boldsymbol{\gamma}^{-1}\mathbb{W}_\circ = \{\lambda : \gamma(\lambda) \in W_{\circ,\lambda}\} \supset \Delta^1(\lambda_\circ). \tag{47.9}$$

Thus for $\lambda \in \Lambda^1(\lambda_\circ)$, the critical value $\gamma(\lambda)$ has the same itinerary through the initial tiling as the critical value $\gamma(\lambda_\circ)$, except for the last moment s, when the former lands somewhere in $Y^{(1+(\mathbf{n}-1)p)}_\lambda$. This extension of $\Delta^1(\lambda_\circ)$ will be used for a priori bounds on the parameter geometry.

47.5. Renormalization of holomorphic families. Let us now consider a generalized quadratic-like family $(\mathbf{g} : \bigcup \mathbb{V}_i \to \mathbb{U}, \mathbf{h})$ over $(\Lambda, \lambda_\circ)$. Let $\mathcal{I}$ stand for the labeling set of tubes $\mathbb{V}_i$. Remember that $\mathcal{I} \ni 0$ and $\mathbb{V}_0 \ni \mathbf{0}$. Let $\mathcal{I}_\#$ stand for the set of all finite sequences $\bar{i} = (i_0, \dots, i_{t-1})$ of non-zero symbols $i_k \neq 0$. For any $\bar{i} \in \mathcal{I}_\#$, there is a tube $\mathbb{V}_{\bar{i}}$ such that

$$\mathbf{g}^k(\mathbb{V}_{\bar{i}}) \subset \mathbb{V}_{i_k},\ k = 0, \dots, t-1 \quad \textit{and} \quad \mathbf{g}^t\mathbb{V}_{\bar{i}} = \mathbb{U}.$$

We call $t = |\bar{i}|$ the *rank* of this tube. The map $\mathbf{g}^t : \mathbb{V}_{\bar{i}} \to \mathbb{U}$ is a holomorphic diffeomorphism which fibers over id, so $g^t_\lambda(V_{\bar{i},\lambda}) = U_\lambda,\ \lambda \in \Lambda$.

Let us lift the holomorphic motion $\mathbf{h}$ of $\mathbb{U}$ to a holomorphic motion $\hat{\mathbf{h}}$ of the $\mathbb{V}_{\bar{i}}$:

$$g^t_\lambda \circ \hat{h}_{\bar{i},\lambda}(z) = h_\lambda(g^t_\circ z), \quad z \in V_{\bar{i},\lambda_\circ}.$$

Note that by (47.5) it coincides with $\mathbf{h}$ on the $\partial\mathbb{V}_i$.

Let $\mathbf{X}_{\bar{i}} \subset \mathbb{V}_{\bar{i}}$ be such a tube that $\mathbf{g}^t(\mathbf{X}_{\bar{i}}) = \mathbb{V}_0$, where $t = |\bar{i}|$. The first landing map $\mathbf{L} : \bigcup \mathbf{X}_{\bar{i}} \to \mathbb{V}_0$ is defined as $\mathbf{L}|\, X_{\bar{i}} = \mathbf{g}^t$. It is a holomorphic diffeomorphism fibered over id. Extend the holomorphic motion $\hat{h}_\lambda$ to the tubes $\mathbf{X}_{\bar{i}}$ by pulling it back from $\mathbb{V}_0$ by $\mathbf{L}$. Then extend it by the λ-lemma to the whole tube $\mathbb{U}$ keeping it unchanged on the boundaries $\partial\mathbb{U}, \cup \bigcup \partial\mathbb{V}_{\bar{i}}$.

Let $\gamma(\lambda) = g_\lambda(0)$ and $\boldsymbol{\gamma}(\lambda) = (\lambda, \gamma(\lambda))$. For a base point $\lambda_\circ$, let $\bar{i}_\circ$ be the itinerary of the critical value $\gamma(\lambda_\circ)$ under iterates of $g_\circ$ through the domains $V_{i,\circ}$, until its first return to $V_{0,\circ}$. In other words, let $g_\circ(0) \in X_{\bar{i}_\circ} \equiv X_\circ$.

Let us now consider the following parameter region around $\lambda_\circ$:

$$\Lambda' \equiv \Lambda'(\lambda_\circ) = \boldsymbol{\gamma}^{-1}(\mathbf{X}_\circ).$$

For $\lambda \in \Lambda'$, the itinerary of the critical value under iterates of g_λ until the first return back to $V_{0;\,\lambda}$ is the same as for $g_\circ$ (that is, $\bar{i}_\circ$). Let us define new tubes $\mathbb{V}'_j \subset \mathbb{V}_0$ as the components of $(\mathbf{g}|\mathbb{V}_0)^{-1}(\mathbf{X}_{\bar{i}}|D')$. Let

$$\mathbf{g}' : \bigcup \mathbb{V}'_j \to \mathbb{V}_0 \,|\, \Lambda' \equiv \mathbb{U}' \tag{47.10}$$

be the first return map of the union of these tubes to $\mathbb{V}_0$.

For $\lambda \in \Lambda'$, the critical value $\boldsymbol{\gamma}(\lambda)$ does not intersect the boundaries of the the tubes $\mathbf{X}_{\bar{i}}$. Hence we can lift the holomorphic motion on $\mathbb{U} \smallsetminus \mathbf{X}_\circ$ to a holomorphic motion $\mathbf{h}'$ on $\mathbb{U}' \smallsetminus \mathbb{V}_0$ over Λ' and extend it by the λ-lemma to the whole tube $\mathbb{U}'$. Thus we obtain a generalized quadratic-like family $(\mathbf{g}', \mathbf{h}')$ over Λ' which will be called the *generalized renormalization* of the family $(\mathbf{g}, \mathbf{h})$ (with base point $\lambda_\circ$).

If $\mathbf{g}$ is a proper family then $\mathbf{g}'$ is clearly proper as well. Moreover, $w(\mathbf{g}') = 1$ if $w(\mathbf{g}) = 1$. Indeed, by the Argument Principle the curve $\boldsymbol{\gamma}|\, La'$ intersects once every leave of $\partial\mathbb{L}_\circ$. Hence it has winding number one about this tube. As the first landing map $\mathbf{L} : \mathbf{X}_\circ \to \mathbb{V}_0$ is a fiber bundles diffeomorphism, it preserves the winding number. Thus the new critical value $\boldsymbol{\gamma}' : \Lambda' \to \mathbb{U}'$, $\boldsymbol{\gamma}' = \mathbf{L} \circ \boldsymbol{\gamma}$, has also winding number one.

Let us summarize the above discussion:

LEMMA 47.9. *Let* $\mathbf{g} : \bigcup \mathbb{V}_i \to \mathbb{U}$ *be a generalized quadratic-like family over* $(\Lambda, \lambda_\circ)$*. Assume it is proper and has winding number one. Then its generalized renormalization* $\mathbf{g}' : \bigcup \mathbb{V}'_j \to \mathbb{U}'$ *over* Λ' *is also proper and has winding number one.*

47.6. Central cascades. In this section we will describe the renormalization of a generalized quadratic-like family through a central cascade, which will be then treated as a single step in the procedure of parameter subdivisions. Let us consider a holomorphic family $(\mathbf{g} : \bigcup \mathbb{V}_i \to \mathbb{U}, \mathbf{h})$ of generalized quadratic-like maps over $(\Delta, \lambda_\circ)$. We will now subdivide Δ according to the combinatorics of the central cascades of maps g_λ (see §31.8). To this end let us first stratify the parameter values according to the length of their central cascade. This yields a nest of parapuzzle pieces

$$\Delta \equiv D \supset D^{'} \supset \cdots \supset D^{(N)} \supset \ldots$$

For $\lambda \in D^{(N)}$, the map g_λ has a central cascade

$$V_\lambda^{(0)} \equiv U_\lambda \supset V_\lambda \equiv V_\lambda^{(1)} \supset \cdots \supset V_\lambda^{(N)} \tag{47.11}$$

of length N, so that $g_\lambda(0) \in V_\lambda^{(N-1)} \setminus V_\lambda^{(N)}$. Note that the puzzle pieces $V_\lambda^{(k)}$ are organized into the tubes $\mathbb{V}^{(k)}$ over $D^{(k-1)}$ (with the convention that $D^{(-1)} \equiv D$).

The intersection of these puzzle pieces, $\bigcap D^{(N)}$, is the little Mandelbrot set $M(\mathbf{g})$ centered at the superattracting parameter value $c = c(\mathbf{g})$ such that $g_c(0) = 0$. Let us call c the *center* of D.

Let $\lambda_\circ \in D^{(N-1)} \setminus D^{(N)}$. Let us consider the Bernoulli map

$$\mathbf{G} : \bigcup \mathbb{W}_j \to \mathbb{U} \tag{47.12}$$

associated with the cascade (47.11). Here the tubes $\mathbb{W}_j$ over $D^{(N-1)}$ are the pull-backs of the tubes $\mathbb{V}_i |\, D^{(N-1)}$, $i \neq 0$, by the covering maps

$$\mathbf{g}^k : (\mathbb{V}^{(k)} \setminus \mathbb{V}^{(k+1)}) |\, D^{(N-1)} \to (\mathbb{U} \setminus \mathbb{V}) | D^{(N-1)}, \quad k = 0, 1 \dots, N-1. \tag{47.13}$$

In the same way as in §47.5, to any string $\bar{j} = (j_0, \dots, j_{t-1})$ corresponds the tube over $D^{(N-1)}$,

$$\mathbb{W}_{\bar{j}} = \{p \in \mathbb{U} |\, D^{(N-1)} : \mathbf{G}^n p \in \mathbb{W}_{j_n},\ n = 0, \dots, t-1\}.$$

Note that $\mathbf{G}^t$ univalently maps each $\mathbb{W}_{\bar{j}}$ onto $\mathbb{U} |\, D^{(N-1)}$. Thus $\mathbb{W}_{\bar{j}}$ contains a tube $\mathbf{X}_{\bar{j}}$ which is univalently mapped by $\mathbf{G}^t$ onto the central tube $\mathbb{V}^{(N)}$. These maps altogether form the first landing map to $\mathbb{V}^{(N)}$,

$$\mathbf{L} : \bigcup \mathbf{X}_{\bar{j}} \to \mathbb{V}^{(N)}. \tag{47.14}$$

Remark. Note that

$$\operatorname{mod}(W_{\bar{j},\lambda} \setminus X_{\bar{j},\lambda}) = \operatorname{mod}(U_\lambda \setminus V_\lambda^{(N)}) \geq \operatorname{mod}(U_\lambda \setminus V_\lambda), \tag{47.15}$$

since G_λ^t univalently maps the annulus $W_{\bar{j},\lambda} \setminus X_{\bar{j},\lambda}$ onto $U_\lambda \setminus V_\lambda^{(N)}$.

Let us now consider the itinerary $\bar{j}_\circ$ of the critical value $\gamma(\lambda_\circ) \equiv g_\circ(0)$ through the tubes W_j until its first return to $V^{(N)}$, so that $\boldsymbol{\gamma}(\lambda_\circ) \in \mathbf{X}_{\bar{j}_\circ} \equiv \mathbf{X}_\circ$. Let $\mathbb{W}_\circ \equiv \mathbb{W}_{\bar{j}_\circ}$ and

$$\Delta^\diamond(\lambda_\circ) = \boldsymbol{\gamma}^{-1}\mathbf{X}_\circ, \quad \Lambda^\diamond(\lambda_\circ) = \boldsymbol{\gamma}^{-1}\mathbb{W}_\circ. \tag{47.16}$$

Thus, the annuli $D^{(N-1)} \setminus D^{(N)}$ are tiled by the parapuzzle pieces $\Delta^\diamond(\lambda)$ according as the itinerary of the critical point through the Bernoulli scheme (47.12) until the first return to $V_\lambda^{(N)}$. Altogether these tilings form the desired new subdivision of Δ. (Note however that the new tiles do not cover the whole domain Δ: the residual set consists of the Mandelbrot set $M(\mathbf{g})$ and of the parameter values $\lambda \in D^{(N-1)} \setminus D^{(N)}$ for which the critical orbit never returns back to $V_\lambda^{(N)}$.)

The affiliated quadratic-like family over $\Delta^\diamond(\lambda_\circ)$ is defined as the first return map to $V_\lambda^{(N)} \equiv U_\lambda^\diamond$. Its domain $\bigcup \mathbb{V}_i^\diamond$ is obtained by pulling back the tubes $\mathbf{X}_{\bar{j}}$ from (47.14) by the double branched covering $\mathbf{g} : \mathbb{V}^{(N)} \to \mathbb{V}^{(N-1)} | \Delta^\diamond(\lambda_\circ)$, and the return map itself is just $\mathbf{L} \circ \mathbf{g}$.

The affiliated holomorphic motion is also constructed naturally. Let us first lift the holomorphic motion $\mathbf{h}$ from the condensator $\mathbb{U} \setminus \mathbb{V}$ to the condensators $(\mathbb{V}^{(k)} \setminus \mathbb{V}^{(k+1)}) |\, D^{(N-1)}$ via the coverings (47.13). This provides us with a holomorphic motion of $(\mathbb{U} \setminus \mathbb{V}^{(N)}, \bigcup \mathbb{W}_j)$ over $D^{(N-1)}$. Extend it through $\mathbb{V}^{(N)}$ by the λ-lemma, lift it to the tubes $(\mathbb{W}_{\bar{j}}, \mathbf{X}_{\bar{j}})$ and then extended again by the λ-lemma to the whole domain $\mathbb{U}$ over D^{N-1}. Let us denote it by $\mathbf{H}$.

Lifting this motion via the fiberwise analytic double covering over $\Delta^\diamond(\lambda_\circ)$,

$$\mathbf{g} : (\mathbb{U}^\diamond \smallsetminus \mathbb{V}^\diamond, \bigcup_{i\neq 0} \mathbb{V}_i^\diamond) \to (\mathbb{V}^{(N-1)} \smallsetminus \mathbf{X}_\circ, \bigcup_{\bar{j}\neq\bar{j}_\circ} \mathbf{X}_{\bar{j}}),$$

we obtain the desired motion of $(\mathbb{U}^\diamond \smallsetminus \mathbb{V}^\diamond, \bigcup_{i\neq 0} \mathbb{V}_i^\diamond)$ over $\Delta^\diamond(\lambda_\circ)$. By the λ-lemma it extends through $\mathbb{V}_0^\diamond$.

47.7. Principal parapuzzle nest.

47.7.1. *Top renormalization level.* Let us now summarize the above discussion. First, we consider the tiling $\mathcal{D}^1$ of a truncated satellite dyadic wake $\mathcal{V}_{\mathrm{tr}}$ as described in §47.4. Each tile $\Delta \in \mathcal{D}^1$ comes together with a generalized quadratic-like family $(\mathbf{g}_\Delta, \mathbf{h}_\Delta)$ over Δ.

Now assume inductively that we have constructed the tiling $\mathcal{D}^l$ of level l. Then the tiling of the next level, $\mathcal{D}^{l+1}$ is obtained by partitioning each tile $\Delta \in \mathcal{D}^l$ by means of the cascade renormalization as described in §47.6.

Let $\Delta^l(\lambda)$ stand for the tile of $\mathcal{D}^l$ containing λ, while $\Delta^l(\lambda) \subset \Lambda^l(\lambda) \subset \Delta^{l-1}(\lambda)$ stand for the other tile defined in (47.16). Each tile $\Delta = \Delta^l(\lambda)$ contains a *central subtile* $\Pi^l(\lambda) = \boldsymbol{\gamma}_\Delta^{-1}\mathbb{V}_0$ corresponding to the central return of the critical point (here $\boldsymbol{\gamma}_\Delta(\lambda) = (\lambda, \gamma_\Delta(\lambda)))$. Note that $\Pi^l(\lambda)$ may or may not contain λ itself.

Let us then consider the sequence of renormalized families $(\mathbf{g}_{l,\lambda}, \mathbf{h}_{l,\lambda})$ over topological discs $\Delta^l(\lambda)$. We call the nest of topological discs $\Delta^0 \supset \Delta^1(\lambda) \supset \Delta^2(\lambda) \supset \dots$ (supplied with the corresponding families) *the principal parapuzzle nest* of λ. This nest is finite if and only if λ is renormalizable. Moreover, in this case the *last generalized renormalization of the nest is a full unfolded ql family that produces a primitive M-copy corresponding to the first DH renormalization of f*. If λ is not renormalizable, then $\bigcap \Delta^l(\lambda)$ is equal to the combinatorial class $\mathcal{C}(\lambda)$. It will be shown in vol III that in fact $\operatorname{diam} \Delta^l(\lambda) \to 0$ in this case, so *non-renormalizable parameters are rigid*.

Let $c_{l,\lambda} \in \Delta^l(\lambda)$ be the centers of the corresponding parapuzzle pieces. Let us call them the *principal superattracting approximations* to λ. If λ is not renormalizable, then $c_{l,\lambda} \to \lambda$ as $l \to \infty$, (since the $\Delta^l(\lambda)$ shrink).

The $\operatorname{mod}(\Delta^l(\lambda) \smallsetminus \Delta^{l+1}(\lambda))$ are called the *principal parameter moduli* for λ.

47.7.2. *Deeper renormalization levels.* The above discussion can be repeated for any M-copy $\mathcal{M}' \subset \mathcal{M}$. (corresponding to n-fold renormalization). It contains the main hyperbolic component Δ_0' bounded by the main cardioid $\mathcal{N}'$ of $\mathcal{M}'$. Attached to $\mathcal{N}'$, are wakes $\mathcal{W}'_{\mathfrak{p}/\mathfrak{q}}$ of $\mathcal{M}'$. The appropriate periodic ray portrait moves holomorphically (under the Böttcher motion) in this wake. so it can be further used to construct the parapuzzle that refines $\mathcal{W}'_{\mathfrak{p}/\mathfrak{q}}$. The corresponding parapuzzle pieces either shrink to little M-copies $\mathcal{M}_i$" corresponding to the $(n+1)$-fold renormalization (including the satellite copy attached at the bifurcation point $c'(\mathfrak{p}/\mathfrak{q})$), or to the combinatorial classes of maps that are exactly n times renormalizable.

If a parameter $c \in \mathcal{M}$ is infinitely renormalizable then its combinatorial class $\mathcal{C}(c)$ is the equal to the intersection of the infinite nest $\mathcal{M}^1 \supset \mathcal{M}^2 \supset \dots$ of little copies corresponding to various renormalizations of f_c. Thus, we ultimately obtain a nest of principal parapuzzle pieces of various renormalization levels shrinking to the combinatorial class of c.

In all cases, we obtain:

THEOREM 47.10. *For any $c \in \mathcal{M}$, there is a canonical nest of parapuzzle pieces $\Delta^{n,l}(c)$ shrinking to either the combinatorial class $\mathcal{C}(c)$ (if c is periodically repelling), or to the little M-copy $\mathcal{M}'(c)$ such that c belongs to the closure of the main component, $c \in \operatorname{cl} \Delta_0'$ (if c has a non-repelling cycle).*

47.7.3. *Parapuzzle for a ql family.* The above construction works well combinatorially, but it does not capture geometry of little M-copies $\mathcal{M}^n$ of deep level. To this end, we need *a priori* bounds for the sequence of ql families $(\mathbf{g}_n, \mathbf{h}_n)$ producing the little copies $\mathcal{M}^n$.

Indeed, due to the straightening, the above discussion can be adapted for any full unfolded ql family $(\mathbf{g}, \mathbf{h})$ instead of the polynomial quadratic family. All what is needed is to select a tubing of the family and to replace the Böttcher motion with *tubing motion*. The corresponding parapuzzle pieces essentially coincide with the combinatorial pieces discussed above (i.e., their intersections with the Mandelbrot set $\mathcal{M}(\mathbf{g})$ are the same).[17] So, they produce the same little M-copies and the combinatorial classes inside $\mathcal{M}(\mathbf{g})$.

In vol III we will show that the geometry of the *maximal primitive* M-copies in $\mathcal{M}(\mathbf{g})$ depends only on the geometry of the family $(\mathbf{g}, \mathbf{h})$. Hence, in the infinitely primitively renormalizable case with *a priori* bounds (for the families $(\mathbf{g}_n, \mathbf{h}_n)$), we obtain a uniform control on the geometry of all the M-copies on all renormalization levels.

47.8. MLC at primitive roots.

THEOREM 47.11. *The Mandelbrot set is perfectly rigid* (*and hence locally connected*) *at any neutral point. Moreover,*

- *There is a single access to any neutral irrational point;*
- *There are two accesses to any parabolic point except the cauliflower* (*which has only one access*).

PROOF. Due to Lemma 37.23, we only have to deal with the primitive roots $\mathfrak{r}$ of hyperbolic components. Let M be the little M-copy rooted at $\mathfrak{r}$. Assume first it is maximal. Then there is a nest of parapuzzle pieces shrinking to M. It follows that the body $\mathcal{B}_\mathfrak{r}$ of $\mathcal{M}$ (i.e., the branch of $\mathcal{M}$ at $\mathfrak{r}$ containing 0) is rigid at $\mathfrak{r}$. By Lemmas 9.19 (iii) and 9.20, $\mathcal{M}$ is rigid at $\mathfrak{r}$.

Moreover, since preperiodic parameter are well branched (Proposition 37.38), $\mathcal{M}$ is perfectly rigid at $\mathfrak{r}$ (see Exercise 9.15).

□

47.9. Combinatorial model for $\mathcal{M}$.

47.9.1. *Combinatorial classification of quadratic polynomials* (*prelimenary version*). Given a hyperbolic component Δ of $\operatorname{int} \mathcal{M}$, we call $\Delta \cup \{\mathfrak{r}_\Delta\}$ the *rooted* hyperbolic component (recall that $\mathfrak{r}_\Delta \in \partial\Delta$ stands for the root of Δ). We also let $\partial_{\mathrm{irr}}\Delta$ be the set of boundary parameters $c \in \partial\Delta$ with irrational rotation number.

Let us summarize our current understanding of the combinatorial classes in the quadratic family:

---

[17]However, the boundary of these parapuzzle pieces can cross the original big Mandelbrot set $\mathcal{M}$ in a wild way.

THEOREM 47.12. *Any combinatorial class $\mathcal{C}(c_\circ) \subset \mathcal{M}$ is of one of the following types:*

- *If $c_\circ$ is not periodically-repelling, then $\mathcal{C}(c_\circ)$ is equal to the rooted hyperbolic component $\Delta_\circ \cup \{\mathfrak{r}_\circ\}$ union $\partial_{\mathrm{irr}}\Delta_\circ$ (containing $c_\circ$).*
- *For postcritically non-recurrent $c_\circ$, we have rigidity: $\mathcal{C}(c_\circ) = \{c_\circ\}$.*
- *For $c_\circ$ periodically repelling, poscritically recurrent, and exactly m times renormalizable ($m \in \mathbb{N}$), the combinatorial class $\mathcal{C}(c_\circ)$ is equal to the interesection of the Principal Parapuzzle Nest*

$$\Upsilon^0(c_\circ) \supset \Upsilon^1(c_\circ) \supset \Upsilon^1(c_\circ) \supset \dots$$

of the M-copy $\mathcal{M}^{[m]}$ containing $c_\circ$.

- *For infinitely renormalizable $c_\circ$, the combinatorial class $\mathcal{C}(c_\circ)$ is equal to the intersection of the nest*

$$\mathcal{M} \equiv M^{[0]} \supset M^{[1]}(c_\circ) \supset M^{[2]}(c_\circ) \supset \dots$$

of the little M-copies containing $c_\circ$.

Moreover, in the two latter cases, $\mathcal{C}(c_\circ)$ is either a hull or a singleton.

47.9.2. *Minor lamination.* The *characteristic* (or *minor) lamination* is a geodesic lamination in $\mathbb{D}$ obtained by taking the closure of the set of all characteristic leaves of all parabolic quadratic polynomials except the cauliflower (or equivalently: of all superattracting polynomials except $f_0(z) = z^2$). This is a lamination indeed since these levels correspond to the cut-lines through all parabolic points in question.

The quotient of $\mathbb{C}$ modulo the characteristic lamination is called the *combinatorial model* for $\mathcal{M}$ (embedded into $\mathbb{C}$). The following result justifies this name:

THEOREM 47.13. *The following properties are equivalent:*

Shrinking*: All nests of parapuzzle pieces around Yoccoz parameters and all nests of little M-copies shrink;*

MLC*: The Mandelbrot set $\mathcal{M}$ is locally connected.*

Under these circumstances, $\mathcal{M}$ is ambiently homeomorphic to its combinatorial model.

PROOF. The shrinking property implies MLC by Corollary 9.9.

Vice versa, assume $\mathcal{M}$ is locally connected. If it is not homeomorphic to the model, then there is a non-singleton combinatorial class $\mathcal{C} \equiv \mathcal{C}(c_\circ)$. Such a class is a continuum. Since $\mathcal{M}$ is locally connected, all points of $\mathcal{C}$ are landing points for some rays, so there exist uncountably many rays accumulating into $\mathcal{C}$. This contradicts Lemma 38.3.

Let us prove the last assertion. Assuming $\mathcal{M}$ is locally connected, we want to approximate any Green cut-line L by a cut-line through a parabolic point. Since parabolic points are the only cut-points on the boundaries of hyperbolic components, we can assume that L cuts $\mathcal{M}$ at a periodically repelling point. Such a cut-line is approximated by vertical boundary cut-lines of puzzle pieces. The latter cut $\mathcal{M}$ through preperiodic parameters. In turn, they are approximated by the cut lines through parabolic parameters (Proposition 43.23).

□

In fact, the combinatorial model is useful even without MLC:

THEOREM 47.14. *There is a continuous surjection $\pi : \mathcal{M} \to \mathcal{M}_{\rm com}$ whose fibers are either singletons or combinatorial classes of periodically repelling points.*

47.9.3. *Landing of rays: summary.* Let us finish by summarizing our current knowledge on the behavior of the parameter rays:

THEOREM 47.15. *Let us consider a parameter ray $\mathcal{R}^\theta_{\rm par} \subset \mathbb{C} \smallsetminus \mathcal{M}$. Then we have the following possibilities:*

(i) *For $\theta \in \mathbb{Q}_{\rm odd}/\mathbb{Z}$, the ray $\mathcal{R}^\theta_{\rm par}$ lands at a parabolic point c_θ. Moreover, for $\theta \neq 0$, there are two rays landing at c_θ. (Of course, $\mathcal{R}^0$ is the only ray landing at the main cusp $c_0 = 1/4$.)*

(ii) *For $\theta \in \mathbb{Q}_{\rm ev}/\mathbb{Z}$, the ray $\mathcal{R}^\theta_{\rm par}$ lands at a critically preperiodic point c_θ. Moreover, there are finitely many rays landing at c_θ.*

(iii) *For $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$ with a non-polygonal lamination $\mathcal{L}_\theta$, the ray $\mathcal{R}^\theta_{\rm par}$ lands at an irrational point c_θ of the boundary of some hyperbolic component. $\mathcal{R}^\theta_{\rm par}$ is the only ray landing at c_θ, and the corresponding lamination is tuned rotational.*

(iv) *For $\theta \in (\mathbb{R} \smallsetminus \mathbb{Q})/\mathbb{Z}$ with a polygonal lamination $\mathcal{L}_\theta$, the ray $\mathcal{R}^\theta_{\rm par}$ accumulates into a combinatorial class $\mathcal{C}om_\theta$ of periodically repelling, but non-Misiurewicz, parameters. Moreover, there are at most two such rays, and the combinatorial model of the corresponding Julia sets, $\mathcal{L}_{\rm com}(f_c)$ for $c \in \mathcal{C}om_\theta$, is equal to $\mathcal{L}_\theta$.*

PROOF. Part (i) is the content of Lemma 37.7 and Theorems 37.8, 47.11 . Part (ii) is the content of Theorem 37.35.

By Theorem 32.10, any periodically repelling combinatorial class $\mathcal{C}om$ has a polygonal lamination $\mathcal{L}$, while by Lemma 38.3, $\mathcal{L} = \mathcal{L}_\theta$ for any ray $\mathcal{R}^\theta$ accumulating on $\mathcal{C}om$, and, for an irrational θ, there are at most two such rays.

On the other hand, by Proposition 37.17, any neutral irrational parameter $c_\circ$ is the landing point of a single ray $\mathcal{R}^\theta_{\rm par}$. Moreover, this ray is approximated near ∞ by rational rays $\mathcal{R}^{\theta_n}_{\rm par}$ (bounding approximating wakes) such that the corresponding dynamical rays $\mathcal{R}^{\theta_n}_c$, $c \in \mathcal{R}^{\theta_n}_{\rm par}$, are rotational. It follows that $\mathcal{R}^\theta_c$ is also rotational for $c \in \mathcal{R}^\theta_{\rm par}$. It generates a tuned rotational (and hence non-polygonal) lamination (see Problem 32.42).

The conclusions follow. □

47.10. Preview of the further theory. In the next volume we will prove the following result:

THEOREM 47.16. *The nest of parapuzzle pieces around any non-renormalizable parameter $c \in \mathfrak{Y}_0$ shrinks.*

This implies MLC at any parameter $c \in \mathfrak{Y}_0$. Refining it further and combining with Theorem 47.11, we will obtain:

THEOREM 47.17. *The Mandelbrot set is locally connected at any parameter $c \in \mathcal{M}$ which is not infinitely renormalizable.*

This will reduce the MLC Conjecture to the following problem:

CONJECTURE 47.18. *Any infinite nest of little M-copies*

$$\mathcal{M} \supset M^{[1]} \supset M^{[2]} \supset \dots$$

shrinks: $\operatorname{diam} M^{[n]} \to 0$.

The corresponding real conjecture is established:

Theorem 47.19. *Any infinite nest of real renormalization intervals*

$$\mathcal{M}_{\mathbb{R}} \supset M_{\mathbb{R}}^{[1]} \supset M_{\mathbb{R}}^{[2]} \supset \dots$$

shrinks: $\operatorname{diam} M_{\mathbb{R}}^{[n]} \to 0$.

This implies the Real Combinatorial Rigidity and Density of Real Hyperbolicity in the quadratic family (see §38.3).

In conclusion, let us mention the measure-theoretic counterpart of Theorem 47.17:

Theorem 47.20. *The set of at most finitely renormalizable parameters $c \in \mathcal{M}$ has zero area.*

This reduces the ABM Conjecture asserting that $\operatorname{area}(\partial M) = 0$ (see §38.2) to the collowing one:

Conjecture 47.21. *The set of infinitely renormalizable parameters $c \in \mathcal{M}$ has zero area.*

Let us give some evidence in favor of this conjecture. As computer experiments indicate, all little Mandlebrot copies M are not much distorted, giving a reason to believe that all of them are uniformly $K-$qc equivalent to the big set $\mathcal{M}$ (after some truncations at their roots, in the satellite case). This would immediately imply shrinking of the copies (Conjecture 47.18), but would also imply that the main hyperbolic component of each M represents a definite gap in ∂M. So, near infinitely renormalizable parameters, $\partial\mathcal{M}$ would have gaps in arbitrary small scales.

48. More of topological and combinatorial fun

File: fun.tex. Under construction

We will finish with some fun stuff which does necessarily not belong to the mainstream theme of this book. We will not be shy applying the deep theory developed so far, though all of the results below can be obtained by elementary means. The reader can further entertain himself by trying to develop an elementary (and more general) theory himself, or to consult other sources.

48.1. Sharkovsky scale. Here is a remarkable order on the set $\mathbb{Z}_+$ discovered by Sharkovsky:

$$\begin{gathered} 3 \succ 5 \succ 7 \succ \dots \\ \succ 2\cdot 3 \succ 2\cdot 5 \succ 2\cdot 7 \succ \dots \\ \succ 4\cdot 3 \succ 4\cdot 5 \succ 4\cdot 7 \succ \dots \\ \dots\dots\dots \\ \vdots \\ \vdots \\ \succ 4 \\ \succ 2 \\ \succ 1. \end{gathered}$$

Sharkovsky Theorem. If a continuous map $f : \mathbb{R} \to \mathbb{R}$ has a periodic point of (exact) period $p \in \mathbb{Z}_+$, then it has a periodic point of all periods that follow p in the Sharkovsky order. Moreover, any set $\mathbb{P}$ of periods with the property[18]

$$p \in \mathbb{P}, \ p \succ q \implies q \in \mathbb{P}$$

can be realized by some continuous map $f : \mathbb{R} \to \mathbb{R}$ (in fact, by a quadratic polynomial).

For instance, the final column of the Sharkovsky scale (with periods 2^n) is represented by quadratic polynomials f_c going through the period doubling bifurcations.

It is striking that factor 2 plays some special role. In fact, one can observe that *the scale is self-similar*: by chopping off the first and the last rows (altogether comprising all odd number) and dividing by 2, we get the original scale back. As we will see momentarily, this feature is directly related to the doubling renormalization.

We also see that period 3 is the strongest one: it forces all other periods to exist. It is famously expressed by the title of the Li-Yorke paper: "Period three implies chaos".

Exercise 48.1. *Prove that the airplane has periodic points of all periods. (Recall that the airplane is the real quadratic polynomial that has a superattracting cycle of period 3.)*

Proof for the quadratic family. Let us orient the parameter interval $[-2, 1/4]$ from $1/4$ to -2, and accordingly write $c \triangleright c'$ if c farther from $1/4$ than c' .

For $p \in \mathbb{Z}_+$, let $c_p \in [-2, 1/4]$ be the first moment when the periodic orbit of period p appears on the real line. We know that it happens through the saddle-node bifurcation, one orbit in a time. Moreover, by Proposition 37.29, once this happened, the orbit stays forever. It follows that the linear order of the points c_p induces the *forcing order* on the periods: If $c_p \triangleright c_q$ then period p implies period q, and this is sharp (at the moment of the first appearence of period q, period p is not present yet).

So, we only need to identify the order $\triangleright$ on the set $\mathcal{C} := \{c_p\}_{p\in\mathbb{Z}_+}$ with the order $\succ$. Let $\mathcal{C}^{[2]} := \{c_{2p}\}_{p\in\mathbb{Z}_+}$

Let us consider the nest of satellite renormalization windows $M^{[n]}_{\mathbb{R}} \equiv M^{[n]}_{\mathbb{R}}(c_F)$ around the Feigenbaum parameter c_F:

$$[-2, 1/4] \equiv \mathcal{M}_{\mathbb{R}} \equiv M^{[0]}_{\mathbb{R}} \supset M^{[1]}_{\mathbb{R}} \supset M^{[2]}_{\mathbb{R}} \supset \cdots \ni \{c_F\},$$

As we know, the doubling renormalization yields a straightening homeomorphism

$$\chi : M^{[1]}_{\mathbb{R}} \to \mathcal{M}_{\mathbb{R}},$$

giving a natural one-to-one correspondence between periodic points for $c \in M^{[1]}_{\mathbb{R}}$ (except for its β-fixed point) and those for $c' := \chi(c)$. Moreover, under this correspondence the periods are divided by 2 (except for the α-fixed point for f_c). It follows that

$$\mathbb{P}(\chi(c)) = \frac{\mathbb{P}(c) \setminus \{1\}}{2}, \quad c \in M^{[1]}_{\mathbb{R}}, \tag{48.1}$$

where $\mathbb{P}(c)$ is the set of periods for f_c. It follows that:

[18] We will say that such a set represents a "Dedekind section" in the Sharkovskii scale.

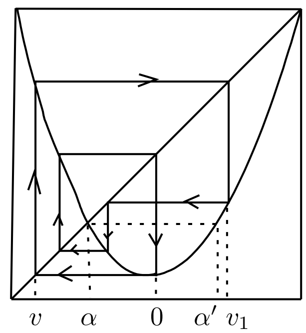

FIGURE 48.1. Superattracting parameter c_5° without period 3.

- $\chi(c_{2p}) = c_p$ for $c_{2p} \in M_{\mathbb{R}}^{[1]}$;
- $\mathcal{C}^{[2]} = \mathcal{C} \cap M_{\mathbb{R}}^{[1]}$;
- $\chi : \mathcal{C}^{[2]} \to \mathcal{C}$ is a bijection.

This reduces the problem to recognizing the ordering of the odd points $c_{2k+1} \in \mathcal{M} \setminus M_{\mathbb{R}}^{[1]}$, $k \geq 1$, as the rest of the Sharkovsky scale is obtained by pulling $\mathcal{C}^{\text{odd}} := \{c_{2k+1}\}_{k\geq 1}$ by the iterated straightenings

$$\chi^n : M_{\mathbb{R}}^{[n]} \setminus M_{\mathbb{R}}^{[n+1]} \to \mathcal{M}_R \setminus M_{\mathbb{R}}^{[1]}$$

To verify the Sharkovsky order for odd periods, we only need to produce for any odd $q > 1$, one parameter c such that the map f_c has an orbit of period q but misses all smaller odd periods $p \in [3, q-2]$. This will be accomplished by a certain sequence c_{2k+1}° of superattracting parameters beginning with the airplane and converging from the left to the Chebyshev tip of the satellite copy $M_{\mathbb{R}}^{[1]}$. Namely, at c_{2k+1}°, the third iterate of 0 lands in $(\alpha, 0)$, and then spins away from α, landing at 0 after $k-1$ "revolutions" (see Figure 48.1), i.e.,

$$0_2 > \alpha', \quad 0_{2m+1} \in (\alpha, 0), \ m = 1, \dots, k-1, \quad 0_{2k+1} = 0.$$

EXERCISE 48.2. (i) *Parameters c_{2k+1}° with this combinatorics exist.*

(ii) *The map $f_k \equiv f_{c_k}$ has no periodic points of odd periods $p \in [3, 2k-1]$.*

(iii) *It has only two periodic orbits of period $2k+1$, the superattracting one and the root orbit of its immediate basin.*

(iv) *The first parameter c_{2k+1} when period $2k+1$ appears is the root of the hyperbolic window centered at c_{2k+1}°.*

This completes a proof that the forcing order on the periods coincides with the Sharkovsky order, and shows that it is sharp in the weak sense: for any period p there is a map f_c that has a periodic orbit of period p but does not have any periodic orbits of periods $q \prec p$. We leave to the reader the following exercise that completes the proof:

EXERCISE 48.3. *Check that any Dedekind section $\mathbb{P}$ in the Sharkovsky scale is realizable by some quadratic polynomial. Find the first and the last realizing parameter.*

QED

Since the quadratic family is full, we immediately obtain a more general result:

COROLLARY 48.4. *Sharkovsky Theorem is valid for arbitrary unimodal maps.*

PROJECT 48.5. *Translate the above proof into purely combinatorial one by considering the kneading model for the quadratic family* (see §33.7). *In this way, deep analytic results about the quadratic family become irrelevant.*

PROJECT 48.6. *Study the Sharkovsky Theorem in full generality.* (See references in the Notes below.)

48.2. Entropy and the route to chaos. Let $M(f)$ be the number of monotonicity intervals of a piecewise monotone interval map.

THEOREM 48.7. *For a real quadratic polynomial f, we have:*

$$h(f) = \lim_{n\to\infty} \frac{1}{n} \log M(f^n) = \lim_{n\to\infty} \frac{1}{n} \log |\mathrm{Per}_n(f)| \tag{48.2}$$

PROOF. Let us denote the above limits μ and π repsectively, and let $h \equiv h(f)$. It is easy to see, using the kneading coding, that $h \leq \min\{\mu, \pi\}$. To prove the opposite inequalities, $\mu \leq \pi \leq h$, we can use approximating hyperbolic sets from §46.2. □

EXERCISE 48.8. *Show that $h(f_{\text{Ч}}) = \log 2$ for the Chebyshev map.*

THEOREM 48.9. *A real quadratic polynomial has zero entropy iff it has only cycles of periods 2^n (so, it belongs to the final segment of the Sharkovsky scale). In the real quadratic family, this corresponds is the interval $[c_F, 1/4]$, where c_F is the Feigenbaum parameter.*

PROOF. For $c \in (c_F, 1/4]$, $\mathrm{Per}(f_c)$ is finite (consisting of two fixed points (coinciding for $c = 1/4$) and cycles of periods 2^n, one for each $n = 2, \dots, N$), so $h(f) = 0$.

For $c = c_F$, $\mathrm{Per}(f_c)$ comprises all of the above periodic points, so $|\mathrm{Per}(f_c^m)| = O(m)$, and $h(f_c) = 0$.

EXERCISE 48.10. *Show that $h(f) = 0$ for the Feigenbaum map in a different way, using its dynamical structure.*

On the left, c_F can be approximated by the tips $\mathfrak{t}_n$ of the satellite renormalization windows $M_{\mathbb{R}}^{[n]}$. At such a parameter, the renormalizatoin $f^{2^n} \,|\, \mathcal{I}^n$ is a Chebyshev map, so it has entropy $\log 2$. It follows that $h(f) \geq \dfrac{\log 2}{2^n}$. By monotonicity, $h(f) > 0$ for all $c \in [-2, c_F)$. □

Since positive entropy is associated with "chaos", the cascade of doubling bifurcations is viewed to be a *route of chaos.*

Since the quadratic family is full, we immediately conclude:

COROLLARY 48.11. *For arbitrary unimodal maps:*

(i) *Entropy formula* (48.2) *is valid;*

(ii) *$h(f) = 0$ iff f has only cycles of periods 2^n.*

48.3. Measure with constant slope and the saw-like model. Let us consider the most interesting particular case:

THEOREM 48.12. *Let $f \in \mathfrak{Q}_{\mathbb{R}}$ be a map of Yoccoz class. Then:*

(i) *It has a quasi-invariant measure λ with constant Jacobian $r > 1$ supported on the invariant exact interval J.*

(ii) *The restriction $f \,|\, J$ is topologically conjugate to a saw-like map with constant slope r.*

PROOF. Let us approximate f by a sequence f_n of once renormalizable superattracting maps f_n: for instance, take the canonical approximands.[19] By Exercise 25.55, they have the quasi-invariant measures λ_n with constant Jacobians $r_n > 1$. Moreover, by the monotonicity of the entropy (Theorem 37.33), the r_n increase. Passing to a limit λ, we obtain a desired quasi-invariant measure λ with a constant slope $r > 1$.

(ii) Exactness easily implies that $\operatorname{supp} \lambda = J$. Hence there is is a homeomorphism $h : J \to \tilde{J}$ such that the push-forward $h_* \mu$ is Lebesgue. Then the map $h \circ f \circ \tilde{h}^{-1}$ has the constant slope r. □

REMARK 48.13. The above proof can be adjusted to become purely combinatorial, without using deep analytic machinery. One can also do it directly for f, without approcimating it with superattracting or cricically preperiodic maps.

48.4. Appendix: Topological entropy.

48.4.1. *Shifts.* The "topological entropy" of the full shift $\sigma : \Sigma_d^+ \to \Sigma_d^+$ is set to be $\log d$. Note that it is the growth rate for the number of cylinders of rank n (which is equal to $d^n = e^{n \log d}$).

More generally, for a Markov shift $\sigma_A : \Sigma_A^+ \to \Sigma_A^+$, the number of *admissible cylinders* of rank n is comparable with $r(A)^n = e^{n \log r(A)}$, where $r(A)$ is the spectral radius of the transit matrix A (see §19.14.1). This makes the topological entropy $h(\sigma_A)$ equal to $\log r(A)$.

Even more generally, let $X \subset \Sigma_d^+$ be a closed shift-invariant subset, and let $\sigma_X : X \to X$ be the restriction of the full shift to X. In this case, the topological entropy can be defined as the growth rate of the number of admissible sequences of length n. Namely, a string $\bar{i} = (i_0 \dots i_{n-1})$ is called *admissible* if it can be extended to an infinite string $(i_0 \dots i_{n-1} \dots) \in X$ (in other words: $\Sigma_{\bar{i}}^n \cap X \neq \emptyset$). Let N_n be the number of admissible strings of length n. Then the topological entropy of σ_X is defined as

$$h(\sigma_X) := \limsup_{n \to \infty} \frac{1}{n} \log N_n.$$

Let us now pass to a general definition.

48.4.2. *General notion.* Let $f : X \to X$ be a continuous map on a (metrizable) compact space, with a metric d. For $\varepsilon > 0$, a set of points x_i is called an (d, ε)-*net* if any point $x \in X$ can be ε-approximated by some x_i. Let $N_d(\varepsilon)$ be smallest number of points in all (d, ε)-nets.

Let us now consider metrics

$$d_n(x, y) := \max_{0 \le k \le n-1} d(f^k x, f^k y),$$

[19] We can also use the canonical critically preperiodic approximands instead: see Exercise 27.13.

and let $N_n(\varepsilon) \equiv N_{d_n}(\varepsilon)$ be the corresponding smallest number of points in (d_n, ε)-nets, i.e., the smallest number of orbits of length m such that any orbit of length n can be ε-*shadowed* by one of those. Let

$$h(f, \varepsilon) := \limsup_{n\to\infty} \frac{1}{n} \log N_n(\varepsilon).$$

Obviously, this function is monotonically increasing as $\varepsilon \searrow 0$, so we can take the limit:

$$h(f) := \lim_{\varepsilon\to 0} h(f, \varepsilon).$$

It is called the *topological entropy* of f.

Exercise 48.14. *Show that:*
(i) *$h(f)$ is independent of the choice of the metric d.*
(ii) *$h(f$ is a topological invariant, i.e., $h(f) = h(g)$ if f and g are topologically conjugate.*
(iii) *More generally, If $f : X \to X$ is semi-conjugate to $g : Y \to Y$ then $h(g) \le h(f)$.*
(iv) *$h(f^n) = n\, h(f)$.*

Exercise 48.15. *If f is an isometry then $h(f) = 0$. In particular, $h(f) = 0$ for a circle rotation or for an adding machine.*

A point $x \in X$ is called *wandering* if it has a weakly wandering neighborhood. So, a point x is *non-wandering* if any neighborhood $U \ni x$ contains a returning point $y \in U$, i.e., $f^n y \in U$ for some $n \in \mathbb{Z}_+$. The set of non-wandering points is denoted by $\Omega \equiv \Omega(f)$. It is a closed invariant subset of X that captures chaos:

Exercise 48.16. *$h(f) = h(f\,|\,\Omega)$.*

The *Minkowski dimension* of X (with respect to the metric d) is defined as

$$\mathrm{MD}(X) := \limsup_{\varepsilon\to 0} \frac{\log N_d(\varepsilon)}{\log \varepsilon^{-1}},$$

so it is the sup of $\delta > 0$ such that $N_d(\varepsilon) = O(\varepsilon^{-\delta})$.

It can happen that $h(f) = \infty$ (give an example!). However, we have:

Kushnirenko Bounds. (i) If f is L-Lipschitz then $h(f) \le \mathrm{MD}(X) \cdot \log L$.
(ii) In particular, if f is a smooth endomorphism of a compact manifold X endowed with a Riemannian metric d, then

$$h(f) \le \mathrm{MD}(X) \cdot \|Df\|_\infty.$$

Exercise 48.17. (i) *Prove the Kushnirenko bounds.*
(ii) *Derive the following asymptotic versions of the bounds*

$$h(f) \le \mathrm{MD}(X) \cdot \lim_{n\to\infty} \frac{1}{n} \log L(f^n); \quad h(f) \le \mathrm{MD}(X) \cdot \lim_{n\to\infty} \frac{1}{n} \log \|Df^n\|_\infty.$$

(including the existence of the limits).

Instead of ε-nets, one can use (d, ε)-*separated* sets, i.e., sets of points x_i such that $d(x_i, x_j) > \varepsilon$ for any two of them Let $S(d, \varepsilon)$ be the maximal number of ε-separated points, and let $S_n(\varepsilon) \equiv S(d_n, \varepsilon)$.

Exercise 48.18. *Show that $h(f) = \lim_{\varepsilon\to 0} \limsup \dfrac{1}{n} \log S_n(\varepsilon)$.*

48.4.3. *Entropy and expansivity.* Recall notion of *expansivity* from §19.9.1.

EXERCISE 48.19. *If f is ε-expansive then:*
(i) $h(f) = h(f, \varepsilon)$;
(ii) $h(f) \geq \limsup_{n\to\infty} \frac{1}{n} \log |\mathrm{Fix}(f^n)|$;
(iii) $h(f) \geq \limsup_{n\to\infty} \frac{1}{n} \log |f^{-n}(x)|$ *for any* $x \in X$.

EXERCISE 48.20. *Let $\sigma_A : \Sigma_{\mathbb{A}} \to \Sigma_A$ be an irreducible topological Markov chain. Then:*
(i) *It is* 1*-expansive in the d-adic metric (generalizing the dyadic metric* (19.5)*;*
(ii) $h(f) = \lim_{n\to\infty} \frac{1}{n} \log |\mathrm{Fix}(f^n)| = \lim_{n\to\infty} \frac{1}{n} \log |\sigma_A^{-n}(x)| = \log r(A)$, *where* $r(A)$ *is the spectral radius of A, x is an arbitrary point of Σ_A.*
(iii) *For an arbitrary subshift σ_X, justify the entropy formula given in §48.4.1.*

48.4.4. *Variational Principle.* By Theorem 46.47 and the Exercise that follows, for any irreducible Markov shift we have:

$$h(\sigma_A) = \max_{\mu \in \mathfrak{M}(\Sigma_A)} h_\mu(\sigma_A),$$

where the maximum is attained on the balanced measure. This property is referred to as a "Variational Principle". Part of it holds under much more general circumstances:

VARIATIONAL PRINCIPLE. For any continuous endomorphism $f : X \to X$ of a compact space,

$$h(f) = \sup_{\mu \in \mathfrak{M}(X)} h_\mu(f).$$

If the sup is attained, then the corresponding measure is naturally called a *measure of maxumal entropy.* (In general, it does not have to exist.) One important particular assertion is easy to verify:

EXERCISE 48.21. *Let us consider any subshift σ_X, $X \subset \Sigma_d^+$. Then for any invariant measure μ, we have: $h_\mu(\sigma_X) \leq h(\sigma_X)$.*

We may return to this theme in one of the forthcoming volumes.

48.4.5. *T.* opological entropy first appeared in print in Adler, Konheim and McAndrew [**AKMcA**]. For a discussion of the Variational Principle, see e.g., [**Bow**].

Notes. Sharkovsky order appeared in [**Sha1**] (1964). One decade later it was partly re-discovered by Li and Yorke in a paper entitled "Period three implies chaos" [**LiY**] that introduced a term "chaos" to dynamics. For more recent approaches and insights, see [**Ste, B3, BMi**].

Entropy formula for interval maps (Theorem 48.7) was proven by Misiurewicz and Szhlenk [**MiS**].

Acknowledgment. The author thanks Sasha Blokh for many useful comments.

Notes

Theorems 46.7 on ergodicity and 46.8 on measure-theoretic attractors are due to Blokh and the author [**BL1, BL2**]. The theory was further refined by Martens [**Ma**]. A different approach was developed by Guckenheimer and Johnson [**GJ**].

Real *a priori* bounds of Theorem 46.3 are due to Martens [**Ma**]. Here we derive them (for quadratic-like maps) as a consequence of complex bounds. Originally, they were proved (for an appropriate class of smooth maps) by purely real methods.

Stochasticity of postcritically non-recurrent maps (Theorem 46.20) was proven by Misiurewicz [**Mi2**] (which prompted the term "Misiurewicz maps"). The corresponding set of parameters $c \in [-2, 1/4]$ has zero measure.

More general conditions for stochasticity appeared in [**CE, NS, MN**]. Jakobson proved in the late 1970s [**Ja1**] that the set of stochastic parameters $c \in [-2, 1/4]$ has positive measure (see also Benedicks and Carleson [**BC**]). It was proven by the author in the 1990s [**L11**] that almost all real Yoccoz parameters are stochastic. (satisfying the Martens-Nowicki stochasticity condition). This result was further refined by Avila and Moreira [**AM1**]. All these issues will be discussed in vol. III

The first example of a real non-stochastic exact map was constructed by Johnson [**Jo**]. Further pathological examples are due to Hofbauer and Keller [**HoK**].

The Principal Parapuzzle Nest (§47) was studied in [**L11**]. Its geometric applications will be discussed in vol. III. It provides us with one more illustration of the Phase-Parameter Relation.

Help Center and Reference guide

49. Hints and comments to the exercises

Preliminaries: Topological background

Point set topology

1.8. Extend the Devil Staircase from $\mathbb{T}$ to $\bar{\mathbb{D}}$ preserving co-centric circles $\mathbb{T}_r$ so that the intervals $r \cdot I_j \subset \mathbb{T}_r$ are shrunk by factor $(1-r)$.

Local connectivity

1.11. Use Exercise 1.10.

1.13. a) Fix some $t \in [0,1]$, and let $\gamma(t) = x$. Consider the decomposition of $\gamma^{-1}(B(x,\varepsilon))$ into connected components I_n and J_k , where $I_n \ni t$ while $J_k \not\ni t$. Show that the paths $\gamma(J_n)$ do not accumulate on x and conclude that γ is weakly lc at x.

1.14. Construct a sequence of polygonal curves γ_n in $\mathbb{R}^n$ connecting x to y such that:

- The vertices of the γ_n belong to K;
- γ_{n+1} is a *refinement* of γ_n, i.e., the vertices of γ_n are also vertices of γ_{n+1};
- $\|\gamma_n - \gamma_{n+1}\| \le 1/2^n$, where $\|\cdot\|$ stands for the uniform norm.

Remark 49.1. In fact, this is true without assuming that K is embedded into $\mathbb{R}^n$. Indeed, any compact metric space X embeds into a Banach space (for instance, by associating to $x \in X$ the distance function $y \mapsto d(x,y)$), where one can repeat the above argument.

1.15. One direction: arc lc is stronger than weak lc. The other (non-trivial) direction: use the argument for Exercise 1.14.

1.16. Use that J is path lc and show that K is such.

1.18. Have fun!

1.27. (i) A properly embedded star of valence one divides the plane into two half-planes (by the Jordan-Schönflies Theorem). Inductively, a properly embedded star of valence $\mathfrak{q}$ divides the plane into d *sectors*. Map them consecutively onto the standard sectors so that the maps match on the common boundary.

(ii) The natural cyclic order on the set of angles $\theta_k := k/\mathfrak{q}$, $k \in \mathbb{Z}/\mathfrak{q}\mathbb{Z}$, can be also defined as follows: θ_{k+1} in the *next* after θ_k if the corresponding rays $\mathcal{R}_k$ and $\mathcal{R}_{k+1}$ bound a complementary sector (call if S_k), and $\mathcal{R}_k$ is positively oriented with respect to S_k. Since h maps complimentary sectors to complimentary sectors (preserving orientation), it preserves this cyclic order.

FIGURE 49.1. Dyadic tree of pairs of pants.

1.33. Otherwise there is a sequence of arcs $\gamma_n \subset U_{i_n}$ whose diameter is bounded away from 0. Take an accumulation point a for the "mid-points" a_n of the γ_n. Then K is not lc at a.

1.34.

1.35. From any point $z \in D$ one can reach the closest to it boundary point $\zeta \in \partial D$ along the straight interval $[z, \zeta]$.

Group actions and foliations

1.50. (i) See Figure ... Note that the two foliations are mutually orthogonal.

1.51. The vertical circle $\{0\} \times \mathbb{R}/\mathbb{Z}$ is a global transversal to the foliation. Its monodromy map is the rotation by α. Apply Exercise 19.31.

Coverings

1.52. Take a path γ in E connecting e to a point x, and consider its image $h \circ p(\gamma)$ in B'. It lifts to a path γ' in E' that begins at e' and ends at some point x'. Let $x' = H(x)$. The π_1-assumption assures that this definition is independent of the choice of γ.

1.56. (i) These coverings correspond to the subgroups $\Gamma_d = d \cdot \mathbb{Z}$, $d \in \mathbb{Z}_+ \cup \{0\}$. (ii) There are only two automorphisms of $\mathbb{Z}$, and both keep each Γ_d invariant. The Lifting Criterion applies.

1.58. Consider the covering corresponding to the Ker of the monodromy action.

Topological surfaces

1.66. Extend h to a homeomorphism $H : \bar{\mathbb{D}} \to \bar{\mathbb{D}}$ (e.g., radially) and let $\bar{\mathbb{D}} \sqcup_h \bar{\mathbb{D}} \to S^2$ be equal to H on the first disk and be equal to id on the second.

1.67. (i) Connect x_i to y_i with disjoint arcs, and construct a homeomorphism supported in small neighborhoods of these arcs. (ii) Construct a homeomorphism supported in small annuli neighborhoods of the boundary circles.

1.73. Use Lemma 1.71 many times. For instance, to check pre-compactness, follow the proof of the lemma to construct an escaping nest of fjords $F_1 \supset F_2 \supset \dots$ each of which contains infinitely many points of a given escaping sequence (x_n).

To check total disconnectedness, find a simplicial compact set K such that two given ends, E_1 and E_2, belong to different unbounded components of $S \smallsetminus K$. Then approximate every unbounded component of $S \smallsetminus K$ with a fjord.

1.99. Represent $\mathbb{D} \smallsetminus K$ as the dyadic tree of pairs of pants: see Figure 49.

1.100. Compare with the circle case: Exercise 1.56. Note that this relation is natural as Cyl and S^1 are homotopically equivalent.

1.106. Otherwise $f : U \smallsetminus f^{-1}(b) \to V \smallsetminus \{b\}$ would be unbranched. Since $U \smallsetminus \{b\}$ is a topological annulus, $U \smallsetminus f^{-1}(b)$ would also be a topological annulus. But then $f^{-1}(b) = \{a\}$ and $\deg_a f = \deg f$.

Appendix 1: Hausdorff metric

1.120 (iii) Approximate the X_n with finite ε_k-nets with $\varepsilon_k \to 0$, and use the diagonal procedure.

Chapter 1. Conformal geometry

Riemann surfaces

2.5. (iii) The semi-direct product of the group of translations with the cyclic group of order 4 and 6, respectively.

(iv) See [**M6**], §5.

2.6. (i) a) The infinite diahedral group (obtained by adding the involution $\sigma : z \mapsto -z$ to the lattice $\mathbb{Z}$) produces the orbifold with signature $(\mathbb{C}, \{2, 2\})$. The corresponding covering map (if appropriately normalized) is cos.
b) The infinite diahedral groups of rank 2 (obtained by adding σ to a lattices $\approx \mathbb{Z}^2$) produce orbifolds with signature $(\mathbb{T}^2, \{2, 2, 2, 2\})$. The corresponding covering maps are Weierstrass $\mathcal{P}$-functions (see §2.10.2).
c) Remaining special orbifolds are obtained by takling quotients of the tori with extra symmetries (corresponding to the tilings of $\mathbb{C}$ by squares or equilateral triangles) by the corresponding symmetry groups (of order 4, 3 and 6). The signatures are $(\mathbb{T}^2; \{2, 4, 4\})$, $(\mathbb{T}^2; \{3, 3, 3\})$, and $(\mathbb{T}^2; \{4, 3, 6\})$.

(ii) All these orbifolds are covered by the cylinder or the torus.

2.9. Use the normal form.

2.10. The space of ε-separated triples of points is compact. The Möbius transformation ϕ depends continuously on the triple $(\alpha, \beta, \gamma) = \phi^{-1}(0, 1, \infty)$ as obvious from the explicit formula

$$\phi(z) = \frac{z - \alpha}{z - \gamma} \cdot \frac{\beta - \gamma}{\beta - \alpha}.$$

(This can also be used to verify equivalence of the two topologies.)

2.13. The signatures of the Platonic orbifods are: $(S^2; \{2, 3, 3\})$ (tetrahedron), $(S^2; \{2, 3, 4\})$ (cube and octahedron), and $(S^2; \{2, 3, 5\})$ (dodecahedron and icosihedron). The orders $\mathfrak{q}_i$ are halves of the valences of the vertices of the corresponding triangulations.

First formula (2.3) follows from the transformation rule for the Euler characteristic under orbifold coverings (see Exercise 1.117). Then the relation $\text{area}(\mathcal{O}) = 2\pi\, \chi(\mathcal{O})$ becomes straighforward. Note that it is a particular case of a general Gauss-Bonnet Formula. It is also a cosequence of the general formula for the area of spherical triangles (Exercise 2.14).

2.14. See [**Sch**, §9.3] for a nice elementary way to verify the formula.

2.17. The curvature of a metric $\rho(z)|dz|$ can be calculated by the formula:

$$\kappa(z) = -\frac{\Delta \log \rho(z)}{\rho(z)^2}.$$

$\mathrm{PSL}(2, \mathbb{R})$-invariance of the hyperbolic metric in the $\mathbb{H}$-model amounts to the identity:

$$\operatorname{Im} \phi(z) = \frac{\operatorname{Im} z}{|cz + d|^2}, \quad \phi(z) = \frac{az + b}{cz + d}.$$

Smooth isometries preserve angles between tangent vectors, and so conformal. In fact, one does not need to impose smoothness *a priori*. Any isometry is quasiconformal (e.g., by Proposition 12.14), and hence conformal by Weyl's Lemma (§13.1).

2.19. The hyperbolic metric is the only invariant metric on $\mathbb{H}$ coinsiding with the Euclidean metric on the tangent plane at i.

2.22. (ii) Hyperbolic R-neighborhoods of a geodesic γ are invariant under $\mathrm{Stab}(\gamma)$.

2.23. (i) ϕ is the composition of a Möbius map and $\sqrt{\cdot}$

(ii) True by the $\mathbb{R}$-symmetry.

(iii) Use (i) and Exercise 2.22.

2.34. Bring two parabolic elements to a normal form

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ \mu & 1 \end{pmatrix}$$

and calculate $\mathrm{tr}\{A, B\} = 2 + \mu^2$.

In the elliptic case, bring A to the diagonal form with multipliers $e(\pm\theta/2)$ and calculate $\mathrm{tr}\{A, B\} = 2(ad + bc\cos\theta)$.

2.38. The universal covering is $\mathbb{H} \to \mathbb{D}^*$, $z \mapsto e(z)$.

2.39. The coverings over $\mathbb{D}^*$ are classified by the subgroups of the cyclic group generated by the translation $z \mapsto z + 1$ of $\mathbb{H}$.

2.40. (i) See Figure 49. $R = \exp(2\pi^2/\log\rho)$. (ii) The equator is the image of the hyperbolic geodesic $i \circ \mathbb{R}_+$; and its hyperbolic length is equal to

$$\log\rho = \frac{2\pi^2}{\log R}.$$

(iii) This is manifestly so in the $\mathbb{T}$-symmetric model $\mathbb{A}(1/\sqrt{R}, \sqrt{R})$ for the annulus.

2.41. It can be done in two ways: a) by means of the Schwarz Reflection Principle or b) by using the covering by an elementary Fuchsian group.

2.46. They correspond to degree two branched coverings $\mathbb{A}(1/R, R) \to \mathbb{D}$ with the covering group $\mathbb{Z}/2\mathbb{Z}$ generated by the involution $z \mapsto 1/z$. So, it has signature $(\mathbb{D}; 2, 2)$.

2.47. (i),(ii and (v)). Consider the quadrilateral $\Delta \cup \delta(\Delta)$, with two ideal vertices and two vertices with angle $2\pi/3$. Rotating it around the latter by the $\omega_\pm^{\pm 1}$, we obtain a 10-gone with two ideal vertices and 8 vertices with angle $2\pi/3$. It is triangulated by translates of Δ. Rotating these translates around the vertices (by appropriate conjugates of $\omega^{\pm 1}$) we construct a converx polygon with ideal vertices and vertices with angle $2\pi/3$. Proceeding this way, we will tessellate the whole hyperbolic plane $\mathbb{H}$ by translates of Δ.

To show that the translates of Δ cover the whole $\mathbb{H}^+$, take any point $z \in \Pi^+$. If $z \notin \mathbb{D}$, then do nothing. Otherwise apply δ to it and then apply γ^n to bring it back to Π^+. We obtain a point $z_1 := \gamma^n(\delta z) \in \Pi^+$ such that $\mathrm{Im}\, z_1 > \mathrm{Im}\, z$. If $z_1 \in \Delta$ then stop. Otherwise repeat the procedure. We obtain a sequence of points $z_n \in \Pi^+$ with increasing $\mathrm{Im}\, z_n$. By discreteness, only finitely many of them can belong to $\mathbb{D}$.

See [**Akh, Iw**] for a background in modular functions.

2.49. See Figure 49.3.

2.51. See [**Be2**], [**Har**, §V.B].

2.54. Use the orthogonal projection to the corresponding geodesic or horocycle as indicated on Figure 2.7.

2.79. They correspond to the tori with extra symmetries (see Exercise 2.4 (iii)). Notice that the order of the fixed points (2 and 3) are two times smaller than the

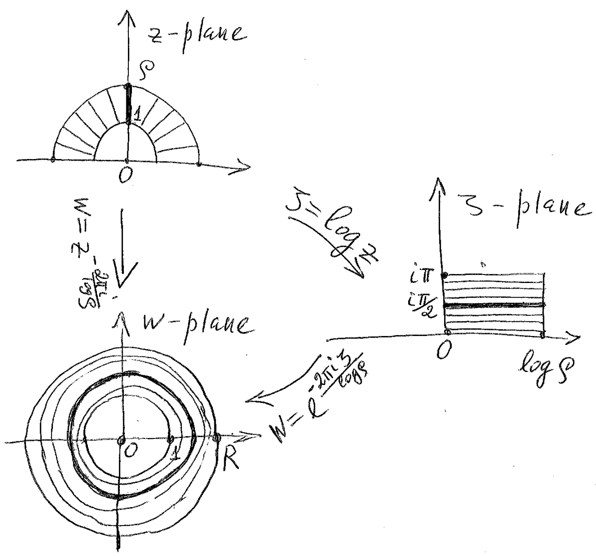

FIGURE 49.2. Covering of an annulus by a cyclic hyperbolic group.

order of the corresponding groups of symmetries (4 and 6), as the latter get factored by the involution $\sigma : z \mapsto -z$ which acts on all tori. (So, the markings $< a, b >$ and $< -a, -b >$ of a lattice $\mathbb{L} = \{ma + nb\}_{m,n \in \mathbb{Z}^2}$ are persistently identified).

2.81. Sum-up all the functions (2.16) comprising the cross-ratio group.

2.101. The cylinder endomorphism $A_n : \mathbb{C}/\mathbb{Z} \to \mathbb{C}/\mathbb{Z}$, $z \mapsto nz$, descends to an endomorphism $\text{Ч}_n : \mathbb{C} \to \mathbb{C}$. [Interprete (2.21) as $\cos \circ A_n = \text{Ч}_n \circ \cos$].

2.103. The torus endomorphism $A_n : \mathbb{C}/\mathbb{L} \to \mathbb{C}/\mathbb{L}$, $z \mapsto nz$, descends to an endomorphism $\mathcal{L}_n : \mathbb{C} \to \mathbb{C}$.

2.105. (iii) This difference is independent of a quadratic differential q since q_1/q_2 is a rational function for any two q_i's. So, any q (e.g., dz^2) can be used for the calculation.

Holomorphic proper maps and branched coverings.

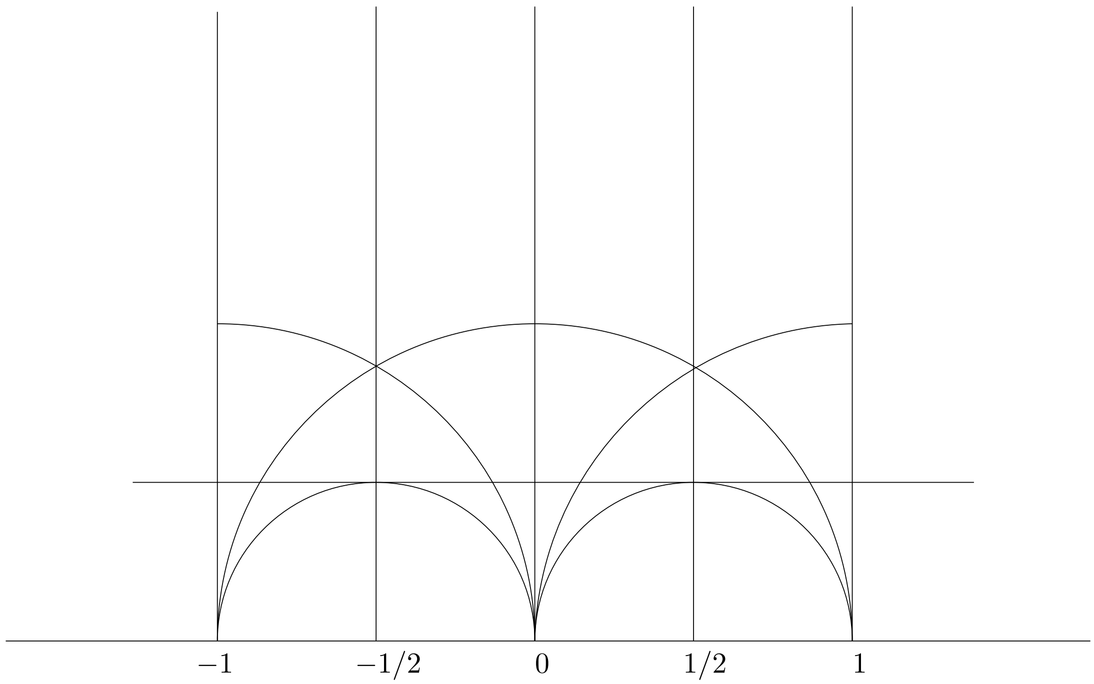

FIGURE 49.3. Tiling of the fundamental triange of the modular group by 6 ideal triangles.

3.2. Consider the unbranched covering over $S' \smallsetminus \{v_i\}$, where v_i are the critical values of f.

3.3. (i) Let a_k be zeros of f, $k = 1, \dots, d$. Then

$$g(z) := f(z) / \prod_{k=0}^{d} \frac{z - a_k}{1 - \bar{a}_k z} \to \partial\mathbb{D} \quad \text{as } z \to \partial\mathbb{D},$$

so it is proper. Since g does not have zeros, it has zero degree, so it is a constant.

Riemann, Montel, Koebe

2.49. (iii) An ideal quadrilateral consisting of two adjacent triangles of the tiling gives us a fundamental domain of λ. In the $\mathbb{H}$-model, we can normalize it so that it is bounded by two vertical lines $x = \pm 1$ and two half-circles $|z \pm 1/2| = 1/2$. Then the boundary identifications are given by two parabolic deck transformations $z \mapsto z + 2$ and $z \mapsto z/(2z + 1)$. They generate the group of deck transformations, on the one hand, and the group Γ_2, on the other.

For (ii), subtract from f a rational map to kill all the poles (which must be simple).

4.3. Let $U_n \Subset U$ be an increasing sequence of domains exhausting U, and let

$$\operatorname{dist}(\phi, \psi) = \sum \frac{1}{2^n} \sup_{z \in U_n} d_s(\phi(z), \psi(z)).$$

4.4. Consider a sequence of holomorphic functions $1/\phi_n(z)$ (which are the original functions written in terms of the local chart $1/z$ near ∞ in the target Riemann sphere). Apply the Hurwitz Theorem on the stability of roots of holomorphic functions.

4.10. Without loss of generality, we can assume that $U = \mathbb{D}$, the functions ψ do not collide in $\mathbb{D}^*$, $\psi_1 \equiv \infty$ and $\psi \equiv \psi_2$ has a pole at 0. Then the functions ϕ_n are holomorphic on $\mathbb{D}$ and form a normal family on $\mathbb{D}^*$. By Exercise 4.7, we can assume that the ϕ_n are either uniformly bounded on each $\mathbb{T}_r$, $r \in (0,1)$, or

$$\phi_n \to \infty \quad \text{uniformly } on\ \mathbb{T}_r. \tag{49.1}$$

In the first case, the Maximal Principle completes the proof, so assume (49.1) occurs. If $\phi_{n(k)}(0) \neq 0$ for a subsequence $n(k)$, then by the Minimum Principle $\phi_{n(k)} \to \infty$ uniformly on $\mathbb{D}_r$, and we are done. So, we can assume that $\phi_n(0) = 0$ for all n. Then the winding number of the curve $\phi_n : \mathbb{T}_r \to \mathbb{C}^*$ around 0 is positive. But by (49.1), the curve $\phi_n - \psi : \mathbb{T}_r \to \mathbb{C}^*$ eventually has the same winding number around 0 (r should be selected so that ψ does have poles on $\mathbb{T}_r$) and hence the equation $\phi_n(z) = \psi(z)$ has a solution in $\mathbb{D}_r$.

Extremal length and width

6.13. (i) Lemma 6.10 implies that there exists $M > 0$ such that $\operatorname{diam} K \leq (1/2) \operatorname{diam} Q$ provided $\operatorname{mod} A \geq M$. Now subdivide A into $n \sim \operatorname{mod} A$ annuli of modulus $\geq M$.

(ii) Normalize K so that $\operatorname{diam} K = 1$ and and let $d := \operatorname{dist}(Q, K)$. Use the Euclidean metric on the d-neighborhood of K to bound from below $\mathcal{L}(\Gamma^{\text{ver}})$.

6.15. Normalize K so that $\operatorname{area} K = \pi$. By the Isoperimetric Inequality, $l_e(\gamma) \geq 2\pi$ for any horizotal curve γ in A. Proceed with an estimate from below for $\mathcal{L}(\Gamma^{\text{hor}})$. (This estimate is due to C. McMullen, see [**BH**, §5.4].)

6.21. Compare Proposition 6.6.

6.23. The path family Γ overflows the half-annulus $\mathbb{A}(1, R) \cap \mathbb{H}$, which implies the lower estimate for $\theta(R)$. Similarly, one can obtain the lower estimate for the dual path family Γ' (connecting $(-\infty, 0]$ to $[1, R]$ in $\mathbb{H}$. This yields the upper estimate for $\theta(R) = 1/\mathcal{L}(\Gamma')$.

Hyperbolic metric and Schwarz Lemma

7.13. Consider the universal covering $\pi : (\mathbb{D}, 0) \to (S', z)$. The lift of S to $\mathbb{D}$ containing 0 covers a disk $\mathbb{D}_t$, where $t = t(r) \to 1$ as $r \to \infty$.

7.20. Triangle inequality.

Carathéodory boundary

8.2. (i) See the hint to Exercise 1.35.

(ii) Take a point $z \in \sigma$. If $D \smallsetminus \sigma$ is connected that there is a Jordan curve $\gamma \subset D$ crossing σ at one point, z. The Jordan disk bounded by γ would contain a limit point $\zeta \in \partial D$ of σ.

8.7. Use the Schwarz Reflection Principle.

8.12. Notice that there is no need to mark a point in A, since only one component of $A \smallsetminus \sigma$ is simply connected. Notice also that the Carathéodory boundary $\partial_e^C D$ literary *coincides* with that of the conformal disk $\hat{\mathbb{C}} \smallsetminus K$, as the notion of a nest of fjords with shrinking cross-cuts is the same in A and in $\hat{\mathbb{C}} \smallsetminus K$ (except for several initial fjords).

8.16. Use Exercise 2.22 and Proposition 7.5.

8.19. Use elementary functions to turn D into a bounded domain. (Compare with the proof of the Riemann Mapping Theorem.)

9.17. For a perfect sector S, the intersection $S \cap K$ is connected by definition. Then use induction in the number of corners (simple and dipole).

9.23. Apply Lemma 1.30, Proposition 9.21, and the Conformal Schönflies Theorem.

Appendix: Potential Theory

10.24 Apply the Index Formula to the gradient vector field ∇G in a region $\{z : 0 < \varepsilon < G(z) < R\}$. (Or apply the Morse theory.)

2.99. Let $\{g_\alpha\}$ be the projective atlas on V. Let us write f in the local parameter $z = g_\alpha(x)$ (i.e., consider the function $f_\alpha = f \circ g_\alpha^{-1}$), and let us take its Schwarzian $Sf_\alpha(z)\, dz^2$. Let $\zeta = g_\beta(x)$ be another local chart (with an overlapping domain), and let $\zeta = A_{\beta\alpha}(z)$ be the transit Möbius map. Then $f_\beta \circ A_{\beta\alpha} = f_\alpha$, and the Chain Rule (2.20) translates into the property that the quadratic differential $Sf_\alpha(z)\, dz^2$ is the pullback of $Sf_\beta(\zeta)\, d\zeta^2$ under $A_{\beta\alpha}$. This means by definition that these local expressions determine a global quadratic differential on V.

Chapter 2. Quasiconformal Geometry

Analytic definition and regularity properties

11.6. If $|\mu| < 1$ then A can be deformed to $z \mapsto az$ through invertible operators. (One can also use formula (11.3) for $\det A$.)

11.9. Since both actions of $\mathrm{PSL}^{\#}(2, \mathbb{R})$ are isometric in the respective metrics it is sufficient to check that $\mathrm{dist}_{\mathrm{T}}(\sigma, \mu) = \mathrm{dist}_{\mathrm{hyp}}(0, \mu)$, where the former distance in measured in $\mathrm{Conf}(\mathbb{C}_{\mathbb{R}})$, while the latter is measured in $\mathbb{D}$. But this is what the first formula of (11.3) tells us.

Quasi-invariance of moduli

12.5. Assume $h : \mathbb{C} \to \mathbb{D}$ is a qc map. Let $A = h(\mathbb{C} \smallsetminus \overline{\mathbb{D}})$. Then $\mathrm{mod}\, A < \infty$ while $\mathrm{mod}(\mathbb{C} \smallsetminus \overline{\mathbb{D}}) = \infty$.

12.12. Consider the points $z_n = x + (z - x)/2^n$, $0 \le n \le N$, where $1/2 < |z_{N+1}| \le 1$, and use $|z_n' - z_{n-1}'| \le L|z_n'|$ inductively.

Further important properties of qc maps

13.14. Combine Corollary 13.12 with the Koebe Distortion Theorem.

Measurable Riemann Mapping Theorem

14.8. Approximate μ_λ by Beltrami differentials with compact support and use that a pointwise limit of (Banach valued) uniformly bounded holomorphic functions is holomorphic.

14.10. Let T be the Riemann surface $(S, f^*\sigma)$. (See §29.1.1 for a discussion of pullbacks of conformal structures by qr maps.)

One-dimensional qs maps, quasicircles and qc welding

15.3. Any interval $J \subset I$ can be sandwitched in between two intervals, $J' \subset J \subset J''$, tessellated by a bounded number of tiles T_k^n of the same level n. Compare Prop. 19.67.

15.5. See [**A2**], Ch IV, Theorem 2.

15.8. The best way is to construct a qc map $h : (\mathbb{C}, K) \to (\mathbb{C}, \tilde{K})$. To this end consider round disks $D_{\bar{\varepsilon}}^n$ based on the generating intervals $I_{\bar{\varepsilon}}^n$ (see §1.1.1) as diameters, and build up h by gluing qc maps between the corresponding pants.

Moduli and Teichmüller spaces of punctured spheres

18.4. A quadratic differential $\phi \in \mathcal{Q}$ can be represented as $\phi(z)dz^2$ where $\phi(z)$ is a holomorphic function on $\hat{\mathbb{C}} \smallsetminus \mathcal{P}$. Since $\int |\phi| < \infty$, this function can have at most simple poles at finite points z_i, $i = 1, \dots, n-1$, and $\phi(z) = O(|z|^{-3})$ near ∞ (which is equivalent to saying that the differential $\phi(z)dz^2$ has a simple pole at ∞). Hence

$$\phi(z) = \sum_{i=1}^{n-1} \frac{\lambda_i}{z - z_i}$$

with $\sum \lambda_i = 0$ and $\displaystyle\sum \lambda_i \sum_{k \neq i} z_k = 0$. These two linear conditions are independent, and in fact, $(\lambda_1, \dots, \lambda_{n-3})$ can be selected as global coordinates on the correspondent subspace (as the the right-most minor of the corresponding $2 \times (n-1)$ matrix is equal to $z_{n-1} - z_{n-2} \neq 1$).

Chapter 3. Dynamical Plane I: basic objects

Glossary of Dynamics

19.1. Use the number-theoretic Möbius Inversion Formula.

19.12. In the invertible case, the density ρ is invariant: $\rho \circ f^{-1} = \rho$ (compare Exercise 19.7).

19.21. Under these circumstances, $\|Df(\alpha)\| < 1$ with respect to some norm in the tangent space $T_\alpha M$. Extending this norm to a Riemannian metric near α, we make f locally contracting.

19.28. (i) Compare with the proof of Theorem 23.4.

(ii) Write a germ $x \mapsto x \pm x^{k+1}$ on I^+ in the coordinate $X = 1/x^k$ (compare with §21.3.2 and Exercise 21.13).

One can use Hölder continuity of qs maps. But it can also be done directly using Exercise 19.28.

19.31. For the unique ergodicity, use Fourier Analysis to check the criterion of Exercise 19.16. [Ergodicity with respect to m can be also deduced from the minimality by means of the Lebesgue Denisty Points Theorem and the isometric property of R_θ.]

circle homeos with irr rot. See [**MvS**]

19.35. Such a homeomorphism is an isometry.

19.36. Show that f is a homeomorphism on $X_\infty := \displaystyle\bigcap_{n \geq 0} f^n(X)$ and consider maximally separated points $x, y \in X_\infty$.

19.37. Otherwise, the restriction $f : X_\infty \to X_\infty$ (see the previous exercise) is a homeomorphism. Hence there exists a $\delta > 0$ such that for any two points $x, y \in X_\infty$ with $d(x, y) < \delta$, we have $d(f^{-1}x, f^{-1}y) < \varepsilon$, where ε is from (19.3). Then $d(f^{-1}x, f^{-1}y) \le \lambda^{-1}d(x, y) < \lambda^{-1}\delta$. Iterating we obtain

$$d(f^{-n}x, f^{-n}y) \le \lambda^{-n}d(x, y) < \lambda^{-n}\delta \quad \forall\, x, y \in X \text{ with } d(x, y) < \delta. \tag{49.2}$$

Cover X with N balls of radius δ. Then by (49.2), X can be covered by N balls of radius $\lambda^{-n}\delta$ for any $n \in \mathbb{N}$, which implies that X is finite.

19.39. (ii) For a tangent vector v, let

$$\|v\|_\rho = \sup_n \max_{1\le i\le 2^n} \lambda^n |Dg_i^{-n}v|,$$

where g_i^{-n} are the inverse branches of g^n.

19.49. More generally, it would be sufficient to assume that $\operatorname{diam}(X^\circ)^n_{\bar{i}} \to 0$ as $n \to \infty$, where $(X^\circ)^n_{\bar{i}}$ are the cylinders of order n with respect to the partition $X^\circ = \bigsqcup X_i^\circ$.

19.60. Compare Exercise 19.53. Condition (i) is used to construct the Devil Staircase: otherwise the Bernoulli conjugacy may not be monotonic. The last assertion uses that h is one-to-one except on the endpoints of gaps, and the latter are pre-fixed points for T_d.

19.65. It can be derived from the symbolic model and Exercise 19.43 or shown directly by noticing that the periodic points and preimages are spread uniformly of the cylinders of rank n.

19.66. Use (19.8).

19.74. Use the Lyapunov metric from Exercise 19.39.

19.83. Adapt the method that was used for Prop. 19.75.

19.84. Follows from the Ergodic Theorem. (See [**Bi**] for these and further nice applications of the Gauss map.)

19.86. Compare §§25.5, 25.9.

19.94. Similar to Proposition 19.75.

19.97. (i) $|\mathrm{Fix}(\sigma^p)|$ is equal to the number of loops in Γ_A of length dividing p. (ii) Follows from the property that for some N any cylinder of rank m contains a periodic point of period $p \le m + N$.

19.104. (i) Show first that the rational nubers form one grand orbit for both Γ (easy) and g (more interesting). Note that any $\mathfrak{p}/\mathfrak{q} \in \hat{\mathbb{Q}}$ is an ideal vertex for some translate of the fundamental domain Δ. (See Exercise 2.47 and the hint to it.)

19.110. To lift an f-invariant measure μ, set

$$\hat{\mu}(\hat{X}) = \mu(U_0 \cap f^{-1}(U_1) \cap \cdots \cap f^{-n}(U_n))$$

for any basic open set $U_0 \times U_1 \times U_n \subset \hat{X}$. See [**KSF**, Ch.X, §4].

19.111. (ii) Invariant measures for both shifts are determined by their cylindrical values, $\mu(\Sigma_{i_m\ldots i_n})$, satisfying compatibility condition and invariant under right translation: $\mu(\Sigma_{i_m\ldots i_n}) = \mu(\Sigma_{i_{m+1}\ldots i_{n+1}})$.

19.113. Take any sequence of positive ε_n with $\sum \varepsilon_n < \infty$, and consider the set $X \subset [0, 1]$ that can be approximated by rational numbers with this rate, i.e.,

$\forall x \in X$ there is an infinite sequence of rational numbers $\mathfrak{p}/\mathfrak{q} \in \mathbb{Q}$ such that

$$|x - \mathfrak{p}/\mathfrak{q}| < \varepsilon_{\mathfrak{q}}.$$

(Compare Proposition 21.38.)

19.115. The size of the gap in $\mathbb{D}(z, \rho)$ depends lower semi-continuous on z.

19.118. Use the Mean Value Theorem to find a point $x_\circ \in D$ such that

$$m((\Delta) = \operatorname{Jac} f(x_\circ) \cdot m(D).$$

19.120. Let p be the maximal common divisor of the lengths of all the loops of Γ. To identify p with the period, show first that if $p = 1$ then Γ is primitive.

19.121. Up to re-labeling of the vertices, A has a triangular form with $a_{ii} = 0$ for $i \in \mathcal{V} \smallsetminus \mathcal{R}$.

Holomorphic dynamics: basic objects

20.1. It follows from the chain rule: $Df^n(z) = \prod_{k=0}^{n-1} Df(f^k z)$.

20.4. (i) Consider fixed points of f and their preimages.

(ii) It is a generality about full sets: a non-trivial loop γ in $\operatorname{int} K$ would break $\hat{\mathbb{C}} \smallsetminus K$ into two pieces. (iii) follows from (ii).

20.17. (iii) Consider separately cases of non-negative and negative multiplier (see Figure 20.8 for the latter). Use that there exists only one cycle of period 2.

Moreover, I_c remains periodic until the "Chebyshev parameter" for $f_c^2 \,|\, I_c$, which is determined from the equation $f_c^2(0) = -\alpha_c$.

20.19. (i) This is related to the doubling formula for cos.

(iii) To show that all points in $\mathbb{C} \smallsetminus \mathcal{I}$ escape to ∞, use Montel's Theorem or (i).

(v) Compare Exercise 19.55.

(vi) μ is the push-forward under the Zhukovsky map of the Lebesgue measure on the unit circle to the interval $[-2, 2]$.

(vii) Interpret dynamically the trigonometric formula $\cos n\theta = P_n(\cos\theta)$, where P_n is the Chebyshev polynomial of degree n (slightly differently normalized).

Periodic motions

21.1. For instance, being simply attracting is equivalent to the existence of an attracting petal described prior to the exercise. (In the interesting direction, it follows from the Schwarz Lemma.)

21.2. For $z \in \mathcal{D}_f(\boldsymbol{\alpha})$, $f^n \to \boldsymbol{\alpha}$ uniformly on a neighborhood of z. Let D be the component of $\operatorname{int} \mathcal{K}(f)$ containing z. Then by normality of the family $(f^n |\, D)$, the $f^n \to \boldsymbol{\alpha}$ uniformly on compact subsets of D.

21.3. (i) $D^\bullet(\alpha)$ is the component of $\{z : f^{pn}(z) \to \alpha \text{ as } n \to \infty\}$ containing α.

(ii) Let $P_\infty = \cup P_n$. Then $f^p(\partial P_\infty) = \partial P_\infty$ since $f^p(\partial P_n) = \partial P_{n-1}$.

21.12. It is particularly clear in Z-coordinate (21.4).

21.13. In $Z = 1/z^k$-coordinate (21.4) we have $Z_n \sim n$. Compare with Exercise 19.28.

21.14. In $Z-$plane (21.4), take a fundamental arc Γ_0 connecting Z and $F(Z)$ and push it forward by the iterates of F.

21.16. By the Leau-Fatou picture, $f\,|\,Q$ has a forward or backward orbit converging to 0, with a power rate. The corresponding hyperbolic orbit would converge exponentially. (Compare with Exercise 21.19(ii))

21.17. Construct a conjugacy starting with the fundamental rectangles/crescents. In fact, one can build up a topological (qc) model for the local dynamics whose local charts are attracting and repelling Fatou coordinates (see §23.7.3). See Camacho [**Ca**].

21.19. (i) Topological and quasisymmetric classifications coincide.

(ii) Exponential and polynomial rates of convergence are not qs compatible (which can be seen, e.g., by considering annuli separating two consecutive points of an orbit from the rest of it). One can also use the Hölder continuity of qs maps. (Compare with Exercise 21.16(ii))

21.21. The dispacement in question is bounded by $\inf\limits_{\operatorname{Re} z>0} \operatorname{dist}(z, z+1)$.

21.22. Use the hyperbolic metric dx/x (near the ends of the strip) to estimate from below the extremal length of the horizontal famiy of curves on the cylinder.

21.24. Here l comes from Theorem 21.11 applied to $f^{p\mathfrak{q}}$; compare also Exercise 21.9.

21.35. Otherwise, the images in question would contain a disk D centered at some $\zeta \in \mathcal{J}$. As the $\operatorname{diam}(f^{n_k}(D))$ are bounded, this would contradict Lemma 21.33.

21.40. Use Exercise 1.106 to show that otherwise all the inverse branches f^{-pn} with $f^{-pn}(\alpha) = \alpha$ are well defined in some neighborhood of α.

21.41. (i) Use normality. (ii) Use Remark 21.6.

(ii) Approximate f by attracting germs $f_n \to f$ and take a Hausdorff limit of the corresponding immediate basins $\mathcal{D}^{\varepsilon}_{f_n}$.

23.41. See [**F3**, Ch. VII, §9], [**M2**].

Remarkable functional equations

23.1. A qc conjugacy $\phi : \mathbb{C} \to \mathbb{C}$ can be written in the polar coordinates (r, θ) as a "power twist map" $r' = r^{\delta},\ \theta' = \theta + \omega(r)$, where

$$\delta = \frac{\log|\rho'|}{\log|\rho|}\,, \quad \omega(r) = \frac{\arg\rho' - \arg\rho}{\log\rho}\,\log r.$$

23.9. It can be written explicitly in the polar coordinates (r, θ) as the logarithmic spiral $r = |\rho|^{\theta/\arg\rho}$.

Alternatively, for any real ρ, we can let $\mathcal{R} := \mathbb{R}_+$, and the general case reduces to this one due to Exercise 23.1.

23.12. See the first proof of Lemma 23.10.

23.13. Linearize map, globalize the curves, and uniformize the strip in between them by a straight strip. The map on the strip will be turned into a translation. (One can also try to implement an elementary topological argument or to make use of the Brouwer Translation Theorem, see [**Fr**]).

23.14. Figure 49.4 illustrates lifts $\tilde{\gamma}_k$, $k \in \mathbb{Z}$, of a cycle of curves with rotation number $\mathfrak{p}/\mathfrak{q} = 2/5$ to the universal covering $\exp : \mathbb{C} \to \mathbb{C}^*$. The curves are labeled so that $\tilde{\gamma}_{k+1}$ lies "above" $\tilde{\gamma}_k$. The branch of $\log\rho$ is selected so that $\tilde{\gamma}_k + \log\rho = \tilde{\gamma}_{k+\mathfrak{p}}$.

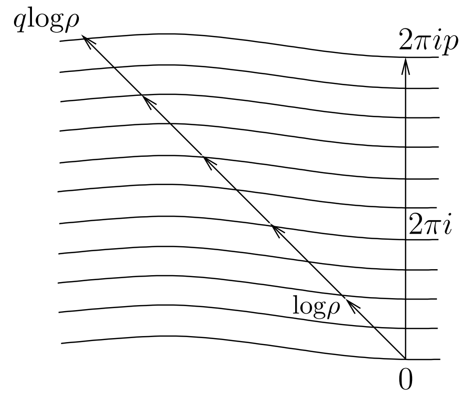

FIGURE 49.4. Lift of a cycle of curves to the universal covering.

23.24. Note that the foliation by round circles is defined dynamically as the closures of the petit orbits $\{z \sim \zeta : \exists n : \ g^n z = g^n \zeta\}$. Hence a germ ϕ commuting with g must respect this foliation. It follows that ϕ is linear (even if it maps just one round circle onto a round circle).

23.30. Two main antipodal needles have tips at $\pm R\, e(\theta)$ corresponding to the critical point (see Theorem 23.29). Other needles are obtained from the main ones by taking preimages under the iterates of $z \mapsto z^2$. See [**LeS**].

23.34. For any $c_\circ$, there exist an $R > 0$ and $\varepsilon > 0$ such that $|f_c(z)| > 2|z|$ for all $c \in \mathbb{D}(c_\circ, \varepsilon)$ and $|z| > R$. Hence $\mathbb{C} \smallsetminus \bar{\mathbb{D}}_R \subset \mathbb{C} \smallsetminus K_c$ for all (c, z) as above.

Now, if $\zeta_\circ \in \mathbb{C} \smallsetminus K_\circ$ then $f_\circ^n(\zeta_\circ) \in \mathbb{C} \smallsetminus \bar{\mathbb{D}}_R$ for some n. By continuity, $f_c^n(\zeta) \in \mathbb{C} \smallsetminus \bar{\mathbb{D}}_R$ for all (c, ζ) sufficiently close to $(c_\circ, \zeta_\circ)$, and the openness follows.

23.42. The realization part is based upon the MRMT, see [**Vo**].

Periodic ray configurations

24.19. The geometry of the ray configuration is controlled by the Böttcher coordinate away from the Julia set, and is controlled by the linearizing coordinates near the repelling points.

Appendix: Rotation sets for the doubling map

24.26. Any minor arc ω_i is disjoint from its image $T(\omega_i)$.

24.28. Construct Θ_θ as a limit of $\Theta_{\mathfrak{p}_n/\mathfrak{q}_n}$ as $\mathfrak{p}_n/\mathfrak{q}_n \to \theta$.

Complementary intervals in Θ are called *gaps*. A gap is called *major* if it has length $\geq 1/2$, and *minor* otherwise.

Key observations:

- The image $T(\omega)$ of any minor gap ω is a minor gap of twice bigger length.
- Any minor gap is eventually mapped onto a major gap, ω_0.
- If $|\omega_0| > 1/2$, then $T(\omega_0)$ would cover the whole circle plus some minor gap, ω_1. Then the end-points of ω_0 would be periodic

24.29. If Θ is contained in a semi-circle, then $T \,|\, \Theta$ preserves the cyclic order.

Chapter 4. Dynamical Plane II: fine structures and models

Hyperbolic maps

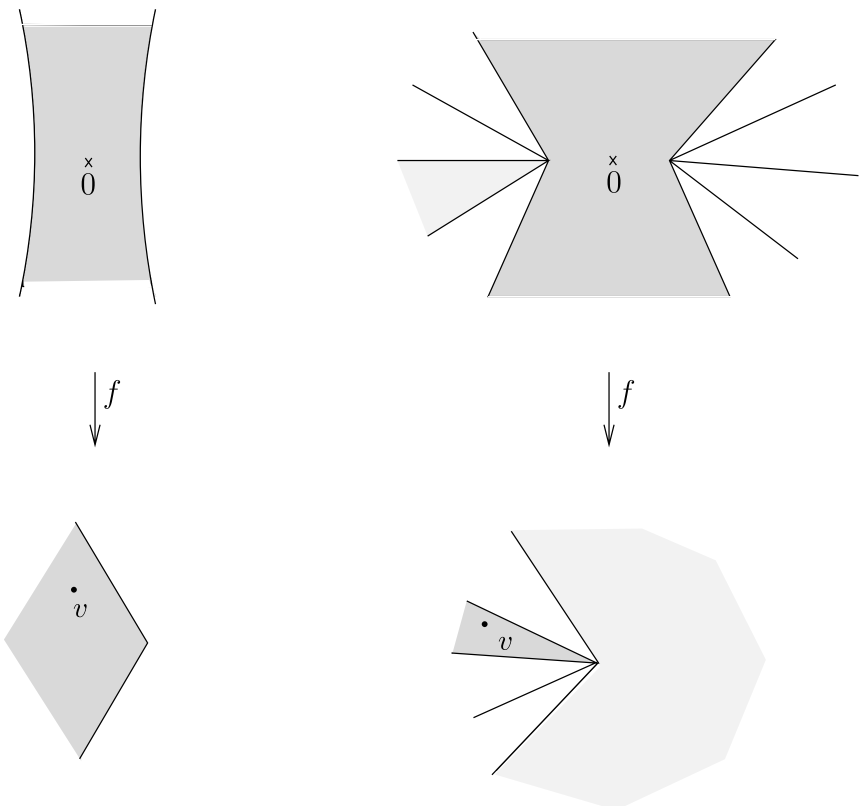

FIGURE 49.5. Preimage of a cut-line and a ray portrait comprising five rays.

25.12. Compare Proposition 41.3.

25.24. The key property is that there exist $\varepsilon > \delta > 0$ and $n \in \mathbb{Z}_+$ such that for any $z \in Q$ the disk $\mathbb{D}(z, \delta)$ is univalently mapped by f^n onto an oval containing $\mathbb{D}(f(z), \varepsilon)$. In particular, any disk $\mathbb{D}(z, \varepsilon)$, $z \in Q$ can be univalently pulled back along any backward orbit $\dots z_{-2} \mapsto z_{-1} \mapsto z$ in Q.

25.30. Follows from Lemma 25.27.

25.32. The permutation of the local branches at α is cyclic, for otherwise there would be a cycle of limbs, from which points 0_k would never return back to 0.

25.34. First, symmetrize $\mathcal{T}$ to obtain $\mathcal{T}^{\mathrm{sym}} \equiv f^{-1}(\mathcal{T})$. For $z \in \mathcal{T}^{\mathrm{sym}} \smallsetminus \{0\}$, let $\mathcal{T}^{\mathrm{sym}}(z)$ be the subtree of $\mathcal{T}^{\mathrm{sym}}$ rooted at z and not containing 0. Let us attach to $\mathcal{T}^{\mathrm{sym}}$ an edge $[z, \beta]$ as follows. Mark consecutive preimages of α' on $\mathcal{T}^{\mathrm{sym}}(\alpha')$,

$$\alpha_{-n} \mapsto \alpha_{-(n-1)} \mapsto \cdots \mapsto \alpha' \equiv \alpha_{-1} \mapsto \alpha, \quad \text{where } \alpha_{-(k+1)} \in \mathcal{T}^{\mathrm{sym}}(\alpha_{-k}),$$

and consider the subtree $X := \mathcal{T}^{\mathrm{sym}}(\alpha_{-n})$. Then $f^n(X) \cap [\alpha, \alpha'] = [\alpha, f^n(z)]$. Moreover, $f^n[z, \beta] \supset [f^n(z), \alpha']$, which determines how $[z, \beta]$ branches off $\mathcal{T}$.

To obtain $\mathcal{T}^e$, attach also the symmetric edge $[z', \beta']$ to $\mathcal{T}^{\mathrm{sym}}(\alpha)$, and let

$$\mathcal{T}^e := \mathcal{T} \cup [\beta, \beta'].$$

25.35. The edge $[z, \beta]$ from the previous Exercise is mapped by f onto a bigger edge $[f(z), \beta]$, and this map is expanding in the hyperbolic metric, implying the desired for all $z \in \sigma^\circ$. Now, make use of Proposition 24.15.

25.40. Compare with Exercise 31.16.

25.43. Any local branch at α stretches under some iterate of $f^{\mathfrak{q}}$ onto the corresponding global branch T. The latter is tessellated by the iterated preimages of the little Hubbard tree $\mathcal{I}^2$ (if exists: see §25.6.10).

25.48. (ii) Compare with the Non-Cutting Assumption from §28.4.1.

25.51. (ii) Use relative exactness.

25.54. Make use of Proposition 19.93.

25.55. (ii) Use only the postcritical points 0_k to produce the Markov tiling.

25.64. Compare with Lemma 32.6.

25.65. Markov property comes form the observation that the boundary of this partition (of the Julia set) is a finite invariant set (compare with Exercise 19.90 and Prop. 25.29).

25.67. Use the hyperbolic metric of $\mathbb{C} \setminus \overline{\mathcal{P}}_f$.

25.68. Compare Exercise 19.86.

25.69. (i) The interval $\mathcal{T}$ already appeared in Exercise 20.18.

25.70. To obtain a Markov family of tiles, remove a small neighborhood of the attracting cycle $\boldsymbol{\alpha}$.

25.71. (ii) Each difference $\mathcal{T}^k \setminus \operatorname{int} \mathcal{T}^{k+1}$ contains an invariant compact set Q^k on which the dynamics is conjugate to an irreducible Topological Markov Chain. Moreover, Q^k is either a Cantor set or a periodic orbit of some period 2^m. Compare with Theorem 30.54. See [**Sha2, JR, vS, B1, B2**].

Parabolic maps

26.11. See the picture.

26.12. Rule out that g is hyperbolic by putting together Exercises 21.21 and 2.22(iii).

26.16. Start with the puzzle piece $P_k \ni \alpha_k$ whose external rays are the same as in the superattracting case (Problem 25.64). Its pullbacks P_k^n under f^p shrink to α_k outside the immediate basin. In the immediate basin use a small thickening.

26.26. Hyperbolic Julia sets do not have cusps (Corollary 25.18), while the Leau-Fatou picture shows that a primitive parabolic Julia set does.

Critically non-recurrent maps

27.1. (i) Use Theorems 21.4, 21.25, and Proposition 21.39 (accompanied with Exercise 21.40) to rule out various types of non-repelling cycles.

(ii) Use Theorem 21.47 to rule out periodic Fatou components (except for $\mathcal{D}(\infty)$) and Lemma 22.1 to rule out wandering components.

(iii) Use Theorem 22.2.

Note that contrary to a common perception, these statements do not need more advanced tools like MRMT or orbifolds.

27.5. The cut-lines in question break $\mathcal{J}$ into an invariant family of pieces on each of which f is univalent and expanding. It follows that they are singletons.

27.7. Compare with the interval dynamics: §§30.6, 30.8.

27.9. (iii) Compare with Exerxice 25.54.

(iv) The matrix A has a hierarchical block structure, with certain blocks of a given level cyclically permuted. All sufficiently high powers A^{pn}, $n \gg 1$, have many positive blocks (which ones?).

27.10–27.11. See [**BFH**] (for polynomials of arbitrary degree).

27.14. (i) Follows from Theorem 22.2.

(ii) For an interval $J \subset \mathcal{A}$, consider a set of points $x \in \mathcal{A}$ that never visit J. Use the dynamical magnification and distortion bounds to show that this set has zero measure. (There is an option of using the Koebe Distortion Theorem or Denjoy distortion estimates in the orbifold metric.)

Alternatively, the assertion follows from ergodicity of the a.c.i.m. μ supported on $\mathcal{A}$ (see Theorem 27.15).

(iii) Use Proposition 19.75 as a model.

27.16. The branches g_i^{-n} near α_k that create poles correspond to backward orbits passing through the critical point. They have a form $g^{-t} \circ g^{-l} \circ g^{-m}$, where g^{-m} fixes $\alpha_k = g^l(0)$. Such a branch contributes density of order $\lambda^{-m}|y - \alpha_k|^{-1/2}$ (see Exercise 19.20), where λ is the multiplier of $\boldsymbol{\alpha}$, so the total contribution has the desider order.

Quadratic-like maps: first glance

28.4. (iv) Use the Argument Principle.

28.12. Take a Jordan curve Γ close to ∂U with winding number 1 around the origin and, look at the curve $g : \Gamma \to \mathbb{C}$, and apply the Argument Principle.

28.19. Since the periodic angles $\theta_\pm$ are repelling under the $p\mathfrak{q}$-fold iterate of the doubling map $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, and α_{ch} is a repelling fixed point for f^p, the line Γ will be "pushed farther away" from α_{ch} under f^p, implying the first assertion.

Moreover, the f^p-orbit of any point $z \in \Omega \smallsetminus W$ escapes from Ω, implying the second assertion.

28.21. In the satellite case, construct an anti- renormalization with period p as follows. Consider the outermost cut-line L of $\mathfrak{R}(\alpha_{\mathrm{ch}})$ through α_{ch} and let W be the component of $\mathbb{C} \smallsetminus \mathcal{R}(\boldsymbol{\alpha})$ attached to L. Show that the return map $f^p : W \dashrightarrow W$ has an f^p-invariant leaf separating α_{ch} from 0. It produces a ql renormalization with period p whose domain contains $\mathfrak{R}(\alpha_{\mathrm{ch}})$. The shadow of $\mathfrak{R}(\alpha_{\mathrm{ch}})$ at infinity is the union of tuned rotation cycles with the same rotation number. Moreover, it the satellite case, these cycles are not boundary points of the shadow intervals I_1, I_2 (coming from the definition of tuned cycles). By Corollary 24.30, there is only one such a cycle, implying that the rays in $\mathfrak{R}(\alpha_{\mathrm{ch}})$ are permuted cyclically under f^p. Compare [**M5**], where this assertion originally appeared (with a different proof).

28.25. Otherwise, the central component of $\mathbb{C} \smallsetminus (\mathfrak{R}(\boldsymbol{\beta}) \cup \mathfrak{R}(\boldsymbol{\beta}'))$ would not be a strip.

28.27. Follows from the inclusion $K \subset \bar{\Pi}_0$.

28.32. (i) Let $\beta_n \in K^n$ be the renormalization β-fixed point. By the Non-Cutting Assumption, all periodic points in K^n except β_n have period at least p_n (under f itself). Hence by Exercise 28.27, all periodic points in K^{n+1} have period at least p_n.

(ii) Use that any non-repelling periodic point is contained in $\omega(0)$.

Topological Dynamics on the Fatou set

29.8. Since $f : D \to D_1$ is a conformal isomorphism, the push-forward $f_*(\mu_0)$ is a conformal structure of D_1 with the same dilatation as μ_0. For the same reason, $f_*^2\mu_0$ is a conformal structure of D_2 with the same dilatation, etc. By pushing it further by all iterates of f, we obtain an invariant measurable conformal structure μ on $\operatorname{orb} D$ with the same dilatation as the original structure on D.

Let us now pull this structure back to preimages of the domains D_n. Of course, one of these preimages can contain the critical point, where the pullback is not well defined. However, it does not cause a problem since a measurable conformal structure needs to be defined only almost everywhere. Since f is locally conformal outside the critical point, the pullback preserves the dilatation of the structure. Iterating this procedure, we obtain an invariant conformal structure on the grand orbit $\operatorname{Orb} D = \bigcup_{m=0}^{\infty} f^{-m}(\operatorname{orb} D)$ (undefined on the critical set $\operatorname{Crit}_f^\infty$ (20.1)) with the same dilatation as the initial structure.

Let us extend this structure to $\hat{\mathbb{C}} \smallsetminus \operatorname{Orb} D$ as the standard one, σ. As σ is invariant in the first place (and has no dilatation), we obtain an invariant measurable structure with bounded dilatation on the whole Riemann sphere.

29.10. The maps H_t do not move periodic points.

Topological dynamics of real quadratic maps

30.6. The action on the basin is proper.

30.7. See Exercise 28.5.

30.9. Compare Exercise 26.30.

30.11. Compare Exercise 25.24.

30.47. See [**Gu1**] and [**L6**].

30.61. See [**CE**].

30.62. See [**MvS**].

Yoccoz puzzle and its Principal Nest

31.3. Otherwise, $Q \cap Q'$ would contain *two* symmetric points.

31.24. Note that $f({Y_{\mathbb{R}}}^{(0)}) = [v, \alpha]$ and $f[v, \alpha] = [\alpha, f^2(0)]$.

For assertion b), note that the map $f : [\alpha', \beta) \to [\alpha, \beta)$ is an orientation preserving diffeomorphism with the property $f(x) < x$.

General combinatorial theory

32.23. The set of points x with itinerary $\bar{\varepsilon}_p^{\text{per}}$ is either a periodic point or a periodic homterval of period p. The period of the limit cycle can be doubled if f^p reverses orientation.

32.32. An important difference is that continuous maps can have non-trivial homtervals that collapse to points in the kneading model (see Exercise 32.29). Statements like Proposition 32.16 should be adjusted accordingly.

However, saw-like maps, being expanding, cannon have homtervals. In particular, they do not have attracting cycles.

32.34. Otherwise the sides of the Q_i would be mapped by $\hat{T}^p$ to diagonals.

Chapter 5. Parameter plane

Definition and first properties

33.1. (iii) Recall the proof of Proposition 20.3.

(iv) We have: $\upsilon_n \to \infty$ locally uniformly on $\mathbb{C} \smallsetminus \mathcal{M}$, and $|\upsilon_n(c)| \leq 2$ on $\mathcal{M}$ (as in Proposition 20.20).

33.6. (Compare with Theorem 21.31.) (i) Since the family of functions υ_n is not normal near $c_* \in \partial\mathcal{M}$, one of the equations $\upsilon_n(c) = 0$ or $\upsilon_n(c) = \pm\sqrt{c}$ should have roots arbitrary close to $c_* \in \partial\mathcal{M}$.

(ii) Consider, for instance, the β-fixed point as a function of c (it branches only at the main cusp 1/4). Then one of the equations $\upsilon_n(c) = \beta(c)$ or $\upsilon_n(c) = \sqrt{\beta(c) - c}$ should have roots arbitrary close to $c_* \in \partial\mathcal{M}$.

Hyperbolic components of $\mathcal{M}$

35.16. (ii) The translation corresponds to the Dehn twist of the torus that keeps the marked loop invariant. It can be dynamically lifted to the attracting basin and then extended by id to the Julia set (something to verify!), yielding an automorphism of $f_\circ$.

35.19. Let us do a Möbius change of variable that moves α to 0 and β to ∞:

$$\zeta = \frac{z-\alpha}{\beta - z} = \frac{\delta + z}{\delta - z}(1 + O(\delta)), \quad \text{where } \delta = \sqrt{b\varepsilon}$$

and the error term is uniform in $z \in \mathbb{D}_{2\delta}$. Conjugating f by this Möbius transformation, we obtain a map of a form $\zeta \mapsto \zeta(1-2\delta)(1+O(\delta^2))$ in the right half-plane $\{\operatorname{Re} z > -1/2\}$ (with the error term being uniform in ζ).

Bringing this half-plane back to the z-plane provides us with a desired disk.

35.28. Compare with Exercises 21.16(ii)) and 26.26.

Structural stability

36.5. Use density of repelling cycles in $\mathcal{J}_c$ and the possibility to launch the active critical orbit so that it lands at any given repelling cycle (by a small perturbation of the map: compare Exercise 33.6(ii)). See **[He1, L7]**.

36.11. For a point $\zeta = z^2 \in \mathbb{A}' = \mathbb{A}[R^2, R^4]$, let $H_c(\zeta) = (H_c(z))^2$. This map is correctly defined (does not depend on the choice of $z = \sqrt{\zeta}$), and is a self-homeomorphism of the annulus $\mathbb{A}'$ identical on $\partial\mathbb{A}'$ and commuting with the group of rotations. Moreover, it commutes with $z \mapsto z^2$ (by definition) and depends holomorphically on c. Now extend it further to $\mathbb{A}[R^4, R^8]$, and so on.

Limbs and wakes of the Mandelbrot set

37.42. (i) Extend the Hubbard tree $\mathcal{T}$ to $\mathcal{T}^e$ (see §25.6.6), and insert it instead of a symmetric small arc $J_\circ \subset \mathcal{T}^e_\circ$ around 0 (by gluing the arc $[\beta, \beta'] \subset \mathcal{T}^e$, appropriately oriented, to $J_\circ$). Spread the inserted tree $\mathcal{T}^e$ around by the dynamics on $\mathcal{T}^e_\circ$. The outcome is $\mathcal{T}^e_f$. [Theory of hyperbolic ql maps can be developed along the lines of the polynomial theory; or else, one can use the Straightening Theorem from §40.2.]

(ii) Use Theorem 37.34.

37.51. (i) Let $\mathrm{Kn}_\circ = (\varepsilon_1 \dots, \varepsilon_{p-1}\, 0)$ and let s be the number of "$-$" in it; let $\mathrm{Kn}_g = (\delta_1\delta_2 \dots)$. Repeat $\mathrm{Kn}_\circ$ periodically and then replace kth "0" with $(-1)^s\delta_k$. (Stop if $\delta_k = 0$.)

(ii) Tune any sequence $\bar{\varepsilon} = (\varepsilon_1 \dots, \varepsilon_{p-1}\, 0)$ of the model for $f_\circ$ by a sequence $\bar{\delta} = (\delta_1\delta_2 \dots)$ of the model for g in the same way as described in (i).

37.53. (i) Start with $(-)$, replace it with $(-+)$, then replace it with $(-+--)$, and proceed inductively as follows: given a sequence of length 2^n, repeat it twice replacing the last symbol to the opposite:

$$(\ominus\oplus - \ominus - + - \oplus - + - - - + - \ominus \dots)$$

(here the 2^nth spots are circled).

Chapter 6. Renormalization, puzzle, and attractors

Straightening

40.5. Since $\partial U'$ is 0-symmetric $\kappa(\delta)$-quasicircle, there is a $L'(\kappa)$-qs homeomorphism $T : \partial U' \to \mathbb{T}_{r^2}$. Since $g : \partial U \to \partial U'$ has a $C(\delta)$-bounded distortion, T lifts to a $L(D, L')$-qc homeomorphism $T : \partial U \to \mathbb{T}_r$. These two qs homeomorphisms can be interpolated by a K-qc homeomorphism $T : A \to \mathbb{A}[r, r^2]$, with K depending only on L', L and bounds for $\operatorname{mod} A$.

40.7. Pull μ back from the fundamental annulus $A = S^2_0 \cap S^2_\infty$ to its preimages $A_n = F^{-n}A$, $\mu|A_n = (F^n)^*(\mu|A)$. Since F is holomorphic in the local chart ϕ_0 (namely, equal to g), all these structures (in this local chart) have the same dilatation as $\mu|A$. Hence they form a single F-invariant measurable conformal structure with bounded dilatation on $S^2 \setminus \phi_0^{-1}K(g)$. Finally, let $\mu = (\phi_0)^*\sigma$ on $\phi_0^{-1}K(g)$.

40.16. See [**McM1**, Theorem 5.11].

Quadratic-like families

42.10. Compare §33.4.

Geometry of Julia sets of Yoccoz class

45.6. It is related to the *König Remark*: If a tree with finite valence of all vertices has arbitrary long branches (at the root) then it has an infinite branch.

45.16. For the last assertion, use Exercise 28.32.

45.15. Let Q_1 be the first kid, and let $f^m : Q_1 \to P$ be the corresponding quadratic-like map. Let $0_n \in Q_1$ be the first landing of $\operatorname{orb} 0$ to Q_1. Since $\mathcal{Y}$ corresponds to the deepest renormalization level, $f^{km}(0_n) \in P \setminus Q_1$ for some $k \in \mathbb{Z}_+$. Pulling P back to 0 under f^{n+km} creates the second kid.

Parapuzzle and its Principal Nest

47.1. The branching valence of the Chebyshev tip of a satellite copy $\mathcal{M}_{\mathtt{p}/\mathtt{q}}$ is equal to $\mathtt{q}$. It is clearly recognizable at the picture.

More fun.

48.1. Use the associated Markov chain.

48.2. Use the first return map

$$g : I^1_- \cup I^1_0 \cup I^1_+ \to [-\alpha, \alpha], \quad \text{where} \quad g \,|\, I^1_0 = f^3, \quad g \,|\, I^1_\pm = f^2,$$

and show that all orbits of odd period must pass through I^1_0.

48.3. For instance, the Feigenbaum parameter represents the final segment of the Sharkovskii scale when all periods 2^n co-exist. The initial segment of the scale (when all even periods co-exist, but no odd periods exist) can be represented e.g., by the Chebyshev tip of the satellite copy $\mathcal{M}^{[1]}$ (which is the last representing parameter).

50. Basic notation, terminology, and conventions

50.1. Logical and set-theoretical symbols and conventions. We use symbol $\equiv$ for tautological equalities (e.g., for modified notations of the same object).

For two real-valued quantities α and β, an implication of sort

$$\alpha \geq \delta \Rightarrow \beta \geq \varepsilon(\delta)$$

means that $\delta = \delta(\varepsilon)$, i.e., "$\forall\ \delta\ \exists\ \varepsilon$".

We say that "property A follows from B, *quantitatively*" if parameters of A depend only on parameters of B. For instance, the statement "a quasiconformal map is quasisymetric, quantitatively" means that the quasisymmetric dilatation of a map in question depends only on its quasiconformal dilatation.

Symbol $\alpha \asymp \beta$ means that $C^{-1} \leq |\alpha/\beta| \leq C$ for some $C > 0$.

Notation $\{x_k\}_{k=1}^n$ stands for an *unordered set* of points $x_k \in X$. If we want to emphasize that the points are *ordered* (also called *colored*), we use notation $(x_k)_{k=1}^n$. Formally speaking, this is a map $\{1, \dots, n\} \to X$. For instance, notation $h : (x_k) \to (y_k)$ means that $h : x_k \mapsto y_k$.

Notation (X, Y) stands for the pair of spaces such that $X \supset Y$. A pair (X, a) of a space X and a "preferred point" $a \in X$ is called a *pointed* space.
Notation $f : (X, Y) \to (X', Y')$ means a map $f : X \to X'$ such that $f(Y) \subset Y'$. In the particular case of pointed spaces $f : (X, a) \to (X', a')$ we thus have: $f(a) = a'$. Similar notations apply to triples, (X, Y, Z), where $X \supset Y \supset Z$, etc.

Notation $f : (X, Y) \to (X', Y')$ is also used for a *germ* near Y i.e., a class of local maps $f : (U_f, Y) \to (X', Y')$ such that any two of them coincide in some neighborhood of Y. In other words, this term serves to emphasize our interest in *local* properties of f near Y. For instance a "holomorphic germ" $f : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ suggests that we are interested in local properties of a holomorphic map near a fixed point (placed at the origin).

A sequence of sets (X_n) forms a (*decreasing*) *nest* if $X_1 \supset X_2 \supset \dots$... It forms an *increasing nest* if $X_1 \subset X_2 \subset \dots$... By default, a "nest" means a decreasing one. It is called *shrinking* (in the metric context) if $\operatorname{diam} X_n \to 0$.

The *intersection multiplicity* of a family of sets X_n is the maximal k such that there is subfamily of k sets X_{n_i}, $i = 1, \dots, k$, with non-empty intersection.

50.2. Complex plane and its affiliates. As usually,
$\mathbb{N} = \{0, 1, 2, \dots\}$ stands for the additive semigroup of natural numbers (with the French convention that zero is natural);
$\mathbb{Z}$ is the group of integers;
$\mathbb{Z}_+$ and $\mathbb{Z}_-$ are the sets of positive and negative integers respectively;
$\mathbb{Q}$ is the field of rationals;
$\mathbb{Q}_{\mathrm{odd}}$ and $\mathbb{Q}_{\mathrm{ev}}$ are the sets of rationals with odd and even denominators respectively (in the irreducible representation), with a convention $0 = 0/1 \in \mathbb{Q}_{\mathrm{odd}}$;
$\mathbb{Q}/\mathbb{Z}$ is the "rational circle", $(\mathbb{Q}/\mathbb{Z})^* = (\mathbb{Q}/\mathbb{Z}) \smallsetminus \{0\}$.

$\mathbb{R}$ stands for the real line;
$\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ is the "real circle";
$\mathbb{R}/\mathbb{Z}$ is the "angular circle", $(\mathbb{R}/\mathbb{Z})^* = (\mathbb{Q}/\mathbb{Z}) \smallsetminus \{0\}$;
$\mathbb{C}$ stands for the complex plane;
and $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ stands for the Riemann sphere;
$\mathbb{C}^* = \mathbb{C} \smallsetminus \{0\}$; $\mathbb{Z}^* = \mathbb{Z} \smallsetminus \{0\}$, etc.;

$\mathbb{C}_{\mathbb{R}} \approx \mathbb{R}^2$ is the decomplexified $\mathbb{C}$ (§11.1.2);
$e(\theta) = e^{2\pi i\theta}$, so $e : (\mathbb{C}, \mathbb{R}) \to (\mathbb{C}^*, \mathbb{T})$ is the exponential covering with $\mathbb{Z}$ serving as the group of deck transformations.

S^2 is a *topological sphere*, i.e., a topological manifold homeomorphic to the standard unit sphere in $\mathbb{R}^3$;
We let $\mathbb{C}_{\mathbb{R}} \approx \mathbb{R}^2$ be the *decomplexified* $\mathbb{C}$ (i.e., $\mathbb{C}$ viewed as 2D real vector space).
For $a \in \mathbb{C}$, $r > 0$, let
(50.1)
$\mathbb{D}_r(a) \equiv \mathbb{D}(a, r) = \{z \in \mathbb{C} : \ |z - a| < r\}; \quad \overline{\mathbb{D}}_r(a) \equiv \overline{\mathbb{D}}(a, r) = \{z \in \mathbb{C} : \ |z - a| \le r\}$

(these notations are also used for balls in $\mathbb{R}^n$), $\mathbb{D}_r \equiv \mathbb{D}(0, r)$, and let $\mathbb{D} \equiv \mathbb{D}_1$ denote the unit disk;
$\mathbb{D}^* = \mathbb{D} \smallsetminus \{0\}$;
Given a real interval $L \Subset \mathbb{R}$, $\mathbb{D}(L)$ is the round disk based upon L is a diameter;
Let $\mathbb{T}(a, r) = \partial\mathbb{D}(a, r)$, $\mathbb{T}_r \equiv \mathbb{T}(0, r)$, and let $\mathbb{T} \equiv \mathbb{T}_1$ denote the unit circle;
S^1 is a *topological circle*;
$\mathbb{A}(r, R) = \{z : \ r < |z| < R\}$ is an open round annulus; the notations $\mathbb{A}[r, R]$ or $\mathbb{A}(r, R]$ for the closed or semi-closed annuli are self-explanatory;
$\partial^i A$ and $\partial^o A$ are the inner and outer boundaries of an annulus $A \subset \mathbb{R}^2$ (§1.7.12);
$\operatorname{mod} A$ is the modulus of an annulus;
Cyl is a topological cylinder (of any type);
$\mathrm{Cyl}^l_h \equiv \mathrm{Cyl}^l(0, h) = (\mathbb{R}/l\mathbb{Z}) \times (0, h)$ is the *flat cylinder* with circumference l and height h (with the affiliated notation for the closed and semi-closed cylinders); $\mathrm{Cyl}_h \equiv \mathrm{Cyl}^1_h$;

$\Pi^l_h = [0, l] \times [0, h]$ is a rectangle; $\Pi^l \equiv \Pi\bar{i}^l_1$.
$\mathbb{H} \equiv \mathbb{H}_+ = \{z : \ \operatorname{Im} z > 0\}$ is the upper half plane;
$\mathbb{H}_- = \{z : \ \operatorname{Im} z < 0\}$ is the lower half-plane;
$\mathbb{H}_h \equiv \mathbb{H}_h(\infty) = \{z : \ \operatorname{Im} z > h\}$, $h > 0$, is a horoball in $\mathbb{H}$ centered at ∞;
$\mathbb{H}_r(a) = \{z \in \mathbb{H} : \ |z - r/2| < r/2\}$, $a \in \mathbb{R}$, is a horoball centered at $a \in \mathbb{R}$;
$\mathbb{L}_h \equiv \mathbb{L}_h(\infty) = \{z : \ \operatorname{Im} z = h\}$, $h > 0$, is a horocycle in $\mathbb{H}$ centered at ∞;
$\mathbb{L}_h(a) = \{z : \ |z - r/2| = r/2\}$, $a \in \mathbb{R}$, is a horocycle centered at $a \in \mathbb{R}$;
$\mathbb{S}_h = \{z : \ 0 < \operatorname{Im} z < h\}$ is a strip, $\mathbb{S} \equiv \mathbb{S}_1$;

$\mathbb{I} = [-1, 1]$ is the standard interval, $\mathbb{I}_\rho = [-\rho, \rho]$;
We use notation $[a, b]_\#$ for a (closed) interval with endpoints a and b, without assuming that $a < b$ (the similar notation is used for open and semi-open intervals);
A similar convention is used for geodesics. For instance, for $a, b \in \partial\mathbb{D} = \mathbb{T}$, notation $[a, b]_\#$ means the closure in $\overline{\mathbb{D}}$ of the non-oriented hyperbolic geodesic $\gamma \subset \mathbb{D}$ with endpoints a and b;
$[-\infty, \infty]$ stands for the natural two-point compactification of $\mathbb{R}$ (oriented);
Given an interval $I \subset \mathbb{R}$, I° stands for its interior rel the real line;
For $\lambda > 0$, we use notation $\lambda \cdot I$ for the λ-scaled interval I (centered at the middle of I;
$\mathbb{C}(I) = \mathbb{C} \smallsetminus (\mathbb{R} \smallsetminus I)$ is the plane slit along two rays (§2.4.5);
$\mathbb{C}_\theta(I)$ is the union of two symmetric disk sectors based on I and meeting $\mathbb{R}$ at angle θ (§7.2.2);
$\mathbb{D}_\theta(I)$ is a symmetric hyperbolic neighborhood of I in $\mathbb{C}(I)$ (§2.4.5);

$U(I) := U \smallsetminus (\mathbb{R} \smallsetminus I)$, where U is an $\mathbb{R}$-symmetric disk containing I (§30.1.4).
A set is called "0-symmetric" (resp, "$\mathbb{R}$-symmetric") if it is invariant under the reflection with respect to the origin (resp.: with respect to the real line).
Given an $\mathbb{R}$-symmetric set $X \subset \mathbb{C}$, $X_{\mathbb{R}} = X \cap \mathbb{R}$ stands its real slice[20];
$r_D(a)$ and $R_D(a)$ are the inner and outer radia of a pointed domain (D, a) (§4.4);
$\mathrm{mod}_{\mathbb{R}}(I : L)$ is the modulus of a pair of intervals $L \subset I \subset \mathbb{R}$ (§6.3.5);
Euc or e is used for the Euclidean metric on various surfaces ($\mathbb{C}$, Cyl, $\mathbb{T}^2$, etc.)

50.3. Spaces and maps.

50.3.1. *Spaces.*

• $\overline{X} \equiv \mathrm{cl}\, X$ denotes the closure of a set X; $\mathrm{int}\, X \equiv X^\circ$ denotes its interior.

• A *neighborhood* of a point x will mean an *open* neighborhood, unless otherwise is explicitly said. For instance, a "closed neighborhood" $P \ni x$ means a closed set such that $\mathrm{int}\, P \ni x$.
• $U \Subset V$ means that U is *compactly contained in* $\mathrm{int}\, V$, i.e., $\overline{U}$ is a compact set contained in $\mathrm{int}\, V$;
• For a domain U (maybe, infinite dimensional), let us say that a subset $K \subset U$ is *strictly contained* in U if $\mathrm{dist}(K, \partial U) > 0$.

• A compact connected space containing more than one point is called a *continuum.*[21]
• A space X is called *totally disconnected* if for any two points $x, y \in X$, there exist disjoint open sets $U \ni x$, $V \ni y$, whose union covers the whole of X.
• A compact space is called *perfect* if it does not have isolated points. Perfect sets are always uncountable. (In particular, continua are uncountable.)

• A precompact subset of a locally compact space is also called *bounded*.
• We say that a sequence $\{z_n\}$ in a locally compact space X *escapes to infinity*, $z_n \to \infty$, if for any compact subset $K \subset X$, only finitely many point z_n belong to K. In other words, $z_n \to \infty$ in the one-point compactification $\hat{X} = X \cup \{\infty\}$ of X.
• Similarly, a sequence of subsets $E_n \subset X$ escapes to infinity, if for any compact subset $K \subset X$, only finitely many sets E_n intersect K. In other words, $E_n \to \infty$ uniformly in $\hat{X}$.
• For a tiling (tessellation) $\mathcal{X}$ of a space $X = \bigcup X_i$, we let $\partial\mathcal{X} := \bigcup \partial X_i$, $\mathrm{diam}\, \mathcal{X} := \max \mathrm{diam}\, X_i$.

50.3.2. *Maps.* • An *embedding* $i : X \hookrightarrow Y$ is a homeomorphism onto the image. An *immersion* $i : X \to Y$ is a continuous (not necessarily injective) map which is locally an embedding.
• A function $f : X \to \mathbb{R}$ is called *upper semicontinuous* at $z \in X$ if

$$f(z) \geq \limsup_{\zeta \to z} f(\zeta).$$

It is called *lower semicontinuous* if $f(z) \leq \liminf_{\zeta \to z} f(\zeta)$.

---

[20] We use notation $X_{\mathbb{R}}$ or $X^{\mathbb{R}}$ for real slices of much more gerneral complex objects.

[21] We hope this will not be confused with the set-theoretical notion of continuum. Note also that our definition slightly differs from the standard definition used in the Continuum Theory, as the latter allows a "continuum" to be a single point.

EXERCISE 50.1. *Show that the set of zeros of a non-negative upper semicontinuous function is of type G_δ.*

• A continuous map $f : X \to Y$ between two locally compact spaces is called *proper* if for any compact set $K \subset Y$, its full preimage $f^{-1}(K)$ is compact. Equivalently, $f(z) \to \infty$ in Y as $z \to \infty$ in X, or in other words, f extends continuously to a map $\hat{f} : \hat{X} \to \hat{Y}$ between the one-point compactifications of X and Y.

EXERCISE 50.2. *An injective proper map $i : X \hookrightarrow Y$ is an embedding.*

In this case we say that X *is properly embedded into* Y.

We have found convenient to adjust this definition to the following situation. Let $(X, \partial X)$ and $(Y, \partial Y)$ be two locally compact spaces with marked boundary, for instance, bordered manifolds (e.g., closed intervals), or closures of domains in a bigger manifold (e.g., in S^2), or cell complexes (e.g., trees) with marked boundaries. Then a continuous map $f : (X, \partial X) \to (Y, \partial Y)$ between the pairs will be called *proper*.

For instance, a path $\gamma : [0, 1) \to Y$ is *proper* if and only if

$$\gamma(t) \in \operatorname{int} Y \text{ for } t \in (0, 1),\ \gamma(0) \in \partial Y \text{ and } \gamma(t) \to \infty \text{ in } \operatorname{int} Y \text{ as } t \to 1.$$

A proper path $\gamma : (0, 1) \to Y$ is called *landing* at a point $a \in \partial Y$ as $t \to 0$ if $\gamma(t) \to a$ as $t \to 0$. It is equivalent to saying that γ extends to a proper path $[0, 1) \to Y$.

• Full preimages of points will also be called its *fibers*. Note that *for a proper map, discrete fibers are finite.*

• $\prod^\circ$ stands for the composition of a chain of maps (§25.8.1).

50.3.3. *Quotients.* An *equivalence relation* $\sim$ on a topological space X is called *closed* if its graph

$$\operatorname{Graph}(\sim) = \{(x, y) \in X : \ x \sim y\}$$

is closed in X^2. In other words, if we have two sequences $x_n \sim y_n$ converging to points x and y respectively, then $x \sim y$.

EXERCISE 50.3. (i) *Show that equivalence classes of a closed relation are closed, but in general, not the other way around.*

(ii) *Show that an equivalence relation $\sim$ is closed iff the quotient space $X/\sim$ is Hausdorff.*

Since an "equivalence relation" is the same as a *partition* of X, we can carry the above terminology over to partitions.

50.3.4. *Metrics.*

• In a metric space , we use notation $B(x, r)$ for an open ball of radius r centered at x (though in $\mathbb{C}$, we also use (50.1)).

• For two sets X and Y in a metric space with metric d, let

$$\operatorname{dist}(X, Y) = \inf_{x \in X,\, y \in Y} d(x, y).$$

If one of these sets is a singleton, say $X = \{x\}$, then we use notation $\operatorname{dist}(x, Y)$ for the distance from X to Y.

• The *diameter* of a set X us defined as

$$\operatorname{diam} X = \sup_{x, y \in X} d(x, y).$$

• $\mathrm{dist}_{\mathrm{H}}(X, Y)$ is the Hausdorff distance between two closed subsets of a compact space Z; $\mathfrak{S}(Z)$ is the space of all closed subsets endowed with this distance.

• For a manifold M, $\mathrm{T}_x M$ stand for its tangent space at x, and $\mathrm{T}M$ stands for its tangent bundle.

• If M is Riemannian and X is a path connected subset of M, then one can induce the metric from M to X in two ways:

– The *chordal metric* is obtained by restricting to X the global metric on M.

– The *path* or *intrinsic metric* is obtained by restricting to X the Riemannian metric on M and then defining the path distance between x and y as the infimum of the length of paths $\gamma \subset X$ connecting x to y (which could be infinite). For instance, one can induce the Euclidean metric from $\mathbb{R}^2$ to the circle $\mathbb{T}$ in these two ways leading to the *chordal* and *angular metrics on* $\mathbb{T}$. A more interesting example is obtained by inducing the Euclidean metric to a Jordan domain $D \subset \mathbb{R}^2$ with fractal boundary.

50.4. Measures and densities. A *measure space* $X \equiv (X, \mu) \equiv (X, \mathfrak{S}, \mu)$ is a set endowed with a (finite or σ-finite measure) μ defined on a σ-algebra $\mathfrak{S}$. A *null-set* in X is a set of zero measure.

An *isomorphism* $h : (X, \mathfrak{S}, \mu) \to (X, \tilde{\mathfrak{S}}, \tilde{\mu})$ between two measure spaces is a measure preserving bijection, i.e.,

$$Y \in \mathfrak{S} \Longleftrightarrow h(Y) \in \tilde{\mathfrak{S}}, \quad \text{and in this case} \quad \mu(Y) = \mu(h(Y)).$$

A (partially defined) map $h : X \dashrightarrow \tilde{X}$ is called an *isomorphism* mod 0 if it restricts to an isomorphism $h : Y \to \tilde{Y}$ between full measure subsets.

More generally, we say that *some property is valid* mod 0 if it is true up to a null-set.

A space X is called *Polish* if it is a separable complete metric[22] space. A measure μ on X is called *Borel* if it is defined on all Borel sets, i.e. on the σ-algebra generated by open (or closed) subsets of X. A *completion* of such a measure amounts to adding to this σ-algebra all subsets of null-sets. By default, *all measures under consideration in this book* are *completed Borel measures on Polish spaces*.

Any Borel measure is *regular* in the sense that any set can be approximated in measure by open sets (from above) and by closed sets (from below): For any measurable set $Z \subset X$ any $\varepsilon > 0$ there exists an open set $U \supset Z$ and a closed set $K \subset Z$ such that $\mu(U \smallsetminus Z) < \varepsilon$ and $\mu(Z \smallsetminus K) < \varepsilon$.

A finite measure space is called a *Lebesgue space* if it isomorphic to the disjoint union of the interval $\mathbb{I}$ (continuous part) and at most countably many atoms (discrete part). (There is an equivalent intrinsic definition which is not directly used in this book.) The main virtue of this category is that the Rokhlin theory of measurable partitions is applicable. Any Polish space (with a finite Borel measure) is Lebesgue.

Let us now briefly dwell on the notion of *absolute continuity*. A functorial way to introduce it is in the category of measurable maps $h : (X, \mu) \to (Y, \nu)$ by requiring that the full preimages of null-sets are null-sets. Equivalently, the push-forward measure $h_* \mu$ is absolutley continuous with respect to ν: $h_* \mu \prec \nu$.

[22]Or rather, metrizable with a complete metric

However, the classical definition for functions of one variable, $h : (0, 1) \to \mathbb{R}$ was the existence of locally integrable ρ which is the derivative of h in the sense of the Newton-Leibnitz Formula:

$$h(y) - h(x) = \int_x^y \rho\, dt \quad \forall\ x, y \in (0, 1)$$

(in other words, h is the distribution function for an absolutely continuous mesure, not necessarily positive). For a 1D homeomorphism, it is equivalent to the absolute continuity of h^{-1} in the above functorial sense. To avoid confusion, we mark the classical notion with the star: *absolute continuity**. The definition extends in an obvious way to functions $h : (0, 1) \to \mathbb{R}^n$, and in particular, to $h : (0, 1) \to \mathbb{C}$. The image of such a maps is a rectifiable curve.

More generally, we can consider absolutely continuous* maps $h : \gamma \to \tilde{\gamma}$ between rectifiable curves (in Euclidean spaces) endowed with the length measures dl and $d\tilde{l}$. Such a map is differentiable a.e. along the curve: for l-a.e. $z_0 \in \gamma$,

$$\exists\ \partial_\gamma h(z_0) := \lim \frac{h(z) - h(z_0)}{z - z_0} \text{ as } z \to z_0,\ z \in \gamma,$$

and the usual change of variable rule is valid: for any $\tilde{\rho} \in L^1(d\tilde{l})$ we have:

$$\int_{\tilde{\gamma}} \tilde{\rho}\, d\tilde{l} = \int_\gamma \rho \circ h\, |\partial_\gamma h|\, dl.$$

When $h : \gamma \to \tilde{\gamma}$ is a homeomorphism then this rule means that $h^*(d\tilde{l}) = |\partial_\gamma h|\, dl$.

Lebesgue measure on $\mathbb{R}^n$ (or a smooth measure on a Riemannian manifold) is usually (in this book) denoted by m;
In 2D we also call it "area", so for $K \subset \mathbb{R}^2$, $\operatorname{area} K \equiv m(K)$;
In 1D, we also use notation $l(K) \equiv \operatorname{length} K \equiv m(K)$;
Similarly, we use notation l_{hyp} for the hyperbolic length;
The densities $\operatorname{dens}(X \mid D)$, $\operatorname{dens}(X \mid a)$, and their upper/lower and one-sided versions are defined in §19.18.

$\nu \prec \mu$ means that μ is absolutely continuous with respect to μ;
Equivalence $\nu \sim \mu$ means that $\nu \prec \mu$ and $\mu \prec \nu$; it is also expressed by saying that μ and ν are in the same *measure class*.

δ_x is the delta-measure supported at the point x.
$\boldsymbol{\Delta} \equiv \boldsymbol{\Delta}^d = \{(p_i) \in \mathbb{R}^{d+1} :\ p_i \geq 0,\ \sum p_i = 1\}$ is the standard probabilistic simplex.

50.5. Dynamical objects.

$\mathrm{R}_\theta : \mathbb{T} \to \mathbb{T}$, $z \mapsto e(\theta)\, z$ is a circle rotation with rotation number $\theta \in \mathbb{R}/\mathbb{Z}$;
Rational rotation numbers are usually denoted $\mathfrak{p}/\mathfrak{q}$ with Gothic fonts (which also applies to other settings: parabolic, combinatorial, etc.).

$\operatorname{orb} z \equiv \operatorname{orb}_f(z) = (z_n \equiv f^n z)_{n=0}^\infty$ is the (forward) orbit of a point z;
$0_n \equiv f^n(0)$ are postcritical points;
$\operatorname{Orb}_-(z) = \bigcup_{n \geq 0} f^{-n}(z)$ is the grand backward orbit;
$\operatorname{Orb}(z) = \bigcup_{n \geq 0} f^{-n}(\operatorname{orb}(z))$ is the grand orbit;
$\omega(z)$ is the limit set of $\operatorname{orb} z$.
$f = f_c : z \mapsto z^2 + c$ is the quadratic family;
$\operatorname{Crit}_f \equiv \operatorname{Crit}(f)$ is set of critical points (§20.9);

$\mathrm{Crit}^n_f = \mathrm{Crit}(f^n)$, $\mathrm{Crit}^\infty_f = \bigcup \mathrm{Crit}^n(f)$ (§§20.1, 20.9);
$\mathrm{Crit}_{\mathbb{R}}(f)$, $\mathrm{Crit}_{\mathbb{R}}(f^n)$ and $\mathrm{Crit}^\infty_{\mathbb{R}}(f)$ are the real counterparts of these sets (§20.7.3);
$\overline{\mathcal{P}}_f$ is the post-valuable set (§§20.1, 20.9).

$\mathcal{D}_f(\infty) \equiv \mathcal{D}_c(\infty)$ basin of ∞;
$\mathcal{K}(f) \equiv \mathcal{K}_c$ is the filled Julia set;
$\mathcal{J}(f) \equiv \mathcal{J}_c$ is the Julia set;
$\mathcal{F}(f) \equiv \mathcal{F}_c$ is the Fatou set;
$\mathcal{K}_{\mathbb{R}} = [\beta, \beta']_\#$ is the real slice of $\mathcal{K}$ (in the case of real f).

$\boldsymbol{\alpha}$ is the cycle of a periodic point α;
If 0 is periodic then $\mathbf{0}$ is its cylce;
$\mathcal{D}(\boldsymbol{\alpha})$ is the attracting or parabolic basin;
$\mathcal{D}_{\mathbb{R}}(\boldsymbol{\alpha})$ is the real basin;
$\mathcal{D}^\bullet(\boldsymbol{\alpha})$ is the immediate basin.

$\mathcal{R}^\theta \equiv \mathcal{R}^\theta_c$ is the external dynamical ray for f_c of angle θ (§23.5.4);
$\mathcal{R}^\theta(t) \equiv \mathcal{R}^\theta(r)$ is the point on the ray $\mathcal{R}^\theta$ whose equipotential level is equal to $t = \log r$;
$\mathcal{R}^\theta[t_1, t_2] = \{R^\theta(t) : \ t_1 \le t \le t_2\}$ be the arc of the ray $\mathcal{R}^\theta$ between equipotentials of level t_1 and t_2 (§24.2).
$\mathcal{E}^t_c \equiv \mathcal{E}^r_c$ is the external dynamical equipotential for f_c of radius r or height t;
$\Sigma_c(r) \equiv \Sigma_c(t)$ is the subpotential disk for f_c of radius r or height t;
$\Omega_c(r) \equiv \Omega_c(t)$ is the superpotential domain for f_c of radius r or height t;
Ω_c is the critical superpotential domain for f_c (bounded by the figure-eight).

$\mathcal{I}_c$, $c \in [-2, 1/4]$ is the maximal invariant interval of f_c;
$\mathcal{T}_c$ is the minimal invariant interval of f_c.

$\mathcal{M}$ is the Mandelbrot set;
$\mathcal{M}_{\mathbb{R}} \equiv \mathcal{M} \cap \mathbb{R} = [-2, 1/4]$ is its real slice (§33.6).

$\mathbf{D} \equiv \mathbf{D}(\infty) \subset \mathbb{C}^2 = \bigcup \mathcal{D}_c(\infty)$ (§23.6.3);
$\boldsymbol{\Omega} = \bigcup \Omega_c \subset \mathbf{D}$ (§23.6.3);
$\mathbf{f}$ is the fibered dynamics (§23.6.3);
$\mathbf{G}$ is the fibered Green function (§23.6.3);
$\mathbf{B}$ is the fibered Böttcher function (§34.4.2);
$\mathbf{Crit}^\infty \subset \mathbb{C}^2$ is the precritical locus og $\mathbf{f}$ (§34.4.3).

$\mathcal{Y}$ is the Yoccoz puzzle, $\mathcal{Y}^{(m)}$ is the puzzle of depth m (§31.1);
$Y^{(m)}(z)$ is the puzzle piece of depth m containing z (maybe, non-elementary);
$\mathcal{Y}[n]$ is the Yoccoz puzzle associated with the nth renormalization level (§31.9).

50.6. Dynamical (and associated) classes.

$\mathfrak{U}_{\mathbb{R}}$ is the class of $\mathbb{R}$-symmetric univalent maps between slit planes (§7.2.1);
$\mathfrak{Q}$ is the space of ql maps (§28.1.2);
$\mathfrak{Q}'$ is the space of perhaps degenerate ql maps (§28.1.2);
$\mathfrak{Q}_{\mathbb{R}}$ is the space of real ql maps (§28.1.5);
$\mathfrak{G}' \equiv \mathfrak{Q}'_{\mathbb{R}}$ is the real slice of the space of perhaps degenerate ql maps (§§28.1.5, §30.1.1);
$\mathfrak{G}$ is the space of maps $f : U \to U'$ of class $\mathfrak{G}'$ such that $U_{\mathbb{R}} \Subset U'_{\mathbb{R}}$ (§30.1.1);

$\mathcal{E}$ is the class of degree two analytic expanding circle maps (§41.1.1);
$\mathfrak{E}$, $\mathfrak{E}'$ are Epstein classes (§30.1.1);
$\mathfrak{Y}$, $\mathfrak{Y}[n]$ are Yoccoz classes (§45.1).

$\operatorname{mod} f$ is the moduls of a ql map or germ;
$\operatorname{mod}_{\mathbb{R}} f$ is the real modulus of a map of class $\mathfrak{G}$ (§30.1.2).

50.7. Functional spaces. $L^p(D, m)$, $p \in [1, \infty]$, stands for the L^p-space on the measure space (D, m), it is abbreviated as $L^p(D)$ when m is the Lebesgue measure, when $D = \mathbb{C}$, it is abbreviated further as L^p;
The corresponding norm is denoted $\|\cdot\|_p$;
$L^p_{\mathrm{loc}}(D, m)$ is the space of locally L^p-functions on (D, m) (with the same abbreviation conventions);
$C^p \equiv C^p(D)$, $p \in \mathbb{N} \subset \{\infty\}$, is the space of C^p-smooth functions on a manifold D (which can have border), with topology of uniform convergence on compact subsets for all derivatives involved, (again if D is not specified, it is assumed to be $\mathbb{C}$);
$C(X) \equiv C^0(X)$ is the space of continuous functions;
For a domain $D \subset \mathbb{C}$, $\mathcal{W}(D)$ is the Sobolev space of bounded continuous functions ϕ on D whose distributional partial derivatives $\partial_x\phi$, $\partial_y\phi$ belong to $L^2_{\mathrm{loc}}(D)$ (§11.5);
$\mathcal{W}^p$, $\mathcal{W}^p_{\mathrm{loc}}$ are similar Sobolev spaces;
$L^p_{\mathrm{comp}}(D, M)$, $C^p_{\mathrm{comp}}(D)$ stand for the corresponding spaces of functions with compact support (where convergent sequences are assumed to be supported on compact subsets);
$\boldsymbol{\mathfrak{M}}(X)$ is the space of finite (comleted Borel) measures on X (§13.7.2);
$\mathfrak{M}(X)$ is the subspace of probability measures on X (endowed with the weak topology) (§19.6.1);
$\mathfrak{M}_f(X)$ is the subspace of invariant measures (§19.6.2).

Given a Banach space $\mathcal{B}$, $\mathcal{B}(a, r)$ stand for the (open) ball of radius r centered at a, $\mathcal{B}_r \equiv \mathcal{B}(0, r)$ (for instance L^∞_1 is the unit ball in L^∞).

$\mathrm{DHomeo}(U, V)$ is the class of homeomorphisms $U \to V$ differentiable a.e. (§11.2);
$\mathrm{DHomeo}^+(U, V)$ is the subclass of orientation preserving homeomorphisms;

50.8. Special groups and homogeneous spaces.
$\mathrm{GL}(V)$ is the group of linear automorphisms of V;
$\mathrm{GL}_+(V)$ is the subgroup of orientation preserving automorphisms;
$\mathrm{SL}(2, \mathbb{R})$ is the group of 2×2 matrices over a ring R with determinant 1 (we will deal with $R = \mathbb{C}$, $\mathbb{R}$, or $\mathbb{Z}$);
$\mathrm{PSL}(2, R) = \mathrm{SL}(2, R)/\{\pm I\}$, where I is the unit matrix;
$\mathrm{SO}(2) \approx \mathbb{T}$ is the group of plane rotations;
$\mathrm{PSO}(2) = \mathrm{SO}(2)/\{\pm I\}$ (this group is actually isomorphic to $\mathrm{SO}(2)$, but it is naturally embedded into $\mathrm{PSL}(2, \mathbb{R})$ rather than $\mathrm{SL}(2, \mathbb{R})$);
$\mathrm{Sim}(2)$ is the group of *similarities* of $\mathbb{R}^2$, i.e., compositions of rotations and scalar operators.
$\mathbb{C}_{\mathbb{R}}$ is naturally embedded into $\mathbb{C}^2$ by $z \mapsto (z, \bar{z})$ (as the reflector for the anti-holomorphic involution $(z, \zeta) \mapsto (\bar{\zeta}, \bar{z})$). Linear operators of $\mathbb{C}^2$ preserving $\mathbb{C}_{\mathbb{R}}$ and

the area therein have the form

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad \alpha, \beta \in \mathbb{C},\ |\alpha|^2 - |\beta|^2 = 1.$$

We let $\mathrm{SL}^{\#}(2, \mathbb{R})$ be the group of these operators (it is another representation of $\mathrm{SL}(2, \mathbb{R})$ in $\mathrm{SL}(2, \mathbb{C})$). Note that it acts on $\mathbb{C}_{\mathbb{R}} \subset \mathbb{C}^2$ by transformations $z \mapsto \alpha z + \beta \bar{z}$. $\mathrm{Möb}(\hat{\mathbb{C}}) \approx \mathrm{PSL}(2, \mathbb{C})$ is the group of Moöbius transformations;

$\mathrm{Aff}(\mathbb{C})$ is the group of complex affine maps;
$\mathrm{Aff}(\mathbb{R})$ is the group of real affine maps (that can act on $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$);
$\mathrm{Euc}(\mathbb{C})$ is the group of orientation preserving Euclidean motions of $\mathbb{C}$.

$\mathrm{Conf}(V)$ is the space of conformal structures on $V \approx \mathbb{R}^2$ (§11.1.1).

For a subgroup $\Gamma \subset G$, we l use the convention that $g \cdot \Gamma$ are *left cosets* while $\Gamma \cdot g$ are *right cosets*. The corresponding homogeneous spaces are denoted G/Γ and $\Gamma \setminus G$, respectively.

50.9. Some elementary facts.

Jensen's Inequality. *Let $\phi(x)$ be a concave function on a real interval I. Then for any points $x_k \in I$, $k = 1, \dots, n$, and any weights $(p_k)_{k=1}^n \in \mathbf{\Delta}^{n-1}$, we have:*

$$\sum p_k\, \phi(x_k) \le \phi\left(\sum p_k x_k\right).$$

A sequence $a_n \in \mathbb{R}_+$ is called *subadditive* if $a_{n+m} \le a_n + a_m$.

Fekete's Lemma. *For a subadditive sequence (a_n)*

$$\exists\ \lim \frac{a_n}{n} = \inf \frac{a_n}{n}.$$

50.10. Abbreviations. We let $e(\theta) := e^{2\pi i \theta}$
1D, 2D, 3D, ... – one-dimensional, two-dimensional, three-dimensional,...
ABM – Area of the Boundary of $\mathcal{M}$
AK – Alexander-Kolmogorov
DH – Douady-Hubbard
IFT – Implicit Function Theorem
IVT – Intermediate Value Theorem
JLC – local connectivity of a Julia set
Leb – Lebesgue
Lip – Lipschitz
MLC – local connectivity of the Mandelbrot set
MRMT – Measurable Riemann Mapping Theorem
a.e. – almost everywhere
acim – absolutely continuous invariant measure
lc – locally connected
ql – quadratic-like
qc – quasiconformal
qs – quasisymmetric
h.o.t. – higher order terms

51. Index

Bibliography

ANALYSIS, GEOMETRY, TOPOLOGY

[Ab] W. Abikoff. The real analytic theory of Teichmüller space. Lecture Notes in Math., v. 820. Springer-Verlag, 1980.

[Akh] N.I. Akhiezer. Elements of the theory of elliptic functions. "Nauka", 1970 (Russian).

[A1] L. Ahlfors. Conformal invariants: Topics in Geometric Function Theory. AMS Chelsea Publ., 2010.

[A2] L. Ahlfors. Lectures on quasiconformal maps, Van Nostrand Co, 1966. 2nd ed.: University Lecture Series, v. 38, 2006. AMS.

[A3] Lars Valerin Ahlfors: collected papers, v. 1–2. Birkhäuser, 1982.

[A4] L Ahlfors. Zur theorie der überlagerungsflächen. Acta Math., v. 65 (1935), 157–194. In: [**A3**], 214–252.

[A5] L. Ahlfors. Quasiconformal mappings, Teichmüller spaces and Kleinian groups. Proceedings of ICM-78, Helsinki. In: [**A3**], 485–498.

[AB] L. Ahlfors and L. Bers. Riemann's mapping theorem for variable metrics. Ann. Math., v. 72 (1960), 385–404. In: [**A3**], 247–271.

[ABeu] L. Ahlfors and A. Beurling. The boundary correspondence under quasiconformal mappings. Acta Math., v. 95 (1956), 125–142. In: [**A3**], 104–122.

[BePes] .P. Belinskii and I.N. Pesin On the closure of the class of continuously differentiable quasiconformal maps. Doklady Akad. Nauk SSSR, v. 102 (1955), 865–866.

[Bel] P.P. Belinskii. General properties of quasiconformal maps (in Russian). Obshie svoistva kvazikonf. otobr. "Nauka", Novosibirsk, 1974.

[BR] L. Bers and H.L. Royden. Holomorphic families of injections. Acta Math., v. 157 (1986), 259–286.

[Bo] B.V. Bojarski. Generalized solutions of systems of differential equations of the first order of elliptic type with discontinuous coefficients. Mat. Sbornik, v. 43 (1957), 451–503. (Russian) English translation is available online as a Report 118 of the University of Jyväskylä (2009), ed.: P. Koskela.

[Ca1] C. Carathéodory. Über die gegenseitige Beziehung der Ränder bei der konformen Abbildung des Inneren einer Jordanschen Kurve auf einen Kreis. [Concerning the correspondence of boundaries under conformal mapping of the interior of a Jordan curve to a disk.] Math. Annalen, v. 73 (1913), 305–320. In: Gesammelte Math. Schriften, v. IV, 3–22.

[Ca2] C. Carathéodory. Über die Begrenzung einfach zusammenhängender Gebiete. [Concerning the boundaries of simply-connected domains.] Math. Annalen, v. 73 (1913), 323–370. In: Gesammelte Math. Schriften, v. IV, 23–80.

[DoC] M.P. Do Carmo. Differential geometry of curves and surfaces. Prentice-Hall, 1976.

[DB] A. Douady, d'après des notes de X. Buff. La théorèm d'intégrabilité des structures presque complexes. In [**TL**], 307–324.

[Doob] J.L. Doob. Probability processes.

[DNF] . B.A. Dubrovin, S.P. Novikov and A.T Fomenko. Modern Geometry: methods and applications, 2nd ed. "Nauka", 1986 (Russian).

[For] F. Forstenrič. Polynomial hulls of sets fibered over the circle. Indiana Univ. Math. J., v. 37 (1988), 869–889.

[Gan] F.R. Gantmacher. Matrix Theory, "Nauka".

[GaL] F.P. Gardiner and N. Lakic. Quasiconformal Teichmüller theory. Math. Surveys and Monographs, v. 76 (2000), AMS.

[GaM] J.B. Garnett and D.E. Marshall. Harmonic measure. New Math. Monographs: 2. Cambridge University Press, 2008.

[Ge] F.W. Gehring. The L^p-integrability of the partial derivatives of a quasiconformal mapping. Acta Math., v. 130 (1973), 265–277.

[GL] I.M. Glazman and Yu.I. Ljubič.[23] Finite dimensional linear analysis (in problems). Nauka, Moscow, 1969 (Russian). Translated by MIT Press (1979) and Dover (2006).

[Gr] H. Gröztsch. [Concerning distortion for univalent non-conformal mappings and a related extension of Picard's Theorem.] Leipz., Math.-Phys. v. 80 (1928), 503—507.

[Ha1] Halmos. Measure theory. Graduate texts in Math., v. 18. Springer.

[Har] P. de la Harpe. Topics in geometric group theory. Chicago Lectures in Math. University of Chicago, 2000.

[Hat] A. Hatcher. Algebraic Topology. Cambridge Univ. Press, 2002.

[He] J. Heinonen. Lectures on analysis on metric spaces. Springer, 2001.

[Hö] A. Hörmander. The analysis of linear partial differential operators I. A series o Comprehensive Studies in Mathematics, v. 256. Springer-Verlag 1990.

[H1] J.H. Hubbard. Teichmüller theory, and applications to geometry, topology and dynamics, vol. 1. Matrix Edition, 2006.

[Iw] H. Iwaniec. Topics in Classical Automorphic Forms. Graduate Studies in Math., v. 17. AMS, 1997.

[KolF] A.N. Kolmogorov and S.V. Fomin. Elements of the Theory of Functions and Functional Analysis. Dover.

[Kol] A.N. Kolmogorov. Les groupes de Betti des espaces localement bicompacts. C.R.Acad.Sci. Paris, v. 202 (1936), 1144–1147.

[KoS] L.B. Koralov and Ya. G. Sinai. Theory of Probability and Random Processes, 2nd edition. Springer 2007.

[Kr] S.L. Krushkal. Quasiconformal mappings and Riemann surfaces. Nauka, 1975 (Russian). English translation: John Wiley, 1979.

[Ku] R. Kühnau. Some historical commentaries on Teichmӥler paper *Extremale quasikonforme Abbildungen und quadratische Differentiale*. In [**Teich**], 537–546.

[IM] T. Ivanec and G. Martin. The Beltrami equation. Memoirs of the AMS, 2008.

[KaC] A. Katok and V. Climenhaga. Lectures on surfaces: (Almost) everything you want to know about them. Student Math. Library, v. 46. AMS. 2008.

[Lav] M.A. Lavrentiev. Sur une classe des representations continues. Mat. Sbornik, v. 42 (1935), 407–434.

[Lew] W. Lewis. The pseudo-arc. Bol. Soc. Mat. Mexicana, v. 5 (1999), 25–77.

[LV] O. Lehto and K.I. Virtanen. Quasiconformal mappings in the plane. Springer, 2nd edition, 1973.

[Lyu] Yuri I. Lyubich. Functional Analysis.

[Moo] R.L. Moore. Concerning upper-semicontinuous collections of continua. Trans. AMS, v. 27 (1925), 416—428.

[Mori] A. Mori. On quasi-conformality and pseudo-analyticity. Trans, AMS, v. 84 (1957), 56–77.

[Mor] C.B. Morrey. On the solutions of quasi-linear elliptic partial differential equations. Trans. Amer. Math. Soc., v. 43 (1938), 126–166.

[Mu1] J. Munkres. Topology. Prentice Hall, 2000.

[Mu2] J. Munkres. The triangulation of locally triangulable spaces. Acta Math., v. 97 (1957), 67–93.

[Na] S.B. Nadler, Jr. Continuum theory. An introduction. Monographs and Textbooks and Pure and Applied Math., v. 158. Marcel Dekker Inc., 1992.

[Pes] I.N. Pesin. Metric properties of Q-quasiconformal mappings. Mat. Sbornik, v. 42 (1956), 281–294 (Russian).

[Rado] T. Radó. Über den Begriff der Riemannschen Fläche. Acta Sci. Math. Szeged, v. 2 (1924–1926), 101–121.

[Re1] L. Rempe-Gillen. On prime ends and local connectivity. arXiv:math/0309022v6 [math.GN], 2014.

[23]"Lyubich" and " Ljubič" are two spellings of the same name. However, Yu. Lyubich and M. Lyubich are two different persons.

[Ro1] V.A Rokhlin. On main concepts of the measure theory. Mat. Sbornik, v. 67 (1949), 107–150 (Russian).
[Ru] W. Rudin. Functional Analysis. McGraw-Hill, 1991.
[SaN] L. Sario and M. Nakai. Classification theory of Riemann surfaces, Springer-Verlag, 1970.
[Sch] Mostly surfaces. AMS, 2010.
[Shab] B.V. Shabat. Complex Analysis.
[Sl] Z. Slodkowski. Extensions of holomorphic motions. Annali della Scuola Normale Superiore di Pisa (classe di scienze, ser, 4), v. 22 (1995), 185–210.
[Spen] E.H. Spenier. Algebraic Topology. McGraw-Hill Book Co, 1966.
[Spiv] Spivak. Comprehensive Introduction to Differential Geometry.
[Ste] E.M. Stein. Singular integrals and differentiability properties of functions. Princeton Univ. Press, 1970.
[STh] D. Sullivan and W. Thurston. Extending holomorphic motions. Acta Math. v. 157 (1986), 243–257.
[Teich] Handbook of Teichmüller theory, v. IV. Ed: A. Papadopoulos. European Math. Soc., 2016.
[Tim] V. Timorin. Moore's Theorem. arXiv:1001.5140v1 (2010).
[To] M. Torhorst. Über den Rand der einfach zusammenhängenden ebenen Gebiete. [Concerning the boundary of simply-connected plane domains.] Math. Z., v. 9 (1921), 44–65.
[U] Uniformisation des surfaces de Riemann: retour sur un théorèm centenaire. ENS Éditions, 2010.
[Ve1] I.N. Vekua. The general representation of functions of two independent variables admitting a derivative in the sense of S.L. Sobolev and the problem of primitives. Dokl. Akad Nauk SSSR, v. 89 (1953), 773–775.
[Ve2] I.N. Vekua. The problem of reduction to canonical form of differential forms of elliptic type and the generalized Cauchy-Riemann system. Dokl. Akad. Nauk SSSR, v. 100 (1955), 197–200.
[V] L.I. Volkovyskii. Quasi-conformal mappings (in Russian). Lvov University, 1954.
[W] J.A. Wolf. Spaces of constant curvature. Berkeley, 1972.
[Zi] H. Zieschang. Surfaces and planar discontinuous groups. Springer, 1980.

GENERAL DYNAMICS and ERGODIC THEORY

[AKMcA] R.L. Adler, A.G. Konheim and M.H. McAndrew. Topological entropy. Trans. AMS, v. 114 (1965), 309–319.
[Ar] D.Z. Arov. To the history of the appearance of the notion of ϵ-entropy of an automorphism of a Lebesgue space and (ϵ, T)-entropy of a dynamical system with continuous time (in Russian). Zap. Nauchn. Sem POMI, v. 436 (2015), 76–100. [Reprint of one chapter from the author's Master thesis, 1957.]
[Bi] P. Billinglsey. Ergodic Theory and Information, 1965. John Wiley and Sons, 1965.
[BDV] C. Bonatti, L. Diaz and M. Viana. Dynamics beyond uniform hyperbolicity. Encycl. of Math. Sci., v. 102 (2005). Springer.
[Bow] R. Bowen. Equilibrium states and the Ergodic Theory for Anosov diffeomorphisms. Lecture Notes in Math., v. 470. Springer, 1975.
[BS] M. Brin and G. Stuck. Introduction to Dynamical Systems. Cambridge University Press, 2002.
[Fr] J. Franks. A new proof of the Brouwer plane translation theorem. Erg. Th. & Dyn. Syst., v. 12 (1992), 217–226.
[Ha2] P. Halmos. Lectures on Ergodic Theory. Tokyo, 1953.
[KaH] A. Katok and B. Hasselblatt. Introduction to the modern theory of dynamical systems. Encyclopedia Math and its Appl., # 54, 1995. Cambridge Univ. Press.
[KY] J. Kennedy and J.A. Yorke. The forced damped pendulum and the Wada property. In: "Continuum Theory and Dynamical Systems" (ed. T. West), 157–181. Dekker.
[K] A.N. Kolmogorov. A new metric invariant of transitive dynamical systems and automorphisms of the Lebesgue space. Dokl. Akad. Nauk SSSR, v. 119 (1958), 861–863.
[KSF] I.P. Kornfeld , Ya.G. Sinai and S.V Fomin. Ergodic Theory. Nauka, 1980.
[Krz] Krzyzewski.
[MaE] N.F.G. Martin and J.W. England. Mathematical theory of entropy. Encyclopedia Math. and Appl., v.12. Addison-Wesley, 1981.

[M1] J. Milnor. On the concept of attractor. Comm. Math. Phys., v. 99 (1985), 177–195. Correction and remarks: ibid. v. 102, 517–519.

[Ni] Z. Nitecki.

[Or] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. Advances Math., v. 4 (1970), 337–352.

[Pa] J. Palis. A global view of dynamics and a Conjecture of the denseness of finitude of attractors. In: "Géometrie Complexe et Systèmes Dynamiques", Asterisque, v. 261 (2000), 335–348. Colloque en l'honneur d'Adrien Douady (Orsay 1995).

[Par] W. Parry. Intrinsic Markov chains.

[Po] M. Pollicott. Lectures on ergodic theory and Pesin theory on compact manifolds. London Math. Soc. Lecture Notes Series, # 180 (1993). Cambridge University Press.

[Ro2] V.A Rokhlin. Lectures on entropy theory of transformations with invariant measure. Uspekhi Mat. Nauk, v. 22 (1967), 3–56 (Russian).

[Si] Ya.G. Sinai. On the notion of entropy of a dynamical system. Dokl. Akad. Nauk SSSR, v. 124 (1959), 268–271.

[Shu] M. Shub. Endomorphisms of compact differentiable manifolds. Amer. J. Math., v. 91 (1969), 175–199.

[VRF] V.G. Vinokurov, B.A. Rubstein and A.L. Fedorov. Lebesgue space and its measurable partitions (in Russian). Tashkent State University, 1985.

[Wa] Walters. Ergodic Theory

HOLOMORPHIC DYNAMICS: CLASSICS and HISTORY

[Al] D.S. Alexander. A history of Complex Dynamics from Schröder to Fatou and Julia. Vieweg, 1994.

[Au] M. Audin. Fatou, Julia, Montel, les grand prix des sciences mathématiques de 1918, et après... Springer, 2009.

[Bot] F.E. Böttcher. Principal laws of convergence of iterates and their application to analysis. Izvestiya Kazan Math-Phys. Obshestva, ser. 2, v. 14 (1904), 155–200 (in Russian).

[Cr] H. Cremer. Über die Haüfigkeit der Nichtzentren. Math. Ann., v. 115 (1938), 573–580.

[F:CR06] P. Fatou. Sur les solutions uniformes des certains équations fonctionnelles. Comp. Rendus Acad. Sci., v. 143 (1906), 546.

[F:CR17] P. Fatou. Sur les substitutions rationnelles. Comp. Rendus Acad. Sci., v. 164 (1917), 806–808, and v. 165 (1917), 992–995.

[F1] P. Fatou. Sur les équations fonctionelles, Premiere Mémoire. Bull. S.M.F., t. 47 (1919), 161–271.

[F2] P. Fatou. Sur les équations fonctionelles, Deuxième Mémoire. Bull. S.M.F., t. 48 (1920), 33–94.

[F3] P. Fatou. Sur les équations fonctionelles, Troisième Mémoire. Bull. S.M.F., t. 48 (1920), 208–314.

[F4] P. Fatou. Notice sure les Travaux Scientifiques de M.P. Fatou, 1929.

[J:CR17] G. Julia. Sur les substitutions rationnelles. Comp. Rendus Acad. Sci., v. 165 (1917), 1098–1100.

[J:CR18] G. Julia. Three notes in Comp. Rendus Acad. Sci., v. 166 (1918), 61–64, 153–156, 599–601.

[J] G. Julia. Mémoire sur l'iteration des fonctions rationnelles. J. Math. Pures et Appl., v. 1 (1918), 47–246.

[Ko] M.G. Koenigs. Recherches sur les integrals de certain equations fonctionelles. Ann. Sci. Éc. Norm. Sup. (3), v. 1 (1884), 1–41.

[Leau] L. Leau. Étude sur les equations fonctionelles a une ou plusieurs variables. Ann. Faculte Sci. Toulouse, v. 11 (1897), E1–E110.

[Mo1] P. Montel. Sur l'integrabilité d'une expression differentiélle. and Sur les suites de fonctions analytiques. C.R. Acad. Sci. Paris, v. 136 (1903), 1233–1236, and v. 138 (1904), 469–471.

[Mo2] P. Montel. Lecons sur les familles normales des fonctions analytiques at leur applications. Gauthier-Villards, Paris, 1927.

[N] V. Nekrashevich. Self-similar groups. Mathematics surveys and monographs, v. 117. AMS, 2005.

[Pap] A. Papadopoulos. Nicolas-Augustine Tissot: A link between cartography and quasiconformal theory. ArXiv:1612.00279v1 [math.HO] (2016).

[Pf] G.A. Pfeiffer. On the conformal mappings of curvilinear angles. The functional equation $\phi[fx] = a_1\phi(x)$. Transactions AMS, v. 18 (1917), 185–198.

[Poi] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. Acta Math., v. 13 (1890), 1–270.

[Si] C.L. Siegel. Iteration of analytic functions. Annals Math., v. 43 (1942), 607–612.

[Schr] M. Schröder. Über unendlich vicle Algorithmen zur Auflösung der Gleichungen. Math. Ann., v. 2 (1870), 317–365.

[UvN] S.M. Ulam and J. von Neumann. On the combinations of stochastic and deterministic processes. Bull. AMS, v. 53 (1947), 1120.

[Va] G. Valiron. Fonctions analytiques. Presses Universitaires de France, 1954.

LOW-DIMENSIONAL DYNAMICS: INTRODUCTIONS, OVERVIEWS and SPECIAL COLLECTIONS

[Be1] A. Beardon. Iteration of rational functions. Graduate texts in mathematics. Springer-Verlag, 1991.

[Bl] P. Blanchard. Complex analytic dynamics on the Riemann sphere. Bull. AMS, v. 11 (1984), 85–141.

[BLS] A. Bonifant, M. Lyubich and S. Sutherland (editors). Frontiers of Complex Dynamics, Princeton Math. Ser., v. 51, 2014. In celebration of John Milnor's 80th birthday. Princeton Univ Press.

[BF] B. Branner and N. Fagella. Quasiconformal surgery in holomorphic dynamics. Cambridge University Press, 2014.

[CG] L. Carleson and W. Gamelin. Complex dynamics. Springer-Verlag, 1993.

[CE] P. Collet and J.-P. Eckmann. Iterated maps of the interval as dynamical systems. Progress in Physics, # 1, 1980. Birkhäuser.

[De1] R. Devaney. An introduction to chaotic dynamical systems, 2nd ed. Addison-Wesley.

[De2] R. Devaney (editor). Complex Dynamical Systems: The mathematics behind the Mandelbrot and Julia sets. Proc. Symp Applied Math., v. 49, 1994. AMS.

[D1] A. Douady. Systèmes dynamiques holomorphes. Séminar Bourbaki, 35^e année, # 599. Astérisque, v. 105/106 (1982-23), 39–63.

[D2] A. Douady. Chirurgie sur les applications holomorphes. In: Proceedings of the International Congress of Mathematicians (Berkeley 1986), 724-738.

[EL1] A. Eremenko and M. Lyubich. Dynamics of analytic transformations. Algebra and Analysis, v. 1 (1989), 1–70.

[FSY] M. Flexor, P. Sentenac and J.-C. Yoccoz (editors). Géométrie Complexe et Systèmes Dynamiques. Astérisque, v. 261, 2000. Colloque en l'honneur d'Adrien Douady (Orsay 1995). Société Math. de France.

[GP] L. Goldberg and A. Phillips. Topological Methods in Modern Mathematics, A Symposium in honor of John Milnor's 60th Birthday". Publish or Perish, 1993.

[HP] P.G. Hjorth and C.L. Petersen (eds). Dynamics on the Riemann sphere. A Bodil Branner Festscrift. Eur. Math. Soc., Zürich (2006).

[L1] M. Lyubich. The dynamics of rational transforms: the topological picture. Russian Math. Surveys, v. 41 (1986), 43–117.

[L2] M. Lyubich. The quadratic family as a qualitatively solvable model of chaos. Notices of the AMS, October 2000.

[L3] M. Lyubich. Six lectures on dynamics of quadratic polynomials, Manuscript 2002. www.math.stonybrook.edu/ mlyubich/papers/index.html .

[L4] M. Lyubich. Forty years of unimodal dynamics: on the occasion of Artur Avila winning the Brin prize. J. of Modern Dynamics., v. 6 (2012), 183–203.

[L5] M. Lyubich. Analytic low dimensional dynamics: from dimension one to two. Proceedings ICM-14 (Seoul), v. I, 443–474.

[Pr] K. Krzyzewski (editor). Dynamical Systems and Ergodic Theory. Banach Center Publ., v. 23, 1989. Polish Scientific Publishers.

[McM1] C. McMullen. Complex dynamics and renormalization. Annals of Math. Studies, v. 135, Princeton University Press, 1994.

[M2] J. Milnor. Dynamics in one complex variable, 3d ed. Annals Math.Studies, # 160. Princeton Univ. Press, 2006.

[MvS] W. de Melo and S. van Strien. One dimensional dynamics. Springer-Verlag, 1993.

[Sch1] D. Schleicher (editor). Complex Dynamics: families and friends. A.K. Peters, 2009. In honor of John Hamal Hubbard.

[SchB] D. Schleicher (in cooperation with Bayani). The dynamics of iterated polynomials and the Mandelbrot set. Book in preparation.

[St] N. Steinmetz. Rational iteration: complex analytic dynamical systems. De Gruyter, 1993.

[vS] S. van Strien. Milnor's conjecture on monotonicity of topological entropy: results and questions. In [**BLS**], 323–337.

[TL] Tan Lei. Similarity between the Mandelbrot set and Julia sets. Comm. Math. Physics, v. 134 (1990), 587–617.

[TL] Tan Lei (editor). The Mandelbrot Set: Theme and Variations. London Math. Soc. Lecture Note Series, v. 274, 2000. Cambridge University Press.

[Sh1] M. Shishikura. Topological, geometric and complex analytic properties of Julia sets. In: Proceedings of the International Congress of Mathematicians (Zürich, 1994), pp. 886-895. Birkhäuser, Basel (1995).

REAL ONE-DIMENSIONAL DYNAMICS

[AM1] A. Avila and C.G. Moreira. Statistical properties of unimodal maps: the quadratic family. Annals Math., v. 161 (2005), 831–881.

[BeC] M. Benedicks and L. Carleson. On iterations of $1 - ax^2$ on (-1,1). Annals Math., v. 122 (1985), 1-25.

[B1] A.M. Blokh. Decomposition of dynamical systems on the interval. Russian Math. Surveys., v. 38 (1983), 133–134.

[B2] A.M. Blokh. On the dynamical systems on one-dimensional branched manifolds, I, II, III. Theory of Functions, Functional Analysis and Applications (Kharkov), v. 46 (1986), 8–18; v. 47 (1987), 67–77; v. 48 (1987), 32–46.

[B3] A. Blokh Rotation numbers, twists and a Sharkovskii-Misiurewicz-type ordering for patterns on the Interval Erg. Th. & Dyn. Syst., v. 15 (1995), 1–14.

[BL1] A. Blokh and M. Lyubich. Attractors of transformations of the interval. Function. Analysis and Appl., v. 21 (1987), 70 - 71. Detailed account: Attractors of maps of the interval. In [**Pr**], 427–442.

[BL2] A. Blokh and M. Lyubich. Measure of solenoidal attractors of unimodal transformations of the interval. Math. Notes., v. 48 (1990), 15 - 20.

[BL3] A. Blokh and M. Lyubich. Measurable dynamics of S-unimodal maps of the interval. Ann. Sci. Éc. Norm Sup., v. 24 (1991), 545-573.

[BMi] A. Blokh, M. Misiurewicz New order for periodic orbits of interval maps Erg. Th. & Dyn. Syst., v. 17 (1997), 565–574.

[BKNS] H. Bruin, G. Keller, T. Nowicki and S. van Strien. Wild Cantor attractors exist. Annals Math., v. 143 (1996), 97-130.

[CE] P. Collet and J-P. Eckmann. On the abundance of aperiodic behavior for maps of the interval. Comm. Math. Phys., v. 73 (1980), 115–160.

[Gu1] J. Guckenheimer. Sensitive dependence to initial conditions for one-dimensional maps. Comm. Math. Phys., v. 70 (1979), 133–160.

[Gu2] J. Guckenheimer. Limit sets of S-unimodal maps with zero entropy. Comm. Math. Physics, v. 110 (1987), 655–659.

[GJ] J. Guckenheimer and S. Johnson. Distortion of S-unimodal maps. Annals Math., v. 132 (1990), 71–130.

[HoK] F. Hofbauer and G. Keller. Quadratic maps without asymptotic measures. Comm. Math. Phys., v. 127 (1990), 319–337.

[Ja1] M.V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Comm. Math. Phys., v. 81 (1981), 39-88.

[Jo] S. Johnson. Singular measures without restrictive intervals. Comm. Math. Phys., v. 110 (1987), 185–190.

[JR] L. Jonker and D. Rand. Bifurcations in one dimension I. The non-wandering set. Inventiones Math., v. 62 (1981), 347–365.

[KN] G. Keller and T. Nowicki. Fibonacci maps re(al)visited. Erg. Th. & Dyn. Syst., v. 15 (1995), 99–120.

[Koz] O.S. Kozlovski. Getting rid of the negative Schwarzian derivative condition. Annals Math., v. 152 (2000), 743–762.

[LiY] T. Li and J.A. Yorke. Period three implies chaos. Amer. Math. Monthly, v. 82 (1975), 985–992.

[Le] F. Ledrappier. Some properties of absolutely continuous invariant measures on an interval. Erg. Th. & Dyn. Syst., v. 1 (1981), 77–94.

[Ma] M. Martens. Distortion results and invariant Cantor sets for unimodal maps. Erg. Th. & Dyn. Syst., v. 14 (1994), 331–349.

[MN] M. Martens and T. Nowicki. Invariant measures for Lebesgue typical quadratic maps. In [**FSY**], 239–252.

[May] R.B. May. Simple mathematical models with very complicated dynamics. Nature, v. 261 (1976), 459–467.

[MeStSt] M. Metropolis, M. Stein, and P. Stein. On finite limit sets for transformations of the unit interval. J. Combinatorial Theory, v. 15 (1973), 25–44.

[Mi1] M. Misiurewicz. Structure of mappings of the interval with zero entropy. Publ. Math. IHES, v. 53 (1981), 17–51.

[Mi2] M. Misiurewicz. Absolutely continuous measures for certain maps of the interval. Publ. Math. IHES., v. 53 (1981), 17–51.

[MiS] M. Misiurewicz and W. Szhlenk. Entropy of piecewise monotone mappings. Astérisque, v. 50 (1977), 299–310.

[MT] J. Milnor and W. Thurston. On iterated maps of the interval. "Dynamical Systems", Proc. U. Md., 1986-87, ed. J. Alexander, Lect. Notes Math., v. 1342 (1988), 465-563.

[MTr] J. Milnor and C. Tresser.

[Myr1] P.J. Myrberg. Sur l'iteration des polynomes réels quadratiques. J. Math. Pures et Appl. (9), v. 41 (1962), 339–351.

[Myr2] P.J. Myrberg. Iteration des reeellen polynome zweiten grades III, ser. A. Annales Academiae Scientiarum Fennicae, v. 336 (1963), 2–18.

[NS] T. Nowicki and S. van Strien. Absolutely continuous measures under a summability condition. Inventiones Math., v. 93 (1988), 619–635.

[O] A. Ognev. Metric properties of a class of mappings of a segment. Math. Notes, v. 30 (1981), 859–866. (Translation from Russian.)

[Rad] A. Radulesku. Thesis, Stony Brook, 2003.

[R] D. Ruelle. Applications conservant une mesure absolutement continue par rapport à dx sur [0,1]. CMP, v. 55 (1977), 47–51.

[Si] D. Singer. Stable orbits and bifurcations of maps of the interval. SIAM J. Applied Math., v. 35 (1978), 260.

[vS] S. van Strien. In the bifurcations creating horseshoes. Lect. Notes Math., v. 898 (1981), 316–351.

[Sha1] A.N. Sharkovskii. Coexistence of cycles of a continuous map of a line into itself. Ukrainian Math. J., v. 16 (1964), 61–71.

[Sha2] A.N. Sharkovskii. Attracting sets that do not contain cycles. Ukrainian Math. J., v. 20 (1968), 136–142.

[ShuS] M. Shub and D. Sullivan Expanding endomorphisms of the circle revisited. Erg. Th. & Dyn. Syst., v. 5 (1985), 285–289.

[Ste] P. Stefan, A theorem of Sharkovskii on the existence of periodic orbits of continuous endomorphisms of the real line. Comm. Math. Phys., v. 54 (1977), 237–248.

HOLOMORPHIC UNICRITICAL DYNAMICS
and its REAL applications

[At] P. Atela. Bifurcations of dynamical rays in complex polynomials of degree two. Erg. Th. & Dyn. Syst., v. 12 (1991), 401-423.

[AKLS] A. Avila, J. Kahn, M. Lyubich, and W. Shen. Combinatorial rigidity for unicritical polynomials. Annals Math., v. 170 (2009).

[AL1] A. Avila and M. Lyubich. The full renormalization horseshoe for unimodal maps of higher degree: exponential contraction along hybrid classes. Publications IHES (2012), p. 171–223.

[ALS] A. Avila, M. Lyubich and W. Shen. Parapuzzle of the Multibrot set and typical dynamics of unimodal maps. Journal of European Math Soc., v. 13 (2011), 27–56.

[AM2] A. Avila and C.G. Moreira. Statistical properties of unimodal maps: periodic orbits, physical measures and pathological laminations. Publications Math IHÉS (2006).

[Bra] B. Branner. The Mandelbrot set. In: "Chaos and fractals", Proceedings of Symposia in Applied Mathematics, v. 39, 75-106.

[BH] B. Branner and J.H. Hubbard. The iteration of cubic polynomials, Part II, Acta Math. v. 169 (1992), 229-325.

[BrMa] R. Brooks and J.P Matelsky. The dynamics of 2-generator subgroups of $\mathrm{PSL}(2, \mathbb{C})$. In: "Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference". Annals of Math. Studies, N 97 (1981). Princeton University Press, eds.: I. Kra and B. Maskit.

[BS] H. Bruin and S. van Strien. Monotonicity of entropy for real multimodal maps. JAMS, v. 28 (2015), 1–61.

[CJY] L. Carleson, P.W. Jones, and J.-C. Yoccoz. Julia and John. Bol. Soc. Brasil. Mat., v. 25 (1994), 1–30.

[D3] A. Douady. Description of compact sets in $\mathbb{C}$. In **[GP]**, 429–466.

[D4] A. Douady. Does a Julia set depend continuously on a polynomial? In: "Complex Dynamical Systems. The mathematics behind the Mandelbrot and Julia sets." Proc. Symp. Applied Math., v. 49. Ed.: R.L Devaney. AMS, 1994, 91–138.

[DH1] A. Douady and J.H. Hubbard. Itérations des polynomes quadratiques complexes. C.R. Acad. Sci. Paris, v. 294 (1982), 123–126.

[DH2] A. Douady and J.H. Hubbard. Étude dynamique des polynômes complexes. Publication Mathematiques d'Orsay, 84-02 and 85-04.

[DH3] A. Douady and J.H. Hubbard. On the dynamics of polynomial-like maps. Ann. Sc. Éc. Norm. Sup., v. 18 (1985), 287-343.

[DH4] A. Douady and J.H. Hubbard, A proof of Thurston's topological characterization of rational functions. Acta Math. v. 171 (1993), 263–297.

[G] L. Goldberg. Fixed points of polynomial maps, Part I: rotation subsets of the circle. Ann. Sci. Éc. Norm Sup., v. 25 (1992), 679–685.

[GS] J. Graczyk and G. Swiatek. Generic hyperbolicity in the logistic family. Annals Math. v. 146 (1997), 1–52.

[Has1] P. Haïssinsky. Chirurgie parabolique. C.R. Acad. Sci. Paris Sér. I Math., v. 327 (1998), 195–198.

[Has2] P. Haïssinsky. Turning hyperbolics into parabolics. In **[BF]**, 385–399.

[He1] M.R. Herman. Exemples de fractions rationnelles ayant une orbit dénse sur la sphere de Riemann. Bull. Soc. Math. France, v. 112 (1984), 93–142.

[HJ] J. Hu and Y. Jiang. The Julia set of the Feigenbaum quadratic polynomial is locally connected. Unpublished preprint, 1993.

[H2] J.H. Hubbard. Local connectivity of Julia sets and bifurcation loci: three theorems of J.-C. Yoccoz. In: "Topological Methods in Modern Mathematics, A Symposium in Honor of John Milnor's 60th Birthday", Publish or Perish, 1993.

[H3] J.H. Hubbard. Teichmüller theory, and applications to geometry, topology and dynamics, vol. 2. Matrix Edition, 2016.

[HSc1] J.H. Hubbard and D. Schleicher. The spider algorithm. In **[De2]**, 155–180.

[J] Y. Jiang. Infinitely renormalizable quadratic Julia sets.

[K] J. Kahn. A priori bounds for some infinitely renormalizable quadratics: I. Bounded primitive combinatorics. Preprint IMS at Stony Brook, # 5 (2006).

[KL1] J. Kahn and M. Lyubich. Quasi-Additivity Law in conformal geometry. Annals of Math., v. 169 (2009), 561–593.

[KL2] J. Kahn and M. Lyubich. Local connectivity of Julia sets for unictritical polynomials. Annals of Math., v. 170 (2009), 413–426.

[KL3] J. Kahn and M. Lyubich. A priori bounds for some infinitely renormalizable quadratics: II. Decorations. Ann. Sci. École Norm. Sup., v. 41 (2008), 57–84.

[LSc] E. Lau and D Schleicher. Internal addresses in the Mandelbrot set and irreducibility of polynomials. Preprint IMS at Stony Brook, # 1994/19.

[Le1] G.M. Levin. On irregular values of parameter for families of polynomial maps. Russian Math.Survey, v. 36 (1981), 189–190. On a sequence of bifurcations of a one-parameter family of maps. Ibid v. 37 (1982), 211–212.

[Le2] G.M. Levin. On Pommerenke's Inequality for eigenvalues of fixed points. Colloq. Math., v. 62 (1991), 167–177.

[LeS] G. Levin and M. Sodin. Polynomials with disconnected Julia sets and Green maps.

[L6] M. Lyubich. Typical behaviour of trajectories of a rational mapping of the sphere. Dokl. Akad. Nauk SSSR, v. 268 (1982), 29 - 32. English translation:

[L7] M. Lyubich. Some typical properties of the dynamics of rational maps, Russian Math. Surveys, v. 38 (1983), 154–155. Detailed exposition: An analysis of stability of the dynamics of rational functions. Teoriya Funk., Funk. Anal. and Prilozh., , v. 42 (1984), 72–91 (Russian). English translation: Selecta Mathematica Sovetica, v. 9 (1990), 69–90.

[L8] M. Lyubich. On the Lebesgue measure of the Julia set of a quadratic polynomial, Preprint IMS at Stony Brook, #1991/10.

[L9] M. Lyubich. Combinatorics, geometry and attractors of quasi-quadratic maps. Annals of Math., v. 140 (1994), 347–404.

[L10] M. Lyubich. Dynamics of quadratic polynomials, I-II. Acta Math., v. 178 (1997), 185–297.

[L11] M. Lyubich. Dynamics of quadratic polynomials, III. Parapuzzle and SBR measures. In [**FSY**], 173–200.

[L12] M. Lyubich. Feigenbaum-Coullet-Tresser Universality and Milnor's Hairiness Conjecture. Annals Math., v. 149 (1999), 319–420.

[L13] M. Lyubich. Almost every real quadratic map is either regular or stochastic. Annals Math., v. 156 (2002), 1–78.

[LM] M. Lyubich and J.Milnor. The unimodal Fibonacci map. JAMS, v. 6 (1993), 425-457.

[Man] B.B. Mandelbrot. Fractal aspects of the iteration $z \mapsto \lambda z(1-z)$ for complex λ and z. Annals New York Acad. Sci., v. 357 (1980), 249–259.

[Mane] R. Mañé. On a theorem of Fatou. Bol. Soc. Brasil. Mat., v. 24 (1993), 1–12.

[MSS] R. Mañé, P. Sad and D. Sullivan. On the dynamics of rational maps. Ann. Sci. École Norm. Sup. 16 (1983), 193–217.

[McM2] C. McMullen. Renormalization and 3-manifolds which fiber over the circle. Annals of Math. Studies, Princeton University Press, 1996.

[McS] C. McMullen and D. Sullivan. Quasiconformal homeomorphisms and dynamics III: The Teichmüller space of holomorphic dynamical systems. Advances Math.

[M3] J. Milnor. Self-similarity and hairiness of the Mandelbrot set. In: "Computers in geometry and topology" Lect. Notes in Pure and Appl. Math., Dekker, 1989.

[M4] J. Milnor. Local connectivity of Julia sets: expository lectures. In [**TL**], 67–116.

[M5] J. Milnor. Periodic orbits, external rays and the Mandelbrot set: An expository account. In [**FSY**], 277–333.

[M6] J. Milnor. On Lattès maps. In [**HP**], 9–43.

[PM1] R. Perez-Marco. Fixed points and circle maps. Acta Math., v. 179 (1997), 243–294.

[PM2] R. Perez-Marco. Hedgehog's dynamics I. Manuscript.

[Pom] C. Pommerenke. On conformal mapping and iteration of rational functions. Complex Variables Theory, v. 5 (1986), 2–4.

[Sc3] D. Schleicher. Rational parameter rays of the Mandelbrot set. In [**FSY**], 405–443.

[Sc4] D. Schleicher. On fibers and local connectivity of Mandelbrot and Multibrot sets. Fractal geometry and applications: a jubillee of Benoit Mandelbrot. Part 1, 477–517. Proc. Symp. Pure Math., v. 72, part 1. AMS, 2004.

[ShTL] M. Shishikura and Tan Lei. An alternative proof of Mañé's theorem on non-expanding Julia sets. In [**TL**], 265–280.

[S1] D. Sullivan. Iteration des functions analytiques complexes. C.R. Acad. Sci. Paris, v. 294 (1982), 301–303. *Detailed accounts*: " Seminar on Conformal and Hyperbolic Geometry", notes by M. Baker and J. Seade. Preprint IHES, 1982, and "Quasiconformal homeomorphisms and dynamics I, solution of the Fatou-Julia problem on wandering domains". Annals Math., v. 122 (1985), 401–418.

[S2] D. Sullivan. " Seminar on Conformal and Hyperbolic Geometry", notes by M. Baker and J. Seade. Preprint IHES, 1982.

[S3] D. Sullivan. Bounds, quadratic differentials, and renormalization conjectures. AMS Centennial Publications, v. 2: Mathematics into Twenty-first Century (1992).

[TL1] Tan Lei. Similarity between the Mandelbrot set and Julia sets. Comm. Math. Physics, v. 134 (1990), 587–617.

[Th1] W. Thurston. On the geometry and dynamics of iterated rational maps. In [**Sch1**], 3–110. (Circulated as a preprint since mid 1980s.) Appendix by D. Schleicher: Laminations, Julia sets, and the Mandelbrot set, 111–138.

FURTHER READING

Neutral phenomenon

A number of references closely related to this section appeared elsewhere, see e.g. **[D4, Sh2, PM1, Ch1]**.

[ABC] A. Avila, X. Buff and A. Cheritat. Siegel disks with smooth boundaries. Acta Math., v. 193 (2004), 1–30.

[Ca] C. Camacho. On the local structure of conformal mappings and holomorphic vector fields. Astérisque, v. 59–60 (1978), 83–94.

[D4] A. Douady. Disques de Siegel et anneaux de Herman. In Séminaire Bourbaki, Astérisque, v. 152–153 (1987), 151–172.

[DLS] D. Dudko, M. Lyubich and N. Selinger. Pacman renormalization and self-similarity of the Mandelbrot set near Siegel parameters. arXiv:1703.01206 [mathDS] (2017).

[Ec] J. Écalle. Théorie itérative: Introduction a la théorie des invariants holomorphes. J. Math. Pure Appl., v. 54 (1975), 183–258.

[FLRT] T. Firsova, M. Lyubich, R. Radu, and R. Tanase. Hedgehogs for neutral dissipative germs of holomorphic diffeomorphisms of $\mathbb{C}^2$. Preprint IMS at Stony Brook, 2016, # 2. To appear in Asterisque.

[He2] M.R. Herman. Sur la conjugations diff/'erentiable des diffeomorphisms du cercle a des rotations. Publications IHES Sci., N° 49 (1979).

[He3] M.R. Herman. Construction of some curious diffeomorphisms of the Riemann sphere. J. London Math. Soc., v. 34 (1986), 375–384.

[ISh] H. Inou and M. Shishikura. The renormalization for parabolic fixed points and their perturbations. Manuscript 2008, www.math.kyoto-u.ac.jp/ mitsu/pararenorm/.

[Lav] P. Lavaurs. Systèmes dynamiques holomorphes: Explosion des points périodiques paraboliques. Thèse, Univ. Paris-Sud, Orsay, 1989.

[MMY] S. Marmi, P. Moussa and J.-C. Yoccoz. The Brjuno functions and their regularity properties. Comm. Math. Phys., v. 186 (1997), 265–293.

[McM3] McMullen. Self-similarity of Siegel disks and Hausdorff dimension of Julia sets. Acta Math., v. 180 (1998), 247–292.

[Pe] C.L. Petersen. Local connectivity of some Julia sets containing a circle with an irrational rotation. Acta Math., v. 177 (1996), 163–224.

[PeZ] C.L. Petersen and S. Zakeri. On the Julia set of a typical quadratic polynomial with a Siegel disk. Annals Math., v. 159 (2004), 1–52.

[Vo] S.M. Voronin. Analytic classification of germs of conformal maps $(\mathbb{C}, 0) \to (\mathbb{C}, 0)$ with identity linear part. Func. Anal. Appl., v. 15 (1981), 1–17.

[Ya1] M. Yampolsky. Siegel Disks and Renormalization Fixed Points. Fields Inst. Comm., v. 53 (2008), 377–393.

[Y] J.-C. Yoccoz. Petits diviseurs en dimension 1. Astérisque, v. 231 (1995).

Holomorphic Ergodic Theory, Measure and dimension of the Mandelbrot set and Julia sets

[ACh] A. Avila and D. Cheraghi. Statistical properties of quadratic polynomials with a neutral fixed point. J. of the EMS, to appear.

[AL2] A. Avila and M. Lyubich. Hausdorff dimension and conformal measures of Feigenbaum Julia sets. J. of the AMS, 21 (2008), 305–383.

[AL3] A. Avila and M. Lyubich. Lebesgue measure of Feigenbaum Julia sets. arXiv: 1504.02986 [mathDS] (2015).

[Br] H. Brolin. Invariant sets under iteration of rational functions. Arkiv Matematik, v 6 (1965), 103–143.

[BC] X. Buff and A. Cheritat. Quadratic Julia sets of positive area. Annals Math., v. 176 (2012), 673–746.

[Ch1] D. Cheraghi. Typical orbits of quadratic polynomials with a neutral fixed point: Brjuno type. Comm. Math. Phys. v. 322 (2013), 999–1035.

[Ch2] D. Cheraghi. Typical orbits of quadratic polynomials with a neutral fixed point: Non-Brjuno type. Ann. Sci. 'Ecole Norm. Sup., to appear.

[DU] M. Denker and M. Urbanski. On Sullivan's conformal measures for rational maps of the Riemann sphere. Nonlinearity, v. 4 (1991), 365–384.

[DS] A. Dudko and S. Sutherland. On the Lebesgue measure of the Feigenbaum Julia set. arXiv:1712.08638 [mathDS] (2017).

[FLM] A. Freire, A. Lopes and R. Mañé. An invariant measure for rational maps. Bol. Soc. Brasil. Math., v. 14 (1983), 45–62.

[GSm] J. Graczyk and S. Smirnov. Non-uniform hyperbolicity in complex dynamics. Inventiones Math., v. 175 (2009), 335–415.

[L14] M. Lyubich. The measure of maximal entropy of a rational endomorphism of the Riemann sphere. Funct. Anal. and Appl., v. 16 (1982), 78–79. *Detailed account*: Entropy properties of rational endomorphisms of the Riemann sphere. Ergod. Th. and Dynam. Syst., v. 3 (1983), 351–386.

[PR] F. Przytycki and S. Rhode. Porosity of Collet-Eckmann Julia sets. Fund. Math., v. 155 (1998), no. 2, 189–199.

[PrU] F. Przytycki and M. Urbanski. Conformal fractals: ergodic theory methods. London Math. Soc. Lecture Notes Series, v. 371, 2010.

[Sh2] M. Shishikura. The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets. Ann. of Math. (2) 147 (1998), no. 2, 225–267.

[S4] D. Sullivan. Conformal dynamical systems. Geometric dynamics (Rio de Janeiro, 1981), 725–752, Lecture Notes in Math., 1007, Springer, Berlin, 1983.

Multimodal, Rational, and Anti-Holomorphic Dynamics

[BFH] B. Bielefield, Y. Fisher, and J.H. Hubbard. The classification of critically preperiodic polynomials as dynamical systems. JAMS, v. 5 (1992), 721–762.

[BL4] A. Blokh and M. Lyubich. Non-existence of wandering intervals and structure of topological attractors of one-dimensional dynamical systems, II. The smooth case. Erg. Th. & Dyn. Syst., v. 9 (1989), 751 - 758.

[BGN] L. Bartholdi, V. Nekrashvich, and R. Grigorchuk. From fractal groups to fractal sets. Fractals in Graz 2001, 25–118. Birkhäuser, 2003.

[BO1] A. Blokh and L. Oversteegen. Wandering triangles exist. C.R. Acad. Sci. Paris, v. 339 (2004), 365–370.

[BO2] A. Blokh, L. Oversteegen, R. Ptacek, and V. Timorin. Combinatorial models for spaces of cubic polynomials. ArXiv:1405.4287v2 [math.DS], 2015.

[BKM] A. Bonifant, J. Kiwi and J. Milnor. Cubic polynomial maps with periodic critical orbits, Part II: Escape regions. Conform Geometry and Dynamics, (Electronic Journal of the AMS). To appear.

[DeM] L. De Marco. Dynamics of rational maps: Lyapunov exponent, bifurcations and capacity. Math. Ann., v. 326 (2003), 43–73.

[Du] R. Dujardin. Bifurcation currents and equidistribution in parameter space. In [**BLS**], 515–566.

[Ep] A. Epstein. Infinitesimal Thurston rigidity and the Fatou-Shishikura inequality. arXiv:math/9902158 (1999).

[Gu3] J. Guckenheimer. Endomorphisms of the Riemann sphere. Global Analysis. Proceedings of Symposium in Pure Math., v. 48 (1970), 95–124.

[Ja2] M.V. Jakobson. On the question of classification of polynomial endomorphisms of the plane. Mat. Sb., v. 80 (1969), 365–387 (Russian).

[Ja3] M.V. Jakobson. Markov partitions for rational endomorphisms of the Riemann sphere. In: "Multicomponent random systems", 381–396. Adv. Prob. and Related Topics, v. 6. Dekker, New York, 1980.

[GM] L. Goldberg and J. Milnor. Part II: Fixed point portraits. Ann. Sc. Éc. Norm. Sup., v. 26 (1993), 51–98.

[HSc2] J.H. Hubbard and D. Schleicher. Multicorns are not path connected. In: [**BLS**], 73–102.

[I] H. Inou. Combinatorics and topology of straightening maps II: Discontinuity. arXiv:0903.4289 (2009).

[IK] H. Inou and J. Kiwi. Combinatorics and topology of straightening maps, I: compactness and bijectivity. Adv. Math., v. 231 (2012), 2666–2733.

[IMu] H. Inou and S. Mukherjee. Discontinuity of straightening in antiholomorphic dynamics. arXiv:1605.08061 (2016).

[Ki] J. Kiwi. Real laminations and the topological dynamics of complex polynomials. Adv. Math., v. 184 (2002), 207–267.

[KSS] O. Kozlovski, W. Shen, and S. van Strien. Rigidity for real polynomials. Annals of Math. (2) 165 (2007), 749–841.

[L6] M. Lyubich. Non-existence of wandering intervals and structure of topological attractors of one-dimensional dynamical systems, I. The case of negative Schwarzian derivative. Erg. Th. & Dyn. Syst., v. 9 (1989), # 4, 737-750.

[M7] J. Milnor. Remarks on iterated cubic maps. Exper. Math., v.1 (1992), 5–25.

[MMS] M. Martens, W. de Melo and S. van Strien. Julia-Fatou-Sullivan theory for real one-dimensional dynamics. Acta Math., v. 168 (1992), 273–318.

[Poir] A. Poirier. Critical portraits for postcritically finite polynomials. Fund. Math., v. 203 (2009), 107–163.

[PS] F. Przytycki and J. Skzrypcszak. Convergence and preimages of limit points for coding trees for iterations of holomorphic maps. Math. Ann., v. 290 (1991), 425–440.

[Re] M. Rees. A partial description of parameter space of rational maps of degree two. Acta Math., v. 168 (1992), 11–87.

[Sh3] M. Shishikura. On the quasiconformal surgery of rational functions. Ann. Sci. Éc. Norm Sup., v. 20 (1987), 1–29.

[TL2] Tan Lei. Accouplements des polynômes quadratiques complexes. C.R. Acad. Sci Paris, Ser I Math., v. 302 (1986), 635–638.

Circle Dynamics

[Ar] V.I Arnold. Small denominators I; on a mapping of the circle into itself. Izvestiya Acad. Nauk, v. 25 (1961), 21–86. *Translations AMS*, v. 46, 213–284.

[Bu] L. Bunimovich. A transformation of the circle. Math. Notes, v. 8 (1970), 587-592.

[Den] A. Denjoy. Sur les courbes définies per des équations différentials á le surface du tore. J. Math. Pures et Appl., v. 11 (1932), 333-375.

[dFdM] E. de Faria and W. de Melo. Rigidity of critical circle mappings, II. J. Amer. Math. Soc. v. 13 (2000), 343–370.

[He3] M. Herman. Sur la conjugaison différentiable des difféomrphisms du cercle a des rotations. Publ. IHES Sci., # 49 (1979), 5–233.

[He4] M. Herman. Conjugaison quasi-symmétrique des difféomorphisms du cercle à des rotations et applications aux disques singuliers de Siegel. Manuscript (1986).

[Sw] G. Swiatek. On critical circle homeomorphisms. Bol. Soc. Bras. Mat., v. 29 (1998), 329–351.

[Poi2] H. Poincaré. Memoire sur les courbes définies per équations différentials, I,II,III,IV. J. Math. Pures et Appl., v. 7 (1881), 375–422; v. 8 (ser. 3) (1882), 251–286; v. 1 (ser. 4) (1885), 167–244; v. 2 (1886), 151–217.

[Ya2] M. Yampolsky. Hyperbolicity of renormalization of critical circle maps. Publ. IHES, v. 96 (2002), 1–41.

[Y2] J.-C. Yoccoz. Il n'y a pas de contre-exemple de Denjoy analytique. C. R. Acad. Sci. Paris, Sér. I Math., v. 298 (1984), 141–144.

[Z] S. Zakeri. Rotation sets.

Transcendental Dynamics

See [**EL1**] for an early survey of this field.

[Ba1] I.N. Baker. The domains of normality of an entire function. Ann. Acad. Sci. Fenn., v. 1 (1975), 277–283.

[Ba2] I.N. Baker. An entire function which has wandering domains. J. Australian Math. Soc., v. 22 (1976), 173–176.

[BaR] I.N. Baker and P.J. Rippon. Iteration of exponential functions. Ann. Acad. Sci. Fenn., v. 9 (1984), 49–77.

[BeL] A. Benini and M. Lyubich. Repelling periodic orbits and landing rays for post-singular bounded exponential maps. Ann. Inst Fourier, v. 64 (2014), 1493–1520.

[Ber] W. Bergweiler. Iteration of meromorphic functions. Bull. AMS, v.29 (1993), 151–188.

[Bi] C. Bishop. Constructing entire functions by quasiconformal folding. Acta Math., v. 214 (2015), 1–60.

[De3] R.L. Devaney. Cantor bouquets, explosions, and Knaster continua: dynamics of complex exponentials. Publ. Math, v. 43 (1999), 27–54.

[DK] R.L. Devaney and M. Krych. Dynamics of Exp z Erg. Th. & Dyn. Syst., v. 4 (1984), 35–52.

[E] A. Eremenko. On the iteration of entire functions. In **[Pr]**, 339–345.

[EL2] A. Eremenko and M. Lyubich. Iterations of entire functions. Dokl. Akad. Nauk SSSR, v. 279 (1984), 25–27.

[EL3] A. Eremenko and M. Lyubich. Examples of entire functions with pathological dynamics. Bull. London Math. Soc., v. 36 (1987), 458–468.

[EL4] A. Eremenko and M. Lyubich. Dynamical properties of some classes of entire functions. Ann. Inst. Fourier, v. 42 (1992), 989–1020.

[F5] P. Fatou. Sur l'itération des fonctions transcendante entirèrs. Acta Math., v. 47 (1926), 337–370.

[GK] L.R. Goldberg and L. Keen. A finiteness theorem for a dynamical class of entire functions. Erg. Th. & Dyn. Syst., v. 6 (1986), 183–192.

[Mayer] J.C. Mayer. The explosion point for the set of endpoints of the Julia set of $\lambda \exp z$. Erg. Th & and Syst., v. 10 (1990), 177–183.

[Re2] L. Rempe. Rigidity of escaping dynamics for transcendental entire functions. Acta Math., v. 203 (2009), 235–267.

[Re3] L. Rempe-Gillen. Arc-like continua, Julia sets of entire functions, and Eremenko's Conjecture. arXiv:1610.06278v2 [mathDS], 2016.

[RSc] L. Rempe and D. Schleicher. Bifurcations in the space of exponential maps. Inventiones Math. (2009).

[Sc2] D. Schleicher. Dynamics of entire functions. In: "Holomorphic Dynamical Systems", Lecture Notes in Math., 1998, 295–339.

Kleinian groups, Fatou-Sullivan Dictionary, and beyond

[Ah2] L.V. Ahlfors. Finitely generated Kleinian groups. Amer. J. Math., v 86 (1964), 413–429.

[Be2] A. Beardon.

[Bers1] L. Bers. Simultaneous uniformization.

[Bers2] L. Bers. On the boundaries of Teichmüller space and on Kleinian groups I. Ann. Math., v. 91 (1970), 570–600.

[Bo] R. Bowen. Hausdorff dimension of quasi-circles. Publications Math. IHES, N° 50 (1979), 11–26.

[BoS] R. Bowen and C. Series. Markov maps associated with Fuchsian groups. Publications Math. IHES, N° 50 (1979), 153–170.

[BCM] J.F. Brock, R.D. Canary and Y.N. Minsky. The classification of Kleinian surface groups, II: The ending lamination conjecture. Annals Math. (2), v. 176 (2012), 1–149.

[BP] S. Bullet and C. Penrose. Mating quadratic maps with the modular group. Invent. Math., v. 115 (1994), 483–511.

[BL] S. Bullet and L. Lomonoco. Mating quadratic maps with the modular group II. Invent. Math., v. 220 (2020), 185–210.

[Can] A. Candel Uniformization of surface laminations. Ann. Sci. Éc. Norm Sup., v. 26 (1993), 489–516.

[G-Mo] X. Gomez-Mont. Transversal holomorphic structures. J. Diff. Geom., v. 15 (1980), 161–185.

[IY] Yu. Ilyashenko and S. Yakovenko. Lecture on analytic differential equations. Grad. Studies in Math., v. 86. AMS, 2007.

[LLMM1] S.-Y. Lee, M. Lyubich, N. Makarov, and S. Mukherjee. Dynamics of Schwarz reflections: the mating phenomena. arXiv:1811.04979 (2018).

[LLMM2] S.-Y. Lee, M. Lyubich, N. Makarov, and S. Mukherjee. Schwarz reflections and anti-holomorphic correspondences. arXiv:1907.09107 (2019).

[LLMM3] R. Lodge, M. Lyubich, S. Merenkov, and S. Mukherjee. On dynamical gaskets generated by rational maps, Kleinian groups, and Schwarz reflections. arXiv:1912.13438 (2019).

[LMMN] M. Lyubich, S. Merenkov, S. Mukherjee, and D. Ntalamkekos. David extension of circle homeomorphisms, mating, and removability. arXiv: (2020).

[LMin] M. Lyubich and Y. Minsky. Laminations in Holomorphic Dynamics.

[Mo] G.D. Mostow. Quasiconformal mappings in n-space and the rigidity of hyperbolic space forms. Publ. IHES Sci., # 34 (1968), 53–104.

[Se] C. Series. The infinite word problem and limit sets for Fuchsian groups. Erg. Th. and Dyn. Syst., v. 1 (1981), 337–360.

[Sc] P. Scott. The geometries of 3-manifolds. Bulletin LMS, v. 15 (1983), 401–487.

[S5] D. Sullivan. Discrete conformal groups and measurable dynamics. Bull. AMS (N.S), v. 6 (1982), 57–73.

[S6] D. Sullivan. In: 'Topological Methods in Modern Mathematics, A Symposium in Honor of John Milnor's 60th Birthday", Publish or Perish, 1993.

[Th2] W. Thurston. The geometry and topology of 3-manifolds.

[Ve] A. Verjovsky. A uniformization theorem for holomorphic foliations. Cont. Math., v. 58, 233-253. AMS, 1987.