POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

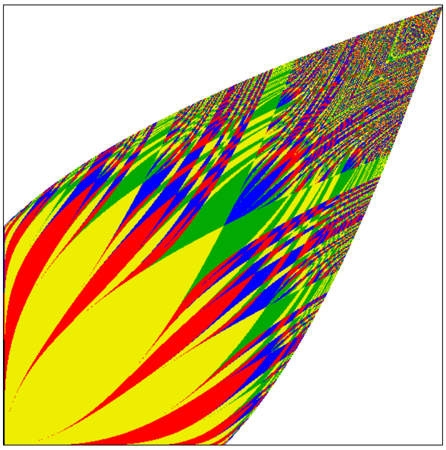
DISSERTATIONES MATHEMATICAE

(ROZPRAWY MATEMATYCZNE)

CCCLXXXII

MATTHIAS ST. PIERRE

Topological and measurable dynamics of Lorenz maps



WARSZAWA 1999

POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

DISSERTATIONES MATHEMATICAE

(ROZPRAWY MATEMATYCZNE)

KOMITET REDAKCYJNY

ANDRZEJ BIAŁYNICKI-BIRULA, BOGDAN BOJARSKI,
ZBIGNIEW CIESIELSKI, LUDOMIR NEWELSKI,
ZBIGNIEW SEMADENI, JERZY ZABCZYK redaktor,
WIESŁAW ŻELAZKO zastępca redaktora

CCCLXXXII

MATTHIAS ST. PIERRE

Topological and measurable dynamics of Lorenz maps

WARSZAWA 1999

Matthias St. Pierre
Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstr. 1 1/2
D-91054 Erlangen, Germany
E-mail: st.pierre@mi.uni-erlangen.de

Published by the Institute of Mathematics, Polish Academy of Sciences
Typeset using TEX at the Institute
Printed and bound by
Publishing House of the Warsaw University of Technology
ul. Polna 50, 00-644 Warszawa

P R I N T E D I N P O L A N D

© Copyright by Instytut Matematyczny PAN, Warszawa 1999

ISSN 0012-3862

Contents

1991 *Mathematics Subject Classification*: Primary 58F03; Secondary 58F11, 58F12, 58F14.

Received 25.11.1998; revised 2.2.1999.

Acknowledgements

This paper is based on the author's doctoral thesis written at the Institute of Mathematics of the Friedrich-Alexander-University Erlangen-Nürnberg under the supervision of Professor Dr. Gerhard Keller. I would like to thank him very much for proposing this interesting subject to me and for all the support he gave me during the making of this thesis. He always had an open ear for my questions and the remarkable capability of tracking down the vital points of my problems extremely quick and then giving me many helpful suggestions and good new ideas.

During his stay in Erlangen, Henk Bruin was an invaluable source of knowledge for me. In particular, I learned many details about the combinatorics of Hofbauer towers from him. He was a very careful reader of early and late versions of this manuscript and made many useful comments and remarks, for which I am very grateful.

I would also like to thank Sebastian van Strien for his hospitality during my visit to the University of Warwick, which was very motivating and encouraging for me. He showed me the striking beauty and simplicity of the Thurston algorithm, which inspired me to write the program that produced many of the figures included in this work.

The work on this thesis was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the Schwerpunktprogramm "Ergodentheorie, Analysis und effiziente Simulation dynamischer Systeme".

1. Introduction

Lorenz maps play an important role in the study of the global dynamics of families of vector fields near homoclinic bifurcations. A typical situation where Lorenz maps are encountered is given by a family $X_\lambda : \mathbb{R}^3 \to \mathbb{R}^3$ of vector fields, where X_0 has a hyperbolic saddle with eigenvalues $\lambda_{ss} < \lambda_s < 0 < \lambda_u$ and with two homoclinic orbits in the configuration of a butterfly connecting the stable and unstable direction, as in the classical Lorenz system [44]. Breaking up the homoclinic loops by changing the parameters it is possible to find vector fields with very complicated chaotic dynamics, where the trajectories of points seem to randomly follow one of two loops near the former homoclinic orbits.

To explain the dynamics on the "strange" attractor of the Lorenz system, Guckenheimer [21] proposed a two-dimensional model for the flow on the attractor, the so-called *geometric Lorenz attractor*. It is obtained, roughly speaking, by forgetting about the strong stable direction and consists of a two-dimensional branched manifold with a hyperbolic saddle as shown in Figure 1.1. Considering the Poincaré map to the cross section $\Sigma = [p, q]$ one obtains an interval map with two monotonic branches: If c denotes the intersection of Σ with the stable manifold of the saddle then all points to the left of c follow the left loop and all points to the right of c follow the right loop until they hit the cross section again.

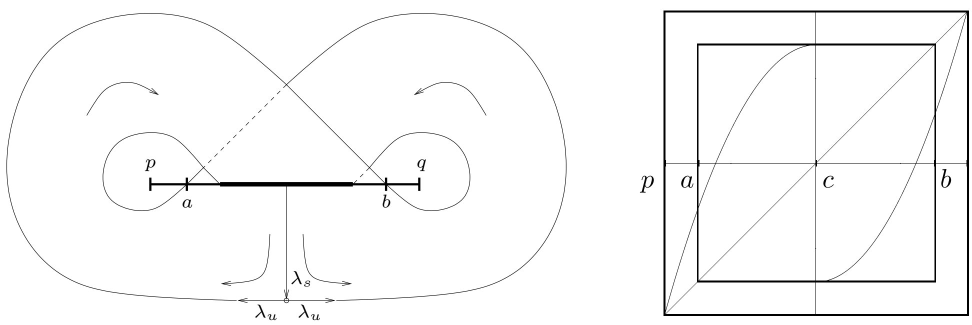


Fig. 1.1. The geometric Lorenz attractor. The left hand side shows a phase portrait of the flow on the branched manifold and the right hand side the first return map to the cross section $\Sigma = [p, q]$.

Hence from a topological viewpoint a Lorenz map is nothing else than a piecewise monotonic interval map with two monotonic branches. From a metrical viewpoint it is smooth to some degree on both branches and of order $\alpha > 0$ at the discontinuity, i.e.,

$f(c \pm \varepsilon) \sim f(c^{\pm}) \pm \varepsilon^{\alpha}$, where the order originates from the local analysis of the linearized flow at the saddle and equals the ratio $|\lambda_s|/|\lambda_u|$ of the stable and unstable eigenvalues. In practice, the smoothness of the Lorenz map is limited by the fact that in order to rigorously justify the geometric Lorenz model, the dynamics of the vector field has to be reduced by means of geometrical methods like the construction of invariant foliations or the existence of a two-dimensional invariant centre manifolds. Even if this can be done, the smoothness of the foliation or the centre manifold depends on certain gap conditions on the eigenvalues of the vector field. In the worst case it can happen that it is only of class $\mathcal{C}^{1+\varepsilon}$, i.e., once differentiable with Hölder continuous derivative (cf. Homburg [33] for an example). Nevertheless, for some of the metrical results we will assume that it is at least of class $\mathcal{C}^2$.

If $\alpha < 1$ then the derivative of f is infinite at the discontinuity. Such maps are typically overall expanding and chaotic. Their topological type is completely determined by the *kneading invariant* of the map, a pair of binary sequences coding the orbits of the two *critical points* c^+ and c^- (i.e., the corresponding one-sided limits of the orbits at the discontinuity c). Moreover, the set of all possible kneading invariants can be characterized by a simple combinatorial condition. Since $\alpha < 1$ holds in the situation of the classical Lorenz system, this kind of Lorenz maps has been studied by many people and their dynamics is well understood (see for example Guckenheimer [21], Rand [54], Guckenheimer & Williams [22], Williams [60], Parry [53], Hubbard & Sparrow [35], and Glendinning & Sparrow [20]).

In this thesis we are mainly concerned with Lorenz maps of exponent $\alpha > 1$. One reason for this is that in this case the derivative of f vanishes at the discontinuity, which means that such maps are typically contracting in some regions and expanding in others, and due to the interplay between contraction and expansion such Lorenz maps can exhibit a much wider spectrum of behaviour. Another reason is that the condition $\alpha > 1$ is compatible with the condition of negative Schwarzian derivative, whereas the condition $\alpha < 1$ is not. Under the assumption of negative Schwarzian derivative one has strong tools to control the distortion on branches of high iterates of the map, just as for smooth interval maps. It will mainly be needed in Chapter 3 where the ergodic properties of Lorenz maps are studied.

Although the theory of vector fields serves as the motivation for studying Lorenz maps, we will focus on what can be said about Lorenz maps from the viewpoint of one-dimensional dynamics. That is, we study the time discrete dynamical system (X, f), where $f : X \to X$ is a Lorenz map defined on some compact interval $X \subset \mathbb{R}$. Given an initial point $x \in X$, its time evolution under the action of f is described by the *orbit* or *trajectory* of x,

$$x_0 \to x_1 \to x_2 \to x_3 \to \dots, \quad x_n := f^n(x).$$

The asymptotic behaviour of x is described by the set of accumulation points of its orbit, called the ω-*limit set* of x. The most important question now is:

What is the asymptotic behaviour of typical points? In general, different points $x \in X$ can have very different asymptotic behaviour. From a physical viewpoint, a certain type of asymptotic behaviour is relevant only if it is observable with positive probability, which means that it is exhibited by a set of initial conditions which has positive Lebesgue measure. This opinion is reflected by the following definition of an attractor which goes back to Milnor [50].

Definition (Attractor). A set $A \subseteq X$ is called *f-invariant* if $f(A) \subseteq A$. For a compact f-invariant set $A \subseteq X$ let $B(A) := \{x \in X \mid \omega(x) \subseteq A\}$ denote the *basin of attraction* of A. Then A is called an *attractor* if the following holds.

1. A attracts a set of positive Lebesgue measure: $m(B(A)) > 0$.
2. Every proper subset $\widetilde{A}$ of A which is compact and f-invariant attracts significantly less points: $m(B(A) \setminus B(\widetilde{A})) > 0$.

An attractor A is called a *global attractor* if $B(A) = X \bmod m$. It is called *indecomposable* if it is not the disjoint union of two smaller attractors, and *minimal* if it does not contain any smaller attractors.

The above question can now be reformulated as follows: What is the global attractor of f and can it be decomposed into minimal attractors?

Even more information about the asymptotic behaviour of a point can be obtained if one not only looks at ω-limit sets but at the empirical distribution of trajectories. More precisely, one considers the empirical distributions $n^{-1}\sum_{k=0}^{n-1}\delta_{x_k}$, where δ_y denotes the Dirac measure at y, and asks whether they converge weakly to some f-invariant measure μ, or more generally, what the set of weak limit points is. Again, the physically most interesting measures are the ones that can be observed for a positive measure set of initial conditions.

Definition (SBR-measure). A measure μ is called a *Sinai–Bowen–Ruelle measure*, or just an *SBR-measure*, if the set $\{x \in X \mid \mu_n(x) := n^{-1}\sum_{k=0}^{n-1}\delta_{x_k} \to \mu\}$ has positive Lebesgue measure.

A trivial example of an SBR-measure is the equidistribution on an attracting periodic orbit. Much more interesting is the case where f has an absolutely continuous invariant probability measure which is ergodic on its support. In this case $\mu_n(x) \to \mu$ holds μ-almost surely by the Ergodic Theorem, whence on a set of positive Lebesgue measure.

In the following we give a description of the contents of this thesis. The theorems stated explicitly below are only meant to sketch the statements of the theorems which they refer to, and have undergone some simplifications. For the precise formulations the reader should consult the referenced theorems.

Markov extensions. Our main tool to study the dynamics of Lorenz maps is the construction of *Markov extensions*. The great advantage of Markov extensions is that they provide a unified framework for the topological and measurable aspects of the dynamics. Markov extensions for piecewise monotonic interval maps were originally introduced by Hofbauer [23–28] as countable state topological Markov chains, which he called *Markov diagrams*. In the situation of Lorenz maps the state space of the Markov chain is a collection $\mathcal{D}$ of subintervals $D \subset X$ which are constructed in such a way that every interval $D \in \mathcal{D}$ is mapped by f onto one or two intervals from $\mathcal{D}$, called the *successors* of D. The successor relation defines the possible transitions for $\mathcal{D}$. Every point from X can be represented by a path in the Markov diagram which symbolically codes the sequence of intervals that are visited by the iterates of x. In this way the original system (X, f) becomes a continuous factor of the Markov chain. This fact can be used to derive statements about topological properties of f like the nature of the nonwandering set or the topological entropy.

Later Keller [37–41] turned the Markov diagram into an even more powerful tool by taking into account that its “states” are more than that, namely intervals which carry a smooth structure and a canonical measure, the Lebesgue measure. He defines the (*canonical*) *Markov extension* of f as a lift of the dynamical system (X, f) to the one-dimensional manifold $\widehat{X} := \bigcup\{\widehat{D} \mid \widehat{D} \in \widehat{\mathcal{D}}\}$ where the sets $\widehat{D} \in \widehat{\mathcal{D}}$ are disjoint copies of the intervals $D \in \mathcal{D}$. The possible transitions of points in the tower with respect to this Markov partition are described by the Markov diagram. The lifted map $\widehat{f}$ respects the Markov diagram and satisfies $\pi \circ \widehat{f} = f \circ \pi$, where $\pi : \widehat{X} \to X$ is the natural projection that maps every set $\widehat{D} \in \widehat{\mathcal{D}}$ identically back to its master copy $D \subset X$. A good way to think of the Markov extension is to imagine $\widehat{X}$ as a tower with infinitely many *levels* $\widehat{D}$ piled up over the *base* X, and π as the vertical projection onto the base. For that reason $\widehat{X}$ is also called the *Hofbauer tower* for f.

In Chapter 2 we explicitly construct the Markov extension and describe the combinatorial properties that characterize its transition diagram. It turns out that the Markov diagram allows only very limited transitions: The levels can be grouped into two different types, denoted by $\widehat{\mathcal{D}}^+ := \{\widehat{D}_n^+ \mid n \in \mathbb{N}\}$ and $\widehat{\mathcal{D}}^- := \{\widehat{D}_n^- \mid n \in \mathbb{N}\}$, such that for every level $\widehat{D}_n^+$ there is an arrow $\widehat{D}_n^+ \to \widehat{D}_{n+1}^+$ (i.e., one can *climb up* one level in the tower) and for some special levels—called the *critical levels*—there is a second arrow $\widehat{D}_n^+ \to \widehat{D}_{\widetilde{n}+1}^-$ to the successor of a certain critical level $\widehat{D}_{\widetilde{n}}^-$ on the other side which lies below $\widehat{D}_n^+$ (i.e., one can *jump down* to a lower level on the other side). A symmetric statement holds for $\widehat{D}_n^-$. The indices of the critical levels are called the *cutting times* and are numbered $(S_k^\pm)_{k\geq 0}$ in increasing order. The transition diagram is completely determined by the *kneading map* $Q = (Q^+, Q^-) : \mathbb{N} \to \mathbb{N}_\infty \times \mathbb{N}_\infty$, where $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$, from which the cutting times and the successors of the critical levels on the other side can be calculated. The characteristic feature of a kneading map Q is that it satisfies the *Hofbauer condition* (Proposition 2.10). The information about the combinatorial peculiarities of the Markov diagram is then used to prove the following decomposition into irreducible components.

Theorem. *The Markov diagram can be decomposed into a finite or infinite chain* $\widehat{\mathcal{T}}_1 < \widehat{\mathcal{X}}_1 < \widehat{\mathcal{T}}_2 < \widehat{\mathcal{X}}_2 < \dots$ *of irreducible components* $\widehat{\mathcal{X}}_i$ *with* (*possibly void*) *transient parts* $\widehat{\mathcal{T}}_i$ *in between, where the ordering is such that there are paths traversing the components from left to right but no paths in the other direction.* (Theorem 2.22) ∎

We introduce two types of renormalizations for Lorenz maps, *proper* and *nonproper* renormalizations. Roughly speaking, a Lorenz map is renormalizable if one can find two branches of some iterates f^n and f^m to the left and to the right of the critical point, respectively, such that the restriction of those two branches to the interval $[f^m(c^+), f^n(c^-)]$

looks like a Lorenz map which has been restricted to its dynamical interval (see Figure 1.1, small box). If the two branches extend to a larger interval $[\widetilde{p},\widetilde{q}]$ such that the restriction to $[\widetilde{p},\widetilde{q}]$ looks like a complete Lorenz map again (see Figure 1.1, large box) then the renormalization is *proper*.

There is a close relation between the above decomposition of the Markov diagram and the possible renormalizations for the Lorenz map.

THEOREM. *Assume that the Lorenz map f has no periodic attractor and let $(\widehat{\mathcal{D}},\rightarrow)$ be its Markov diagram. Then the following holds.*

1. *The proper renormalizations of f correspond to the irreducible components of $\widehat{\mathcal{D}}$.*
2. *The nonproper renormalizations of f correspond to the transient parts of $\widehat{\mathcal{D}}$ between the irreducible components.*

The Markov diagram of the renormalized map can essentially be obtained from the Markov diagram of the original map by removing finitely many links from the beginning of the above chain. (Theorem 2.32) ■

Hopf decompositions and attractors. The smooth structure on the Markov extension makes it possible to study the time evolution of smooth densities under the action of $\widehat{f}$ by means of the *transfer operator* $\mathrm{P}_{\widehat{f}}$ (also known as the *Perron–Frobenius operator*) on the tower. It is defined with respect to a reference measure $\widehat{m}$ which is a natural lift of the Lebesgue measure onto the levels of the tower. The transfer operator $\mathrm{P}_{\widehat{f}}$ on the tower projects down to the transfer operator P_f for the original system (X,f) in the sense that $\mathrm{P}_f = \mathrm{P}_\pi \circ \mathrm{P}_{\widehat{f}}$, where P_π is the transfer operator for the canonical projection $\pi : \widehat{X} \to X$.

The transport of mass by $\mathrm{P}_{\widehat{f}}$ is mainly governed by two influences, on the one hand by the structure of the Markov diagram and on the other hand by the distortion of mass due to the nonlinearity of the map $\widehat{f}$. In order to control the distortion we will assume that the map has negative Schwarzian derivative, i.e., that f is a $\mathcal{C}^3$-Lorenz map with $f' > 0$ and $Sf := f'''/f' - \frac{3}{2}(f''/f')^2 < 0$ on both sides of the discontinuity.

For Lorenz maps with negative Schwarzian derivative the Koebe Principle gives good estimates for the distortion on monotonic branches as long as one stays away from the endpoints of branches (Lemma 3.3). This fact is reflected by the presence of a large positive cone of smooth *regular densities* $\widehat{\psi}$ which is invariant by $\mathrm{P}_{\widehat{f}}$ and for which the quotients $\widehat{\psi}(\widehat{x})/\widehat{\psi}(\widehat{y})$ are bounded uniformly in $\widehat{\psi}$ on every compact subset of the tower (Proposition 3.7). As a consequence it possible to derive strong statements about the conservative and dissipative part and the existence of absolutely continuous invariant measures for $\widehat{f}$. Based on the results of Keller [39] we prove that the Markov extension has the following Hopf decomposition.

THEOREM. *Either $\widehat{f}$ is purely dissipative or the Markov diagram has a maximal irreducible component $\widehat{\mathcal{X}}_{\mathrm{m}}$ and $\widehat{f}$ is conservative on the union $\widehat{X}_{\mathrm{m}}$ of all levels that belong to $\widehat{\mathcal{X}}_{\mathrm{m}}$. If $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then it is ergodic w.r.t. the lifted Lebesgue measure $\widehat{m}$ and there is a unique σ-finite absolutely continuous invariant measure $\widehat{\mu}$ for $\widehat{f}$ which is equivalent to $\widehat{m}$ on $\widehat{X}_{\mathrm{m}}$. Even more, if $\widehat{\mu}$ is finite then the measure preserving system $(\widehat{f},\widehat{\mu})$ is the product of an exact system with a finite rotation.* (Theorem 3.14) ■

The Hopf decomposition already provides a lot of information about the asymptotic behaviour of typical points on the tower. For example, if $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then $\omega(\widehat{x}) = \widehat{X}_{\mathrm{m}}$ for $\widehat{m}$-a.e. point $\widehat{x} \in \widehat{X}$ and if $\widehat{f}$ is purely dissipative then for $\widehat{m}$-a.e. point $\widehat{x} \in \widehat{X}$ the distance $r_n(\widehat{x})$ of the nth iterate $\widehat{x}_n$ to the endpoints of its level $\widehat{D}[\widehat{x}_n]$ tends to zero as $n \to \infty$ (Theorem 3.27). This information can be projected down to the original system and one obtains the following information about the global attractor of a Lorenz map f with negative Schwarzian derivative.

THEOREM. *If f is a Lorenz map with negative Schwarzian derivative then f has a unique global attractor A which is the union of one or two minimal attractors and for m-almost every point x the ω-limit set* coincides *with a minimal attractor. More precisely, one of the following three cases applies.*

1. *If f has an attracting periodic orbit then A is the union of one or two attracting periodic orbits and each attracts at least one of the critical values c_1^+ and c_1^-.*

Now assume that f has no attracting periodic orbits. Then $\omega(x) = A$ for m-a.e. x and one of the following holds.

2. *If f is infinitely often renormalizable then $A = \omega(c_1^+) = \omega(c_1^-)$. If f is infinitely often properly renormalizable then A is a Cantor set.*
3. *If f is finitely often renormalizable then either*

(a) *A is a finite union of intervals*

or A is a nowhere dense set which is of one of the following types:

(b) $A = \omega(c_1^+) = \omega(c_1^-)$.
(c) $A = \omega(c_1^\pm) \supset \mathrm{cl}(\mathrm{orb}(c_1^\mp))$ *with* $c_1^\pm \in \omega(c_1^\pm)$.
(d) $A = \mathrm{cl}(\mathrm{orb}(c_1^\pm)) \supset \mathrm{cl}(\mathrm{orb}(c_1^\mp))$ *with* $c_1^\pm \notin \omega(c_1^\pm)$.

The only possibility where two attractors can coexist is the case of two attracting periodic orbits. (Theorem 3.44) ∎

Here all symbols "±" and "∓" should be replaced either by the upper signs or by the lower signs simultaneously.

If $\widehat{f}$ has a finite absolutely continuous invariant measure $\widehat{\mu}$ then it can be projected from the tower to the base and one obtains an absolutely continuous ergodic invariant measure μ of positive entropy for f (Theorem 3.55). It was already said that such a measure is an SBR-measure. If this is not the case then still the following can be said about the weak limit sets of the empirical distributions for typical points.

THEOREM. *If f is a Lorenz map with negative Schwarzian derivative which has no absolutely continuous invariant probability measure of positive entropy then $\omega^*(x)$ is contained in the convex closure of $\omega^*(c_1^+) \cup \omega^*(c_1^-)$ for a.e. $x \in X$.* (Theorem 3.61) ∎

The reason for this is that the orbit of a typical point $\widehat{x}$ spends more and more of its time climbing very high up in the tower, which implies that on average there are longer and longer blocks $x_{m+1}, \ldots, x_{m+n}$ where the orbit of x shadows one of the critical orbits $c^\pm$ in the sense that x_{m+i} and $c_i^\pm$ lie on the same side of the critical point for $i = 1, \ldots, n$. Although this does not mean that the distance $\mathrm{dist}(x_{m+i}, c_i^\pm)$ is small during the entire shadow—even more since one cannot exclude the existence of wandering intervals for Lorenz maps in general—it is possible to show that at least it is small most of the time when the length of the shadow is large enough (Proposition 3.42).

Kneading theory. The most important property that distinguishes one-dimensional dynamical systems from higher-dimensional systems is the presence of an order structure on $\mathbb{R}$. This order structure is heavily exploited by kneading theory: Since the order between two points x and y is preserved by a Lorenz map under iteration as long as the iterates x_n and y_n lie on the same side of the discontinuity, the information about the relative position of the iterates of a point x with respect to the discontinuity is of great importance. This information can be encoded in a binary sequence $\zeta(x) := (\zeta_n(x))_{n\geq 0}$ which is called the *itinerary* of x. The set of possible itineraries is completely determined by the *kneading invariant* $\nu := (\nu^+, \nu^-)$, a pair of binary sequences consisting of the left hand side and right hand side limit of the itineraries at the discontinuity. In Section 4.1 the kneading invariant is introduced and a necessary and sufficient *admissibility condition* is given for an arbitrary pair of binary sequences to occur as the kneading invariant of some Lorenz map.

The kneading invariant and the Markov diagram are two equivalent methods to describe the combinatorial behaviour of a Lorenz map in the sense that it is possible to translate the combinatorial information contained in the kneading invariant into the language of the Markov diagram, and vice versa. The key to this equivalence is the fact that the cutting times of the Hofbauer tower and hence the kneading map can be determined through a splitting technique from the kneading invariant (see Section 4.2). The splitting technique was already used by Hofbauer and then systematized by Bruin [4, 5] and Sands [55].

As a result we obtain a reformulation of the *admissibility condition* for kneading invariants in the language of Markov diagrams, i.e., a necessary and sufficient combinatorial condition for an abstract graph to occur as the Markov diagram of some Lorenz map.

Theorem. *A pair $(\nu^+, \nu^-) = (10*, 01*)$ of binary sequences is realizable as kneading invariant of a Lorenz map if and only if it satisfies one of the following equivalent conditions*:

1. $\sigma\nu^+ \leq \sigma^n\nu^\pm \leq \sigma\nu^-$ *for every* $n \in \mathbb{N}$, *where* σ *is the shift map.*

2. *The differences between consecutive cutting*$^\pm$ *times are cutting*$^\mp$ *times*, *i.e., there exist integers* $Q^\pm(k)$, $k \geq 1$, *such that* $S^\pm_k - S^\pm_{k-1} = S^\mp_{Q^\pm(k)}$ *and the kneading map* $Q^\pm$ *satisfies the* Hofbauer condition:

$$(Q^\pm(k+j))_{j\geq 1} \geq (Q^\pm(Q^\mp Q^\pm(k)+j))_{j\geq 1} \quad \forall k \geq 1,$$

where the ordering $\leq$ *is just the lexicographical ordering of sequences.*

(Theorems 4.12 and 4.19) ■

Theorem 4.19 contains more equivalent characterizations of admissibility which are useful for various different purposes. We refrained from stating them here, since it would have required a lot of extra notation. Although at first sight the formulation of admissibility in terms of the kneading map looks more complicated than the other one, it is in practice much more powerful, since it contains the essential information about the topological properties of the Lorenz map in a condensed form. This makes it very easy to construct Lorenz maps with prescribed topological properties, e.g., renormalizable maps or maps where the critical points are recurrent, respectively not recurrent, and so on.

Although the combinatorial classification of all admissible Markov diagrams is the foundation for the results in Chapter 2, the kneading theory is postponed until Chapter 4, since the arguments are rather technical and the details of the proofs are not required to understand the analysis of the Markov diagram.

Families of Lorenz maps. In order to obtain a more global picture of the dynamics of Lorenz maps we study families of Lorenz maps which depend smoothly on some parameters. Such families occur naturally in the study of homoclinic bifurcations of families of vector fields with singularities. We ask the following questions:

Does every admissible kneading invariant occur in a $\mathcal{C}^1$-family of Lorenz maps?

DEFINITION (Full family). A family $\mathcal{F}$ of Lorenz maps is called a *full family* if every admissible kneading invariant ν occurs as the kneading invariant of some map $f \in \mathcal{F}$.

An obvious requirement for a Lorenz family which contains maps of all possible combinatorial types is that it should at least contain maps with all different kinds of branches, i.e., by adjusting the parameters it should be possible to tune the length of both branches independently over the whole range—from "short" (i.e., not critical) branches to "long" (i.e., surjective) branches. To achieve this we assume that the family depends $\mathcal{C}^1$ smoothly on two real parameters and impose some reasonable conditions on the family in order to guarantee that it is rich enough and sufficiently well behaved (see Definition 5.1). Prototypes of such $\mathcal{C}^1$-families of Lorenz maps are the families

$$f_{a,b} : x \mapsto \begin{cases} -a + |x|^\alpha & \text{if } x > 0, \\ b - |x|^\alpha & \text{if } x < 0, \end{cases} \tag{1.1}$$

with a fixed constant $\alpha > 1$, in particular the *quadratic Lorenz family*:

$$f_{a,b} : x \mapsto \begin{cases} -a + x^2 & \text{if } x > 0, \\ b - x^2 & \text{if } x < 0. \end{cases} \tag{1.2}$$

The answer to the above question now is: Every $\mathcal{C}^1$-Lorenz family is a full family—well, almost. There is an exceptional set of kneading invariants which do not necessarily occur, namely the ones where one of the critical itineraries, say ν^+, is a shift of the other one (i.e., $\nu^+ = \sigma^k \nu^-$ for some $k > 0$) but ν^+ is not periodic. Whenever a Lorenz map has such a kneading invariant, a one-sided neighbourhood of the critical point is necessarily a wandering interval or contained in the basin of an inessential periodic attractor (cf. Remark 4.11). Such kneading invariants are definitely missing in families of maps with negative Schwarzian derivative like the quadratic Lorenz family (cf. Lemma 3.36 and Remark 3.46).

The following theorem shows that this is the only exception. We call it the *Full Family Theorem*, regardless of its imperfections.

Theorem. *Let* $(f_{a,b})_{(a,b)\in J}$ *be a* $\mathcal{C}^1$*-Lorenz family and let* $\nu = (\nu^+, \nu^-)$ *be an admissible kneading invariant such that either*

(i) ν *is expansive or*

(ii) *at least one of* ν^+ *and* ν^- *is periodic.*

Then there exists a parameter $(a, b) \in \operatorname{cl} J$ *such that* $f_{a,b}$ *has the kneading invariant* ν.

(Theorem 5.3) ∎

In Section 5.1 we present the classical approach for the proof of the Full Family Theorem using the Thurston map, which solves the finite version of the problem posed in the above theorem: Given a kneading invariant as in the theorem with the additional property that both itineraries ν^+ and ν^- are preperiodic, find a parameter where the Lorenz map is *post-critically finite* ([1]) and has the required kneading invariant. The problem is then reformulated as a fixed point problem for the Thurston map which is a continuous map from a finite-dimensional simplex into itself. This method was used by de Melo & van Strien [13] to prove fullness of continuous multimodal maps, and we demonstrate that it can easily be adapted to discontinuous maps like the Lorenz maps. Here we follow essentially the arguments of de Melo & van Strien [13] with some modifications due to Martens & de Melo [47].

The treatment of the Thurston map is a little bit independent of the rest of the chapter, since our proof of the full family is based on the analysis of the parameter dependence of the kneading invariant (see below). Nevertheless, we include it here for two reasons: First, the Thurston algorithm provides a convenient method to find Lorenz maps with specific combinatorial properties in the quadratic Lorenz family which can easily be implemented on a computer. To be honest, one has to cheat a little bit: The algorithm is based on the assumption that the Thurston map for the quadratic family is a contraction with respect to some suitably chosen metric and the fixed point is found by iterating an arbitrary initial value. Unfortunately, we are not able to show that this is indeed the case, but the fact that the Thurston algorithm works very nicely and reliable in practice (it was used to make some of the figures in this thesis) encourages this belief. Second, the Thurston map enters naturally when we study the question of monotonic dependence of the kneading invariant for the quadratic family in Section 5.7.

Where in parameter space does such an admissible sequence occur? Although an affirmative answer to the question whether there are full families is of independent theoretical interest, since it shows that the maps of such a family serve as prototypes for all possible types of Lorenz maps (at least from a topological viewpoint), we are not going to treat this question in isolation, but answer it as part of a global analysis of the parameter dependence of the kneading invariant. In order to do this we describe the "bifurcation diagram" of a $\mathcal{C}^1$-Lorenz family, which consists of a refining sequence of partitions of its parameter space that is obtained by distinguishing longer and longer initial parts of the kneading invariants.

([1]) This means that the set $\operatorname{orb}(c^+) \cup \operatorname{orb}(c^-)$ is finite, i.e., both critical points are preperiodic.

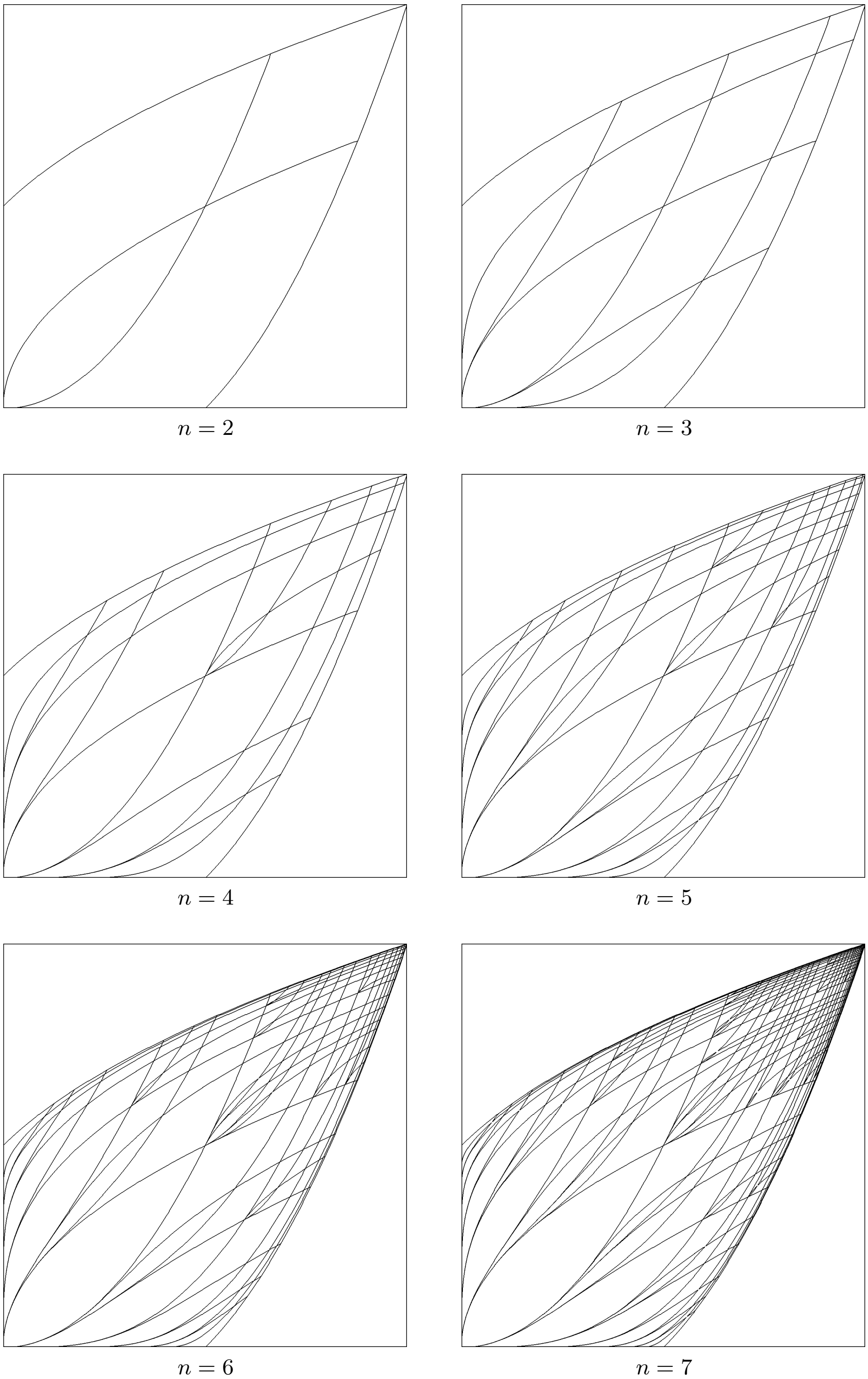


Fig. 1.2. The bifurcation diagram for the quadratic Lorenz family

Definition (Bifurcation diagram). Let $(f_{a,b})_{(a,b)\in J}$ be a $\mathcal{C}^1$-Lorenz family. For every $n \in \mathbb{N}$ and every admissible kneading invariant $\nu = (\nu^+, \nu^-)$ let $J_n(\nu)$ be the set of parameters (a, b) such that

1. the critical points $c^+(a, b)$ and $c^-(a, b)$ of the map $f_{a,b}$ are not periodic of any period less than or equal to n, and
2. the kneading invariant $\nu(a, b)$ of $f_{a,b}$ coincides up to the nth digit with ν.

The collection of partitions $\mathcal{J}_n := \{J_n(\nu) \mid \nu \text{ admissible}\}$ is called the *bifurcation diagram* of the Lorenz family $\mathcal{F}$.

The bifurcation diagram of the quadratic Lorenz family is shown in Figure 1.2 (see also the front cover of this book). The region shaped like an almond is the parameter space J, which is contained in the square $[0, 2] \times [0, 2]$. Every time when n is increased, some pieces of the partition break into smaller parts—which can be of two or four different combinatorial types—while other pieces remain unchanged. The underlying mechanism that lets the pieces break apart can be explained elegantly using the cutting and co-cutting times defined in Chapter 2. The Full Family Theorem above is now an immediate consequence of the following theorem.

Theorem. *Let $(f_{a,b})_{(a,b)\in J}$ be a $\mathcal{C}^1$-Lorenz family and let $\nu = (\nu^+, \nu^-)$ be an admissible kneading invariant. Then there is a decreasing sequence $(G_n(\nu))_{n\in\mathbb{N}}$ of nonvoid simply connected open regions $G_n(\nu) \subseteq J_n(\nu)$. The boundary of each set $G_n(\nu)$ is contained in the union of four differentiable arcs where one of the critical points is either periodic of some period less than or equal to n (which is determined by the cutting and co-cutting times) or is mapped onto one of the repelling fixed points in the boundary.*

For every parameter (a, b) in $G_\infty(\nu) := \bigcap_{n\in\mathbb{N}} G_n(\nu)$ the critical points are not periodic and the Lorenz map $f_{a,b}$ has the kneading invariant ν. If ν is expansive or periodic then $G_\infty(\nu)$ is nonvoid. (Theorem 5.38) ∎

This approach is inspired by Hofbauer & Keller [30], where a one-dimensional analogue of this theorem was used to prove fullness of $\mathcal{C}^1$-unimodal families. Although their proof is probably not so widely known than the two standard proofs—which are based on an intermediate value argument (see Collet & Eckmann [9]) and on the Thurston map (see de Melo & van Strien [13]), respectively—we think that this approach provides an interesting alternative and it sheds some additional light on the kneading theory developed in Chapter 4.

Does the kneading invariant depend monotonically on the parameters? If $(f_a)_{a\in\mathbb{R}}$ is a $\mathcal{C}^1$-family of *unimodal* maps then one says that the kneading invariant depends monotonically on the parameter if the map $a \mapsto \nu(a)$ is a monotonic map from the real line to the shift space $\{-, +\}^{\mathbb{N}}$ endowed with the signed lexicographical order. The following definition is a natural generalization of this notion for families depending on more than one parameter.

Definition (Monotonicity of the kneading invariant). Let $\mathcal{F} = (f_{a,b})_{(a,b)\in J}$ be a $\mathcal{C}^1$-Lorenz family. We say that the kneading invariant of $\mathcal{F}$ *depends monotonically on the* *parameters* if for every admissible kneading invariant ν and every n the set $J_n(\nu)$ is simply connected.

For the quadratic unimodal family $f_a : x \mapsto a - x^2$ it is was shown by Douady, Hubbard & Sullivan (2) that the kneading invariant depends monotonically on the parameter. They showed this by proving that the Thurston map is a contraction with respect to a suitable metric using deep results from Teichmüller theory. Recently, Tsujii [58, 59] gave a simplified proof for the monotonicity in the quadratic family based on the following observations. First, the question whether the kneading invariant depends monotonically on the parameters can be reduced to a local problem, namely to showing that

$$\frac{\mathrm{D}_a c_{n+1}(a)}{(f^n)'(c_1(a))} > 0$$

whenever the critical point is periodic of period $n+1$. The interpretation of this property is that whenever the critical point is periodic and the parameter a is increased, $c_{n+1}(a)$ moves towards the side of c corresponding to the larger kneading invariant. Second, there is a connection between the expression on the left hand side of the inequality and the Thurston map corresponding to the periodic orbit $c_0, \dots, c_n$. If $\mathrm{D}T(c_0, \dots, c_n)$ denotes the linearization of the Thurston map at the fixed point $(c_0, \dots, c_n)$ then

$$\frac{\mathrm{D}_a c_{n+1}(a)}{(f^n)'(c_1(a))} = \det(\mathrm{I} - \mathrm{D}T(c_0, \dots, c_n)).$$

Since the characteristic polynomial $\det(\lambda \mathrm{I} - \mathrm{D}T(c_0, \dots, c_n))$ is a real polynomial in λ which diverges to $+\infty$ as $\lambda \to +\infty$, the above inequality follows if the spectrum of $\mathrm{D}T$ is contained strictly inside the unit disk, i.e., if the linearized Thurston map is a contraction.

In Section 5.2 we give a condition analogous to the above inequality and show that the kneading invariant of $\mathcal{C}^1$-Lorenz families satisfying this condition—which we call *monotonic Lorenz families*—depends monotonically on the parameters. This condition has a dynamical interpretation, too: If both critical points of a map f_{a_0,b_0} are periodic then these two periodic orbits can be unfolded in an orientation preserving way by small perturbations of the parameters (cf. Lemma 5.25). In Section 5.7 we derive an equation similar to the one above which enables us to show that there is a close connection between local monotonicity of the kneading invariant and the question whether the Thurston map is locally a contraction. Unfortunately, we are not able to prove the missing link, namely the local contraction of the Thurston map. But our numerical studies of the bifurcation diagram for the quadratic family provide strong indications that the local monotonicity seems to hold everywhere in the parameter space, which strengthens our belief that the Thurston map is indeed a contraction.

(2) Unpublished. For a proof see the monograph of Milnor & Thurston [51].

2. Markov extensions

In this chapter we are going to explicitly construct the Markov extension of a Lorenz map $f : X \to X$ and describe the properties of its transition diagram. But first, let us give a precise definition of the class of maps we are considering.

2.1. Lorenz maps

2.1. Definition (Lorenz map). Let three points $p < c < q$ on the real line be given. A map $f : [p, q] \to [p, q]$ is called a *Lorenz map* if it satisfies the following conditions:

1. f is continuous and strictly increasing on (p, c) and (c, q).
2. f is discontinuous at c with one-sided limits $f(c^+) < c$ and $f(c^-) > c$.
3. $f(p) = p$, $f(q) = q$, and f has no other fixed points in (p, q).

The point c is called the *critical point* of f.

2.2. Definition. A Lorenz map $f : X \to X$ is a *$\mathcal{C}^2$-Lorenz map of order* α if there exist $\mathcal{C}^2$-diffeomorphisms ϕ^- and $\phi^+ : \mathbb{R} \to \mathbb{R}$ with $\phi^-(c) = \phi^+(c) = 0$ such that $f(x) = a + |\phi^+(x)|^\alpha$ for $x > c$ and $f(x) = b - |\phi^-(x)|^\alpha$ for $x < c$.

Denote by $\mathrm{orb}(x) := (x_n)_{n\in\mathbb{N}} := (f^n(x))_{n\in\mathbb{N}}$ the *orbit* of a point $x \in [p, q]$ and by $\omega(x)$ the *ω-limit set* of x, i.e., the set of accumulation points of $\mathrm{orb}(x)$. Since the interval $[a, b] := [f(c^+), f(c^-)]$ is invariant by the Lorenz map f and all points except for the fixed points p and q in the boundary enter $[a, b]$ after finitely many iterations (cf. Figure 1.1), it is sufficient to restrict attention to this invariant interval in order to study the asymptotic behaviour of the dynamical system. The interval $[a, b]$ is called the *dynamical interval* of f. This observation suggests the following alternative definition of Lorenz maps:

2.3. Definition (Lorenz map). Let three points $a < c < b$ on the real line be given. A map $f : [a, b] \to [a, b]$ is called a *Lorenz map* if it satisfies the following conditions:

1. f is continuous and strictly increasing on (a, c) and (c, b).
2. f is discontinuous at the critical point c with $f(c^+) = a$ and $f(c^-) = b$.
3. f has no fixed points in (a, b).

The point c is called the *critical point* of f.

The two definitions of Lorenz maps are more or less equivalent, since the important part of the dynamics takes place on the dynamical interval $[a, b]$ anyway, and since a Lorenz map $f : [a, b] \to [a, b]$ which is defined only on its dynamical interval can be extended arbitrarily to a Lorenz map $f : [p, q] \to [p, q]$ with fixed boundary points.

The situation here resembles the one for unimodal maps where there are also two common normalizations. In general, it will not be necessary to distinguish between the two normalizations for Lorenz maps and we just write $f : X \to X$, where X equals either $[a, b]$ or $[p, q]$. The only place where the distinction becomes important is Section 2.5, where renormalizations of Lorenz maps are studied. There the Lorenz maps according to Definition 2.1 are called *proper* Lorenz maps.

The reader may have noticed that the map f was left undefined at the critical point. This implies that objects like the orbit or the omega limit set are undefined for all *precritical points*, i.e., for all points which are mapped to the critical point c by some iterate of f. There are two ways to overcome this situation—both having advantages and disadvantages.

The first one is to replace the space X by a new space X' which is obtained by doubling all precritical points in X topologically and to define f on X' by taking one-sided limits. This method is very convenient as long as one is concerned with topological dynamics, in particular since X' is a compact metrizable space and the one-sided shift map on the space of f-itineraries is a continuous factor of the system (X', f) (1).

If one is interested in the measurable dynamics of smooth Lorenz maps, this method is not so ideal, since the doubling of points changes the metric of the underlying space, an effect which is not desirable. However, in measurable dynamics one is only interested in the behaviour of Lebesgue-typical points, and since the set of precritical points is countable, it is perfectly safe to leave the map undefined at the critical point and consider only orbits of points which are not precritical. Although the map is undefined at c it still makes sense to define $\mathrm{orb}(c^{\pm})$ by taking one-sided limits and $\omega(c^{\pm})$ as the set of limit points of $\mathrm{orb}(c^{\pm})$ with respect to the unmodified Euclidean metric. Or, even simpler, one just considers $\omega(c_1^{\pm})$ instead of $\omega(c^{\pm})$, which is well defined if $c^{\pm}$ is not periodic.

We prefer the former method when considering the topological dynamics and the latter when considering the measurable dynamics. In both cases we will just write $f : X \to X$, since it is clear from the context what the underlying space X is.

2.2. The Hofbauer tower. Fix a Lorenz map f and assume that none of its critical points c^+ and c^- is preperiodic (2). In this case the construction of the Markov extension is very simple and the idea behind it is most lucid. The adjustments for the preperiodic case will be discussed afterwards.

2.4. Definition (Hofbauer tower). Let $\mathcal{Z} := \{Z^-, Z^+\} := \{(p, c), (c, q)\}$ denote the partition of the interval $X = (p, q)$ into maximal open intervals of monotonicity of f and let $\mathcal{Z}_n := \bigvee_{n \in \mathbb{N}} f^{-n}(\mathcal{Z})$ be the partition into *n-cylinder sets* of f. Now let $\mathcal{D} \supseteq \mathcal{Z}$ be the smallest family of intervals satisfying

$$f(D \cap Z^{\pm}) \in \mathcal{D} \quad \text{whenever} \quad D \in \mathcal{D} \text{ and } D \cap Z^{\pm} \neq \emptyset \tag{2.1}$$

and take a collection $\widehat{\mathcal{D}} := \{\widehat{D} \mid D \in \mathcal{D}\}$ of disjoint copies $\widehat{D} := D \times \{D\}$ of the intervals

(1) For more details see Chapter 4.

(2) “Preperiodic” always means “periodic or strictly preperiodic”.

$D \in \mathcal{D}$. Their union $\widehat{X} := \bigcup_{\widehat{D} \in \widehat{\mathcal{D}}} \widehat{D}$ is called the *Hofbauer tower* associated with the map f and the sets $\widehat{D}$ are called the *levels* of the tower.

Note that the tower depends on the map f, although this is not indicated by the notation.

The terms *tower* and *level* originate from the fact that one can imagine all these intervals piled up one upon the other, in this way forming the floors of an infinite tower built on the base $X = (p, q)$. There is a natural "vertical" projection $\pi : \widehat{X} \to X$ from the tower onto its base, given by $\pi(\widehat{x}) := x$ for $\widehat{x} = (x, D) \in \widehat{D}$.

2.5. Definition (Markov extension). The (canonical) *Markov extension* for f is the dynamical system $(\widehat{X}, \widehat{f})$ where $\widehat{f} : \widehat{X} \to \widehat{X}$ is a lift of f to a map on the tower satisfying $\pi \circ \widehat{f} = f \circ \pi$, which is defined in the following way:

$$\widehat{f}(\widehat{x}) := (f(x), f(D \cap Z[x])) \quad \text{for } \widehat{x} = (x, D) \in \widehat{D}. \tag{2.2}$$

Here $Z[x]$ denotes the element of the partition $\mathcal{Z}$ containing x. The value $\widehat{f}(\widehat{c})$ is left undefined for all *critical points* $\widehat{c} \in \pi^{-1}(c)$.

The partition into the n-cylinder sets for $\widehat{f}$ equals $\pi^{-1}(\mathcal{Z}_n) \vee \widehat{\mathcal{D}}$ and is denoted by $\widehat{\mathcal{Z}}_n$. For $\widehat{x} \in \widehat{X}$ let $\widehat{D}[\widehat{x}] \in \widehat{\mathcal{D}}$ be the level and $\widehat{Z}_n[\widehat{x}] \in \widehat{\mathcal{Z}}_n$ be the n-cylinder set to which $\widehat{x}$ belongs.

Note that the orbit of a point is only well defined if $\widehat{x}$ is not a precritical point. Therefore, whenever we are talking of the orbit of a point $\widehat{x}$ we will implicitly assume that this is the case.

2.6. Remark. If $\widehat{x}$ and $\widehat{x}'$ are two points from the same fibre $\pi^{-1}(x)$ then the iterates $\widehat{x}_n$ and $\widehat{x}'_n$ belong to the same fibre $\pi^{-1}(x_n)$ but in general to different levels in the tower. However, usually the orbits of $\widehat{x}$ and $\widehat{x}'$ will coalesce after some iterations: From (2.2) it follows by induction that $f^n(D \cap Z_n[x]) \in \mathcal{D}$ and that

$$\widehat{x}_n := \widehat{f}^n(\widehat{x}) = (f^n(x), f^n(D \cap Z_n[x])) \quad \text{for } \widehat{x} = (x, D) \in \widehat{D}, \tag{2.3}$$

where $Z_n[x] \in \mathcal{Z}_n$ is the n-cylinder set containing x. In particular, if $Z_n[x] \subseteq D$ then $f^n(D \cap Z_n[x]) = f^n(Z_n[x])$, whence $\widehat{D}[\widehat{x}_n]$ does not depend on $\widehat{D}[\widehat{x}]$ but only on $Z_n[x]$. Note that $Z_n[x] \subseteq D$ holds if and only if $\pi(\widehat{Z}_n[\widehat{x}]) = Z_n[x]$. This leads to the following definition: A cylinder set $\widehat{Z}_n \in \widehat{\mathcal{Z}}_n$ is called a *complete* cylinder set if $\pi(\widehat{Z}_n) \in \mathcal{Z}_n$, and an *incomplete* cylinder set if $\pi(\widehat{Z}_n) \notin \mathcal{Z}_n$. From what was just said before it follows that two complete n-cylinder sets $\widehat{Z}_n$ and $\widehat{Z}'_n$ with $\pi(\widehat{Z}_n) = \pi(\widehat{Z}'_n) = Z_n$ are mapped onto the same level in the tower by $\widehat{f}^n$. In particular, if $\widehat{x} \in \widehat{Z}_n$ and $\widehat{x}' \in \widehat{Z}'_n$ are two points of the same fibre $\pi^{-1}(x)$ then $\widehat{f}^n(\widehat{x}) = \widehat{f}^n(\widehat{x}')$.

Every level $\widehat{D}$ contains at most two incomplete n-cylinder sets, namely the leftmost and the rightmost of all n-cylinder sets inside $\widehat{D}$. For every point $\widehat{x} \in \widehat{D}$ one has the following alternative: Either there exists an integer such that $\widehat{Z}_n[\widehat{x}]$ is complete or one of the components of $\widehat{D} \setminus \{\widehat{x}\}$ is a homterval ([3]). If the Lorenz map is of class $\mathcal{C}^2$ and has no periodic attractors then it can be shown that homtervals of this type cannot exist ([4]),

([3]) An interval which contains no precritical points (see Section 3.4).
([4]) Cf. Lemma 3.36 and Corollary 3.37(6).

whence the former alternative must hold. So in this case the orbits of any two points $\widehat{x}, \widehat{x}'$ from the same fibre $\pi^{-1}(x)$ coalesce after finitely many iterations, in particular $\widehat{x}$ and $\widehat{x}'$ have the same asymptotic behaviour. $\diamond$

By the very definition of the Markov extension the collection $\widehat{\mathcal{D}}$ is a Markov partition for $\widehat{f}$, since $\widehat{f}$ maps every level $\widehat{D} \in \widehat{\mathcal{D}}$ precisely onto one or two levels of the tower. The latter happens exactly when the critical point c is contained in $D = \pi(\widehat{D})$. Such levels are called *critical levels*. The possible transitions of points in the tower with respect to this Markov partition are described by the Markov diagram.

2.7. Definition (Markov diagram). The *Markov diagram* or *Markov graph* for $(\widehat{X}, \widehat{f})$ is the directed graph $(\widehat{\mathcal{D}}, \to)$ with edges given by the following relation:

$$\widehat{C} \to \widehat{D} :\Leftrightarrow \widehat{D} \subseteq \widehat{f}(\widehat{C}) \quad (\widehat{C}, \widehat{D} \in \widehat{\mathcal{D}}). \tag{2.4}$$

If $\widehat{C} \to \widehat{D}$ then $\widehat{D}$ is called a *successor* of $\widehat{C} \in \widehat{\mathcal{D}}$.

It will turn out in a moment that the Markov diagram of a Lorenz map has a very specific structure and allows only very limited transitions: The levels can be grouped into two different types, denoted by $\widehat{\mathcal{D}}^+ := \{\widehat{D}_n^+ \mid n \in \mathbb{N}\}$ and $\widehat{\mathcal{D}}^- := \{\widehat{D}_n^- \mid n \in \mathbb{N}\}$, such that for every level $\widehat{D}_n^\pm$ there is an arrow $\widehat{D}_n^\pm \to \widehat{D}_{n+1}^\pm$, and for some of the levels—namely for the critical ones—there is an additional arrow $\widehat{D}_n^\pm \to \widehat{D}_{\widetilde{n}+1}^\mp$ to the successor of a certain critical level $\widehat{D}_{\widetilde{n}}^\mp$ on the other side, which lies below $\widehat{D}_n^\pm$ (i.e., $\widetilde{n} < n$).

So instead of thinking of a single tower it is even better to imagine two towers standing next to each other—like the World Trade Centre in New York—where the level $\widehat{D}_n^+$ is the nth floor of the left tower and the level $\widehat{D}_n^-$ is the nth floor of the right tower ([5]). On either side one can move up from every level to the next higher one, and whenever one reaches a critical level, one can also jump down to a level on the other side. The sets D_n^+ and D_n^- are recursively defined by

$$D_0^\pm := Z^\pm, \qquad D_{n+1}^\pm := f(D_n^\pm \cap Z[c_n^\pm]) \tag{2.5}$$

where $Z[x]$ is the element of the partition $\mathcal{Z}$ containing the point x. If $D_n^\pm$ is not critical then $D_{n+1}^\pm$ is just $f(D_n^\pm)$.

2.8. Remark. It is necessary to pause here for a general remark. In the following we will frequently make use of the symbols "$\pm$" and "$\mp$" in order to minimize the amount of repetition for statements and expressions that are symmetric with respect to exchanging "$+$" and "$-$" signs. To avoid misunderstandings, let us agree upon the following convention: *All statements, expressions, theorems, etc., containing one of the symbols "$\pm$" and "$\mp$" are intended to be read twice, once using the upper sign and once using the lower sign for all symbols simultaneously.* In particular, we tried never to use the symbols "$\pm$" and "$\mp$" in the sense of "$+$ *or* $-$", even in cases where it was very tempting. A similar agreement holds for "$\frac{1}{0}$", "$\frac{0}{1}$", "$\lessgtr$", "left/right", etc. $\diamond$

([5]) The terms "next to each other", "left", and "right" should not be taken too literally, the intervals D_n^+ do not lie to the left of the intervals D_n^-. The terms are meant only schematically (cf. Figures 2.3 and 2.4).

2.9. Definition (Cutting$^\pm$ times). The integers n for which the levels $\widehat{D}_n^\pm$ are critical are called the *cutting$^\pm$ times* and denoted by $S_k^\pm$, $k \geq 0$, in increasing order. If there are only finitely many cutting$^\pm$ times $S_0^\pm, \ldots, S_{\gamma^\pm-1}^\pm$ then let $S_k^\pm := \infty$ for $k \geq \gamma^\pm$, and let $S_\infty^\pm := \infty$. The set of cutting$^\pm$ times is denoted by $\mathcal{S}^\pm := \{S_k^\pm \mid k \geq 0\}$.

From (2.5) it follows by induction that $D_n^\pm = f^n(Z_n^\pm)$, where $Z_n^\pm := Z_n[c^\pm]$ is the *central n-cylinder* on the right/left hand side of c, and that $Z_{n+1}^\pm$ is strictly contained in $Z_n^\pm$ if and only if n is a cutting$^\pm$ time. It follows that

$$D_n^+ = (c_n^+, c_{n-S^+\langle n\rangle}^-) \quad \text{and} \quad D_n^- = (c_{n-S^-\langle n\rangle}^+, c_n^-) \quad \text{for } n > 1, \tag{2.6}$$

where $S^\pm\langle n\rangle$ denotes the last cutting$^\pm$ time before and not including n. Comparing (2.5) to (2.1) and (2.4) it is evident that all sets $D_n^\pm$ belong to $\mathcal{D}^\pm$, whence $\mathcal{D}^+ \cup \mathcal{D}^- \subseteq \mathcal{D}$, and that $\widehat{D}_n^\pm \to \widehat{D}_{n+1}^\pm$ holds for all n. Not so obvious is the above statement that the second successor of a critical $\mathcal{D}^\pm$-level is a $\mathcal{D}^\mp$-level, which implies the reverse inclusion $\mathcal{D} \subseteq \mathcal{D}^+ \cup \mathcal{D}^-$. It is the consequence of an important property of the cutting$^\pm$ times, namely that the differences between consecutive cutting$^\pm$ times are always cutting$^\mp$ times. More precisely, the following holds:

2.10. Proposition. *There exists a map $Q = (Q^+, Q^-) : \mathbb{N} \to \mathbb{N}_\infty \times \mathbb{N}_\infty$ such that*

$$S_k^\pm - S_{k-1}^\pm = S_{Q^\pm(k)}^\mp \quad \textit{for all } k \geq 1. \tag{2.7}$$

The map Q is called the kneading map *of f. It satisfies the* Hofbauer condition:

$$(Q^\pm(k+j))_{j\geq 1} \geq (Q^\pm(Q^\mp Q^\pm(k)+j))_{j\geq 1} \quad \textit{for all } k \geq 1, \tag{2.8}$$

where the ordering is the lexicographical ordering of sequences in $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$.

Proof. Cf. Theorem 4.19 in Section 4.3. ∎

2.11. Remark. To the reader familiar with kneading maps for unimodal maps we point out that in contrast to the situation there (cf. Bruin [5]) the kneading map here does *not* have the property that $Q^\pm(k) < \infty$ implies $Q^\pm(k) < k$. In particular, there is no inequality of the type $S_k^\pm \leq 2S_{k-1}^\pm$.

Instead, the following holds: If $k, l \in \mathbb{N}$ is a pair such that $Q^+(k) < \infty$ and $Q^-(l) < \infty$ then either $Q^+(k) < l$ or $Q^-(l) < k$ (cf. part 4 of Lemma 2.24). In particular, this implies that if S_{k-1}^+ and S_{l-1}^- are cutting times then $n := S_{k-1}^+ + S_{l-1}^-$ is the "last chance to cut", i.e., if there is no further cutting$^+$ or cutting$^-$ time until time n (including n) then $S_k^+ = S_l^- = \infty$.

Another useful property is the following: Whenever $Q^\pm(k) < \infty$ then $Q^\mp Q^\pm(k) < k$. To see why this holds apply (2.7) twice to obtain $S_k^\pm > S_{Q^\pm(k)}^\mp > S_{Q^\mp Q^\pm(k)}^\pm$. Now since the cutting$^\pm$ times are indexed in increasing order, $S_{Q^\mp Q^\pm(k)}^\pm < S_k^\pm$ implies $Q^\mp Q^\pm(k) < k$. ◇

Property (2.7), together with (2.6), implies that

$$D_{S_k^+}^+ = (c_{S_k^+}^+, c_{S_k^+ - S_{k-1}^+}^-) = (c_{S_k^+}^+, c_{S_{Q^+(k)}^-}^-) \quad \text{and} \quad D_{S_k^-}^- = (c_{S_k^- - S_{k-1}^-}^+, c_{S_k^-}^-) = (c_{S_{Q^-(k)}^+}^+, c_{S_k^-}^-),$$

whence

$$f(D^{\pm}_{S^{\pm}_k} \cap Z^{\mp}) = D^{\pm}_{S^{\pm}_k+1} \qquad \text{by (2.5), and}$$
$$f(D^{\pm}_{S^{\pm}_k} \cap Z^{\pm}) = D^{\mp}_{S^{\mp}_{Q^{\pm}(k)}+1} \qquad \text{by (2.5) and (2.7).}$$

The kneading map completely determines the Markov diagram of the Hofbauer tower: First, it determines the cutting times through the recursion (2.7), and second, the arrows are given by

$$\widehat{D}^{\pm}_n \to \widehat{D}^{\pm}_{n+1} \quad \text{and} \quad \widehat{D}^{\pm}_{S^{\pm}_k} \to \widehat{D}^{\mp}_{S^{\mp}_{Q^{\pm}(k)}+1} \quad \text{for all } n,k \in \mathbb{N}. \tag{2.9}$$

The significance of the Hofbauer condition (2.8) for the transition diagram is not so easily unveiled. For the moment let us just mention that it is not only a necessary but also a sufficient condition for a map $Q : \mathbb{N} \to \mathbb{N}_\infty \times \mathbb{N}_\infty$ to occur as the kneading map of a Lorenz map. The Hofbauer condition will be treated in more detail in Chapter 4.

Since the critical points were assumed not to be preperiodic, all sets $D^{\pm}_n$ are distinct. This means that the collection $\widehat{\mathcal{D}}^{\pm} = \{\widehat{D}^{\pm}_n = \widehat{D}^{\pm}_n \times \{\widehat{D}^{\pm}_n\} \mid D \in \mathcal{D}^{\pm}\}$ can be identified naturally with the collection $\widehat{\mathcal{D}}^{\pm} = \{\widehat{D}^{\pm}_n := D^{\pm}_n \times \{(\pm, n)\} \mid n \in \mathbb{N}\}$. Now if the critical point $c^{\pm}$ is preperiodic then the two definitions are not equivalent any more, because the former definition of $\widehat{\mathcal{D}}^{\pm}$ yields only finitely many $\widehat{\mathcal{D}}^{\pm}$-levels. Although Definition 2.4 is the most natural way to define the tower and although it is preferable to have a finite tower on one side or both, the necessity of having to distinguish between finite and infinite towers adds unnecessary and annoying complications to the proofs. Another, more weighty reason is that we intend to relate the transition diagram of the Markov extension to the kneading invariant of f in Chapter 4. For the two-sided infinite tower the Markov diagram can be determined solely from the kneading invariant, and vice versa, whereas in general it is not possible to tell from a preperiodic kneading invariant alone whether a critical point is preperiodic or only attracted to a periodic attractor.

Since the case where one or both critical points is preperiodic in general requires only trivial modifications, we will always assume that the tower is infinite on both sides. This could also be justified by the following more technical definition of the Markov extension which yields a two-sided infinite tower and contains the Markov extension of Definition 2.5 as a natural factor.

2.12. Definition (Markov extension). The (canonical) *Markov extension* for f is the dynamical system $(\widehat{X}, \widehat{f})$ consisting of the *Hofbauer tower* $\widehat{X} := \bigcup_{n\in\mathbb{N}}(\widehat{D}^{+}_n \cup \widehat{D}^{-}_n)$, where $\widehat{D}^{\pm}_n := D^{\pm}_n \times \{(\pm, n)\}$ and the sets $D^{\pm}_n$ are defined recursively by equation (2.5), and the map $\widehat{f} : \widehat{X} \to \widehat{X}$, given by

$$\widehat{f}(\widehat{x}) := \begin{cases} (f(x), \pm, n+1) \in \widehat{D}^{\pm}_{n+1} & \text{if } c \notin (c^{\pm}_n, x), \\ (f(x), \mp, S^{\mp}_{Q^{\pm}(k)} + 1) \in \widehat{D}^{\mp}_{S^{\mp}_{Q^{\pm}(k)}+1} & \text{if } c \in (c^{\pm}_n, x) \text{ and } n = S^{\pm}_k, \end{cases}$$

for $\widehat{x} := (x, \pm, n) \in \widehat{D}^{\pm}_n$. If $x = c$ then $f(\widehat{x})$ is left undefined. The natural projection $\pi : \widehat{X} \to X$ is given by $\pi(x, \pm, n) := x$.

From a graph-theoretical viewpoint the noncritical levels are not very interesting because they have precisely one successor, and one could restrict oneself to the set $\widehat{\mathcal{C}}$ of critical levels $\widehat{C}_k^{\pm} := \widehat{D}^{\pm}_{S_k^{\pm}}$ and consider the *induced Markov diagram* $(\widehat{\mathcal{C}}, \rightsquigarrow)$, where the arrows are given by $\widehat{C} \rightsquigarrow \widehat{C}'$ iff there exists a path $\widehat{C} \to \cdots \to \widehat{C}'$ which is contained entirely in $\widehat{\mathcal{D}} \setminus \widehat{\mathcal{C}}$ except for the endpoints $\widehat{C}$ and $\widehat{C}'$. From every level $\widehat{C}_k^{\pm}$ of the *induced tower* $\widehat{X}' := \bigcup_{\widehat{C} \in \widehat{\mathcal{C}}} \widehat{C}$ there are two arrows ([6]) leading to

$$\widehat{C}_k^{\pm} \rightsquigarrow \widehat{C}_{k+1}^{\pm} \quad \text{and} \quad \widehat{C}_k^{\pm} \rightsquigarrow \widehat{C}^{\mp}_{Q^{\pm}(k)+1}.$$

The induced Markov diagram is nothing else than the Markov diagram for the *induced map* $\widehat{F} : \widehat{X}' \to \widehat{X}'$, $\widehat{x} \mapsto \widehat{f}^{n(\widehat{x})}(\widehat{x})$, where $n(\widehat{x}) := \min\{n > 1 \mid \widehat{f}^n(\widehat{x}) \in \widehat{X}'\}$. Induced towers will be very convenient in Section 2.5 where renormalizations of Lorenz maps are studied, because the induced Markov tower for the renormalized map can be obtained from the induced tower of the original map simply by removing finitely many levels.

We conclude the section with a remark on the relation between the levels $D_n^{\pm}$ and the monotonic branches of iterates of f, and give a simple geometrical explanation of the fact that the difference between consecutive cutting$^{\pm}$ times is a cutting$^{\mp}$ time.

2.13. Definition (Closest preimage). A point $x \in X$, $x \neq c$, is called a *closest preimage* to the left/right of c if it is the endpoint of the central n-cylinder set $Z_n^{\mp} := Z_n[c^{\mp}]$ for some integer n.

It was already mentioned that $D_n^{\mp}$ is the image $f^n(Z_n^{\mp})$ of the left/right central branch of f^n. This implies that the cutting$^{\mp}$ times are precisely the integers where a part of the right/left central cylinder is "cut off", i.e., where $Z_{n+1}^{\mp} \subset Z_n^{\mp}$. Consequently, the left/right closest preimages are precisely the preimages of order $S_k^{\mp}$ which lie closest to c. Denote the closest preimage of order $S_k^{\mp}$ on the left/right hand side by $z_k^{\mp}$. Figure 2.1 shows how the critical preimages are located. If there are infinitely many cutting$^{\mp}$ times then the sequence of closest preimages converges on the left/right hand side of c, but not necessarily against c itself. This is the case if and only if $\bigcap_{n\in\mathbb{N}} Z_n^{\mp} = \emptyset$ ([7]).

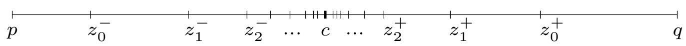


Fig. 2.1. The closest preimages to c

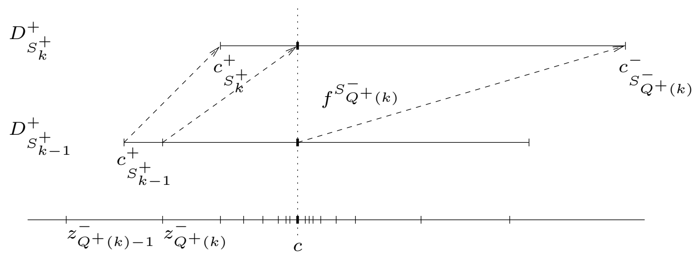


Fig. 2.2. The difference between two cutting$^+$ times

([6]) Except for $\widehat{C}^{\pm}_{\gamma^{\pm}}$ if $\gamma^{\pm} < \infty$, which has only one successor.

([7]) For example, if f has negative Schwarzian derivative and no periodic attractors (cf. Section 3.4).

Now if $D^{\pm}_{S^{\pm}_{k-1}}$ is a critical level then the cutting time $S^{\pm}_k$ is determined by the number of times the interval $D^{\pm}_{S^{\pm}_{k-1}} \cap Z^{\mp}$ can be mapped by f until it contains the critical point. This number equals $f^{S^{\mp}_l}$ if $z^{\mp}_l$ denotes the closest preimage inside $D^{\pm}_{S^{\pm}_{k-1}} \cap Z^{\mp}$ which has the lowest index l. It follows that $S^{\pm}_k - S^{\pm}_{k-1} = S^{\mp}_l$, whence $Q^{\pm}(k) = l$ (cf. Figure 2.2). In other words, if $A^-_k := (z^-_{k-1}, z^-_k)$ and $A^+_k := (z^+_k, z^+_{k-1})$ then condition (2.7) is equivalent to the statement that $c^{\pm}_{S^{\pm}_{k-1}} \in A^{\mp}_{Q^{\pm}(k)}$ holds for all integers k.

2.3. The extended Hofbauer tower. The recurrence behaviour of the critical points c^+ and c^- plays an important role for the dynamics of a Lorenz map. Let us see how it is visible in the Hofbauer tower.

2.14. Definition. We call c^+_n a *closest return* of c^+ to c if $c^+_m, c^-_m \notin (c^+_n, c)$ for all $m < n$ ([8]). If $c^+_n < c$ (resp. $c^+_n > c$) then it is called a closest return on the left (resp. right) hand side. Closest returns of c^- are defined symmetrically.

If c^+_n is a closest return on the left hand side then $D^+_n = (c^+_n, c^-_{n-S^+\langle n\rangle})$ has to contain the critical point, otherwise $c^-_{n-S^+\langle n\rangle}$ would be contained in the interval (c^+_n, c^-) which is excluded by the very definition of a closest return. This implies that n is a cutting$^+$ time. Now since c^+_n is very close to c^-, the iterates of those two points will remain close together for a long time, which means that it will take a long time until the next cutting time occurs. Unfortunately, the same does not hold for closest returns of c^+ on the right hand side. They do not trigger a cutting$^+$ time and thus are not visible in the combinatorial structure of the tower. The reason for this is that the interval D^+_n is just a one-sided neighbourhood of the point c^+_n. This problem was caused by the fact that the tower was constructed following the orbit of the critical points c^+ and c^-. The natural solution to this problem is to follow the critical values c^+_1 and c^-_1 instead. Let

$$E^{\pm}_1 := Z^{\mp}, \quad E^{\pm}_{n+1} := f(E^{\pm}_n \cap Z[c^{\pm}_n]). \tag{2.10}$$

Having (2.1) in mind it is clear that one does not obtain any new intervals with this procedure. They are only labelled differently. But now $D^{\pm}_1 \cap Z[c^{\pm}_1]$ is strictly contained in $E^{\pm}_1$ and by comparing the recursions (2.5) and (2.10), we see that for $n \geq 2$ every set $D^{\pm}_n$ is strictly contained in $E^{\pm}_n$, which is a two-sided neighbourhood of $c^{\pm}_n$. Thus the tower built from the sets $E^{\pm}_n$ can be regarded as an extension of the tower built from the sets $D^{\pm}_n$, and therefore it is called the *extended Hofbauer tower*.

The integers n where the level $\widehat{E}^{\pm}_n$ is critical can be divided in two groups: The ones for which $\widehat{D}^{\pm}_n$ is also critical, which are just the cutting$^{\pm}$ times defined in the previous section, and the times where this is not the case. These are called the co-cutting$^{\pm}$ times.

2.15. Definition (Co-cutting$^{\pm}$ times). The integers n for which the level $\widehat{E}^{\pm}_n$ is critical but the level $\widehat{D}^{\pm}_n$ is not are called the *co-cutting*$^{\pm}$ *times* and denoted by $\widetilde{S}^{\pm}_k$, $k \geq 0$, in increasing order. If there are only finitely many co-cutting$^{\pm}$ times $\widetilde{S}^{\pm}_0, \ldots, \widetilde{S}^{\pm}_{\widetilde{\gamma}^{\pm}-1}$ then let $\widetilde{S}^{\pm}_k := \infty$ for $k \geq \widetilde{\gamma}^{\pm}$, and let $\widetilde{S}^{\pm}_{\infty} := \infty$. The set of co-cutting$^{\pm}$ times is denoted by $\widetilde{\mathcal{S}}^{\pm} := \{\widetilde{S}^{\pm}_k \mid k \geq 0\}$.

([8]) The notation (c^+_n, c) is used regardless of whether $c^+_n < c$ or $c^+_n > c$.

The first co-cutting$^\pm$ time $\widetilde{S}_0^\pm$ coincides with the minimal integer $n > 0$ such that $c_n^\pm$ is contained in $Z^\pm$, and by induction one obtains

$$E_n^+ = (c^+_{n-\widetilde{S}^+\langle n\rangle}, c^-_{n-S^+\langle n\rangle}) \quad \text{and} \quad E_n^- = (c^+_{n-S^-\langle n\rangle}, c^-_{n-\widetilde{S}^-\langle n\rangle})$$

for $n > \widetilde{S}_0^\pm$, where $\widetilde{S}^\pm\langle n\rangle$ denotes the last co-cutting$^\pm$ time before n. For $n \le \widetilde{S}_0^\pm$ one of the boundary points of $E_n^\pm$ is just the repelling fixed point p resp. q instead of $c^\pm_{n-\widetilde{S}^\pm\langle n\rangle}$. Setting

$$\widetilde{D}_n^+ := (c^+_{n-\widetilde{S}^+\langle n\rangle}, c_n^+) \quad \text{and} \quad \widetilde{D}_n^- := (c_n^-, c^-_{n-\widetilde{S}^-\langle n\rangle}),$$

such that $E_n^\pm = D_n^\pm \cup \{c_n^\pm\} \cup \widetilde{D}_n^\pm$, the cutting$^\pm$ times are the times when $D_n^\pm$ is critical and the co-cutting times are the times when $\widetilde{D}_n^\pm$ is critical. In Lemma 4.16 it will be shown that the differences between consecutive co-cutting$^\pm$ times are cutting$^\pm$ times, i.e.,

$$\widetilde{S}_k^\pm - \widetilde{S}_{k-1}^\pm = S^\pm_{\widetilde{Q}^\pm(k)}, \tag{2.11}$$

where the *co-kneading map* $\widetilde{Q} = (\widetilde{Q}^+, \widetilde{Q}^-)$ is defined analogously to the kneading map. Note the difference between equations (2.7) and (2.11): For the cutting times the "$\pm$"-sign changes, whereas for the co-cutting times it does not.

Now in the extended tower all closest returns of c^+ and c^- are registered:

2.16. PROPOSITION. *Every closest return c_n^+ of c^+ to c^- is a cutting$^+$ time and every closest return c_n^+ of c^+ to c^+ is a co-cutting$^+$ time.*

PROOF. The first part was already shown at the beginning of the section. The second part follows similarly. ∎

Cutting and co-cutting times are special cases of the following more general cutting times, which can be defined for arbitrary (non-precritical) points $x \in X$ and determine the range $D_n[x] := f^n(Z_n[x])$ of the monotonic branches of f^n at x:

2.17. DEFINITION (Cutting$^{\lessgtr}$ times). Let $Z_n[x]$ be the n-cylinder set which contains the point x. Whenever $Z_{n+1}[x] \subset Z_n[x]$, an interval $Z_n[x] \setminus Z_{n+1}[x]$ is "cut off" to the left or to the right of x. Denote the integers where this happens in increasing order by $S_k^<(x)$ and $S_k^>(x)$, respectively. The numbers $S_k^{\lessgtr}(x)$ are called the *left/right cutting times* or just the *cutting$^{\lessgtr}$ times* of x, respectively.

As before it can be shown by induction that

$$D_n[x] = (c^+_{n-S^<\langle n\rangle(x)}, c^-_{n-S^>\langle n\rangle(x)}) \quad \text{for } n > \max(S_0^<(x), S_0^>(x)),$$

where $S^{\lessgtr}\langle n\rangle(x)$ denotes the last left/right cutting time before n, and that the left/right cutting times are precisely the integers n where the left/right component of $D_n[x] \setminus \{x_n\}$ contains c. The relation to the cutting$^\pm$- and co-cutting$^\pm$-times is given by

$$S_k^\pm = S_k^{\gtrless}(c_1^\pm) + 1 \quad \text{and} \quad \widetilde{S}_k^\pm = S_k^{\lessgtr}(c_1^\pm) + 1.$$

2.4. The decomposition of the Markov diagram. Since the Markov diagram is just a countable state topological Markov chain, it is near at hand to search for invariant subsets and irreducible components in order to study the long time behaviour of paths in the tower. We will do this in a purely combinatorial fashion, i.e., we will not use the information that there is a “real” Lorenz map behind the Markov diagram, but only our knowledge about the possible transitions which is collected in the following definition.

2.18. Definition (Lorenz graph). Let $\widehat{\mathcal{D}} = \widehat{\mathcal{D}}^{+} \cup \widehat{\mathcal{D}}^{-}$ be the disjoint union of two countably infinite sets $\widehat{\mathcal{D}}^{\pm} = \{\widehat{D}_n^{\pm} \mid n \in \mathbb{N}\}$. A directed graph $(\widehat{\mathcal{D}}, \to)$ is called a *Lorenz graph* if there exist numbers $S_k^{\pm} \in \mathbb{N}_\infty$ and a map $Q = (Q^+, Q^-) : \mathbb{N} \to \mathbb{N}_\infty \times \mathbb{N}_\infty$ satisfying the Hofbauer condition (2.8) such that the arrows are given by (2.9).

It was already mentioned that these conditions have to be necessarily satisfied by the Markov diagram of a Lorenz map. The converse, namely that every Lorenz graph can be realized as Markov diagram of some Lorenz map, will be shown in Chapter 4.

2.19. Definition. Let $\widehat{A}, \widehat{B} \in \widehat{\mathcal{D}}$ be two levels in the Hofbauer tower. If there exists a *path* $\widehat{A} = \widehat{D}_0 \to \widehat{D}_1 \to \ldots \to \widehat{D}_n = \widehat{B}$ (*of length* n) that connects $\widehat{A}$ with $\widehat{B}$ then this is denoted by $\widehat{A} \overset{n}{\twoheadrightarrow} \widehat{B}$ or just by $\widehat{A} \twoheadrightarrow \widehat{B}$. Sometimes $\widehat{A} \twoheadrightarrow \widehat{B}$ also denotes the connecting path itself. Similarly, for two subsets $\widehat{\mathcal{A}}$ and $\widehat{\mathcal{B}}$ of $\widehat{\mathcal{D}}$ the notation $\widehat{\mathcal{A}} \overset{n}{\twoheadrightarrow} \widehat{\mathcal{B}}$ indicates that there is a path of length n starting from a level $\widehat{A} \in \widehat{\mathcal{A}}$ and ending in a level $\widehat{B} \in \widehat{\mathcal{B}}$.

2.20. Definition (Irreducible component). The following two relations on $\widehat{\mathcal{D}}$:

$$\begin{aligned} \widehat{C} \preccurlyeq \widehat{D} &:\Leftrightarrow \widehat{C} = \widehat{D} \text{ or } \widehat{C} \twoheadrightarrow \widehat{D}, \\ \widehat{C} \asymp \widehat{D} &:\Leftrightarrow \widehat{C} \preccurlyeq \widehat{D} \text{ and } \widehat{C} \succcurlyeq \widehat{D}, \end{aligned}$$

define a quasi-order and its associated equivalence relation, respectively. The latter induces equivalence classes $[\widehat{C}]$, $[\widehat{D}]$ which can be ordered by $[\widehat{C}] \le [\widehat{D}] :\Leftrightarrow \widehat{C} \preccurlyeq \widehat{D}$.

A level $\widehat{C}$ is called *recurrent* if $\widehat{C} \twoheadrightarrow \widehat{C}$, and *transient* if $\widehat{C} \not\twoheadrightarrow \widehat{C}$. If $\widehat{C}$ is recurrent then its equivalence class $\mathcal{X}_i := [\widehat{C}]$ is called an *irreducible component* of the Markov diagram. An irreducible component $\mathcal{X}_i$ is called *maximal* if it is maximal w.r.t. the ordering $\le$ in the set of all equivalence classes.

In analogy to Markov chains one defines the *period* p of an irreducible component $\widehat{\mathcal{X}}_i$ as the greatest common divisor of the lengths of all closed paths in $\widehat{\mathcal{X}}_i$. If the period is one then $\widehat{\mathcal{X}}_i$ is called *aperiodic*. Otherwise, $\widehat{\mathcal{X}}_i$ can be partitioned into sets $\widehat{\mathcal{X}}_{i,1}, \ldots, \widehat{\mathcal{X}}_{i,p}$ such that $\widehat{\mathcal{X}}_{i,j} \overset{n}{\twoheadrightarrow} \widehat{\mathcal{X}}_{i,j'}$ implies $n \equiv j' - j \bmod p$.

Before stating the next theorem we need to introduce some more notation.

2.21. Definition. Let $\widehat{\mathcal{X}}_i$ be an irreducible component of period p and $k, l \in \mathbb{N}_\infty$. Set

$$\begin{aligned} \widehat{X}_i &:= \bigcup \{\widehat{D} \mid \widehat{D} \in \widehat{\mathcal{X}}_i\}, \\ \widehat{X}_{i,j} &:= \bigcup \{\widehat{D} \mid \widehat{D} \in \widehat{\mathcal{X}}_{i,j}\} \quad \text{for } 1 \le j \le p, \\ \widehat{\mathcal{X}}_i^{\wedge} &:= \{\widehat{D} \in \widehat{\mathcal{D}} \setminus \widehat{\mathcal{X}}_i \mid \exists\, \widehat{C} \in \widehat{\mathcal{X}}_i : \widehat{C} \twoheadrightarrow \widehat{D}\}, \\ \widehat{\mathcal{D}}_{kl}^{\wedge} &:= \{\widehat{D}_i^{+} \mid i > S_k^{+}\} \cup \{\widehat{D}_j^{-} \mid j > S_l^{-}\}. \end{aligned}$$

For a subset $\widehat{\mathcal{A}}$ of $\widehat{\mathcal{D}}$ let $\sigma(\widehat{\mathcal{A}}) := \{\widehat{D} \in \widehat{\mathcal{D}} \mid \exists\, \widehat{A} \in \widehat{\mathcal{D}} : \widehat{A} \to \widehat{D}\}$ denote the set of all successors of levels contained in $\widehat{\mathcal{A}}$. The set $\widehat{\mathcal{A}}$ is called *invariant* if $\sigma(\widehat{\mathcal{A}}) \subseteq \widehat{\mathcal{A}}$.

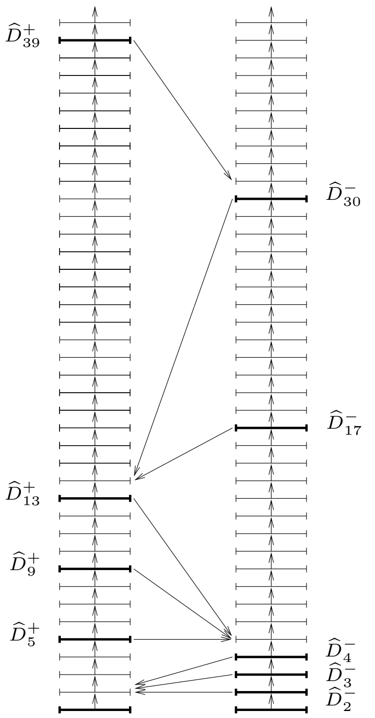


Fig. 2.3. A Markov diagram which is entirely transient

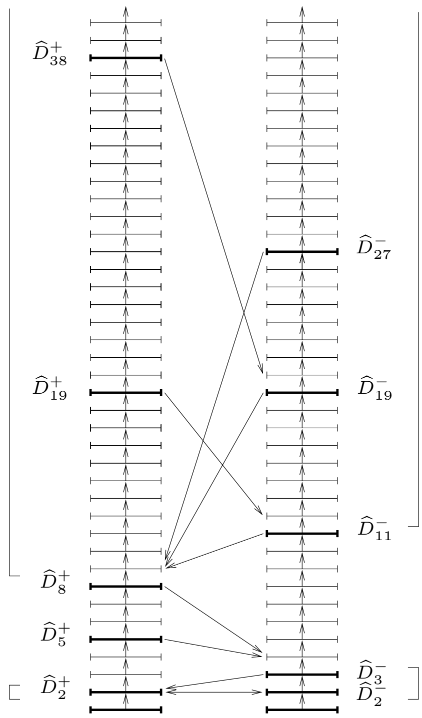


Fig. 2.4. A Markov diagram with two irreducible components

2.22. Theorem. *The Markov diagram* $(\widehat{\mathcal{D}}, \to)$ *of a Lorenz map can be decomposed as follows: The ground levels* $\widehat{D}_1^+$ *and* $\widehat{D}_1^-$ *are always transient. All other levels can be grouped into a finite chain* $\widehat{\mathcal{T}}_1 < \widehat{\mathcal{X}}_1 < \widehat{\mathcal{T}}_2 < \widehat{\mathcal{X}}_2 < \ldots < \widehat{\mathcal{X}}_m < \widehat{\mathcal{T}}_{m+1}$ *or an infinite chain* $\widehat{\mathcal{T}}_1 < \widehat{\mathcal{X}}_1 < \widehat{\mathcal{T}}_2 < \widehat{\mathcal{X}}_2 < \ldots$ *of disjoint subsets of* $\widehat{\mathcal{D}} \setminus \{\widehat{D}_1^+, \widehat{D}_1^-\}$ *which are of the following type:*

- $\widehat{\mathcal{X}}_i$ *is an irreducible component.*
- $\widehat{\mathcal{T}}_i$ *is either void or contains only transient levels:* $\widehat{D} \in \widehat{\mathcal{T}}_i \Rightarrow \widehat{D} \not\twoheadrightarrow \widehat{D}$.

Here $\widehat{\mathcal{T}}_i < \widehat{\mathcal{X}}_i < \widehat{\mathcal{T}}_{i+1}$ *means that* $\widehat{\mathcal{T}}_i \twoheadrightarrow \widehat{\mathcal{X}}_i \twoheadrightarrow \widehat{\mathcal{T}}_{i+1}$ *but* $\widehat{\mathcal{T}}_{i+1} \not\twoheadrightarrow \widehat{\mathcal{X}}_i \not\twoheadrightarrow \widehat{\mathcal{T}}_i$.

Since every irreducible component must occur in this chain, there is at most one maximal irreducible component. Before proving the theorem we give a more detailed description of the irreducible components.

2.23. Proposition. *For every irreducible component* $\widehat{\mathcal{X}}_i$ *of* $\widehat{\mathcal{D}}$ *the following holds.*

1. $\widehat{\mathcal{X}}_i$ *contains levels from both towers. The lowest levels on each side are successors of critical levels and the highest levels—if they exist—are critical levels. The set* $\widehat{\mathcal{X}}_i^{\wedge}$ *is invariant and* $\widehat{\mathcal{X}}_i^{\wedge} = \widehat{\mathcal{D}}_{kl}^{\wedge}$ *for suitably chosen indices* $k, l \in \mathbb{N}_\infty$.

2. *The $\widehat{\mathcal{D}}^{\pm}$-sets belonging to $\widehat{\mathcal{X}}_i$ form an "interval" in the tower*, *i.e.*,

$$\widehat{\mathcal{X}}_i \cap \widehat{\mathcal{D}}^{\pm} = \{\widehat{D}^{\pm}_{m^{\pm}}, \widehat{D}^{\pm}_{m^{\pm}+1}, \dots, \widehat{D}^{\pm}_{n^{\pm}}\} \quad or \quad \widehat{\mathcal{X}}_i \cap \widehat{\mathcal{D}}^{\pm} = \{\widehat{D}^{\pm}_{m^{\pm}}, \widehat{D}^{\pm}_{m^{\pm}+1}, \dots\}$$

for suitably chosen integers $m^{\pm}$ and $n^{\pm}$.

3. *$\widehat{\mathcal{X}}_i < \widehat{\mathcal{X}}_j$ if and only if all $\widehat{\mathcal{D}}^{\pm}$-sets belonging to $\widehat{\mathcal{X}}_i$ have indices lower than the $\widehat{\mathcal{D}}^{\pm}$-sets belonging to $\widehat{\mathcal{X}}_j$.*

Proof. Every closed path $\widehat{D}^{\pm}_n \twoheadrightarrow \widehat{D}^{\pm}_n$ must contain at least two jumps: The first jump leads from a critical level above or equal to $\widehat{D}^{\pm}_n$ to a level on the other side and the last jump leads back from the other side to the successor of a critical level below $\widehat{D}^{\pm}_n$. This implies statement 1 except for the invariance of $\widehat{\mathcal{X}}_i^{\wedge}$. Statements 2 and 3 follow from the fact that $\widehat{D}^{\pm}_n \to \widehat{D}^{\pm}_{n+1}$ holds for every $n \in \mathbb{N}$. If $\widehat{\mathcal{X}}_i^{\wedge}$ is not invariant, say, there exists a set $\widehat{D}^{+}_m \in \sigma(\widehat{\mathcal{X}}_i^{\wedge}) \setminus \widehat{\mathcal{X}}_i^{\wedge}$ then $\widehat{D}^{+}_m$ cannot be contained in $\widehat{\mathcal{X}}_i$ because $\widehat{\mathcal{X}}_i$ is a complete equivalence class. But $\widehat{D}^{+}_m$ cannot lie below $\widehat{\mathcal{X}}_i$ either, because part 1 implies that the path $\widehat{D}^{+}_m \to \widehat{D}^{+}_{m+1} \to \dots$ enters $\widehat{\mathcal{X}}_i$. ∎

The proof of Theorem 2.22 will proceed by induction based on the following lemma.

2.24. Lemma. *A set $\widehat{\mathcal{D}}^{\wedge}_{kl}$ is invariant if and only if $Q^+(j) \ge l$ for all integers $j > k$ and $Q^-(j) \ge k$ for all integers $j > l$. If it is invariant then there are four possibilities*:

1. *If $Q^+(k+1) = l$ and $Q^-(l+1) = k$ then $\widehat{C}^{+}_{k+1}$ and $\widehat{C}^{-}_{l+1}$ belong to the same irreducible component which contains at least $\widehat{\mathcal{D}}^{\wedge}_{kl} \setminus \widehat{\mathcal{D}}^{\wedge}_{k+1,\,l+1}$.*
2. *If $Q^+(k+1) = l$ and $Q^-(l+1) > k$ then $\widehat{\mathcal{D}}^{\wedge}_{k+1,\,l}$ is invariant.*
3. *If $Q^+(k+1) > l$ and $Q^-(l+1) = k$ then $\widehat{\mathcal{D}}^{\wedge}_{k,\,l+1}$ is invariant.*
4. *If $Q^+(k+1) > l$ and $Q^-(l+1) > k$ then $Q^+(k+1) = \infty$ and $Q^-(l+1) = \infty$ and $\widehat{\mathcal{D}}^{\wedge}_{kl}$ contains only non-critical levels.*

Proof. The first statement of the lemma is obvious.

1) All levels in $\widehat{\mathcal{D}}^{\wedge}_{kl} \setminus \widehat{\mathcal{D}}^{\wedge}_{k+1,\,l+1}$ belong to one closed loop

$$\widehat{C}^{+}_{k+1} \rightsquigarrow \widehat{C}^{-}_{Q^+(k+1)+1} = \widehat{C}^{-}_{l+1} \rightsquigarrow \widehat{C}^{+}_{Q^-(l+1)+1} = \widehat{C}^{+}_{k+1},$$

which implies that $\widehat{\mathcal{D}}^{\wedge}_{kl} \setminus \widehat{\mathcal{D}}^{\wedge}_{k+1,\,l+1}$ is part of an irreducible component ([9]).

2) We have to show that $Q^-(j) > k$ for all $j \ge l+2$. Assume to the contrary that there exists an index $j \ge l+2$ such that $Q^-(j) = k$ and choose j minimal. The Hofbauer condition (2.8) implies $Q^-(j) = Q^-((j-1)+1) \ge Q^-(Q^+Q^-(j-1)+1)$. Since $\widehat{\mathcal{D}}^{\wedge}_{kl}$ is invariant and since $Q^-(j-1) > k$, it follows that $Q^+Q^-(j-1)+1 > l$. On the other hand, $Q^+Q^-(j-1)+1 < j$ holds by Remark 2.11. Now $l < Q^+Q^-(j-1)+1 < j$ implies $Q^-(Q^+Q^-(j-1)+1) > k$, a contradiction. The proof of case 3) is analogous.

4) As in cases 2) and 3) it follows that $Q^+(i) > l$ and $Q^-(j) > k$ for all $i > k$ and $j > l$. Now if $Q^+(k+1)$ were finite this would imply $k < Q^-Q^+(k+1)$, contradicting Remark 2.11. Similarly, $Q^-(l+1)$ cannot be finite either. ∎

2.25. Remark. The proof of Lemma 2.24 would have been slightly simpler if we had assumed that there is a Lorenz map behind the Markov diagram. For example, if

([9]) Recall that $\widehat{C}^{\pm}_k := \widehat{D}^{\pm}_{S^{\pm}_k}$.

$Q^+(k+1) > l$ and $Q^-(l+1) > k$ then this implies that the interval (c_k^+, c^-) is mapped monotonically into the interval (c^+, c_l^-), and vice versa. It follows that c^+ and c^- are attracted to a periodic orbit of period $S_k^+ + S_l^-$ and there are no further cutting times.

Nevertheless, we prefer to give a purely combinatorial answer to a purely combinatorial question without referring to the underlying geometry, even more since it gives us an opportunity to illustrate the "mysterious" Hofbauer condition. $\diamond$

Proof of Theorem 2.22. Assume by induction that the set $\widehat{\mathcal{D}}$ of levels can be decomposed into a chain $\widehat{\mathcal{T}}_1 < \widehat{\mathcal{X}}_1 < \ldots < \widehat{\mathcal{X}}_n < \widehat{\mathcal{X}}_n^\wedge$, where the meaning of the sets $\widehat{\mathcal{T}}_k$ and $\widehat{\mathcal{X}}_k$ is explained in the statement of the theorem. The starting point of the induction is $n = 0$, in the sense that the above chain consists of a single link, namely $\widehat{\mathcal{X}}_0^\wedge := \widehat{\mathcal{D}}$. If $\widehat{\mathcal{X}}_n^\wedge$ is void or contains only levels from one of the two towers then $\widehat{\mathcal{X}}_n^\wedge$ cannot contain any further critical levels, therefore every level in $\widehat{\mathcal{X}}_n^\wedge$ must be transient. Let $\widehat{\mathcal{T}}_{n+1} := \widehat{\mathcal{X}}_n^\wedge$ and terminate the induction. Now for the remaining induction step assume that $\widehat{\mathcal{X}}_n^\wedge$ contains both $\widehat{\mathcal{D}}^+$- and $\widehat{\mathcal{D}}^-$-levels. By Proposition 2.23, $\widehat{\mathcal{X}}_n^\wedge$ equals $\widehat{\mathcal{D}}_{kl}^\wedge$ for suitably chosen integers $k,\ l$. We distinguish four cases according to the cases in Lemma 2.24:

In case 1 the levels $\widehat{C}_{k+1}^+$ and $\widehat{C}_{l+1}^-$ belong to a common irreducible component $\widehat{\mathcal{X}}_{n+1} := [\widehat{C}_{k+1}^+] = [\widehat{C}_{l+1}^-]$ inside $\widehat{\mathcal{X}}_n^\wedge$ and there are no transient levels between $\widehat{\mathcal{X}}_n$ and $\widehat{\mathcal{X}}_{n+1}$. Let $\widehat{\mathcal{T}}_{n+1} := \emptyset$ and continue with the induction.

In cases 2 and 3 the set $\widehat{\mathcal{D}}_{k_1 l_1}^\wedge := \widehat{\mathcal{D}}_{k+1,\, l}^\wedge$, respectively $\widehat{\mathcal{D}}_{k_1 l_1}^\wedge := \widehat{\mathcal{D}}_{k,\, l+1}^\wedge$, is a proper invariant subset of $\widehat{\mathcal{D}}_{k_0 l_0}^\wedge := \widehat{\mathcal{D}}_{kl}^\wedge$. Now Lemma 2.24 can be applied to $\widehat{\mathcal{D}}_{k_1 l_1}^\wedge$ again, and so on, and one obtains a chain $\widehat{\mathcal{D}}_{k_0 k_0}^\wedge \supset \widehat{\mathcal{D}}_{k_1 k_1}^\wedge \supset \ldots$ until neither case 2 nor 3 applies any more. If they apply infinitely often then the remaining set $\bigcap_{i\in\mathbb{N}} \widehat{\mathcal{D}}_{k_i l_i}^\wedge$ is void or contains only noncritical levels ([10]), in particular no recurrent levels. Let $\widehat{\mathcal{T}}_{n+1} := \widehat{\mathcal{D}}_{kl}^\wedge$ and terminate the induction. Otherwise, if j is the minimal index such that neither case 2 nor 3 applies to $\widehat{\mathcal{D}}_{k_j l_j}^\wedge$ then let $\widehat{\mathcal{T}}_{n+1} := \widehat{\mathcal{D}}_{k_0 l_0}^\wedge \setminus \widehat{\mathcal{D}}_{k_j l_j}^\wedge$ and consider two subcases: If case 1 of Lemma 2.24 applies to $\widehat{\mathcal{D}}_{k_j l_j}^\wedge$ then the two bottom levels $\widehat{\mathcal{D}}_{k_j l_j}^\wedge$ are contained in a common irreducible component. Denote this irreducible component by $\widehat{\mathcal{X}}_{n+1}$ and continue with the induction. If case 4 applies then $\widehat{\mathcal{D}}_{k_j l_j}^\wedge$ contains no critical levels. Let $\widehat{\mathcal{T}}_{n+1} := \widehat{\mathcal{D}}_{k_j l_j}^\wedge$ and terminate the induction.

In case 4 the set $\widehat{\mathcal{D}}_{kl}^\wedge$ contains no critical levels and hence only transient levels. Let $\widehat{\mathcal{T}}_{n+1} := \widehat{\mathcal{D}}_{kl}^\wedge$ and terminate the induction. ∎

2.26. Remark. The proof of the theorem also shows that the Markov diagram is entirely transient (i.e., there is no irreducible component in the tower) if and only if no integer is a cutting$^+$ and a cutting$^-$ time simultaneously. Indeed, if $\widehat{D}_n^+$ and $\widehat{D}_n^-$ are both critical then they belong to the same irreducible component in the tower. Conversely, if $\widehat{\mathcal{T}}_1$ is finite and if k and l denote the minimal indices such that $\widehat{C}_{k+1}^+$ and $\widehat{C}_{l+1}^-$ belong to $\widehat{\mathcal{X}}_1$ then $Q^+(k+1) = l$ and $Q^-(l+1) = k$. This implies that

$$S_{k+1}^+ = S_k^+ + S_{Q^+(k+1)}^- = S_k^+ + S_l^- = S_{Q^-(l+1)}^+ + S_l^- = S_{l+1}^-. \ \diamond$$

([10]) The latter happens precisely if only one of the two cases 2 and 3 occurs infinitely often.

The following lemma completes the picture given in Lemma 2.24 and will be needed in the next section.

2.27. Lemma. *If $\widehat{\mathcal{D}}^{\wedge}_{kl}$ is invariant then there is an irreducible component $\widehat{\mathcal{X}}_i$ immediately below $\widehat{\mathcal{D}}^{\wedge}_{kl}$ (i.e., $\widehat{\mathcal{D}}^{\wedge}_{kl} = \widehat{\mathcal{X}}^{\wedge}_i$) iff $Q^-(Q^+(k)+1) < k$ and $Q^+(Q^-(l)+1) < l$. If there is no such irreducible component then either $Q^+(k) = l$ or $Q^-(l) = k$.*

Proof. Obviously, $\widehat{\mathcal{D}}^{\wedge}_{kl} = \widehat{\mathcal{X}}_i$ for some irreducible component $\widehat{\mathcal{X}}_i$ iff $\widehat{C}^+_k \asymp \widehat{C}^-_l$. If both inequalities hold then there is a loop $\widehat{C}^+_k \twoheadrightarrow \widehat{C}^-_{Q^+(k)+1} \twoheadrightarrow \widehat{C}^+_{Q^-(Q^+(k)+1)+1} \twoheadrightarrow \widehat{C}^+_k$ ([11]) and a similar loop starting at $\widehat{C}^-_l$. This proves that $\widehat{C}^+_k$ and $\widehat{C}^-_l$ are recurrent. Since there are no periodic attractors, one of the inequalities $Q^+(k) < l$ and $Q^-(l) < k$ must hold, which implies that $\widehat{C}^+_k \twoheadrightarrow \widehat{C}^-_l$ or $\widehat{C}^-_l \twoheadrightarrow \widehat{C}^+_k$. Conversely, if $Q^-(Q^+(k)+1) \geq k$ then $\widehat{C}^+_k$ is a transient level: The level $\widehat{C}^+_k$ has two successors, namely $\widehat{D}^+_{S^+_k+1}$ and $\widehat{D}^-_{S^-_{Q^+(k)}+1}$. The first successor $\widehat{D}^+_{S^+_k+1}$ is contained in $\widehat{\mathcal{D}}^{\wedge}_{kl}$, whence $\widehat{D}^+_{S^+_k+1} \not\twoheadrightarrow \widehat{C}^+_k$. From the other successor $\widehat{D}^-_{S^-_{Q^+(k)}+1}$ one has to climb at least to the next critical level $\widehat{C}^-_{Q^+(k)+1}$ before being able to return. But $S^-_{Q^+(k)+1} = S^-_{Q^+(k)} + S^+_{Q^-(Q^+(k)+1)} > S^+_{Q^-(Q^+(k)+1)} \geq S^+_k$, since $Q^-(Q^+(k)+1) \geq k$. This implies that $\widehat{C}^-_{Q^+(k)+1} \in \widehat{\mathcal{D}}^{\wedge}_{kl}$, whence $\widehat{C}^-_{Q^+(k)+1} \not\twoheadrightarrow \widehat{C}^+_k$.

Now assume that $\widehat{C}^+_k \not\asymp \widehat{C}^-_l$. This implies that $Q^+(k) \geq l$ or $Q^-(l) \geq k$, otherwise there would be a loop connecting $\widehat{C}^+_k$ and $\widehat{C}^-_l$. Assume w.l.o.g. that $Q^+(k) \geq l$. If $Q^+(k) > l$ then $S^+_k > S^-_{Q^+(k)} \geq S^-_{l+1} > S^+_{Q^-(l+1)}$, whence $Q^-(l+1)+1 < k$. This implies that the successor of $\widehat{C}^-_{l+1} \in \widehat{\mathcal{D}}^{\wedge}_{kl}$ on the other side of the tower does not belong to $\widehat{\mathcal{D}}^{\wedge}_{kl}$, which is impossible. It follows that $Q^+(k) = l$. ∎

2.5. Renormalization

2.28. Definition (Renormalizable). Let $Z^{\pm}_n$ denote the left/right central n-cylinder set. Fix two integers m and n. The restrictions of f^n to Z^-_n and f^m to Z^+_m are called the two *central branches*. Assume that $c^+_m \in Z^-_n$ and $c^-_n \in Z^+_m$, Then we say that the Lorenz map f is

1. (m,n)-*renormalizable* if both central branches are critical and if the restriction of the two central branches to $[c^+_m, c^-_n]$ is a Lorenz map in the sense of Definition 2.3,
2. *properly* (m,n)-*renormalizable* if there are points $\widetilde{p}$ and $\widetilde{q}$ in Z^-_n and Z^+_m, respectively, such that the restriction of the two central branches to $[\widetilde{p}, \widetilde{q}]$ is a proper Lorenz map in the sense of Definition 2.1.

The renormalized map will be denoted by $\mathcal{R}f$.

The difference between a proper and a nonproper renormalizable map is that in the former case both central branches must have fixed points. If this is the case then the boundary points of the renormalization interval are the innermost fixed points on either side. If the map has negative Schwarzian derivative then it is easily seen that each of the central branches has precisely one fixed point, which is a hyperbolic repeller. In Figures 2.6 and 2.7 a proper and a nonproper renormalization are shown.

([11]) Omit the last arrow if $Q^-(Q^+(k)+1) = k-1$.

2.29. Remark. If the renormalization is proper then both central branches are critical automatically. For the nonproper renormalizations it was explicitly assumed in order to exclude a certain type of renormalization which is even more "degenerate" than a nonproper renormalization ([12]). However, this assumption can be made without loss of generality, because if the restriction of the two central branches to $[c_m^+, c_n^-]$ is a nonproper Lorenz map and one or both of the central branches is not critical—say, the left one is critical but the right one is not—then either $c_n^+ < c^+_{S^+\langle n\rangle} < c^-$, in which case one can replace f^n by $f^{S^+\langle n\rangle}$ on the right hand side, or the interval $(c^+_{S^+\langle n\rangle}, c^-)$ is mapped monotonically into itself, which implies that both critical points are attracted to a periodic attractor. $\diamond$

Note that we did not assume that $\mathcal{R}f$ is the first return map to the renormalization interval, as is done for example in Martens & de Melo [47]. Indeed, for nonproper renormalizations it is possible that the intermediate iterates of the intervals $f^i(c_m^+, c^-)$, $i < n$, and $f^j(c^+, c_n^-)$, $j < m$, are not disjoint, in which case the renormalized map is obviously not a first return map ([13]). The following proposition shows that this phenomenon cannot occur if the renormalization is proper.

2.30. Proposition. *If f is properly (m,n)-renormalizable, then the renormalized map coincides with the first return map to the interval $(\widetilde{p}, \widetilde{q})$.*

Proof. Step (i): The minimal periods of $\widetilde{q}$ and $\widetilde{p}$ are m and n, respectively: Assume by contradiction that the period $\widetilde{m}$ of $\widetilde{q}$ is less than m. Since the interval $(c^+, \widetilde{q})$ is contained in Z_m^+ and since $\widetilde{m} < m$, the interval $f^{\widetilde{m}}(c^+, \widetilde{q}) = (c^+_{\widetilde{m}}, \widetilde{q})$ cannot contain the critical point, whence $c^+_{\widetilde{m}} \geq c^+$. Consequently, $f^{\widetilde{m}}$ maps $(c^+, \widetilde{q})$ monotonically into itself and it follows that $c_m^+ \geq c^+$, because m is a multiple of $\widetilde{m}$. But this is a contradiction to Definition 2.28.

Step (ii): The points $\widetilde{p}$ and $\widetilde{q}$ are a *nice pair*, i.e., the iterates of $\widetilde{p}$ and $\widetilde{q}$ never enter the interval $(\widetilde{p}, \widetilde{q})$: We proof this for $\widetilde{q}$: First of all note that $\widetilde{q}$ cannot be precritical since it belongs to a repelling periodic orbit. Any iterate of $\widetilde{q}$ is a fixed point of f^m, so none of them is allowed to enter $(c^+, \widetilde{q})$, since f^m has no fixed points there by Definition 2.28. But also the interval $(\widetilde{p}, c^-)$ must be avoided by iterates of $\widetilde{q}$, because every point in $(\widetilde{p}, c^-)$, except for precritical points, enters $(c^+, \widetilde{q})$ after finitely many iterations.

Step (iii): $\mathcal{R}f$ is the first return map: Assume w.l.o.g. that there exists $x \in (c^+, \widetilde{q})$ such that $f^{\widetilde{m}}(x)$ is contained in $(\widetilde{p}, \widetilde{q})$ for some $\widetilde{m} < m$.

Since $f^{\widetilde{m}}(\widetilde{q}) \notin (\widetilde{p}, \widetilde{q}]$ by steps (i) and (ii), we get $\widetilde{q} \in (f^{\widetilde{m}}(x), f^{\widetilde{m}}(\widetilde{q}))$, which implies $f^{m-\widetilde{m}}(\widetilde{q}) \in (f^m(x), f^m(\widetilde{q})) \subseteq (\widetilde{p}, \widetilde{q})$. This is impossible since $\widetilde{q}$ is nice. ∎

2.31. Remark. The end points of the interval $(\widetilde{p}, \widetilde{q})$ were excluded in the proposition on purpose, since it can happen that $\widetilde{p}$ and $\widetilde{q}$ belong to a common periodic orbit. A typical example for that is a $(2,2)$-renormalization, better known as the *period doubling renormalization*. $\diamond$

([12]) To visualize such a degenerate case have a look at Figure 2.7 and "bend down" the right end of the right branch until it does not intersect the horizontal axis any more.

([13]) Typical examples therefore are trivial renormalizations for overlap maps, cf. Example 2.37.

The following theorem shows how the property of being renormalizable is reflected in the structure of the Markov diagram.

2.32. Theorem. *Assume that the Lorenz map f has no periodic attractor and let $(\widehat{\mathcal{D}}, \to)$ be its Markov diagram. Then the following holds.*

1. *The map f is (S_k^+, S_l^-)-renormalizable if and only if $\widehat{\mathcal{D}}_{kl}^{\wedge}$ is invariant.*
2. *The map f is properly (S_k^+, S_l^-)-renormalizable if and only if $\widehat{\mathcal{D}}_{kl}^{\wedge}$ is invariant and $\widehat{\mathcal{D}}_{kl}^{\wedge}$ lies immediately above an irreducible component $\widehat{\mathcal{X}}_i$ (i.e., iff $\widehat{\mathcal{D}}_{kl}^{\wedge} = \widehat{\mathcal{X}}_i^{\wedge}$).*

The kneading map $\mathcal{R}Q$ of the renormalized map is given by "shifting" the kneading map Q, i.e., $\mathcal{R}Q^+(j-k) = Q^+(j) - l$ for $j > k$ and $\mathcal{R}Q^-(j-l) = Q^-(j) - k$ for $j > l$. In other words, the induced tower of the renormalized map is obtained from the induced tower of f by removing all levels below $\widehat{C}_k^+$ and $\widehat{C}_l^-$, respectively.

Proof. 1) The invariance of $\widehat{\mathcal{D}}_{kl}^{\wedge}$ is equivalent to $Q^+(j) \geq l$ for all $j \geq k$ and $Q^-(j) \geq k$ for all $j \geq l$ (cf. Lemma 2.24). It is easily checked that invariance of $\widehat{\mathcal{D}}_{kl}^{\wedge}$ is necessary for renormalizability.

Since $Q^-(l+1) \geq k$, the interval $D_{S_l^-}^- \cap Z^+ = (c^+, c_{S_l^-}^-)$ is mapped monotonically onto $D_{S_l^-+S_k^+}^- = (c_{S_k^+}^+, c_{S_l^-+S_k^+}^-)$ by $f^{S_k^+}$. We claim that $c_{S_l^-+S_k^+}^- \leq c_{S_l^-}^-$ (cf. Figure 2.5). Assume by contradiction that $c_{S_l^-+S_k^+}^- > c_{S_l^-}^-$. Then $D_{S_l^-+S_k^+}^- \supset (c_{S_k^+}^+, c_{S_l^-}^-)$, which implies that $D_{S_l^-+S_k^+}^-$ is critical, i.e., $S_{l+1}^- = S_l^- + S_k^+$, and that $D_{S_{l+1}^-}^- \cap Z^+ \supset D_{S_l^-}^- \cap Z^+$.

Because also $Q^-(l+2) \geq k$, the intervals $D_{S_{l+1}^-}^- \cap Z^+ \supset D_{S_l^-}^- \cap Z^+$ are mapped monotonically onto $D_{S_{l+1}^-+S_k^+}^- \supset D_{S_l^-+S_k^+}^- = D_{S_{l+1}^-}^-$ by $f^{S_k^+}$. It follows that $S_{l+2}^- = S_l^- + 2S_k^+$ and that $D_{S_{l+2}^-}^- \supset D_{S_{l+1}^-}^-$. Repeating this argument over and over we see that the following cutting$^-$ times are precisely $S_{l+j}^- = S_l^- + j \cdot S_k^+$, $j \in \mathbb{N}$, and that $f^{S_k^+}$ maps $(c^+, c_{S_{l+j}^-}^-)$ monotonically onto the strictly bigger interval $(c^+, c_{S_{l+j+1}^-}^-)$. This implies that the sequence of intervals $(c_{S_{l+j}^-}^-, c_{S_{l+j+1}^-}^-)$ is increasing and converges to an attracting fixed point of $f^{S_k^+}$, a contradiction. Similarly, one shows that $c_{S_k^++S_l^-}^+ \geq c_{S_k^+}^+$.

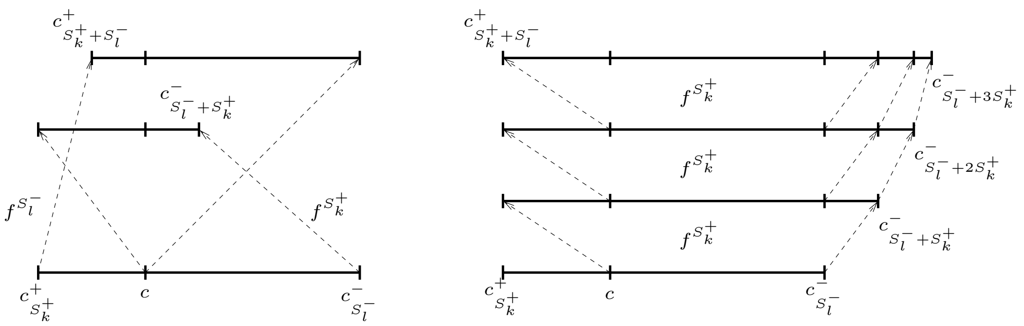


Fig. 2.5. An invariant subgraph implies renormalizability. If $(\widehat{\mathcal{D}}_{kl}^{\wedge}, \to)$ is an invariant subgraph of the Markov diagram then either the interval $(c_{S_k^+}^+, c_{S_l^-}^-)$ is a renormalization interval (l.h.s.), or there exists a periodic attractor. E.g., if $f^{S_k^+}(c^+, c_{S_l^-}^-) \not\subseteq (c_{S_k^+}^+, c_{S_l^-}^-)$ then the points $c_{S_l^-+jS_k^+}^-$, $j \in \mathbb{N}$, converge to an attracting fixed point for $f^{S_k^+}$ (r.h.s.).

2) Recall that $z_k^{\mp}$ denote the closest preimages to the left/right of c (cf. Figure 2.1) and that condition (2.7) is equivalent to the statement that $c^{\pm}_{S^{\pm}_k} \in A^{\mp}_{Q^{\pm}(k+1)}$ for all integers k, where $A_k^- := (z_{k-1}^-, z_k^-)$ and $A_k^+ := (z_k^+, z_{k-1}^+)$. In Figures 2.6 and 2.7 the two central branches are shown for a properly and a nonproperly renormalizable Lorenz map.

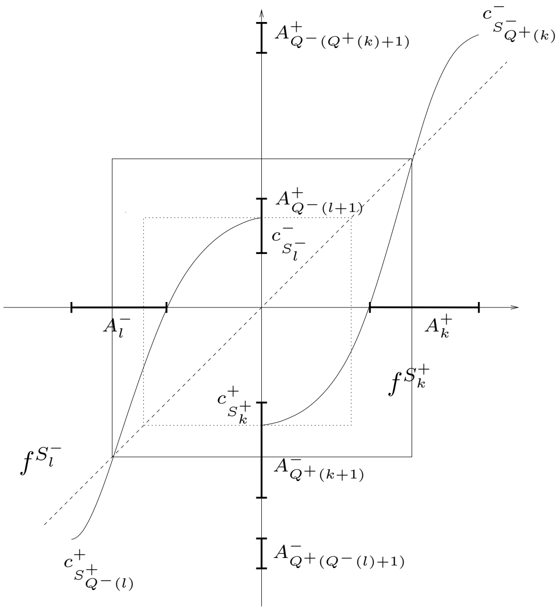


Fig. 2.6. A properly (S_k^+, S_l^-)-renormalizable Lorenz map

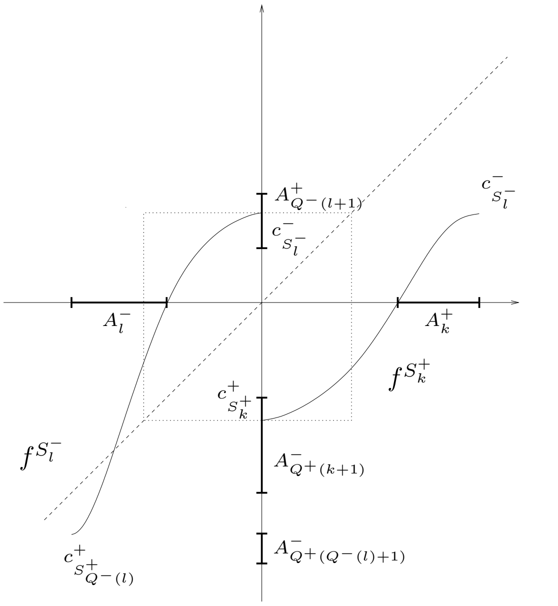


Fig. 2.7. A nonproperly (S_k^+, S_l^-)-renormalizable Lorenz map

The image of the right central branch is $f^{S_k^+}(c^+, c^-_{-S^+_{k-1}}) = (c^+_{S_k^+}, c^-_{S^-_{Q^+(k)}})$ and $c^-_{S^-_{Q^+(k)}}$ is contained in $A^+_{Q^-(Q^+(k)+1)}$.

In Lemma 2.27 it was shown that there is an irreducible component immediately below $\widehat{\mathcal{D}}_{kl}^{\wedge}$ iff $Q^-(Q^+(k)+1) < k$ and $Q^+(Q^-(l)+1) < l$. Now we prove that the right central branch has a fixed point iff $Q^-(Q^+(k)+1) < k$ (the corresponding statement for the left central branch follows by analogy):

The condition $Q^-(Q^+(k)+1) < k$ implies that the interval $A^+_{Q^-(Q^+(k)+1)}$ lies to the right of A_k^+, whence $f^{S_k^+}$ maps A_k^+ monotonically onto the interval $(c, z^+_{k-1}) \supset A_k^+$. It follows that there is a fixed point of $f^{S_k^+}$ in A_k^+ [14]. Conversely, if $Q^-(Q^+(k)+1) \geq k$ then $f^{S_k^+}$ maps A_k^+ into $(c, z^+_{k-1}) \supset A_k^+$. Moreover, the right central branch of $f^{S_k^+}$ cannot have a fixed point $\widetilde{q}$, since this would imply that the interval $(\widetilde{q}, z^-_{k-1})$ is mapped monotonically into itself, which would imply the existence of a periodic attractor.

This shows that $\widehat{D}^+_{S_k^+} \asymp \widehat{D}^-_{S_l^-}$ if and only if both central branches have fixed points and proves part 2. The last statement about the renormalized kneading map follows by induction. The proof is not difficult but lengthy and tedious to write down, and will be omitted. ∎

2.33. Definition. If f is (m,n)-renormalizable then the renormalization is called *minimal* if $m+n \leq m'+n'$ for every pair (m',n') such that f is (m',n')-renormalizable. It is called a *trivial renormalization* if $(m,n) \in \{(1,2),(2,1)\}$ and a *period doubling renormalization* if $(m,n) = (2,2)$.

Putting together Theorem 2.32 and Theorem 2.22 we obtain the following corollary:

2.34. Corollary. *Every renormalization $\mathcal{R}f$ of a Lorenz map f can be decomposed uniquely into minimal renormalizations*, *i.e.*, *there are integers (m_k,n_k), $k \leq N$, such that $\mathcal{R}^k f$ is the (m_k,n_k)-renormalization of $\mathcal{R}^{k-1}f$ and $\mathcal{R}^0 f = f$, $\mathcal{R}^N f = \mathcal{R}f$ and the renormalization is minimal. The minimal renormalizations are proper if and only if they are not trivial.*

Proof. A minimal renormalization corresponds to removing the smallest number of levels at the bottom of the induced tower such that the rest is an invariant subgraph. If D_2^+ and D_2^- are both critical then $[D_2^+] = [D_2^-]$ is removed and the renormalization is proper. If only one of those two levels is critical then the map is trivially renormalizable and the noncritical level is removed by the renormalization. ∎

2.35. Example. Abbreviate the decomposition of the renormalization in the corollary by $(m,n) = (m_N,n_N) * \cdots * (m_0,n_0)$. In Figure 2.3 the map is $(39,30)$-renormalizable with decomposition $(39,30) = (1,2)^2 * (2,1)^3 * (1,2)^3$ and in Figure 2.4 the map is $(8,11)$-renormalizable with decomposition $(8,11) = (1,2) * (2,1)^2 * (2,3)$. This can be read directly off the figure without any calculations just by searching for the invariant subsets and irreducible components of the Markov diagram. For example in Fig-

[14] If f has negative Schwarzian derivative then there is precisely one fixed point and it is repelling.

ure 2.4 the $(2,3)$-renormalization removes all levels below levels D_2^+ and D_3^-, then a $(2,1)$-renormalization removes all $\widehat{\mathcal{D}}^+$-levels below level $\widehat{D}_5^+$ and another one removes all $\widehat{\mathcal{D}}^+$-levels below level $\widehat{D}_8^+$. Finally, the $(1,2)$-renormalization removes all $\widehat{\mathcal{D}}^-$-levels below level $\widehat{D}_1^-$. $\diamond$

2.36. Remark. If a renormalizable Lorenz map is of class $\mathcal{C}^2$ and if there are no attracting or neutral periodic orbits outside the renormalization interval then it follows from a classical result of Mañé (cf. [13, Section III.5]) that the set of points that never enter the renormalization interval is a hyperbolic repeller on which f is conjugate to a subshift of finite type. The subshift is just given by the set of paths in the Markov diagram that start at $\widehat{D}_0^+$ or $\widehat{D}_0^-$ and never enter $\widehat{\mathcal{D}}_{kl}^{\wedge}$. In particular, if the renormalization is minimal then the repeller is 1) void if the renormalization is trivial, 2) a periodic orbit of period two and its preimages if it is a period doubling renormalization, and 3) a Cantor set if the renormalization is proper and not a period doubling renormalization. If the Lorenz map is several times renormalizable then the above argument applies to every step of the decomposition into minimal renormalizations.

If f is not of class $\mathcal{C}^2$ then cases 1–3 still hold modulo homtervals. $\diamond$

The essence of Theorem 2.22 is that proper renormalizations are related to the existence of nonmaximal irreducible components in the tower and the nonproper renormalizations have to do with the presence of transient parts in between. Moreover, it also shows that every nonproper renormalization can either be decomposed entirely into trivial renormalizations or into one proper renormalization of maximal length followed by a sequence of trivial renormalizations.

2.37. Example. The prototypes of maps where trivial renormalizations are frequently encountered are Lorenz maps which satisfy $c_2^- \leq c_2^+$. They can be considered in a natural way as injective circle maps on their dynamical interval $[a,b] = [c_1^+, c_1^-]$ by gluing together the two endpoints a and b. If $c_2^- = c_2^+$ then one obtains a circle homeomorphism, but in the general case where $c_2^- < c_2^+$ the map f is not surjective on $[a,b]$, because there is a *gap* in the image, namely the interval (c_2^-, c_2^+). For that reason such maps are called *gap maps*. Since such maps can be considered as the inverse of a monotonic and continuous circle map of degree one with a constant segment, they have a well defined rotation number. In Section 4.4 we are going to show how the rotation number of f is related to the structure of the transition diagram. For the moment, let us just briefly explain why such maps are several times trivially renormalizable.

If f is a gap map then obviously at most one of the intervals D_2^+ and D_2^+ is critical. If neither D_2^+ nor D_2^- is critical then D_1^+ is mapped into D_1^- and vice versa, which implies that every point is asymptotic to a periodic orbit of period two (and the rotation number equals $1/2$). Otherwise, the map is trivially renormalizable and it is easily checked that the renormalized map is again an injective circle map. By induction it follows that either the map is infinitely often trivially renormalizable or after a finite number of trivial renormalizations one obtains a Lorenz map which maps the left hand side into the right hand side, and vice versa. The decomposition of the Markov diagram for such a gap map consists of a single infinite transient chain $\widehat{\mathcal{T}}_1$.

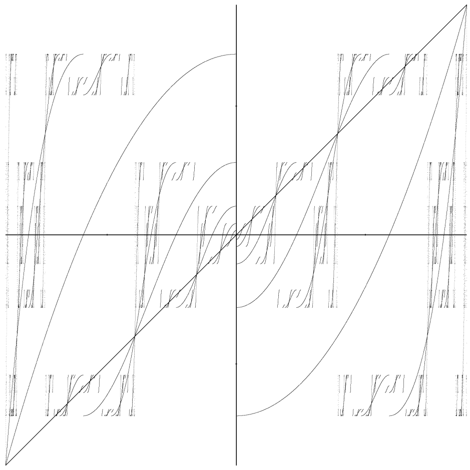


Fig. 2.8. The Feigenbaum map. The figure shows the iterates 1, 2, 4, 8, and 16 of the Feigenbaum map in the quadratic Lorenz family. It is infinitely often properly $(2,2)$-renormalizable.

If f is an *overlap map* (i.e., if $c_2^- > c_2^+$) but the overlap interval (c_2^+, c_2^-) does not contain the critical point then f is still trivially renormalizable (possibly even more than once). It is easily checked that in this case the renormalized map is *not* the first return map. $\diamond$

3. Hopf decompositions and attractors

In the previous chapter the Markov extension was considered merely as a topological Markov chain. In the following we are going to make the model more stochastical by introducing the transfer operator $\mathrm{P}_{\widehat{f}}$ in order to study the time evolution of mass distributions on the tower under the influence of $\widehat{f}$.

3.1. Transfer operators. The Lebesgue measure m has a natural lift to a σ-finite measure $\widehat{m}$ on $\widehat{X}$ given by $\widehat{m}(\widehat{A}) := \sum_{\widehat{D}\in\widehat{\mathcal{D}}} m(\pi(\widehat{A}\cap\widehat{D}))$ and $\widehat{f}$ is nonsingular with respect to $\widehat{m}$, i.e., $\widehat{f}(\widehat{m}) \ll \widehat{m}$, where $\widehat{f}(\widehat{m}) := \widehat{m}\circ\widehat{f}^{-1}$, which makes it possible to define a transfer operator for $\widehat{f}$.

3.1. Definition (Transfer operator). The *transfer operator* associated with $\widehat{f}$ is the map $\mathrm{P}_{\widehat{f}} : \mathrm{L}^1_{\widehat{m}}(\widehat{X}) \to \mathrm{L}^1_{\widehat{m}}(\widehat{X})$ defined implicitly by $\widehat{f}(\widehat{\psi}\cdot\widehat{m}) = \mathrm{P}_{\widehat{f}}(\widehat{\psi})\cdot\widehat{m}$ (1). It is also called the *Perron–Frobenius operator*. Its explicit form is given by

$$\mathrm{P}_{\widehat{f}}(\widehat{\psi}) = \sum_{\widehat{Z}\in\widehat{\mathcal{Z}}} \mathrm{P}_{\widehat{f}}(\widehat{\psi}_{\widehat{Z}}) = \sum_{\widehat{Z}\in\widehat{\mathcal{Z}}} (\widehat{\psi}\cdot\widehat{\phi})\circ\widehat{f}^{-1}_{\widehat{Z}}, \tag{3.1}$$

where $\widehat{\psi}_{\widehat{Z}} := \widehat{\psi}\cdot 1_{\widehat{Z}}$, $\widehat{\phi} := 1/|\widehat{f}'|$ and $(\widehat{\psi}\cdot\widehat{\phi})\circ\widehat{f}^{-1}_{\widehat{Z}}$ stands for $(\widehat{\psi}\cdot\widehat{\phi})|_{\widehat{Z}}\circ(\widehat{f}|_{\widehat{Z}})^{-1}$ on $\widehat{f}(\widehat{Z})$ and 0 elsewhere. For convenience we will tacitly identify $\widehat{\psi}_{\widehat{Z}}$ with $\widehat{\psi}|_{\widehat{Z}}$ and $\widehat{f}^{-1}_{\widehat{Z}}$ with $(\widehat{f}|_{\widehat{Z}})^{-1}$, and so on.

Using the explicit form (3.1) the transfer operator can also be defined for arbitrary nonnegative measurable functions. Talking of *densities* or *mass distributions* we will always mean positive measurable functions $\widehat{\psi}$, without assuming that they are integrable or normalized. Similarly, one obtains transfer operators $\mathrm{P}_f : \mathrm{L}^1_m(X) \to \mathrm{L}^1_m(X)$ and $\mathrm{P}_\pi : \mathrm{L}^1_{\widehat{m}}(\widehat{X}) \to \mathrm{L}^1_m(X)$ from the nonsingular maps $f : X \to X$ and $\pi : \widehat{X} \to X$, respectively. In the following, we use the abbreviations $\mathrm{P} := \mathrm{P}_f$ and $\widehat{\mathrm{P}} := \mathrm{P}_{\widehat{f}}$. The operators P and $\widehat{\mathrm{P}}$ are related by $\mathrm{P}\circ\mathrm{P}_\pi = \mathrm{P}_\pi\circ\widehat{\mathrm{P}}$.

By construction, the densities of absolutely continuous invariant measures are precisely the fixed points of the transfer operator: $\widehat{f}(\widehat{h}\cdot\widehat{m}) = \widehat{h}\cdot\widehat{m}$ if and only if $\widehat{\mathrm{P}}(\widehat{h}) = \widehat{h}$, and $f(h\cdot m) = h\cdot m$ if and only if $\mathrm{P}(h) = h$. Moreover, if $\widehat{h}$ is an invariant probability density for $\widehat{f}$ then $h := \mathrm{P}_\pi(\widehat{h})$ is an invariant probability density for f.

As in the theory of countable state Markov chains the topological Markov structure already yields a lot of information about the existence and location of absolutely con-

(1) That is, the image measure (w.r.t. $\widehat{f}$) of a measure with density $\widehat{\psi}$ (w.r.t. $\widehat{m}$) has density $\mathrm{P}_{\widehat{f}}(\widehat{\psi})$ (w.r.t. $\widehat{m}$).

tinuous invariant measures, since the mass can only be transported along the allowed transitions in the tower. Additionally, due to the nonlinearity of the map $\widehat{f}$, the mass will be compressed in some parts of the tower and thinned out in other parts. How big this effect is depends on the distortion of the map. In order to control the distortion we will assume that the map f has negative Schwarzian derivative.

3.2. Definition (Schwarzian derivative). For a map $g : I \to \mathbb{R}$ which is piecewise $\mathcal{C}^3$ with $g' \neq 0$ except for finitely many points the *Schwarzian derivative* Sg is

$$Sg := \frac{g'''}{g'} - \frac{3}{2}\left(\frac{g''}{g'}\right)^2,$$

which is well defined for all points where the first derivative does not vanish. We say that the map has *negative Schwarzian derivative* and write $Sg < 0$ if $Sg(x) < 0$ holds for all $x \in I$ such that $g'(x) \neq 0$. A similar agreement holds for $Sg > 0$, etc.

We briefly state some properties of maps with negative Schwarzian derivative which will be needed in the following. More information can be found in [13]. By the chain rule for the Schwarzian derivative, $S(g \circ f) = Sg \circ f \cdot (f')^2 + Sf$, it follows that the composition of maps with negative Schwarzian derivative again has negative Schwarzian derivative. In particular, if g maps I into I and has negative Schwarzian derivative then all iterates of g have negative Schwarzian derivative, too. The same holds for positive Schwarzian derivative.

A straightforward calculation yields the equality $(|g'|^{-1/2})'' = -\frac{1}{2}|g'|^{-1/2}Sg$ which shows that $Sg < 0$ if and only if $|g'|^{-1/2}$ is strictly convex and $Sg > 0$ if and only if $|g'|^{-1/2}$ is strictly concave ([2]). An immediate consequence of this is the *Minimum Principle*: If the map g has negative Schwarzian derivative then $|g'|$ cannot have a positive local minimum. It also yields an elementary proof of the Koebe Principle which is an important tool to estimate the distortion on the branches of g.

3.3. Lemma (Koebe Principle). *Let $I \subseteq Z$ be two intervals and let $g : Z \to g(Z)$ be a diffeomorphism with negative Schwarzian derivative. If $g(Z)$ contains a δ-scaled neighbourhood of $g(I)$, i.e., if the length of both components of $g(Z) \setminus g(I)$ is at least $\delta|g(I)|$, then the distortion $\sup_{x,y\in I} |f'(x)/f'(y)|$ on I is bounded by $((1+\delta)/\delta)^2$.*

Proof. Let $h := g^{-1}$. Then $|\widehat{h}'|^{-1/2}$ is positive and concave on $g(Z)$, which implies

$$\left(\frac{\delta}{1+\delta}\right)^2 \leq \left(\frac{g(y)-g(b)}{g(x)-g(b)}\right)^2 \leq \left|\frac{h'(g(x))}{h'(g(y))}\right| \leq \left(\frac{g(y)-g(a)}{g(x)-g(a)}\right)^2 \leq \left(\frac{1+\delta}{\delta}\right)^2$$

for all points $a < x < y < b \in Z$ such that $(f(a), f(b))$ is a δ-scaled neighbourhood of $(f(x), f(y))$. Since $|g'(f(x))/g'(f(y))| = |f'(y)/f'(x)|$, the lemma follows. ∎

A more general version of the Koebe Principle can be found in [13, Section IV.1]. Intuitively, the Koebe Principle states that maps with negative Schwarzian derivative behave very much like linear maps on every branch if one stays away far enough from the ends of the branch. That makes it sound reasonable that densities are not distorted

([2]) In fact, in what follows the condition of negative Schwarzian derivative can be weakened to the requirement that g is $\mathcal{C}^1$ and $|g'|^{-1/2}$ is strictly convex.

too much when transported along such branches—at least this should hold for the part of the mass that is located away from the boundary of the branch. There is an elegant way to make this reasoning rigorous which is due to Misiurewicz [52]. It is based on the following observation.

The proof of the Koebe Principle is based on the fact that the inverse branch of g has positive Schwarzian derivative, which implies that $|(g^{-1})'|^{-1/2}$ is concave. Now if one writes (3.1) in the form $\widehat{\mathrm{P}}(\widehat{\psi}) = \sum_{\widehat{Z}\in\widehat{\mathcal{Z}}} \widehat{\psi} \circ \widehat{f}_{\widehat{Z}}^{-1} \cdot |(\widehat{f}_{\widehat{Z}}^{-1})'|$ then it becomes clear that it is precisely the derivatives of such inverse branches which determine the push forward of densities. Having this in mind it does not come as a surprise that the following family of densities is invariant by $\widehat{\mathrm{P}}$ and its elements have nice regularity properties.

3.4. Definition (Regular densities). For every level $\widehat{D} \in \widehat{\mathcal{D}}$ let $\mathcal{C}(\widehat{D})$ be the set of continuous functions that vanish outside $\widehat{D}$ and let

$$\mathcal{H}(\widehat{D}) := \{\widehat{\psi} \in \mathcal{C}(\widehat{D}) \mid \widehat{\psi} > 0 \text{ on } \widehat{D} \text{ and } \widehat{\psi}^{-1/2} \text{ is concave on } \widehat{D}\} \cup \{0\}.$$

Denote by $\mathcal{H} := \mathcal{H}(\widehat{X})$ the set of all functions $\widehat{\psi} \in \mathcal{C}(\widehat{X})$ such that $\widehat{\psi}_{\widehat{D}} \in \mathcal{H}(\widehat{D})$ for all levels $\widehat{D} \in \widehat{\mathcal{D}}$. The elements of $\mathcal{H}$ are called *regular densities*.

3.5. Remark. Regular densities were introduced by Misiurewicz [52] for continuous maps of the interval with negative Schwarzian derivative satisfying the Misiurewicz condition and studied intensively by Keller [39] for abstract Markov systems. $\diamond$

Regular densities can also be characterized as follows, which shows that they come in very naturally when transfer operators for maps with negative Schwarzian derivative are studied.

3.6. Lemma. *A function $\widehat{\psi} \in \mathcal{C}^2(\widehat{D})$ belongs to $\mathcal{H}_2(\widehat{D}) := \mathcal{H}(\widehat{D}) \cap \mathcal{C}^2(\widehat{D})$ if and only if there is an interval $\widehat{C}$ and a $\mathcal{C}^3$-homeomorphism $\widehat{g} : \widehat{C} \to \widehat{D}$ with $S\widehat{g} \le 0$ such that $\widehat{\psi} = \mathrm{P}_{\widehat{g}}(1)$.*

Proof. If $\widehat{\psi} = \mathrm{P}_{\widehat{g}}(1) = (1/|\widehat{g}'|) \circ g^{-1} = |(\widehat{g}^{-1})'|$ then $\widehat{\psi}^{-1/2}$ is concave, because $S(\widehat{g}^{-1}) \ge 0$. Conversely, if $\widehat{\psi} \in \mathcal{C}^2(\widehat{D})$ is positive on $\widehat{D} = (\widehat{a}, \widehat{b})$ then let $\widehat{\Psi}(\widehat{x}) := \int_{\widehat{a}}^{\widehat{x}} \widehat{\psi}\, d\widehat{m}$, $\widehat{C} := \widehat{\Psi}(\widehat{D})$ and $\widehat{g} := \widehat{\Psi}^{-1}$ on $\widehat{C}$. ∎

Now we are ready to state the counterpart of the Koebe Principle in the context of transfer operators.

3.7. Proposition. (i) *$\mathcal{H}$ is a positive cone which is closed in the compact-open topology and $\mathcal{H} - \mathcal{H} = \mathcal{C}(\widehat{X})$ is dense in $\mathrm{L}^1_{\widehat{m}}(\widehat{X})$. The cones $\mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}(\widehat{X})$ and $\mathcal{H}_2 \cap \mathrm{L}^1_{\widehat{m}}(\widehat{X})$ are invariant by $\widehat{\mathrm{P}}$.*

(ii) *For every $\delta > 0$ there exists a constant $C(\delta) > 0$ with the following property: Whenever $\widehat{D} \in \widehat{\mathcal{D}}$ contains a δ-scaled neighbourhood of $\widehat{J} \subset \widehat{D}$ and $\widehat{\psi} \in \mathcal{H}$ is positive on $\widehat{D}$ then*

$$\left| \frac{\widehat{\psi}(\widehat{x})}{\widehat{\psi}(\widehat{y})} - 1 \right| \le \frac{C(\delta)}{|\widehat{J}|} |\widehat{x} - \widehat{y}| \quad \text{for all } \widehat{x}, \widehat{y} \in \widehat{J}.$$

In particular, for every compact set $\widehat{K} \subset \widehat{X}$ there is a uniform bound for the Lipschitz norm of the family $\{\log \widehat{\psi}|_{\widehat{K}} \mid \widehat{\psi} \in \mathcal{H},\ \widehat{\psi}|_{\widehat{K}} > 0\}$ which depends only on the relative Koebe space around $\widehat{K}$ in $\widehat{X}$.

3.8. Corollary. *For every compact interval $\widehat{I} \subset \widehat{D}$ there exists a constant $C_{\widehat{I}}$ such that*
$$C_{\widehat{I}}^{-1}\widehat{\psi}(\widehat{x}) \le \frac{1}{\widehat{m}(\widehat{B})}\int_{\widehat{B}}\widehat{\psi}\,d\widehat{m} \le C_{\widehat{I}}\,\widehat{\psi}(\widehat{x})$$
for every $\widehat{x} \in \widehat{I}$ and every $\widehat{B} \subseteq \widehat{I}$ of positive measure. In particular, for every $\widehat{x} \in \widehat{X}$ there is a constant $C_{\widehat{x}}$ such that $\widehat{\psi}(\widehat{x}) \le C_{\widehat{x}}\int\widehat{\psi}\,d\widehat{m}$ for all $\widehat{\psi} \in \mathcal{H}$. ∎

Proof of Proposition 3.7. (ii) Since $\widehat{\psi}^{-1/2}$ is positive and concave,
$$\left(\frac{\tau}{1+\tau}\right)^2 \le \left(\frac{y-b}{x-b}\right)^2 \le \frac{\widehat{\psi}(\widehat{x})}{\widehat{\psi}(\widehat{y})} \le \left(\frac{y-a}{x-a}\right)^2 \le \left(\frac{1+\tau}{\tau}\right)^2$$
for all $a < x < y < b$ such that $D = (a,b)$ is a τ-scaled neighbourhood of (x,y). Now if D is a δ-scaled neighbourhood of J then it is a τ-scaled neighbourhood of (x,y), where $\tau = \delta|J|/|y-x|$. Since $(\tau/(\tau+1))^2 - 1$ and $((\tau+1)/\tau)^2 - 1$ are of order $1/\tau$ for $\tau \to \infty$, part (ii) follows.

(i) $\mathcal{H}$ *is a positive cone*: Obviously, $\mathcal{H}$ is positively homogeneous. Let us show that $\widehat{\psi}_1 + \widehat{\psi}_2 \in \mathcal{H}$ for $\widehat{\psi}_1, \widehat{\psi}_2 \in \mathcal{H}$. Choose a level $\widehat{D}$ on which $\widehat{\psi}_1$ and $\widehat{\psi}_2$ are positive and points $\widehat{a} < \widehat{c} < \widehat{b} \in \widehat{D}$, $\widehat{c} = t\widehat{a} + (1-t)\widehat{b}$. Let $\widehat{\phi}_i$ be affine functions interpolating $\widehat{\psi}_i^{-1/2}$ in $\widehat{a}$ and $\widehat{b}$. Then $\widehat{\phi}_i(\widehat{c}) \le \widehat{\psi}_i^{-1/2}(\widehat{c})$ and one obtains $(\widehat{\psi}_1(\widehat{c}) + \widehat{\psi}_2(\widehat{c}))^{-1/2} \ge \chi(\widehat{c})$ with $\chi(\widehat{x}) := (\widehat{\phi}_1^{-2}(\widehat{x}) + \widehat{\phi}_2^{-2}(\widehat{x}))^{-1/2}$. A straightforward calculation shows that $\chi'' < 0$ on $[\widehat{a}, \widehat{b}]$, in particular $\chi(\widehat{c}) \ge t\chi(\widehat{a}) + (1-t)\chi(\widehat{b})$. Since $\chi(\widehat{a}) = (\widehat{\psi}_1(\widehat{a}) + \widehat{\psi}_2(\widehat{a}))^{-1/2}$ and $\chi(\widehat{b}) = (\widehat{\psi}_1(\widehat{b}) + \widehat{\psi}_2(\widehat{b}))^{-1/2}$, it follows that $(\widehat{\psi}_1 + \widehat{\psi}_2)^{-1/2}$ is concave.

$\mathcal{H}$ *is closed in the compact-open topology*: Let $(\widehat{\psi}_n) \in \mathcal{H}$ be a sequence of functions from $\mathcal{H}$ converging uniformly on compact sets to $\widehat{\psi} \in \mathcal{H}$. If $\widehat{\psi} > 0$ on $\widehat{D}$ then also the sequence $(\widehat{\psi}_n^{-1/2})$ converges uniformly on compact subsets of $\widehat{D}$ to $\widehat{\psi}^{-1/2}$ and $\widehat{\psi}^{-1/2}$ is concave on $\widehat{D}$. The only thing to show is that whenever $\widehat{\psi}(\widehat{x}) = 0$ for some $\widehat{x} \in \widehat{D}$ then $\widehat{\psi} = 0$ on $\widehat{D}$. Indeed, if $\lim\widehat{\psi}_n(\widehat{x}) = 0$ and $\lim\widehat{\psi}_n(\widehat{y}) > 0$ for $\widehat{x}, \widehat{y}$ in $\widehat{D}$ then $\lim\widehat{\psi}_n^{-1/2}(\widehat{x}) = \infty$ and $\lim\widehat{\psi}_n^{-1/2}(\widehat{y}) < \infty$, and the concavity of $\widehat{\psi}_n^{-1/2}$ implies that $\lim\widehat{\psi}_n^{-1/2} = -\infty$ on the side of $\widehat{D} \setminus \{\widehat{y}\}$ that does not contain $\widehat{x}$. This is clearly impossible.

$\mathcal{H} - \mathcal{H} = \mathcal{C}(\widehat{X})$: If $\widehat{\psi} \in \mathcal{C}^2(\widehat{D})$ is positive and satisfies
$$\widehat{\psi}'' - \frac{3}{2}\frac{(\psi')^2}{\psi} \ge \varepsilon > 0$$
on $\widehat{D}$ then $\widehat{\phi} \in \mathcal{H}_2$ for all functions $\widehat{\phi}$ close to $\widehat{\psi}$ in the $\mathcal{C}^2$-norm, which is obvious from the equality
$$(\phi^{-1/2})'' = -\frac{1}{2}\phi^{-3/2}\left(\phi'' - \frac{3}{2}\frac{(\phi')^2}{\phi}\right).$$
It follows that $\mathcal{H}_2(\widehat{D}) - \mathcal{H}_2(\widehat{D})$ contains a small ε-ball of $\mathcal{C}^2(\widehat{D})$-functions around the origin with respect to the $\mathcal{C}^2$-norm. Since $\mathcal{H}_2(\widehat{D})$ is positively homogeneous, it follows that $\mathcal{H}_2(\widehat{D}) - \mathcal{H}_2(\widehat{D}) = \mathcal{C}^2(\widehat{D})$, whence $\mathcal{H}(\widehat{D}) - \mathcal{H}(\widehat{D}) = \mathcal{C}(\widehat{D})$.

$\mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}(\widehat{X})$ *is invariant by* $\widehat{\mathrm{P}}$: First we show that $\widehat{\mathrm{P}}(\widehat{\psi}_{\widehat{Z}}) \in \mathcal{H}$ for every $\widehat{\psi} \in \mathcal{H}$ and $\widehat{Z} \in \widehat{\mathcal{Z}}$. Choose $\widehat{g} : \widehat{C} \to \widehat{Z}$ with $S\widehat{g} \le 0$ such that $\widehat{\psi}_{\widehat{Z}} = \mathrm{P}_{\widehat{g}}(1)$. It follows that $\widehat{\mathrm{P}}(\widehat{\psi}_{\widehat{Z}}) = \mathrm{P}_{\widehat{f}} \circ \mathrm{P}_{\widehat{g}}(1) = \mathrm{P}_{\widehat{f}\circ\widehat{g}}(1)$ and $S(\widehat{f} \circ \widehat{g}) \le 0$, whence $\widehat{\mathrm{P}}(\widehat{\psi}_{\widehat{Z}}) \in \mathcal{H}$ by Lemma 3.6.

Now the family of finite partial sums $\{\sum_{Z\in\widehat{\mathcal{W}}}\widehat{\mathrm{P}}(\widehat{\psi}_{\widehat{Z}}) \mid \widehat{\mathcal{W}}\subset\widehat{\mathcal{Z}} \text{ finite}\}$ belongs to the positive cone $\mathcal{H}$, whence it is equicontinuous on every compact interval by part (ii). Since $\int\widehat{\mathrm{P}}(\widehat{\psi})\,d\widehat{m} = \int\widehat{\psi}\,d\widehat{m} < \infty$, the sum $\sum_{Z\in\widehat{\mathcal{Z}}}\widehat{\mathrm{P}}(\widehat{\psi}_{\widehat{Z}})$ is finite everywhere on $\widehat{X}$ and hence converges uniformly on compact sets to $\widehat{\mathrm{P}}(\widehat{\psi})\in\mathcal{H}$. ■

3.2. The Hopf decomposition. Let us briefly recall the definition of the Hopf decomposition which is used to split the Markov extension into a *recurrent* (= conservative) and a *nonrecurrent* (= dissipative) part. Most of the definitions and results stated below can be found in the monograph of Krengel [42].

3.9. Definition. A measurable set $\widehat{A}\subseteq\widehat{X}$ is a *wandering set* if $\widehat{f}^n(\widehat{A})\cap\widehat{A}=\emptyset$ for all $n>0$. The map $\widehat{f}$ is called *conservative* if it has no wandering sets of positive measure. If it does have wandering sets of positive measure then it is called *dissipative*.

3.10. Theorem (Hopf decomposition). *There is a partition $\widehat{X}=\widehat{X}_c\cup\widehat{X}_d$ into two disjoint measurable subsets $\widehat{X}_c$ and $\widehat{X}_d$ which are characterized uniquely modulo null sets by the following properties*:

1. *$\widehat{X}_c$ is invariant under $\widehat{f}$ and the restriction of $\widehat{f}$ to $\widehat{X}_c$ is conservative.*
2. *$\widehat{X}_d$ is a countable union of wandering sets.*

The sets $\widehat{X}_c$ and $\widehat{X}_d$ are called the conservative *and* dissipative *part of $\widehat{f}$, respectively.*

Proof. The set $\widehat{X}_d$ can easily be obtained as a union of wandering sets of positive measure using an exhaustion argument and $\widehat{X}_c := X\setminus X_d$ (see [42, §1.3]). ■

By a similar argument one can split $\widehat{X}_c$ further into two sets $\widehat{X}_+$ and $\widehat{X}_0 := \widehat{X}_c\setminus\widehat{X}_+$, where $\widehat{X}_+$ is the largest set supporting a finite absolutely continuous invariant measure. $\widehat{X}_+$ is called the *positively recurrent* part and $\widehat{X}_0$ is called the *null recurrent* part of $\widehat{\mathrm{P}}$.

3.11. Definition. For a nonnegative measurable function $\widehat{\psi}$ let

$$\widehat{\mathrm{S}}_n\widehat{\psi} := \sum_{k=0}^{n-1}\widehat{\mathrm{P}}^k\widehat{\psi} \quad\text{and}\quad \widehat{\mathrm{A}}_n\widehat{\psi} := \frac{1}{n}\sum_{k=0}^{n-1}\widehat{\mathrm{P}}^k\widehat{\psi}$$

and denote by $\widehat{\mathrm{S}}_\infty\widehat{\psi}$ and $\widehat{\mathrm{A}}_\infty\widehat{\psi}$ the corresponding limits as $n\to\infty$ if they exist.

Roughly speaking, for a given initial distribution $\widehat{\psi}$ the sum $\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x})$ measures the total amount of mass that passes through the point $\widehat{x}$ in n time steps when the distribution evolves under the influence of $\widehat{f}$, and $\widehat{\mathrm{A}}_n\widehat{\psi}(\widehat{x})$ measures the average amount. Consequently, the behaviour of $\widehat{\mathrm{S}}_n\widehat{\psi}$ as $n\to\infty$ can be used to distinguish the conservative and dissipative part, and the behaviour of $\widehat{\mathrm{A}}_n\widehat{\psi}$ as $n\to\infty$ can be used to distinguish the null recurrent from the positively recurrent part.

3.12. Proposition. *The Hopf decomposition is determined uniquely modulo null sets by the following property*: *If $\widehat{\psi}$ is a positive integrable function then $\widehat{X}_c=\{\widehat{\mathrm{S}}_\infty\widehat{\psi}=\infty\}$ a.s. and $\widehat{X}_d=\{\widehat{\mathrm{S}}_\infty\widehat{\psi}<\infty\}$ a.s.*

Proof. See [42, §3.1]. ■

3.13. Remark. Everything said up to now holds of course for an arbitrary nonsingular map on a σ-finite measure space and we are going to use analogous notions (without the hat sign) for the original system (X, f). $\diamond$

If one starts with an integrable regular density $\widehat{\psi}$ then all functions $\widehat{\mathrm{S}}_n\widehat{\psi}$ are again integrable regular densities. Because of Proposition 3.7 the sequence $(\log \widehat{\mathrm{S}}_n\widehat{\psi})_{n\in\mathbb{N}}$ is equicontinuous on compact subsets of $\widehat{X}$, which makes it sound reasonable that the behaviour (convergence or divergence) of the partial sums $(\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x}_0))_{n\in\mathbb{N}}$ at a particular point $\widehat{x}_0$ determines the behaviour of the partial sums $(\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x}))_{n\in\mathbb{N}}$ for all other points $\widehat{x}$ that belong to the same irreducible component as $\widehat{x}_0$. This is the core of the following theorem, which we adopt from Keller [39] to our setting.

3.14. Theorem (Keller [39]). *Fix a density $\widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$ and a point $\widehat{x}_0 \in \widehat{D} \in \widehat{\mathcal{D}}$, and let $s_n := \widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x}_0)$ for $n \in \mathbb{N}_\infty$. Then $s_n = O(n)$ and if $s_\infty > 0$ then the following holds.*

1. *If $(s_n)_{n\in\mathbb{N}}$ is bounded then $\widehat{f}$ is dissipative on $\widehat{D}$ and on every level $\widehat{C} \twoheadrightarrow \widehat{D}$. The sequence $(\widehat{\mathrm{S}}_n\widehat{\psi})_{\widehat{C}}$ converges uniformly on compact sets to $(\widehat{\mathrm{S}}_\infty\widehat{\psi})_{\widehat{C}} \in \mathcal{H}$. If $(s_n)_{n\in\mathbb{N}}$ is unbounded then $[\widehat{D}] = \widehat{\mathcal{X}}_{\mathrm{m}}$ and $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$.*

2. *If $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then it is $\widehat{m}$-ergodic and the sequence $\frac{1}{s_n}\widehat{\mathrm{S}}_n\widehat{\psi}$ converges uniformly on compact sets to a function $\widehat{h} \in \mathcal{H}$ which is strictly positive on $\widehat{X}_{\mathrm{m}}$ and vanishes everywhere else. The function $\widehat{h}$ is the unique* (3) *positive measurable function which is invariant under $\widehat{f}$. For every $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}$,*

$$\lim_{n\to\infty} \frac{1}{s_n}\widehat{\mathrm{S}}_n\widehat{\phi} = \frac{\int \widehat{\phi}\, d\widehat{m}}{\int \widehat{\psi}\, d\widehat{m}} \cdot \widehat{h} \quad \widehat{m}\text{-a.e. on } \widehat{X}_{\mathrm{m}}. \tag{3.2}$$

3. *If $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then $s_n/n \to \gamma$ for some constant $\gamma \geq 0$ and $\widehat{h}$ is integrable if and only if $\gamma > 0$. In particular, for every $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}$,*

$$\lim_{n\to\infty} \frac{1}{n}\widehat{\mathrm{S}}_n\widehat{\phi} = \frac{\int \widehat{\phi}\, d\widehat{m}}{\int \widehat{h}\, d\widehat{m}} \cdot \widehat{h} \quad \widehat{m}\text{-a.e. on } \widehat{X}_{\mathrm{m}}, \tag{3.3}$$

where the right hand side is zero if $\int \widehat{h}\, d\widehat{m} = \infty$.

(a) *If $\int \widehat{h}\, d\widehat{m} < \infty$ then the measure preserving system $(\widehat{f}, \widehat{h} \cdot \widehat{m})$ is the product of an exact system with a finite rotation.*

(b) *If $\int \widehat{h}\, d\widehat{m} = \infty$ then $\widehat{\mathrm{P}}^n\widehat{\psi}$ tends to zero uniformly on compact sets for every $\widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$ with $\widehat{\psi} \leq \widehat{h}$.*

Before proving the theorem we give some explanatory definitions and state some results needed during the proof in order to keep it self-contained.

3.15. Theorem (Chacon–Ornstein). *For $\widehat{\phi}, \widehat{\psi} \in \mathrm{L}^1_{\widehat{m}}$ with $\widehat{\psi} \geq 0$, $\widehat{\mathrm{S}}_n\widehat{\phi}/\widehat{\mathrm{S}}_n\widehat{\psi}$ converges a.e. on $\{\widehat{\mathrm{S}}_\infty\widehat{\psi} > 0\}$ to a finite limit. Moreover,*

(3) In this context "unique" always means "unique up to sets of measure zero and multiplication by a constant".

$$\lim_{n\to\infty} \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} = \frac{d(\mathrm{P}_{\widehat{F}}(\widehat{\phi})\cdot\widehat{m})|_{\widehat{\mathcal{J}}}}{d(\mathrm{P}_{\widehat{F}}(\widehat{\psi})\cdot\widehat{m})|_{\widehat{\mathcal{J}}}} \quad \widehat{m}\text{-a.e. on } \{\widehat{\mathrm{S}}_\infty\widehat{\psi}>0\}\cap\widehat{X}_c. \tag{3.4}$$

The expression on the right hand side is the Radon–Nikodym derivative, $\widehat{\mathcal{J}}$ denotes the σ-algebra of $\widehat{f}$-invariant subsets of $\widehat{X}_c$, and $\mathrm{P}_{\widehat{F}}$ is the transfer operator associated with the first entry map $\widehat{F}$ to $\widehat{X}_c$.

Proof. See Theorem 2.7 and Theorem 3.4 in [42, Chapter 3]. ∎

3.16. Remark. Intuitively, the operator $\mathrm{P}_{\widehat{F}}$ maps the part of the mass of $\widehat{\psi}$ which is located in the basin $\widehat{B}_c := \bigcup_{n\in\mathbb{N}} \widehat{f}^{-n}(\widehat{X}_c)$ of $\widehat{X}_c$ to the place where it first arrives in $\widehat{X}_c$ under iteration. If $\widehat{\phi}$ and $\widehat{\psi}$ vanish on the complement of the conservative part—which we denote by $\widehat{\phi},\widehat{\psi}\in \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$—then one can replace $\mathrm{P}_{\widehat{F}}(\widehat{\phi})$ and $\mathrm{P}_{\widehat{F}}(\widehat{\psi})$ by $\widehat{\phi}$ and $\widehat{\psi}$, respectively. ◇

3.17. Definition. The map $\widehat{f}$ is called *ergodic* (w.r.t. $\widehat{m}$) if $\widehat{f}(\widehat{A})=\widehat{A}$ implies $\widehat{A}=\emptyset$ modulo $\widehat{m}$ or $\widehat{A}=\widehat{X}$ modulo $\widehat{m}$ for every measurable set $\widehat{A}$. If $\widehat{Y}$ is an $\widehat{f}$-invariant subset of $\widehat{X}$ then $\widehat{f}$ is *ergodic on $\widehat{Y}$* (w.r.t. $\widehat{m}$) if it is ergodic w.r.t. the measure $1_{\widehat{Y}}\cdot\widehat{m}$.

The map $\widehat{f}$ is ergodic on its conservative part if and only if the σ-algebra $\widehat{\mathcal{J}}$ of $\widehat{f}$-invariant subsets of $\widehat{X}_c$ is trivial modulo $\widehat{m}$. If this is the case then the limit in the Chacon–Ornstein theorem becomes a very simple form on $\widehat{X}_c$.

3.18. Theorem (Chacon–Ornstein). *If $\widehat{f}$ is ergodic on the conservative part $\widehat{X}_c$ then for $\widehat{\phi},\widehat{\psi}\in\mathrm{L}^1_{\widehat{m}}$ with $\widehat{\psi}\ge 0$,*

$$\lim_{n\to\infty} \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} = \frac{\int_{\widehat{X}_c}\mathrm{P}_{\widehat{F}}(\widehat{\phi})\,d\widehat{m}}{\int_{\widehat{X}_c}\mathrm{P}_{\widehat{F}}(\widehat{\psi})\,d\widehat{m}} = \frac{\int_{\widehat{B}_c}\widehat{\phi}\,d\widehat{m}}{\int_{\widehat{B}_c}\widehat{\psi}\,d\widehat{m}} \quad \widehat{m}\text{-a.e. on } \{\widehat{\mathrm{S}}_\infty\widehat{\psi}>0\}\cap\widehat{X}_c, \tag{3.5}$$

where $\widehat{B}_c := \bigcup_{n\in\mathbb{N}}\widehat{f}^{-n}(\widehat{X}_c)$ denotes the basin of $\widehat{X}_c$. ∎

3.19. Theorem (Hopf). *If there is a nonnegative real-valued measurable function $\widehat{h}$ with $\{\widehat{h}>0\}=\widehat{X}_c$ and $\widehat{\mathrm{P}}(\widehat{h})=\widehat{h}$, then $\widehat{\mathrm{A}}_n\widehat{\psi}$ converges a.e. for every function $\widehat{\psi}\in\mathrm{L}^1_{\widehat{m}}$. The limit vanishes on the complement of the positively recurrent part $\widehat{X}_+$. For every $\widehat{\psi}\in\mathrm{L}^1_{\widehat{m}}(\widehat{X}_+)$ the sequence $(\widehat{\mathrm{A}}_n\widehat{\psi})_{n\in\mathbb{N}}$ converges in $\mathrm{L}^1_{\widehat{m}}$.*

Proof. See Theorem 3.12 in [42, Chapter 3]. The last statement follows from the previous ones because there is an integrable invariant density $\widetilde{h}$ with $\{\widetilde{h}>0\}=\widehat{X}_+$: This implies that $\int(|\widehat{\mathrm{A}}_n\widehat{\psi}|-c\widetilde{h})^+\,d\widehat{m}\le\int\widehat{\mathrm{A}}_n(|\widehat{\psi}|-c\widetilde{h})^+\,d\widehat{m}=\int(|\widehat{\psi}|-c\widetilde{h})^+\,d\widehat{m}$, which holds for any $c>0$. The right hand side becomes arbitrarily small as $c\to\infty$, whence the sequence $(\widehat{\mathrm{A}}_n\widehat{\psi})_{n\in\mathbb{N}}$ is uniformly integrable. ∎

3.20. Definition. A measure preserving system $(\widehat{f},\widehat{\mu})$ is called *exact* if it has a trivial tail field mod $\widehat{\mu}$. The tail field is $\widehat{\mathcal{B}}_\infty := \bigcap_{n\ge 0}\widehat{f}^{-n}\widehat{\mathcal{B}}$, where $\widehat{\mathcal{B}}$ denotes the σ-algebra of Borel subsets of $\widehat{X}$.

3.21. Proposition (Lin [43]). *A measure preserving system $(\widehat{f},\widehat{\mu})$ is exact if and only if $\|\widehat{\mathrm{P}}^n_{\widehat{\mu}}\widehat{\psi}-\int\widehat{\psi}\,d\widehat{\mu}\|_{1,\widehat{\mu}}\to 0$ for every $\widehat{\psi}\in\mathrm{L}^1_{\widehat{\mu}}$, i.e., if and only if $(\widehat{\mathrm{P}}^n_{\widehat{\mu}}\widehat{\psi})_{n\in\mathbb{N}}$ converges in $\mathrm{L}^1_{\widehat{\mu}}$ to a constant for every $\widehat{\psi}\in\mathrm{L}^1_{\widehat{\mu}}$.*

Here $\widehat{\mathrm{P}}_{\widehat{\mu}}$ denotes the transfer operator with respect to the reference measure $\widehat{\mu}$ instead of $\widehat{m}$. If $\widehat{\mu} = h\widehat{m}$ for a function $\widehat{h}$ with $\{\widehat{h} > 0\} = \widehat{X}_c$ then $\widehat{\mathrm{P}}_{\widehat{\mu}}(\widehat{\psi}) = \widehat{\mathrm{P}}(\widehat{\psi h})/\widehat{h}$ holds $\widehat{\mu}$-a.s. and $\widehat{m}$-a.s. on $\widehat{X}_c$. It follows that exactness of $\widehat{\mu}$ is equivalent to the fact that $\|\mathrm{P}^n\,\widehat{\psi} - \int \widehat{\psi}\,d\widehat{m} \cdot \widehat{h}\|_{1,\widehat{m}} \to 0$ holds for every $\widehat{\psi} \in \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$.

Proof of Theorem 3.14. Since $\widehat{\psi}$ is integrable, $\widehat{\mathrm{P}}^n\widehat{\psi}$ and $\widehat{\mathrm{S}}_n\widehat{\psi}$ are contained in $\mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$ for all $n \in \mathbb{N}$. Corollary 3.8 implies $\widehat{\mathrm{P}}^k\widehat{\psi}(\widehat{x}) \le C_{\widehat{x}} \int \widehat{\mathrm{P}}^k\widehat{\psi}\,d\widehat{m} = C_{\widehat{x}} \int \widehat{\psi}\,d\widehat{m} =: \widetilde{C}_{\widehat{x}}$ for all $k \in \mathbb{N}$ and $\widehat{x} \in \widehat{X}$, in particular, $\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x}) = O(n)$.

1. For two points $\widehat{x}$ and $\widehat{y}$ contained in the same level $\widehat{D} \in \widehat{\mathcal{D}}$ there is a constant $C_{\widehat{x},\widehat{y}}$ such that

$$C^{-1}_{\widehat{x},\widehat{y}} \le \frac{\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x})}{\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{y})} \le C_{\widehat{x},\widehat{y}} \quad \text{whenever } \widehat{\mathrm{S}}_n\widehat{\psi} > 0 \text{ on } \widehat{D}. \tag{3.6}$$

It follows that every level $\widehat{D}$ in the tower is contained entirely in one of the three sets $\{\widehat{\mathrm{S}}_\infty\widehat{\psi} = 0\}$, $\{\widehat{\mathrm{S}}_\infty\widehat{\psi} = \infty\}$, and $\{0 < \widehat{\mathrm{S}}_\infty\widehat{\psi} < \infty\}$. In the third case the sequence $(\widehat{\mathrm{S}}_n\widehat{\psi})_{\widehat{D}}$ converges pointwise and hence uniformly on compact sets to $(\widehat{\mathrm{S}}_\infty\widehat{\psi})_{\widehat{D}} \in \mathcal{H}$ by Proposition 3.7.

If $\widehat{C}$ and $\widehat{D}$ are two levels with $\widehat{C} \twoheadrightarrow \widehat{D}$ then one can find points $\widehat{x} \in \widehat{D}$, $\widehat{y} \in \widehat{C}$ and an integer j such that $\widehat{x} = \widehat{f}^j(\widehat{y})$. From

$$\widehat{\mathrm{P}}^{k+j}\widehat{\psi}(\widehat{x}) = \widehat{\mathrm{P}}^j(\widehat{\mathrm{P}}^k\widehat{\psi})(\widehat{x}) = \sum_{\widehat{u} \in \widehat{f}^{-j}(\widehat{x})} \frac{\widehat{\mathrm{P}}^k\widehat{\psi}(\widehat{u})}{|(\widehat{f}^j)'(\widehat{u})|} \ge \frac{\widehat{\mathrm{P}}^k\widehat{\psi}(\widehat{y})}{|(\widehat{f}^j)'(\widehat{y})|} \tag{3.7}$$

one obtains $\widehat{\mathrm{P}}^k\widehat{\psi}(\widehat{y}) \le |(\widehat{f}^j)'(\widehat{y})|\widehat{\mathrm{P}}^{k+j}\widehat{\psi}(\widehat{x})$ and consequently

$$\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{y}) \le |(\widehat{f}^j)'(\widehat{y})|\widehat{\mathrm{S}}_{n+j}\widehat{\psi}(\widehat{x}) \le |(\widehat{f}^j)'(\widehat{y})|(\widehat{\mathrm{S}}_n\widehat{\psi}(\widehat{x}) + j\widetilde{C}_{\widehat{x}}), \tag{3.8}$$

where $\widetilde{C}_{\widehat{x}}$ is the constant from the beginning of the proof. This implies that the set $\{\widehat{\mathrm{S}}_\infty\widehat{\psi} = \infty\}$ is $\widehat{f}$-invariant. Now statement 1 follows, except for the last sentence. But if $\widehat{\mathrm{P}}$ is conservative on $\widehat{D}$ then $\widehat{D}$ contains no wandering set of positive measure, which implies $\widehat{D} \twoheadrightarrow \widehat{D}$ and $[\widehat{D}] = \widehat{X}_{\mathrm{m}} = \widehat{X}_c$.

2. Assume that $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}_{\mathrm{m}}$, i.e., $\widehat{X}_c = \widehat{X}_{\mathrm{m}}$. First we show that there is at most one density $\widehat{h}$ in $\mathcal{H}$ such that $\widehat{\mathrm{P}}(\widehat{h}) = \widehat{h}$ and $\widehat{h}(\widehat{x}_0) = 1$. If $\widehat{h}_0, \widehat{h}_1 \in \mathcal{H}$ are two such densities then $\widehat{\mathrm{P}}(\widehat{h}_0 \wedge \widehat{h}_1) \le \widehat{h}_0 \wedge \widehat{h}_1$ by the positivity of $\widehat{\mathrm{P}}$, which implies $\widehat{\mathrm{P}}(\widehat{h}_0 \wedge \widehat{h}_1) = \widehat{h}_0 \wedge \widehat{h}_1$, because $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}_c$ and $\widehat{h}_i = 0$ on $\widehat{X}_d$. Now the functions $\widehat{g}_i^+ := (\widehat{h}_i - \widehat{h}_{1-i})^+ = \widehat{h}_i - \widehat{h}_0 \wedge \widehat{h}_1$ $(i = 0, 1)$ are nonnegative, continuous and invariant. (3.7) and the fact that preimages of all points $\widehat{x} \in \widehat{X}_c$ are dense in $\widehat{X}_c$ (cf. Lemma 3.24) imply that $\widehat{g}_i^+ = 0$ or $\widehat{g}_i^+ > 0$ on $\widehat{X}_c$. Since $\widehat{g}_1^+(\widehat{x}_0) = \widehat{g}_2^+(\widehat{x}_0) = 0$, it follows that $\widehat{g}_1^+ = \widehat{g}_2^+ = 0$, i.e., $\widehat{h}_1 = \widehat{h}_2$ on $\widehat{X}_c$.

Since $s_n \to \infty$ it follows that $\overline{\psi}_n := (1/s_n)\widehat{\mathrm{S}}_n\widehat{\psi}$ converges to 0 pointwise on $\widehat{X}_d$ and hence also uniformly on compact subsets of $\widehat{X}_d$. Because $\overline{\psi}_n(\widehat{x}_0) = 1$ for all n it follows from (3.6) and (3.8) that $0 < \liminf_{n\to\infty} \overline{\psi}_n(\widehat{x}) \le \limsup_{n\to\infty} \overline{\psi}_n(\widehat{x}) < \infty$ for every $\widehat{x} \in \widehat{X}_c = \widehat{X}_{\mathrm{m}}$. The sequence $(\overline{\psi}_n)_{n\in\mathbb{N}}$ is equicontinuous and bounded on compact sets, so there is a subsequence $\overline{\psi}_{n_j}$ converging uniformly on compact sets to a function $\overline{\psi} \in \mathcal{H}$

with $\{\overline{\psi} > 0\} = \widehat{X}_c$. From

$$\widehat{\mathrm{P}}\overline{\psi}(\widehat{x}) = \widehat{\mathrm{P}}(\lim_{j\to\infty} \overline{\psi}_{n_j})(\widehat{x}) \overset{(4)}{\leq} \liminf_{j\to\infty} \widehat{\mathrm{P}}\overline{\psi}_{n_j}(\widehat{x}) \leq \liminf_{j\to\infty} \frac{1}{s_{n_j}} \widehat{\mathrm{S}}_{n_j+1}\widehat{\psi}(\widehat{x}) \leq \liminf_{j\to\infty} \frac{1}{s_{n_j}} (\widehat{\mathrm{S}}_{n_j}\widehat{\psi}(\widehat{x}) + \widetilde{C}_{\widehat{x}}) = \overline{\psi}(\widehat{x}) \tag{3.9}$$

and from the conservativity on $\widehat{X}_c$ it follows that $\widehat{\mathrm{P}}\overline{\psi} = \overline{\psi}$. This not only proves the existence of a $\widehat{\mathrm{P}}$-invariant density but also that $\overline{\psi} = \widehat{h}$, where $\widehat{h} \in \mathcal{H}$ is the unique density from the beginning of step 2, because $\overline{\psi}_n(\widehat{x}_0) = 1$ for all n. Since the above argument applies to *any* subsequence of $\overline{\psi}_n$ which converges uniformly on compact sets, it follows that the sequence $(\overline{\psi}_n)_{n\in\mathbb{N}}$ itself converges uniformly on compact sets to $\widehat{h}$.

Claim. *$\widehat{f}$ is ergodic w.r.t. the Lebesgue measure $\widehat{m}$.*

First of all, it is sufficient to prove ergodicity on the conservative part, since almost every point enters $\widehat{X}_c$ eventually, i.e., $\widehat{B}_c := \bigcup_{n\in\mathbb{N}} f^{-n}(\widehat{X}_c) = \widehat{X}$ mod m: Indeed, if n is chosen large enough such that $\widehat{D}_n^+ \subseteq \widehat{X}_c$ and $\widehat{D}_n^- \subseteq \widehat{X}_c$, and if $\widehat{K} \subset \widehat{X}_d$ is the compact set obtained from $\widehat{X}_d$ by removing the leftmost and rightmost n-cylinder set of every level contained in $\widehat{X}_d$ then $\widehat{X} \setminus \widehat{K} \subseteq \widehat{f}^{-n}(\widehat{X}_c)$. For every level $\widehat{D} \in \widehat{\mathcal{D}}$ it follows from step 1 that $(\widehat{\mathrm{S}}_\infty 1_{\widehat{D}})_{\widehat{X}_d} \in \mathcal{H}$, since $1_{\widehat{D}} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$. Now $\sum_{n\geq 0} \widehat{m}\{\widehat{x} \in \widehat{D} \mid \widehat{f}^n\widehat{x} \in \widehat{K}\} = \int_{\widehat{K}} \widehat{\mathrm{S}}_\infty 1_{\widehat{D}}\, d\widehat{m} < \infty$ by Proposition 3.7, and the Borel–Cantelli Lemma implies that $\widehat{m}$-almost every $\widehat{x} \in \widehat{X}$ visits $\widehat{K}$ only finitely often and hence enters $\widehat{X}_c$ eventually.

Restricting $\widehat{f}$ to $\widehat{X}_c$ if necessary, we may assume for the proof of the claim that $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}$. Let $\widehat{\phi}$ be another integrable regular density and let $\overline{\phi}_n := (1/s_n)\widehat{\mathrm{S}}_n\widehat{\phi}$. The Chacon–Ornstein theorem (Theorem 3.15) implies that

$$\overline{\phi} := \lim_{n\to\infty} \overline{\phi}_n = \lim_{n\to\infty} \frac{1}{s_n} \widehat{\mathrm{S}}_n\widehat{\psi} \cdot \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} \quad \text{exists and is finite } \widehat{m}\text{-a.e.},$$

which in turn implies that $\overline{\phi}_n \to \overline{\phi}$ uniformly on compact sets and $\overline{\phi} \in \mathcal{H}$. Replacing $\widehat{\psi}$ by $\widehat{\phi}$ in (3.9) and using the conservativity it follows that $\overline{\phi}$ is invariant. Consequently, $\overline{\phi} = c \cdot \widehat{h}$ for some constant c. Since $\widehat{\psi}$ can also be chosen arbitrarily, this proves that

$$\lim_{n\to\infty} \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} \quad \text{is constant for any } \widehat{\phi}, \widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}},\ \widehat{\psi} \neq 0.$$

This shows that $(\widehat{f}, \widehat{m})$ is ergodic, using the subsequent Lemma 3.23, applied with $\mathcal{C} := (\mathcal{H} - \mathcal{H}) \cap \mathrm{L}^1_{\widehat{m}}$ and $\widehat{\psi} := \sum_{k\in\mathbb{N}} 2^{-k}(1_{\widehat{D}_k^+} + 1_{\widehat{D}_k^-}) \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}} \cap \mathrm{L}^\infty_{\widehat{m}}$. This finishes the proof of the claim. From now on it is no longer assumed that $\widehat{X}_c = \widehat{X}$.

Since the σ-algebra of f-invariant subsets of $\widehat{X}_c$ is trivial in $\widehat{X}_c$ and $\widehat{B}_c = \widehat{X}$ mod $\widehat{m}$, the Chacon–Ornstein theorem (Theorem 3.18) yields (3.2), because for every function $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}$,

$$\lim_{n\to\infty} \frac{1}{s_n} \widehat{\mathrm{S}}_n\widehat{\phi} = \lim_{n\to\infty} \frac{1}{s_n} \widehat{\mathrm{S}}_n\widehat{\psi} \cdot \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} = \frac{\int \widehat{\phi}\, d\widehat{m}}{\int \widehat{\psi}\, d\widehat{m}} \cdot \widehat{h} \quad \widehat{m}\text{-a.e. on } \widehat{X}_c.$$

(4) Cf. Remark 3.22.

3. By the ergodicity there is at most one measurable invariant density, so it must coincide with the regular density $\widehat{h}$ constructed in step 2. Note also that the positively recurrent part $\widehat{X}_+$ can only equal $\widehat{X}_c$ or $\emptyset$ $\widehat{m}$-a.s., since $\widehat{X}_+$ belongs to $\widehat{\mathcal{J}}$, and that the two cases correspond to $\int \widehat{h}\, d\widehat{m} < \infty$ and $\int \widehat{h}\, d\widehat{m} = \infty$, respectively. If $\int \widehat{h}\, d\widehat{m} < \infty$ then $\widehat{X}_+ = \widehat{X}_c$ and the Hopf Theorem (Theorem 3.19) implies that for every density $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$,

$$\lim_{n\to\infty} \widehat{\mathrm{A}}_n \widehat{\phi} = \frac{\int \widehat{\phi}\, d\widehat{m}}{\int \widehat{h}\, d\widehat{m}} \cdot \widehat{h} \quad \widehat{m}\text{-a.e. and in } \mathrm{L}^1_{\widehat{m}}.$$

In particular, $\widehat{\mathrm{A}}_n \widehat{\psi} \to (\int \widehat{\psi}\, d\widehat{m} / \int \widehat{h}\, d\widehat{m})\widehat{h}$. Together with $\overline{\psi}_n \to \widehat{h}$ this implies $s_n/n \to \gamma := \int \widehat{\psi}\, d\widehat{m} / \int \widehat{h}\, d\widehat{m}$. If $\int \widehat{h}\, d\widehat{m} = \infty$ then $\widehat{X}_+ = \emptyset$ and Theorem 3.19 implies that for every density $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}$ the sequence $\widehat{\mathrm{A}}_n \widehat{\phi}$ converges to 0 $\widehat{m}$-a.e. Applied to $\widehat{\psi}$ this yields $s_n/n \to 0$.

For the proof of 3(a) and 3(b) we only consider the case where the graph $(\widehat{\mathcal{X}}_{\mathrm{m}}, \to)$ is aperiodic. The case where the period p is greater than one can be reduced to the former in the same way as it is done for ordinary Markov chains by decomposing $\widehat{X}_c$ into disjoint sets $\widehat{X}_{c,1}, \dots, \widehat{X}_{c,p}$ which are cyclically permuted by $\widehat{f}$ and considering $\widehat{f}^p$ on every set $\widehat{X}_{c,i}$ (cf. Definition 2.20). Choose $\widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$ and let $\widetilde{\psi} := \limsup_{n\to\infty} \widehat{\mathrm{P}}^n \widehat{\psi}$. Since $\widehat{\mathrm{P}}^n \widehat{\psi} \le \widehat{h}$ for all n and $\widehat{h} \in \mathrm{L}^1_{\mu_{\widehat{x},1}}$ (5) for every $\widehat{x} \in \widehat{X}$, it follows that $\widehat{\mathrm{P}}(\widetilde{\psi}) \ge \limsup_{n\to\infty} \widehat{\mathrm{P}}(\widehat{\mathrm{P}}^n \widehat{\psi}) = \widetilde{\psi}$. Now $\widehat{h} - \widetilde{\psi}$ is nonnegative, vanishes on $\widehat{X}_d$, and $\widehat{\mathrm{P}}(\widehat{h} - \widetilde{\psi}) \le \widehat{h} - \widetilde{\psi}$, which implies $\widehat{\mathrm{P}}(\widehat{h} - \widetilde{\psi}) = \widehat{h} - \widetilde{\psi}$. It follows that $\widehat{\mathrm{P}}\widetilde{\psi} = \widetilde{\psi}$, i.e., $\widetilde{\psi} = c\widehat{h}$ for some constant c.

Claim. $\int \widetilde{\psi}\, d\widehat{m} \le \int \widehat{\psi}\, d\widehat{m}$.

Before proving the claim we finish the proof of the theorem with this information. If $\int \widehat{h}\, d\widehat{m} = \infty$ then $\int \widetilde{\psi}\, d\widehat{m} \le \int \widehat{\psi}\, d\widehat{m} < \infty$ implies $c = 0$, i.e., $\limsup_{n\to\infty} \widehat{\mathrm{P}}^n \widehat{\psi} = 0$. If $\int \widehat{h}\, d\widehat{m} < \infty$ then $\widehat{h}$ serves as an integrable majorant for $(\widehat{\mathrm{P}}^n \widehat{\psi})_{n\in\mathbb{N}}$ and the Fatou Lemma yields the reverse inequality which shows that $\int \widetilde{\psi}\, d\widehat{m} = \int \widehat{\psi}\, d\widehat{m} = \int \widehat{\mathrm{P}}^n \widehat{\psi}\, d\widehat{m}$ for all n. Moreover, $\int (\widehat{\mathrm{P}}^n \widehat{\psi} - \widetilde{\psi})^+ \, d\widehat{m} \to 0$ by the Dominated Convergence Theorem, which in view of the previous equality also implies $\int (\widehat{\mathrm{P}}^n \widehat{\psi} - \widetilde{\psi})^- \, d\widehat{m} \to 0$. This shows that $\widehat{\mathrm{P}}^n \widehat{\psi}$ converges to $\widetilde{\psi} = c\widehat{h}$ in $\mathrm{L}^1_{\widehat{m}}$.

Proof of the claim. Let $\widehat{K}$ be a compact subset of $\widehat{X}_c$. Choose an arbitrary point $\widehat{x} \in \widehat{K}$ and a sequence $(n_k)_{k\in\mathbb{N}}$ such that $\widetilde{\psi}(\widehat{x}) = \lim_{k\to\infty} \widehat{\mathrm{P}}^{n_k} \widehat{\psi}(\widehat{x})$. Now fix $N \in \mathbb{N}$ and let $\widetilde{\psi}' := \limsup_{k\to\infty} \widehat{\mathrm{P}}^{n_k - N} \widehat{\psi}$. Because

$$\widehat{\mathrm{P}}^N \widetilde{\psi}(\widehat{x}) = \widetilde{\psi}(\widehat{x}) = \lim_{k\to\infty} \widehat{\mathrm{P}}^N (\widehat{\mathrm{P}}^{n_k - N} \widehat{\psi}(\widehat{x})) \le \widehat{\mathrm{P}}^N (\limsup_{k\to\infty} \widehat{\mathrm{P}}^{n_k - N} \widehat{\psi}(\widehat{x})) = \widehat{\mathrm{P}}^N \widetilde{\psi}'(\widehat{x})$$

and because $\widetilde{\psi}' \le \widetilde{\psi}$, it follows that $\widetilde{\psi}(\widehat{y}) = \widetilde{\psi}'(\widehat{y})$ for $\widehat{y} \in \widehat{f}^{-N}(\widehat{x})$.

For any $\varepsilon > 0$ there is a $\delta > 0$ such that $|\widehat{\mathrm{P}}^n \widehat{\psi}(\widehat{y}) / \widehat{\mathrm{P}}^n \widehat{\psi}(\widehat{z})| \le 1 + \varepsilon$ for all $n \in \mathbb{N}$ and all $\widehat{y}, \widehat{z} \in \widehat{K}$ with $d(\widehat{y}, \widehat{z}) \le \delta$ by Proposition 3.7. In Lemma 3.24 it will be shown that there is a δ-cover for $\widehat{K}$ consisting of finitely many points from $\widehat{f}^{-N}(\widehat{x})$ if N is chosen large enough. Now $\widetilde{\psi}(\widehat{y}) = \widetilde{\psi}'(\widehat{y}) \le (1+\varepsilon)\widehat{\mathrm{P}}^{n_k - N} \widehat{\psi}(\widehat{y})$ for those finitely many points if k is large enough, which implies $\widetilde{\psi} \le (1+\varepsilon)^2 \widehat{\mathrm{P}}^{n_k - N} \widehat{\psi}$ on $\widehat{K}$. It follows that for n large

(5) Cf. Remark 3.22.

$$\int_{\widehat{K}} \widetilde{\psi}\, d\widehat{m} \le (1+\varepsilon)^2 \int_{\widehat{K}} \widehat{\mathrm{P}}^{n_k - N}\widehat{\psi}\, d\widehat{m} \le (1+\varepsilon)^2 \int \widehat{\psi}\, d\widehat{m}.$$

Because $\widehat{K} \subset \widehat{X}_c$ and $\varepsilon > 0$ were chosen arbitrarily this proves the claim. ■

3.22. Remark. In the proof we frequently swapped a limit operation with some iterate $\widehat{\mathrm{P}}^k$ of the transfer operator. This can easily be done pointwise if one considers $\widehat{\mathrm{P}}^k(\cdot)(\widehat{x})$ as a discrete measure $\mu_{\widehat{x},k} := \sum_{\widehat{y}\in \widehat{f}^{-k}(\widehat{x})} |(\widehat{f}^k)'(\widehat{y})|^{-1}\delta_{\widehat{y}}$, where $\delta_{\widehat{y}}$ denotes the Dirac measure in $\widehat{y}$, and applies the Fatou Lemma or the Monotone or Dominated Convergence Theorem. ◇

The following lemma does not use any specific properties of the Markov extension. It holds for an arbitrary nonsingular map $\widehat{f}$ on a σ-finite measure space $(\widehat{X}, \widehat{m})$.

3.23. Lemma. *Assume that $\widehat{f}$ is conservative. If there is a function $0 < \widehat{\psi} \in \mathrm{L}^1_{\widehat{m}} \cap \mathrm{L}^\infty_{\widehat{m}}$ and a dense subset $\mathcal{C}$ of $\mathrm{L}^1_{\widehat{m}}$ such that for every $\widehat{\phi} \in \mathcal{C}$,*

$$\lim_{n\to\infty} \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} \quad \textit{is constant } \widehat{m}\textit{-a.e.,} \tag{3.10}$$

then $(\widehat{f}, \widehat{m})$ is ergodic.

Proof. Assume w.l.o.g. that $\int \widehat{\psi}\, d\widehat{m} = 1$ and let $\widetilde{m} := \widehat{\psi}\cdot\widehat{m}$. Then $\mathrm{L}^1_{\widehat{m}} \subseteq \mathrm{L}^1_{\widetilde{m}}$ because $\widehat{\psi}$ is bounded $\widehat{m}$-a.e. Since $\widehat{\psi} \in \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$ is a probability density, for any $\widehat{\phi} \in \mathrm{L}^1_{\widetilde{m}}(\widehat{X}_c)$, in particular for any $\widehat{\phi} \in \mathrm{L}^1_{\widehat{m}}(\widehat{X}_c)$, (3.4) can also be stated as follows:

$$\lim_{n\to\infty} \frac{\widehat{\mathrm{S}}_n\widehat{\phi}}{\widehat{\mathrm{S}}_n\widehat{\psi}} = \mathrm{E}_{\widetilde{m}}[\widehat{\phi} \mid \widehat{\mathcal{J}}] \quad \widehat{m}\text{-a.e. on } \widehat{X}_c. \tag{3.11}$$

where $\mathrm{E}_{\widetilde{m}}[\,\cdot \mid \widehat{\mathcal{J}}]$ denotes the conditional expectation with respect to the σ-algebra $\widehat{\mathcal{J}}$ in $\widehat{X}_c$. Having a look at (3.11) it is clear that condition (3.10) is equivalent to the requirement that $\mathrm{E}_{\widetilde{m}}[\widehat{\phi} \mid \widehat{\mathcal{J}}] = \mathrm{E}_{\widetilde{m}}\left[\widehat{\phi}\right]$ for every $\widehat{\phi} \in \mathcal{C}$. But $\mathcal{C}$ is dense in L^1_{m} with respect to the $\mathrm{L}^1_{\widetilde{m}}$-norm, because $\|\cdot\|_{1,\widetilde{m}} \le \|\widehat{\psi}\|_\infty \|\cdot\|_{1,\widehat{m}}$ and $\mathrm{L}^1_{\widehat{m}}$ is dense in $\mathrm{L}^1_{\widetilde{m}}$, because the sequence $\widehat{g}_n := \widehat{g}\cdot 1_{\{\psi \ge 1/n\}} \in \mathrm{L}^1_{\widehat{m}}$ converges to $\widehat{g}$ in $\mathrm{L}^1_{\widetilde{m}}$. It follows that $\mathrm{E}_{\widetilde{m}}[\widehat{\phi} \mid \mathcal{J}] = \mathrm{E}_{\widetilde{m}}\left[\widehat{\phi}\right]$ holds for every $\widehat{\phi} \in \mathrm{L}^1_{\widetilde{m}}$, which implies that $\mathcal{C}$ is trivial modulo $\widetilde{m}$ and hence also modulo $\widehat{m}$. ■

3.24. Lemma. *Assume that $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$ and that the graph $(\widehat{\mathcal{X}}_{\mathrm{m}}, \to)$ has period p. Let $\widehat{X}_{\mathrm{m},j}$, $1 \le j \le p$, be the cyclic decomposition of $\widehat{X}_{\mathrm{m}}$. Then for every $\widehat{x} \in \widehat{X}_{\mathrm{m},j}$, every compact set $\widehat{K} \subset \widehat{X}_{\mathrm{m},j}$ and every $\delta > 0$ there is an integer n such that the δ-neighbourhood of the set $\widehat{f}^{-np}(\widehat{x})$ covers $\widehat{K}$. In particular, the preiterates of $\widehat{x}$ are dense in $\widehat{X}_{\mathrm{m}}$.*

Proof. Assume w.l.o.g. that $p = 1$. It is sufficient to prove the following: For every interval $\widehat{I} \subset \widehat{X}_{\mathrm{m}}$ there is an integer $n(I)$ such that $I \cap f^{-n}(x) \ne \emptyset$ for all $n \ge n(I)$. The lemma then follows using a compactness argument.

Since there are no homtervals ([6]) inside $\widehat{X}_{\mathrm{m}} = \widehat{X}_c$, the interval $\widehat{I}$ contains an entire cylinder set $\widehat{Z} \in \widehat{\mathcal{Z}}_{n_1}$ if n_1 is chosen large enough. Let $\widehat{C} := \widehat{f}^{n_1}(\widehat{Z})$ and $\widehat{D} := \widehat{D}[\widehat{x}]$.

([6]) The existence of homtervals contradicts conservativity. For more information, cf. Section 3.4.

Because $(\widehat{\mathcal{X}}_{\mathrm{m}}, \to)$ is aperiodic, there is an integer n_2 such that $\widehat{C} \overset{n}{\twoheadrightarrow} \widehat{D}$ for all $n \geq n_2$. This implies that $\widehat{I} \cap f^{-n}(\widehat{x}) \neq \emptyset$ for all $n \geq n_1 + n_2$. ■

3.25. Definition. If $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then $\widehat{f}$ is called (*essentially*) *conservative*. Otherwise, $\widehat{f}$ is called (*purely*) *dissipative*. If $\widehat{f}$ is essentially conservative then the absolutely continuous invariant measure is denoted by $\widehat{\mu} := \widehat{h}\widehat{m}$. If $\widehat{h}$ is $\widehat{m}$-integrable then $\widehat{f}$ is said to be *positively recurrent* (on $\widehat{X}_{\mathrm{m}}$) and if it is not $\widehat{m}$-integrable then $\widehat{f}$ is said to be *null recurrent* (on $\widehat{X}_{\mathrm{m}}$). In the positively recurrent case let $\widehat{h}$ be normalized such that $\widehat{\mu}$ becomes a probability measure.

3.3. The asymptotic behaviour of points on the tower. In this section we will see that typical points behave quite differently depending on whether $\widehat{f}$ is essentially conservative or not. In the former case their orbit is everywhere dense in $\widehat{X}_{\mathrm{m}}$, while in the latter case the distance of the points $\widehat{x}_n$ to the boundary of the current level $\widehat{D}[\widehat{x}_n]$ approaches zero. In the conservative case one can say a bit more: If $\widehat{f}$ is positively recurrent then the orbits of $\widehat{m}$-typical points are distributed according to the invariant probability measure $\widehat{\mu}$ (in particular, every subset of $\widehat{X}_{\mathrm{m}}$ of positive Lebesgue measure is visited with positive frequency), otherwise the orbit of an $\widehat{m}$-typical point spends most of the time near the endpoints of the levels. In order to make this more precise we need to introduce some notation.

3.26. Definition. For every point $\widehat{x} \in \widehat{X}$ which is not precritical denote by

$\widehat{Z}_n[\widehat{x}] \in \widehat{\mathcal{Z}}_n$	the cylinder set of order n containing $\widehat{x}$,
$\widehat{D}[\widehat{x}] \in \widehat{\mathcal{D}}$	the level in the tower containing $\widehat{x}$,
$\widehat{D}_n[\widehat{x}] := \widehat{f}^n(\widehat{Z}_n[\widehat{x}]) = \widehat{D}[\widehat{x}_n]$	the level containing the nth iterate $\widehat{x}_n$,
$r(\widehat{x}) := \mathrm{dist}(\widehat{x}, \partial\widehat{D}[\widehat{x}])$	the distance of $\widehat{x}$ to the endpoints of its level,
$r_n(\widehat{x}) := r(\widehat{x}_n)$	the distance of $\widehat{x}_n$ to the endpoints of its level.

Finally, let $\widehat{X}_\varepsilon := \{\widehat{x} \in \widehat{X} \mid r(\widehat{x}) \geq \varepsilon\}$ and $\widehat{D}_\varepsilon := \widehat{D} \cap \widehat{X}_\varepsilon$ for every level $\widehat{D} \in \widehat{\mathcal{D}}$.

3.27. Theorem. *Let f be a Lorenz map with negative Schwarzian derivative and let $(\widehat{X}, \widehat{f})$ be its canonical Markov extension. Then the following holds.*

1. *If $\widehat{f}$ is purely dissipative then $r_n(\widehat{x}) \to 0$ for $\widehat{m}$-a.e. $\widehat{x} \in \widehat{X}$.*
2. *If $\widehat{f}$ is essentially conservative then $\omega(\widehat{x}) = \widehat{X}_{\mathrm{m}}$ for $\widehat{m}$-a.e. $\widehat{x} \in \widehat{X}$. Moreover,*
   - (a) *if $\int \widehat{h}\, d\widehat{m} = 1$ then $n^{-1}\sum_{k=0}^{n-1} \delta_{\widehat{x}_k} \to \widehat{\mu}$ weakly as $n \to \infty$ for $\widehat{m}$-a.e. $\widehat{x}$,*
   - (b) *if $\int \widehat{h}\, d\widehat{m} = \infty$ then $\lim_{n\to\infty} n^{-1}\,\mathrm{card}\{k < n \mid r_k(\widehat{x}) \geq \varepsilon\} = 0$ for $\widehat{m}$-a.e. $\widehat{x}$ and every $\varepsilon > 0$.*

Proof. 1) Since $1_{\widehat{D}} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$ and $\widehat{\mathrm{P}}$ is dissipative, $\widehat{\mathrm{S}}_\infty 1_{\widehat{D}} \in \mathcal{H}$ by Theorem 3.14(1). Now $\sum_{n\geq 0} \widehat{m}\{\widehat{x} \in \widehat{D} \mid \widehat{x}_n \in \widehat{X}_\varepsilon\} = \sum_{n\geq 0} \int 1_{\widehat{D}} \cdot 1_{\widehat{X}_\varepsilon} \circ \widehat{f}^n\, d\widehat{m} = \int_{\widehat{X}_\varepsilon} \widehat{\mathrm{S}}_\infty 1_{\widehat{D}}\, d\widehat{m}$ and the Borel–Cantelli Lemma implies that almost every point $\widehat{x}$ in $\widehat{D}$ visits $\widehat{X}_\varepsilon$ only finitely many times, provided we are able to show that $\widehat{\mathrm{S}}_\infty 1_{\widehat{D}}$ is integrable on $\widehat{X}_\varepsilon$. This will be achieved in Lemma 3.31.

2) Let $\widehat{I} \subseteq \widehat{X}_c$ be a nondegenerate interval and $\widehat{B}_{\widehat{I}}$ the set of points that enter $\widehat{I}$ infinitely often. Then $\widehat{f}(\widehat{B}_{\widehat{I}}) = \widehat{B}_{\widehat{I}}$ and $\widehat{m}(\widehat{B}_{\widehat{I}}) \geq \widehat{m}(\widehat{I}) > 0$, because $\widehat{f}$ is infinitely

recurrent on $\widehat{X}_c$. Since $\widehat{f}$ is ergodic, it follows that $\widehat{B}_{\widehat{I}} = \widehat{X} \bmod \widehat{m}$. Because $\widehat{I}$ was arbitrary, this proves that $\omega(\widehat{x}) = \widehat{X}_{\mathrm{m}}$ for $\widehat{m}$-a.e. point $\widehat{x} \in \widehat{X}$.

Since $\widehat{\mu}$ is ergodic, statements 1 and 2 follow from Birkhoff's Ergodic Theorem together with Lemma 3.30, which shows that $\widehat{\mu}(\widehat{X}_\varepsilon) < \infty$ even if $\widehat{\mu}$ is not finite. ■

3.28. Remark. As in the proof of part 1 of the lemma it follows from the Borel–Cantelli Lemma that every compact set $\widehat{K} \subset \widehat{X}_d$ is visited only finitely often by almost every point, since $\widehat{\mathrm{S}}_\infty 1_{\widehat{D}}$ is integrable on $\widehat{K}$ for every level $\widehat{D} \in \widehat{\mathcal{D}}$. ◇

The following lemmas complete the proof of Theorem 3.27.

3.29. Lemma. *Let $\widehat{g} \in \mathcal{H}$ be $\widehat{\mathrm{P}}$-subinvariant, K a compact subset of $Z \in \mathcal{Z}_n$ and let $\widehat{K} := \bigcup\{\pi^{-1}K \cap \widehat{D} \mid \widehat{D} \in \widehat{\mathcal{D}},\ \pi^{-1}Z \cap \widehat{D} \in \widehat{\mathcal{Z}}_n$ is complete* (7)*$\}$. Then $\int_{\widehat{K}} \widehat{g}\, d\widehat{m} < \infty$.*

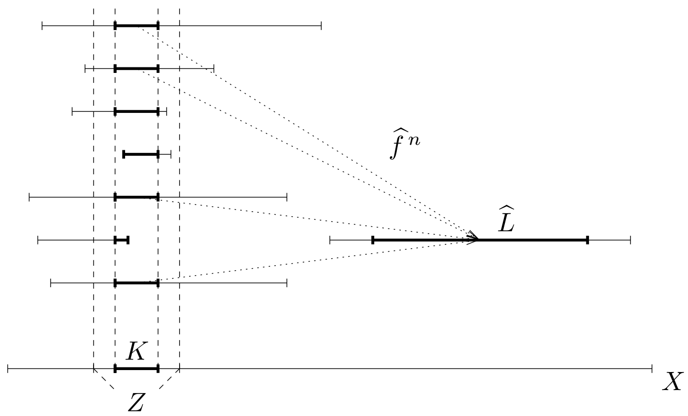


Fig. 3.1. Why $\widehat{g}$ is integrable on $\widehat{K}$. The compact subinterval K of the cylinder set $Z \in \mathcal{Z}_n$ and above it some of the levels that intersect components of $\pi^{-1}(K)$ (thick lines). Four of them belong to $\widehat{K}$ and are mapped to $\widehat{L}$ by $\widehat{f}^{\,n}$.

Proof. By definition of $\widehat{K}$, whenever $\widehat{K} \cap \widehat{D}$ is nonvoid it is contained in a complete n-cylinder set $\widehat{Z} \subseteq \pi^{-1}Z$. This implies that all these sets are mapped by $\widehat{f}^n$ onto the same compact subset $\widehat{L}$ on the same level of the tower (cf. Figure 3.1). Since $\widehat{g}$ is bounded on $\widehat{L}$, it follows from Proposition 3.7 that

$$\int_{\widehat{K}} \widehat{g}\, d\widehat{m} \le \int_{\widehat{f}^{-n}\widehat{L}} \widehat{g}\, d\widehat{m} = \int_{\widehat{L}} \widehat{\mathrm{P}}^n \widehat{g}\, d\widehat{m} \le \int_{\widehat{L}} \widehat{g}\, d\widehat{m} < \infty. \ ■$$

3.30. Lemma. *Assume that $\widehat{f}$ is essentially conservative and let $\widehat{h}$ be the invariant density for $\widehat{f}$ (which need not be integrable). Then $\int_{\widehat{X}_\varepsilon} \widehat{h}\, d\widehat{m} < \infty$ for every $\varepsilon > 0$.*

Proof. It is enough to estimate the integral on the conservative part because $\widehat{h}$ vanishes on $\widehat{X}_d$. Since $\widehat{X}_c$ cannot contain any homtervals, the length of all cylinder sets $Z \in \mathcal{Z}_n$ which are contained in $\pi(\widehat{X}_c)$ tends to 0 uniformly in n as $n \to \infty$. Divide the set

(7) Cf. Remark 2.6.

$\pi(\widehat{X}_c)$ into finitely many intervals $I_1, \dots, I_l$ of length less than $\delta < \varepsilon/2$ and let n be large enough such that every interval I_k contains a cylinder set $Z_k \in \mathcal{Z}_n$. Choose compact subintervals $K_k \subset Z_k$ and let $\widehat{K}_k := \bigcup\{\pi^{-1}K_k \cap \widehat{D} \mid \widehat{D} \in \widehat{\mathcal{D}},\ \pi^{-1}I_k \cap \widehat{D}_\varepsilon \neq \emptyset\}$ and $\widehat{K} := \bigcup_{i=1}^l \widehat{K}_k$. Observe that $\widehat{K} \subseteq \widehat{X}_\delta$, since $\delta < \varepsilon/2$. Because $\widehat{h}$ is invariant, it follows from Lemma 3.29 that $\int_{\widehat{K}} \widehat{h}\, d\widehat{m} = \sum_{k=1}^l \int_{\widehat{K}_k} \widehat{h}\, d\widehat{m} < \infty$.

Now we fix a level $\widehat{D} \in \widehat{\mathcal{D}}$ such that $\widehat{D}_\varepsilon \neq \emptyset$ and want to estimate the integral of $\widehat{h}$ over $\widehat{D}_\varepsilon$. According to Corollary 3.8, there is a constant $C = C(\delta)$ such that

$$C^{-1}h(\widehat{x}) \le \frac{1}{\widehat{m}(\widehat{B})} \int_{\widehat{B}} \widehat{h}\, d\widehat{m} \le Ch(\widehat{x})$$

for every $\widehat{x} \in \widehat{D}_\delta$ and every subset $\widehat{B} \subseteq \widehat{D}_\delta$ of positive measure. Consequently,

$$\int_{\widehat{D}_\varepsilon} \widehat{h}\, d\widehat{m} \le C^2 \frac{\widehat{m}(\widehat{D}_\varepsilon)}{\widehat{m}(\widehat{K} \cap \widehat{D})} \int_{\widehat{K}\cap\widehat{D}} \widehat{h}\, d\widehat{m} \le \frac{C^2}{m} \int_{\widehat{K}\cap\widehat{D}} \widehat{h}\, d\widehat{m}$$

where $m := \min_{1\le k\le l} m(K_k)$. Now $\int_{\widehat{X}_\varepsilon} \widehat{h}\, d\widehat{m} \le (C^2/m) \int_{\widehat{K}} \widehat{h}\, d\widehat{m} < \infty$ follows. ∎

3.31. Lemma. *Assume that $\widehat{f}$ is purely dissipative. Then $\int_{\widehat{X}_\varepsilon} \widehat{\mathrm{S}}_\infty \widehat{\psi}\, d\widehat{m} < \infty$ for every $\widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$.*

Proof. Because $\widehat{\psi} \in \mathcal{H} \cap \mathrm{L}^1_{\widehat{m}}$ and $\widehat{\mathrm{P}}$ is dissipative, $\widehat{\mathrm{S}}_\infty \widehat{\psi} \in \mathcal{H}$ by Theorem 3.14. Since $\widehat{\psi} + \widehat{\mathrm{P}}(\widehat{\mathrm{S}}_n \widehat{\psi}) = \widehat{\mathrm{S}}_{n+1} \widehat{\psi}$ for all n, we get $\widehat{\psi} + \widehat{\mathrm{P}}(\widehat{\mathrm{S}}_\infty \widehat{\psi}) = \widehat{\psi} + \lim_{n\to\infty} \widehat{\mathrm{P}}(\widehat{\mathrm{S}}_n \widehat{\psi}) = \widehat{\mathrm{S}}_\infty \widehat{\psi}$ ([8]). This proves that $\widehat{\mathrm{S}}_\infty \widehat{\psi}$ is subinvariant.

Divide the base $X = \pi(\widehat{X})$ into finitely many intervals $I_1, \dots, I_l$ of equal length $\delta < \varepsilon/2$. Obviously, every interval I_k either contains a cylinder set of order n_k for n_k sufficiently large or it is the union of one or two homtervals. Hence there is an integer n such that every interval I_k either contains a cylinder set $Z_k \in \mathcal{Z}_n$ or a homterval J_k. With every interval I_k we associate an interval $K_k \subset I_k$ in the following way:

1. If I_k contains a cylinder set $Z_k \in \mathcal{Z}_n$ then let K_k be a compact subinterval of Z_k.
2. Otherwise, I_k contains a homterval J_k. It is easily checked that either all iterates $f^n(J_k)$ are disjoint or J_k is contained in the basin of a periodic attractor. In the former case let $K_k = I_k$ and in the latter case choose a fundamental basin for the periodic attractor and let $K_k \subset J_k$ be an interval with constant transfer time to the fundamental basin. Now in both cases all iterates $f^n(K_k)$ of K_k are disjoint.

Let $\widehat{K}_k := \bigcup\{\pi^{-1}K_k \cap \widehat{D} \mid \widehat{D} \in \widehat{\mathcal{D}},\ \pi^{-1}I_k \cap \widehat{D}_\varepsilon \neq \emptyset\}$ and $\widehat{K} := \bigcup_{i=1}^l \widehat{K}_k$. Observe that $\widehat{K} \subseteq X_\delta$, since $\delta < \varepsilon/2$.

We claim that $\int_{\widehat{K}} \widehat{\mathrm{S}}_\infty \widehat{\psi}\, d\widehat{m} < \infty$. Indeed, if K_k is chosen according to case 1 then $\int_{\widehat{K}_k} \widehat{\mathrm{S}}_\infty \widehat{\psi}\, d\widehat{m} < \infty$ by Lemma 3.29 and if K_k is chosen as in case 2 then the sets $\widehat{f}^{-n}(\widehat{K}_k)$, $n \in \mathbb{N}$, are disjoint. Hence

$$\int_{\widehat{K}_k} \widehat{\mathrm{S}}_\infty \widehat{\psi}\, d\widehat{m} = \sum_{n\ge 0} \int_{\widehat{f}^{-n}\widehat{K}_k} \widehat{\psi}\, d\widehat{m} = \int_{\bigcup_{n\ge0} f^{-n}\widehat{K}_k} \widehat{\psi}\, d\widehat{m} \le \int \widehat{\psi}\, d\widehat{m} < \infty.$$

([8]) Cf. Remark 3.22.

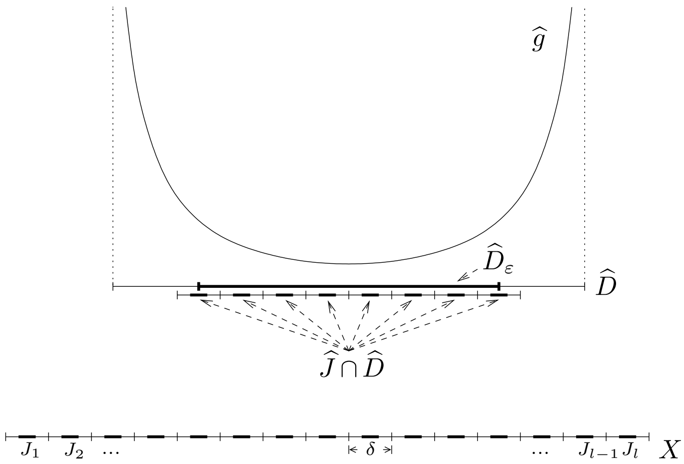


Fig. 3.2. Estimating the integral of $\widehat{g}$ on $\widehat{D}_\varepsilon$. The density $\widehat{g}$ plotted over the level $\widehat{D}$, where $\widehat{g} := \widehat{h}$ in Lemma 3.30 and $\widehat{g} := \widehat{\mathrm{S}}_\infty\widehat{\psi}$ in Lemma 3.31. To estimate the integral of $\widehat{g}$ over $\widehat{D}_\varepsilon$, its integral over $\widehat{J} \cap \widehat{D}$ is used as reference.

Taking the finite sum over the integers k proves the claim. Now as in the proof of Lemma 3.30 it follows that $\int_{\widehat{X}_\varepsilon} \widehat{\mathrm{S}}_\infty\widehat{\psi}\, d\widehat{m} \le (C^2/m)\int_{\widehat{K}} \widehat{\mathrm{S}}_\infty\widehat{\psi}\, d\widehat{m} < \infty$, where $m := \min_{1\le k\le l} m(K_k)$. ■

3.4. Wandering intervals. In the next section we will relate the results obtained for the Markov extension $(\widehat{X}, \widehat{f})$ to the original system (X, f). But before doing this we make a little excursion to homtervals and wandering intervals for Lorenz maps.

3.32. Definition (Wandering interval). An open interval J is called a *homterval* ([9]) if $f^n : J \to f^n(J)$ is a homeomorphism for every n, or equivalently, if it contains no precritical points. If J is a homterval which is not contained in a larger homterval then it is called a *maximal* homterval. A maximal homterval J is called

1. a *preperiodic* interval if there are indices $0 \le m < n$ such that $J_m \cap J_n \ne \emptyset$, and
2. a *wandering* interval if the intervals J_n, $n \ge 0$, are pairwise disjoint.

An arbitrary homterval J is called *preperiodic* resp. *wandering* if the maximal homterval which contains it is preperiodic resp. wandering. Note that preperiodic does not mean strictly preperiodic here.

3.33. Remark. From the definition it is obvious that every homterval J is either preperiodic or wandering. It can be easily checked that J is preperiodic if and only if every point in J converges to a periodic orbit, and J is wandering if and only if no point in J converges to a periodic orbit. Note also that a nonmaximal interval whose iterates are pairwise disjoint need not be wandering. For example, take a periodic attractor and let J be a component of its fundamental basin. ◇

([9]) The word *homterval* is an amalgamation of the words *home*omorphism and in*terval*.

For smooth interval maps with nonflat critical points it is well known that there are no wandering intervals (cf. de Melo & van Strien [13, Chapter IV]). This is a very pleasant property which makes "life" a lot easier in the continuous case. For example, it implies that if such a map has no periodic attractors then the length of the n-cylinder sets $Z \in \mathcal{Z}_n$ tends to 0 uniformly as $n \to \infty$ and the Contraction Principle can be applied. With Lorenz maps we are not in such a lucky situation. Indeed, there are cases where wandering intervals occur inevitably, namely if the Lorenz map is renormalizable to a gap map with irrational rotation number.

Recall that a gap map is a Lorenz map which is not surjective on its dynamical interval $I := [c_1^+, c_1^-]$, or equivalently for which $c_2^- < c_2^+$ holds (cf. Example 2.37), and denote the gap by $J := (c_2^-, c_2^+) = I \setminus f(I)$. Since f is injective on I it follows that $f^n(J) = f^n(I) \setminus f^{n+1}(I)$, whence the iterates of J are pairwise disjoint. It is not hard to see that the gap J is a wandering interval if and only if the rotation number is irrational (cf. Section 4.4).

We say that a Lorenz map f has a *Cherry attractor* ([10]) if it is renormalizable to a gap map with irrational rotation number (including the case where f itself is such a gap map). If this is the case then the gap of the renormalized map and all its preimages are wandering intervals. However, such wandering intervals can be considered as relatively good-natured, because the dynamics of an injective circle map is not chaotic. Moreover, the presence of a Cherry attractor requires the Markov diagram of the Hofbauer tower—or, equivalently, the kneading invariant—to be of a very particular combinatorial type (cf. Section 4.5). Hence in most cases the existence of "Cherry type" wandering intervals can be immediately excluded for combinatorial reasons.

It was shown in Berry & Mestel [1] that Lorenz maps with bounded nonlinearity ([11]) have no wandering intervals unless they have a Cherry attractor, in which case the wandering interval is a preimage of such a gap. Although this result does not apply to the type of Lorenz maps we are interested in, because the derivative vanishes at the discontinuity, there is some hope that a similar result can be shown for Lorenz maps with negative Schwarzian derivative or even for $\mathcal{C}^2$-Lorenz maps.

3.34. Conjecture. *If a $\mathcal{C}^2$-Lorenz map of order $\alpha > 1$ has a wandering interval J then the map is renormalizable to a gap map with irrational rotation number and some iterate of J is mapped into the gap.*

3.35. Remark. The proof in the case of bounded nonlinearity is based on the Schwarz Lemma and gets along with very little knowledge of the underlying dynamics: If J is a maximal wandering interval then its iterates J_n must accumulate on both sides of c. Let $J_{n(k)}$ be the sequence of closest approach to c and let $J_{l(k)}$ and $J_{r(k)} \in \{J_j \mid j < n(k)\}$ denote the immediate neighbours of $J_{n(k)}$ on both sides. Then $D_{n(k)}[J] := f^{n(k)}(Z_{n(k)}[J])$ has to cover one of these neighbours, otherwise f would be renormalizable to a circle map.

([10]) Similar maps occur as Poincaré maps for Cherry flows on the torus (cf. Martens *et al.* [48]).

([11]) More precisely, Lorenz maps for which $\log f'$ is Lipschitz continuous on both branches.

The Schwarz Lemma now implies that $|J_{n(k)}| \geq \min(|J_{l(k)}|, |J_{r(k)}|)$ for large k, which contradicts the fact that $\sum_{n \in \mathbb{N}} |J_n| < \infty$.

Maybe it is possible to prove the conjecture for $\mathcal{C}^2$-Lorenz maps in a similar fashion, replacing the Schwarz Lemma with the Macroscopic Koebe Principle, but perhaps only under some combinatorial restrictions on the recurrence behaviour of the wandering interval at the critical point. For example, if eventually all intervals $D_{n(k)}$ cover both neighbours $J_{l(k)}$ and $J_{r(k)}$ then the situation is just as simple as the one above: Either $|J_{n(k)}| \geq \min(|J_{l(k)}|, |J_{r(k)}|)$ eventually or $D_{n(k)}[J]$ is infinitely often a 1-scaled neighbourhood of $J_{n(k)}$. This space can be pulled back to yield a strictly larger wandering interval $J' \supset J$, a contradiction. The difficult part, however, is to deal with the cases where $D_{n(k)}$ only covers one of its neighbours and maybe one needs some additional conditions to guarantee that this does not happen too often. $\diamond$

Fortunately, the situation is not completely hopeless. There is a special type of wandering intervals which can be excluded immediately by showing that their existence would imply wandering intervals for smooth interval maps.

3.36. LEMMA. *If f is a $\mathcal{C}^2$-Lorenz map then every wandering interval accumulates on both sides of c. In particular, a wandering interval cannot contain any interval of the form $(c_k^{\pm}, c_k^{\pm} \pm \varepsilon)$.*

PROOF. Let J be a wandering interval whose iterates J_n do not accumulate, say, on the right hand side of c and choose $\varepsilon > 0$ such that $(c, c+\varepsilon) \cap J_n = \emptyset$. Then one can modify the map f on $(c, c+\varepsilon)$ in order to obtain a continuous bimodal map $\widetilde{f}$ which is $\mathcal{C}^2$ except for a turning point of order $\alpha > 1$ at c and another turning point of order 1 at $c+\varepsilon$. Since this change does not affect the iterates of J, it is also a wandering interval for $\widetilde{f}$, which is impossible (cf. [13, Chapter IV.2]). ∎

This lemma has some simple but very useful consequences which we state below as a corollary.

3.37. COROLLARY. *Let f be a $\mathcal{C}^2$-Lorenz map without periodic attractors and let $(\widehat{X}, \widehat{f})$ be its Markov extension. Then the following holds:*

1. *The length of the central cylinders $Z_n[c^-]$ and $Z_n[c^+] \in \mathcal{Z}_n$ tends to zero as $n \to \infty$.*
2. *$c_1^{\mp} \in \omega(c_1^{\pm})$ iff the kneading map $Q^{\pm}$ is unbounded, and $c_1^{\pm} \in \omega(c_1^{\pm})$ iff the co-kneading map $\widetilde{Q}^{\pm}$ is unbounded.*
3. *For every level $\widehat{D} \in \widehat{\mathcal{D}}$ the length of the left- and the rightmost cylinder set of order n inside $\widehat{D}$ tends to zero as $n \to \infty$.*
4. *The iterates of every nonprecritical point $\widehat{x} \in \widehat{X}$ visit both sides of the tower infinitely often.*
5. *For every nonprecritical point $\widehat{x} \in \widehat{X}$ there is an integer n such that $\widehat{Z}_n[\widehat{x}]$ is a complete n-cylinder set.*
6. *For every nonprecritical point $x \in X$ and every $\widehat{x}, \widehat{x}' \in \pi^{-1}(x)$ there is an integer n such that $\widehat{f}^n(\widehat{x}) = \widehat{f}^n(\widehat{x}')$.*
7. *If $x \in X$ is not precritical and $\widehat{x} \in \pi^{-1}(x)$ then $r_n(x) = r_n(\widehat{x})$ for all but finitely many n, where $r_n(x) := \operatorname{dist}(x_n, \partial D_n[x])$ and $D_n[x] := f^n(Z_n[x])$.*

8. *A measurable set* $\widehat{A} \in \widehat{\mathcal{B}}$ *satisfies* $\widehat{f}^{-1}(\widehat{A}) = \widehat{A} \bmod \widehat{m}$ *iff* $\widehat{A} = \pi^{-1}(A) \bmod \widehat{m}$ *and* $f^{-1}(A) = A \bmod m$, *where* $A := \pi(\widehat{A})$.

PROOF. Trivial. ∎

The question whether there are wandering intervals or not will become crucial in Section 3.6, where the shadowing of critical points is studied.

3.38. DEFINITION (ε-shadow). A block $x_{m+1}, x_{m+2}, \ldots, x_{m+n}$ of the orbit of x is called a *shadow* (of length n) of the orbit of $c_1^{\pm}$ if none of the intervals $(c_i^{\pm}, x_{m+i})$, $i = 1, \ldots, n$, is critical. The shadow is called an ε-*shadow* if $\max\{\operatorname{dist}(c_i^{\pm}, x_{m+i}) \mid i = 1, \ldots, n\} \le \varepsilon$.

If a shadow is very long then the points x_{m+i} are close to $c_i^{\pm}$ in a combinatorial sense as long as i is not too large, since their itineraries coincide during the rest of the shadow. If f does not have homtervals, the diameter of the n-cylinder sets tends to zero uniformly in n, which implies that closeness in a combinatorial sense also means closeness in a metric sense. More precisely, if f has no homtervals then the following Shadowing Principle applies to f.

3.39. DEFINITION (Shadowing Principle). For every $\varepsilon > 0$ there is an integer N such that the following holds: If the block $x_{m+1}, \ldots, x_{m+n}$ is a shadow of the orbit of $c_1^{\pm}$ then the block $x_{m+1}, \ldots, x_{m+(n-N)}$ is an ε-shadow of the orbit of $c_1^{\pm}$.

As an immediate consequence of this Shadowing Principle one obtains a second one, which assures uniform approximation along shadows: For every $\varepsilon > 0$ there is a $\delta > 0$ such that the following holds: If $x_{m+1}, \ldots, x_{m+n}$ is a shadow of the orbit of $c_1^{\pm}$ and if $\operatorname{dist}(x_{m+n}, c_n^{\pm}) \le \delta$ then $x_{m+1}, \ldots, x_{m+n}$ is an ε-shadow of the orbit of $c_1^{\pm}$. This follows from the Shadowing Principle, because an interval of length ε cannot become arbitrarily small when it is mapped at most N times, since f was assumed to be strictly monotonic on both branches. The second Shadowing Principle is of course just a special case of the well known Contraction Principle which holds for Lorenz maps without homtervals.

CONTRACTION PRINCIPLE. For every $\varepsilon > 0$ there is a $\delta > 0$ such that $|f^n(I)| \le \delta$ implies $|I| \le \varepsilon$ for every interval I on which f^n is a homeomorphism.

In contrast to the case of continuous interval maps it is necessary to assume that f^n is a homeomorphism, which makes it a quite trivial statement.

In general, one can only exclude the existence of homtervals of the type $(c_k^{\pm}, c_k^{\pm} \pm \varepsilon)$, which only yields the following nonuniform versions of the two Shadowing Principles.

3.40. PROPOSITION. *Let f be a $\mathcal{C}^2$-Lorenz map with the property that $c_1^{\pm}$ is not attracted to a periodic orbit. Then for every $\varepsilon > 0$ and every $k > 0$ there exists*

1. *an integer $N_k > k$ such that the following holds: If $x_{m+1}, \ldots, x_{m+n}$ is a shadow of the orbit of $c_1^{\pm}$ of length $n \ge N_k$ then* $\operatorname{dist}(x_{m+k}, c_k^{\pm}) \le \varepsilon$,

2. *a number $\delta_k > 0$ such that the following holds: If $x_{m+1}, \ldots, x_{m+n}$ is a shadow of the orbit of $c_1^{\pm}$ and if* $\operatorname{dist}(x_{m+n}, c_n^{\pm}) \le \delta_k$ *then $n > k$ and* $\operatorname{dist}(x_{m+k}, c_k^{\pm}) \le \varepsilon$.

In particular, $\liminf_{n\to\infty} r_n^{\pm}(x) = 0$ *implies $c_k^{\pm} \in \omega(x)$ for all $k \in \mathbb{N}$.* ∎

The following proposition illustrates what goes wrong if the Shadowing Principle is violated. It also confirms our statement that wandering intervals of Cherry type can be considered as rather harmless.

3.41. Proposition. *If f is a $\mathcal{C}^2$-Lorenz map which has no periodic attractors and if the Shadowing Principle is violated then there is a wandering interval whose preiterates accumulate at c. In particular, if Conjecture* 3.34 *is true then the Shadowing Principle applies to f.*

Proof. Assume that the Shadowing Principle is violated for c_1^+ and some $\varepsilon > 0$. Then it follows that for every $k \in \mathbb{N}$ there is a one-sided neighbourhood $(c^+, c^+ + \gamma_k)$ of c and there are integers $j_k < l_k$ with $l_k - j_k \to \infty$ as $k \to \infty$ such that f^{l_k} is a homeomorphism on $(c^+, c^+ + \gamma_k)$ and $|f^{j_k}(c^+, c^+ + \gamma_k)| \geq \varepsilon$.

Taking a subsequence along which the intervals $f^{j_k}(c^+, c^+ + \gamma_k)$ converge one finds a wandering interval J of length $\varepsilon/2$ which is contained in infinitely many of these intervals. Since $\gamma_k \to 0$ by Proposition 3.40, it follows that the preimages of J accumulate on the right hand side of c.

Now if Conjecture 3.34 holds then f is renormalizable to a gap map with irrational rotation number and we may assume w.l.o.g. that f itself is a gap map. But then all wandering intervals of f are *images* ([12]) of the gap (c_2^-, c_2^+) and the gap itself has no preimage inside the dynamical interval (c_1^+, c_1^-). Hence there is no wandering interval whose preimages accumulate at the critical point. ∎

So the reason why the Shadowing Principle may fail to hold is because the intervals $(x_{m+i}, c_i^{\pm})$ can contain large wandering intervals. Even worse, it is possible that the same wandering interval is contained in the interval $(x_{m+i}, c_i^{\pm})$ for many indices i. But fortunately, the average amount of time where a shadow is not an ε-shadow becomes negligible when the length of the shadow tends to infinity. This is the content of the following proposition.

3.42. Proposition (Weak Shadowing Principle). *For every $\varepsilon > 0$ and every $\delta > 0$ there is an integer N such that for every $n \geq N$ the following holds: If the block $x_{m+1}, \ldots, x_{m+n}$ is a shadow of the orbit of $c_1^{\pm}$ then*

$$\frac{1}{n} \operatorname{card}\{j \leq n \mid \operatorname{dist}(x_{m+j}, c_j^{\pm}) \geq \varepsilon\} \leq \delta.$$

Proof. For a fixed $\varepsilon > 0$ let $\mathcal{W}(\varepsilon)$ be the collection of the finitely many wandering intervals of length $\geq \varepsilon$ ([13]). Since $|f^k(J)| \to 0$ as $k \to \infty$ for every interval $J \in \mathcal{W}(\varepsilon)$, there is an integer n_0 such that $(1/n_0) \operatorname{card}\{k \leq n_0 \mid f^k(J) \in \mathcal{W}(\varepsilon)\} \leq \delta/2$. Let $\mathcal{W}(\varepsilon, n_0) := \{f^k(J) \mid J \in \mathcal{W}(\varepsilon),\ k \leq n_0\}$ and choose $n_1 \in \mathbb{N}$ such that (i) every n_1-cylinder set contains at most one interval from $\mathcal{W}(\varepsilon, n_0)$ and (ii) the diameter of every n_1-cylinder set which does not contain an interval from $\mathcal{W}(\varepsilon)$ is less than ε.

([12]) This does not contradict Conjecture 3.34, because a gap map with irrational rotation number is infinitely often (trivially) renormalizable.

([13]) All wandering intervals are meant to be maximal, i.e., not contained in a larger wandering interval.

Finally, fix $N \in \mathbb{N}$ such that $(n_0+n_1)/N \le \delta/2$ and let $x_{m+1},\dots,x_{m+n}$ be a shadow of length $n \ge N$. Now whenever $j \le n-(n_0+n_1)$ and $\operatorname{dist}(x_{m+j}, c_j^{\pm}) \ge \varepsilon$, the interval $(x_{m+j}, c_j^{\pm})$ must be contained in an n_1-cylinder that contains an interval J from $\mathcal{W}(\varepsilon)$, and consequently $(x_{m+j+k}, c_{j+k}^{\pm})$ is contained in the n_1-cylinder set that contains $f^k(J)$ for $k = 0,\dots,n_0$. It follows that $(1/n_0)\operatorname{card}\{k \le n_0 \mid \operatorname{dist}(x_{m+j+k}, c_{j+k}^{\pm}) \ge \varepsilon\} \le \delta/2$, whence

$$\frac{1}{n}\operatorname{card}\{j \le n-(n_0+n_1) \mid \operatorname{dist}(x_{m+j}, c_j^{\pm}) \ge \varepsilon\} \le \frac{\delta}{2}.$$

Since the cardinality of the remaining indices $j > n-(n_0+n_1)$ is bounded by $\delta n/2$, the claim follows. ∎

3.5. Attractors and invariant measures. The results obtained for the Markov extension $(\widehat{X}, \widehat{f})$ can now be translated back to the original system (X, f) where they allow detailed statements about the possible metric attractors for a Lorenz map with negative Schwarzian derivative. Let us recall the definition of an attractor from the introduction.

3.43. Definition (Attractor). A set $A \subseteq X$ is called *f-invariant* if $f(A) \subseteq A$. For a compact f-invariant set $A \subseteq X$ let $B(A) := \{x \in X \mid \omega(x) \subseteq A\}$ denote the *basin of attraction* of A. Then A is called an *attractor* if the following holds.

1. A attracts a set of positive Lebesgue measure: $m(B(A)) > 0$.
2. Every proper subset $\widetilde{A}$ of A which is compact and f-invariant attracts significantly less points: $m(B(A) \setminus B(\widetilde{A})) > 0$.

An attractor A is called a *global attractor* if $B(A) = X \bmod m$. It is called *indecomposable* if it is not the disjoint union of two smaller attractors and *minimal* if it does not contain any smaller attractors.

3.44. Theorem. *If f is a Lorenz map with $Sf < 0$ then f has a unique global attractor A which is the union of one or two minimal attractors and for m-almost every point x the ω-limit set* coincides *with a minimal attractor. More precisely, one of the following three cases applies.*

1. *If f has an attracting periodic orbit then A is the union of one or two attracting periodic orbits and each attracts at least one of the critical values c_1^+ and c_1^-.*

Now assume that f has no attracting periodic orbits. Then $\omega(x) = A$ for m-a.e. x and one of the following holds.

2. *If f is infinitely often renormalizable then $A = \omega(c_1^+) = \omega(c_1^-)$. If f is infinitely often properly renormalizable then A is a Cantor set which is the intersection of its invariant neighbourhoods.*
3. *If f is finitely often renormalizable then either*

(a) *A is a finite union of intervals*

or A is a nowhere dense set which is of one of the following types.

(b) $A = \omega(c_1^+) = \omega(c_1^-)$.
(c) $A = \omega(c_1^{\pm}) \supset \operatorname{cl}(\operatorname{orb}(c_1^{\mp}))$ *with* $c_1^{\mp} \in \omega(c_1^{\pm})$.

(d) $A = \mathrm{cl}(\mathrm{orb}(c_1^{\pm})) \supset \mathrm{cl}(\mathrm{orb}(c_1^{\mp}))$ *with* $c_1^{\pm} \notin \omega(c_1^{\pm})$.

In cases 2 *and* 3 *the attractor contains the critical point* c. *The only possibility where two attractors can coexist is the case of two attracting periodic orbits.*

In order to keep the proof more readable it is split into a sequence of lemmas. Before starting we remark that a periodic attractor is obviously a minimal attractor and in the other cases the global attractor A is a minimal attractor because $\omega(x) = A$ m-a.s.

3.45. Lemma. *If f has an attracting periodic orbit then for m-a.e. x the ω-limit set is an attracting periodic orbit which attracts at least one of the critical values.*

Proof. First of all note that the periodic orbit need not be hyperbolic attracting. However, a neutral periodic orbit is always weakly attracting on one side and weakly repelling on the other side by the Minimum Principle (cf. page 38), because f has negative Schwarzian derivative. For the same reason every periodic attractor must contain at least one of the points c_1^+ and c_1^- in its immediate basin. Indeed, if $f^m(p) = p$ and $\mathrm{orb}(p)$ does not contain c_1^+ or c_1^- itself then by the Minimum Principle p attracts one or both of the endpoints of $Z_m[p]$, which are precritical points. In particular, there are at most two periodic attractors.

Assume for definiteness that $\mathrm{orb}(p^-)$ is a periodic attractor of period m which contains c_1^- in its immediate basin and $p^- < c_1^-$ is the point of the periodic orbit closest to c_1^-. Since $\widehat{f}$ must be purely dissipative in the presence of a periodic attractor, $r_n(x) \to 0$ for a.e. $x \in X$ by Theorem 3.27 and Corollary 3.37(8). Let $(c_{l(n)}^+, c_{r(n)}^-) := D_n[x]$, $r_n^+(x) := \mathrm{dist}(c_{l(n)}^+, x_n)$, and $r_n^-(x) := \mathrm{dist}(x_n, c_{r(n)}^-)$.

We claim that $\omega(x) \neq \mathrm{orb}(p^-)$ implies $r_n^+(x) \to 0$. This is trivial if the periodic orbit is attracting on both sides, because then its immediate basin is a two-sided neighbourhood of $\mathrm{orb}(p^-)$ which is forbidden for all iterates x_n but contains the endpoints $c_{r(n)}^-$. It follows that $\liminf r_n^-(x) > 0$, whence $\limsup r_n^+(x) = 0$. The remaining case is when p^- is repelling on the left hand side. Assume by contradiction that $\limsup r_n^+(x) > 0$. Since $r_n(x) \to 0$, there exists some $\varepsilon > 0$ and arbitrarily large integers n such that $D_n[x]$ contains p^-, $r_n^+(x) \geq \varepsilon$ and $r_n^-(x) \leq \varepsilon/2$ (cf. Figure 3.3). Choose ε small enough that $g'(y) > 1$ for $y \in (p^- - \varepsilon, p^-)$, where $g := f^m$, and choose k minimal such that $g^k(x_n) \leq p^- - \varepsilon/2$. Then $g(p^- - \varepsilon, p^-) \subseteq D_{n+jm}[x] = (c_l^+, c_{r+jm}^-)$ for $j = 1, \ldots, k$. It follows that $r_{n+km}^-(x) \geq \mathrm{dist}(x_{n+km}, p^-) \geq \varepsilon/2$ and $r_{n+km}^+(x) \geq \mathrm{dist}(x_{n+km}, g(p^- - \varepsilon)) \geq \mathrm{dist}(x_{n+(k-1)m}, p^- - \varepsilon) \geq \varepsilon/2$, whence $r_{n+km}(x) \geq \varepsilon/2$. Since this happens infinitely often, $r_n(x)$ does not converge to 0, a contradiction. This finishes the proof of the claim.

If c_1^+ is also contained in the basin of a periodic attractor $\mathrm{orb}(p^+)$—which is not necessarily distinct from $\mathrm{orb}(p^-)$—then $\lim r_n^+(x) = 0$ implies that x is attracted to the same periodic orbit. So for almost every x either $\omega(x) = \mathrm{orb}(p^-)$ or $\omega(x) = \mathrm{orb}(p^+)$.

If c_1^+ does not converge to a periodic attractor then infinitely many of the intervals $(c_{l_n(x)}^+, x_n)$ must be critical, because otherwise they would eventually be wandering intervals, which is impossible. But whenever $c \in (c_{l_n(x)}^+, x_n)$, $r_n^+(x)$ has to be at least $|p^-|$, because otherwise $c_{l_n(x)}^+$ would be attracted to $\mathrm{orb}(p^-)$. So in this case $r_n^+(x) \not\to 0$ and it follows that $\omega(x) = \mathrm{orb}(p^-)$ for almost every x. ∎

3.46. Remark. A periodic attractor is called *essential* if it contains one of the critical values c_1^+ and c_1^- in its immediate basin, and *inessential* if this is not the case. The previous lemma shows in particular that Lorenz maps with negative Schwarzian derivative do not have inessential periodic attractors. If there is only one periodic attractor then there are three possible situations: (i) the periodic orbit contains c_1^+ and c_1^- in its immediate basin, (ii) the periodic orbit contains only $c_1^\pm$ in its immediate basin but also attracts $c_1^\mp$, and (iii) the periodic orbit contains $c_1^\pm$ in its immediate basin and does not attract $c_1^\mp$. (These three cases are distinguishable by the kneading invariant of f, cf. Remark 4.7.) $\diamond$

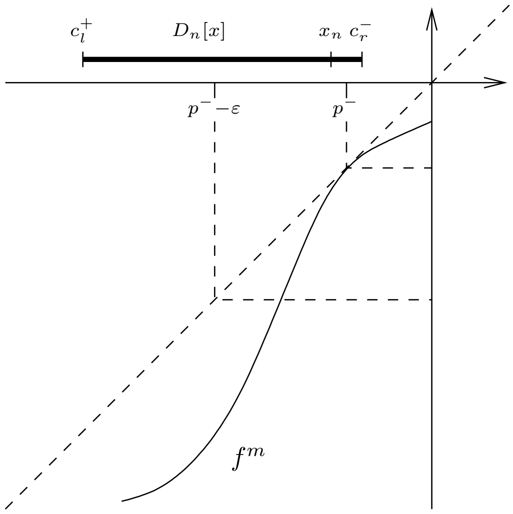


Fig. 3.3. A nonhyperbolic periodic attractor. If $\operatorname{orb} p^-$ is a non-hyperbolic periodic orbit that attracts c_1^- but not c_1^+ then $\omega(x) \neq \operatorname{orb}(p^-)$ implies $r_n^-(x) \not\to 0$, because x_n is infinitely often close to the left hand side of p^- and then pushed away from the right endpoint of the level.

3.47. Lemma. *If f is infinitely often renormalizable then for m-a.e. x the ω-limit set equals $A = \omega(c_1^+) = \omega(c_1^-)$. Moreover, if f is infinitely often properly renormalizable then A is a Cantor set which is the intersection of its invariant neighbourhoods.*

Proof. If f is infinitely renormalizable then $\widehat{f}$ is purely dissipative and it follows from Theorem 3.27 and Corollary 3.37 that $r_n(x) \to 0$ m-almost surely. This immediately implies that $\omega(x) \subseteq \operatorname{cl}(\operatorname{orb}(c_1^+)) \cup \operatorname{cl}(\operatorname{orb}(c_1^-))$ m-a.s. On the other hand, if $r_n(x) \to 0$ then at least one of the limits $\liminf r_n^+(x)$ and $\liminf r_n^-(x)$ is zero. But $\liminf r_n^\pm(x) = 0$ implies $c_1^\pm \in \omega(x)$ by Proposition 3.40, whence

$$\operatorname{cl}(\operatorname{orb}(c_1^\pm)) \subseteq \omega(x) \subseteq \operatorname{cl}(\operatorname{orb}(c_1^+)) \cup \operatorname{cl}(\operatorname{orb}(c_1^-)) \qquad m\text{-a.s.} \tag{3.12}$$

Since f is infinitely renormalizable the cutting$^+$ and cutting$^-$ times tend to infinity, which implies that $c_1^+ \in \omega(c_1^-)$ and $c_1^- \in \omega(c_1^+)$, whence the left hand and the right hand side in (3.12) are equal to $\omega(c_1^+) = \omega(c_1^-)$. This proves that $\omega(x) = \omega(c_1^+) = \omega(c_1^-)$ m-a.s.

If f is infinitely often properly renormalizable then choose integers $m_k, n_k \to \infty$ such that f is (m_k, n_k)-renormalizable for all k. Then

$$A = \bigcap_{k\in\mathbb{N}} A_k, \quad \text{where every} \quad A_k := \bigcup_{j=0}^{m_k-1} f^j[c^+_{n_k}, c^-] \cup \bigcup_{j=0}^{n_k-1} f^j[c^+, c^-_{m_k}] \tag{3.13}$$

is invariant and consists of two cycles of intervals. Since the renormalized map is the first return in the properly renormalizable case, it follows that the intervals from the two cycles that belong to A_k are pairwise disjoint. Since $m_k, n_k \to \infty$ and since there are no periodic attractors in the infinitely renormalizable case, it follows from Proposition 3.40 that the length of all intervals $f^j[c^+_{n_k}, c^-]$ and $f^j[c^+, c^-_{m_k}]$ tends to zero—not necessarily uniformly w.r.t. j—if k goes to infinity, because they are mapped monotonically into $[c^+_{n_k}, c^-_{m_k}]$ by some iterate of f. And, last but not least, every component of A_k contains at least two components of A_{k+1}. Consequently, A is a Cantor set. ∎

3.48. Remark. In the case where only finitely many of the renormalizations are proper the attractor can also be represented as in (3.13), but the intervals of the cycles can touch or overlap. Consequently, the attractor can be very large. A typical example of such a situation is the rigid rotation $f : x \mapsto x + \alpha \bmod 1$ with irrational angle α, where $A = \omega(c_1^+) = \omega(c_1^-) = [0, 1]$.

In the case where f is infinitely often properly renormalizable it is not known whether the Cantor attractor can have positive Lebesgue measure or not. ◇

3.49. Lemma. *If $\widehat{f}$ is essentially conservative then for m-a.e. x the ω-limit set equals $A = \mathrm{cl}(X_{\mathrm{m}})$, where $X_{\mathrm{m}} = \pi(\widehat{X}_{\mathrm{m}})$, and A is a finite union of intervals.*

Proof. If $\widehat{f}$ is essentially conservative then there is a maximal irreducible component $\widehat{X}_{\mathrm{m}}$ and $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}_{\mathrm{m}}$. By Theorem 3.27, $\omega(x) = \mathrm{cl}(X_{\mathrm{m}})$ almost surely. If f is not renormalizable then $\widehat{X}_{\mathrm{m}} = \widehat{X} \setminus \{\widehat{D}_1^+, \widehat{D}_1^-\}$ and $\mathrm{cl}(X_{\mathrm{m}}) = [c_1^+, c_1^-]$. If f is renormalizable then let $\widehat{D}_{S_k^+}^+$ and $\widehat{D}_{S_l^-}^-$ be the highest critical levels below $\widehat{X}_{\mathrm{m}}$. Then the map f is (S_l^+, S_m^-)-renormalizable by Theorem 2.32 and the above considerations apply to the renormalized map, i.e., the attractor of the renormalized map equals $[c^+_{S_k^+}, c^-_{S_l^-}]$. It follows that

$$A = \mathrm{cl}(X_{\mathrm{m}}) = \bigcup_{j=0}^{S_l^- -1} f^j[c^+_{S_k^+}, c^-] \cup \bigcup_{j=0}^{S_k^+ -1} f^j[c^+, c^-_{S_l^-}]. \ ∎$$

3.50. Lemma. *Assume that f has no periodic attractors and is finitely often renormalizable. If $\widehat{f}$ is purely dissipative then there is a compact invariant set A such that $\omega(x) = A$ for m-a.e. x. The set A is of one of the types* 3(b)–(d) *in Theorem* 3.44 *and either $A = \mathrm{cl}(X_{\mathrm{m}})$ or A is nowhere dense in* $\mathrm{cl}(X_{\mathrm{m}})$. *In particular*, *$A$ is nowhere dense in cases* 3(c) *and* 3(d).

Proof. From Theorem 3.27 and Corollary 3.37 it follows that $r_n(x) \to 0$ almost surely. This immediately implies that $\omega(x) \subseteq \mathrm{cl}(\mathrm{orb}(c_1^+)) \cup \mathrm{cl}(\mathrm{orb}(c_1^-))$. On the other hand, if $r_n(x) \to 0$ then at least one of the limits $\liminf r_n^+(x)$ and $\liminf r_n^-(x)$ is zero. But $\liminf r_n^\pm(x) = 0$ implies $c_1^\pm \in \omega(x)$ by Proposition 3.40, whence

$$\mathrm{cl}(\mathrm{orb}(c_1^\pm)) \subseteq \omega(x) \subseteq \mathrm{cl}(\mathrm{orb}(c_1^+)) \cup \mathrm{cl}(\mathrm{orb}(c_1^-)) \quad \text{if} \quad \liminf r_n^\pm(x) = 0. \tag{3.14}$$

We claim that $c_1^+ \in \omega(c_1^-)$ or $c_1^- \in \omega(c_1^+)$. If this is not the case then there exists an $\varepsilon > 0$ such that $(c^+, c^+ + \varepsilon) \cap \mathrm{orb}(c^-) = (c^- - \varepsilon, c^-) \cap \mathrm{orb}(c^+) = \emptyset$ and $|D_n^+|, |D_n^-| \geq \varepsilon$

for all n ([14]). If N is large enough such that $r_n(x)\sup|f'| \le \varepsilon/3$ for all $n \ge N$ and if, say, $r_N(x) = r_N^+(x)$ then one can easily show by induction that $r_n(x) = r_n^+(x)$ for all $n \ge N$ and that none of the intervals $(c_{l(n)}^+, x_n)$ can be critical for $n \ge N$. Hence they must be wandering intervals, a contradiction.

If $c_1^- \in \omega(c_1^+)$ and $c_1^+ \in \omega(c_1^-)$ then the set A is of type 3(b), because (3.14) implies that $\omega(x) = \mathrm{cl}(\mathrm{orb}(c_1^+)) = \mathrm{cl}(\mathrm{orb}(c_1^-)) = \omega(c_1^+) = \omega(c_1^-)$ almost surely. If this is not the case, say, if $c_1^- \in \omega(c_1^+)$ but $c_1^+ \notin \omega(c_1^-)$, then $\mathrm{cl}(\mathrm{orb}(c_1^-)) \subset \mathrm{cl}(\mathrm{orb}(c_1^+))$ and inequality (3.14) simplifies to $\mathrm{cl}(\mathrm{orb}(c_1^-)) \subseteq \omega(x) \subseteq \mathrm{cl}(\mathrm{orb}(c_1^+))$. In order to show that $\omega(x) = \mathrm{cl}(\mathrm{orb}(c_1^+))$ we prove $\limsup r_n^-(x) > 0$, which implies $\liminf r_n^+(x) = 0$. Indeed, if $\lim r_n^-(x) = 0$ then for all sufficiently large n the right component of $D_n[x] \setminus \{x_n\}$ cannot contain the critical point, because $c_1^+ \notin \omega(c_1^-)$. Consequently, these intervals must be wandering intervals, contradicting Corollary 3.37. This proves that A is of type 3(c) or 3(d), depending on whether $c_1^+ \in \omega(c_1^+)$ or $c_1^+ \notin \omega(c_1^+)$.

To complete the proof we show that either $\omega(c_1^\pm) = \mathrm{cl}(X_\mathrm{m})$ or $\omega(c_1^\pm)$ is nowhere dense in $\omega(c_1^\pm)$. Assume that $\omega(c_1^\pm)$ contains an open interval $I \subset X_\mathrm{m}$. Then there exists a level $\widehat{D} \in \widehat{\mathcal{X}}_\mathrm{m}$ and a point $\widehat{x} \in \widehat{D}$ such that $x := \pi(\widehat{x}) \in I$. Since $\mathrm{orb}(c_1^\pm)$ is dense in I, Lemma 3.36 implies that I contains no homtervals, whence $\bigcap_{n\in\mathbb{N}} Z_n[x] = \{x\}$. Choose n large enough such that $Z_n[x] \subset J$. Then $\widehat{f}^n(\widehat{Z}_n[x])$ is a level from the maximal irreducible component and $f^n(Z_n[x]) \subseteq \omega(c_1^\pm)$. Since $\widehat{\mathcal{X}}_\mathrm{m}$ is irreducible, it follows that $\omega(c_1^\pm) = \mathrm{cl}(X_\mathrm{m})$. ∎

3.51. Remark. Note that one obtains an interval attractor (type 3(a)) not only in the case where $\widehat{f}$ is essentially conservative (Lemma 3.49) but also in the case where $\widehat{f}$ is purely dissipative and $\omega(c_1^+) = \omega(c_1^-) = \mathrm{cl}(X_\mathrm{m})$ (Lemmas 3.47 and 3.50). ◇

The proof of Theorem 3.44 is now complete.

In Theorem 3.44 all possible types of attractors which can occur are listed but nothing is said about whether they actually do occur. In the following we give some examples of Lorenz maps with attractors of types 1–3(b). Indeed, it is not hard to find examples in the class of symmetric Lorenz maps, because there is a close connection between symmetric Lorenz maps and unimodal maps and because attractors of these types have unimodal "counterparts", as the following example shows.

3.52. Example. A (proper) Lorenz map $f : [p,q] \to [p,q]$ is called *symmetric* if there exists a continuous involution $\tau \neq \mathrm{id}$ with $\tau(c) = c$ and $f \circ \tau = \tau \circ f$. A Lorenz map f is symmetric if and only if there exists a unimodal ([15]) map $g : [p,q] \to [p,q]$ with critical point c such that $g(p) = g(q) = p$ and

$$f(x) := \begin{cases} g(x) & \text{if } x < c, \\ \tau \circ g(x) & \text{if } x > c, \end{cases} \tag{3.15}$$

where τ is the involution defined by $g \circ \tau = g$. (τ is well defined since a unimodal map is two-to-one except at the critical point.) For a point x let $x_n := f^n(x)$ and $\widetilde{x}_n := g^n(x)$,

([14]) The latter follows from the former—possibly with a smaller ε—using Proposition 3.40.

([15]) A continuous interval map with two montonic branches and a unique global maximum in between, called the *critical point*. For a precise definition and related results, see [13].

and denote by $\omega(x)$ and $\widetilde{\omega}(x)$ the corresponding ω-limit sets of x. Since $g \circ f = g \circ g$, the map g is a semiconjugacy between f and g which collapses precisely the sets $\{x, \tau(x)\}$, $x \in [p, q]$. This immediately implies that $x_n \in \{\widetilde{x}_n, \tau(\widetilde{x}_n)\}$ for all x and all n, whence

$$\omega(x) \subseteq \widetilde{\omega}(x) \cup \tau(\widetilde{\omega}(x)) \quad \text{and} \quad \widetilde{\omega}(x) \subseteq \omega(x) \cup \omega(\tau(x)). \tag{3.16}$$

More precisely, $x_n = \widetilde{x}_n$ if g^n is increasing at x, and $x_n = \tau(\widetilde{x}_n)$ if g^n is decreasing at x. In particular, if x is an orientation preserving n-periodic point for g then x and $\tau(x)$ are n-periodic points for f which belong to different periodic orbits ([16]) and if x is an orientation reversing n-periodic point for g then x and $\tau(x)$ are $2n$-periodic points for f which belong to the same periodic orbit.

For unimodal maps one has the following classification of attractors which is due to Blokh and Lyubich (cf. Theorem 1.3 in [13, Chapter V]).

THEOREM. *Let $g : [p, q] \to [p, q]$ be a unimodal map with nonflat critical point and with negative Schwarzian derivative. Then g has a unique attractor $\widetilde{A}$, $\widetilde{\omega}(x) = \widetilde{A}$ for m-a.e. x, and $\widetilde{A}$ is of one of the following types.*

1. *If g has an attracting periodic orbit then $\widetilde{A}$ is this periodic orbit and it attracts the critical point.*
2. *If g is infinitely often renormalizable then $\widetilde{A} = \widetilde{\omega}(c)$.*
3. *If g is finitely often renormalizable then either*
   (a) *$\widetilde{A}$ is a finite union of intervals which contains c in its interior*, *or*
   (b) *$\widetilde{A}$ is a Cantor set and $A = \widetilde{\omega}(c)$.*

The numbering of the cases corresponds to the one in Theorem 3.44. All attractors except for type 3(b) occur in the quadratic unimodal family $x \mapsto b - x^2$ and attractors of type 3(b) occur in the family $x \mapsto b - |x|^\alpha$ if $\alpha > 1$ is large enough (see Bruin *et al.* [8]).

Let A be the global attractor of f. Since $B(\tau(A)) = \tau(B(A))$ and since the map τ is absolutely continuous in the situation of the above theorem, it follows that A is symmetric w.r.t. τ and together with (3.16) one obtains $A = \widetilde{A} \cup \tau(\widetilde{A})$. Now it is easily checked that the type of A (w.r.t. the classification of Theorem 3.44) is precisely the same as the type of $\widetilde{A}$ (w.r.t. the above classification). For type 1 observe that an orientation reversing n-periodic attractor for g yields a symmetric $2n$-periodic attractor for f, and an orientation preserving n-periodic attractor for g yields two distinct n-periodic attractors for f which are mirror images of each other. For cases 2 and 3 it is easily checked that the unimodal map g is n-renormalizable if and only if the Lorenz map f is properly (n, n)-renormalizable, and that the renormalized maps $\mathcal{R}g$ and $\mathcal{R}f$ again satisfy (3.15) with the same involution τ. $\diamond$

Naturally, we are more interested in asymmetric examples, since symmetric Lorenz maps are rather rare and since the above example does not yield any new insight. Asymmetric examples for types 1–3(a) can easily be found in the quadratic Lorenz family (5.2). This follows from the fact that quadratic Lorenz family is a full family of maps (cf. Chapter 5) with negative Schwarzian derivative. For type 3(a) one only has to observe that it is sufficient to construct examples such that both critical points are preperiodic but

([16]) Note that the g-orbit of x cannot be τ-symmetric.

not periodic. Then f has no periodic attractors and $\widehat{f}$ is essentially conservative, because $c^+ \notin \omega(c^-)$ and vice versa (cf. the proof of Lemma 3.50). We conjecture that asymmetric examples of type 3(b) can be found in the families (5.1) for large α. However, the proofs in the unimodal case rely a lot on the local symmetry near the critical point and the arguments break down in the case of Lorenz maps. Attractors of types 3(c) and 3(d) have no unimodal counterpart and it is not clear whether they actually occur. At least all types of attractors are thinkable from a topological viewpoint: At the end of Section 4.3 we are going to give examples of Lorenz maps using kneading theory which meet the minimal topological requirements in order to have an attractor of one of these types, i.e., for which the omega limit sets $\omega(c_1^+)$ and $\omega(c_1^-)$ have the required properties.

3.53. Example. Figures 3.4–3.6 show Lorenz maps with attractors of the types listed in parts 1 and 2 of Theorem 3.44. The attractor is visualized by plotting an approximation of $\omega(x)$ as function of x together with the map. $\diamond$

We conclude the discussion of attractors with a brief remark on attractors of type 3(d). If the attractor is of type 3(d), say, if $A = \mathrm{cl}(\mathrm{orb}(c_1^+))$, then A is the disjoint union of two sets A_α and A_ω, where A_α is a finite or countably infinite set of points that are isolated in A and $A_\omega = \omega(A_\alpha)$ is a nowhere dense set which accumulates on the left hand side but not on the right hand side of c. This implies that such maps exhibit the following kind of intermittent behaviour: The orbit of a typical point x will spend most of its time near A_ω until some iterate of x happens to land very close to c on the right hand side, where it is separated from A_ω in the next step and has to start all over again following A_α, and so on.

The rest of this section is devoted to the question to what extent the results on the Hopf decomposition and the existence of invariant measures obtained for the Markov extensions can be translated back to the original system again. Unfortunately, this is not always possible, in particular if the omega limit set $\omega(c_1^+) \cup \omega(c_1^-)$ is large. The problems which one encounters here are precisely the ones which made it more convenient to study the transfer operator for $\widehat{f}$ instead of f in the first place: It is one of the main advantages of the tower construction that the forward iterates of the critical points—the endpoints of the levels in the tower—do not accumulate in the tower but instead escape to infinity, because these are the points where the iterates $\widehat{\mathrm{P}}^n \widehat{\psi}$ of smooth densities develop poles if $f'(c^\pm) = 0$. Another problem is caused by the fact that the reference measure $\widehat{m}$ is only σ-finite, and that the projection π is not finite-to-one.

There are two implications that hold trivially, because $\mathrm{P} = \mathrm{P}_\pi \circ \widehat{\mathrm{P}}$: If $\widehat{\mathrm{P}}$ is conservative on $\widehat{X}_{\mathrm{m}}$ then also P is conservative on $X_{\mathrm{m}} := \pi(\widehat{X}_{\mathrm{m}})$, and if $\widehat{\mathrm{P}}$ has an integrable invariant density $\widehat{h}$ then P also has an integrable invariant density, namely $h := \mathrm{P}_\pi \widehat{h}$. The converse does not necessarily hold. A symptomatic counterexample is the rigid rotation $f : x \mapsto x + \alpha \bmod 1$ with irrational angle α. It preserves the Lebesgue measure, yet the map on the tower is dissipative because it has no closed paths. Note that (f, m) is ergodic but has zero entropy. It was shown in Keller [37, Theorem 3] that every ergodic invariant measure of positive entropy lifts onto the tower.

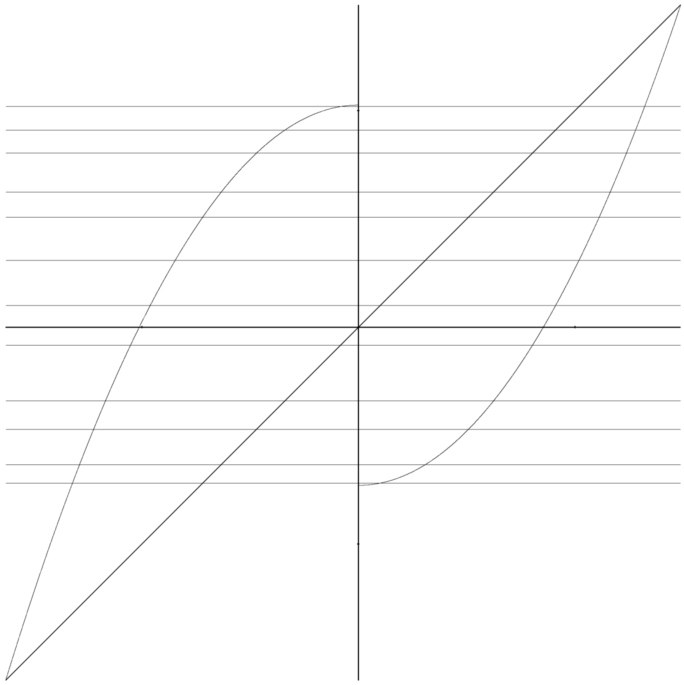


Fig. 3.4. A Lorenz map with one periodic attractor

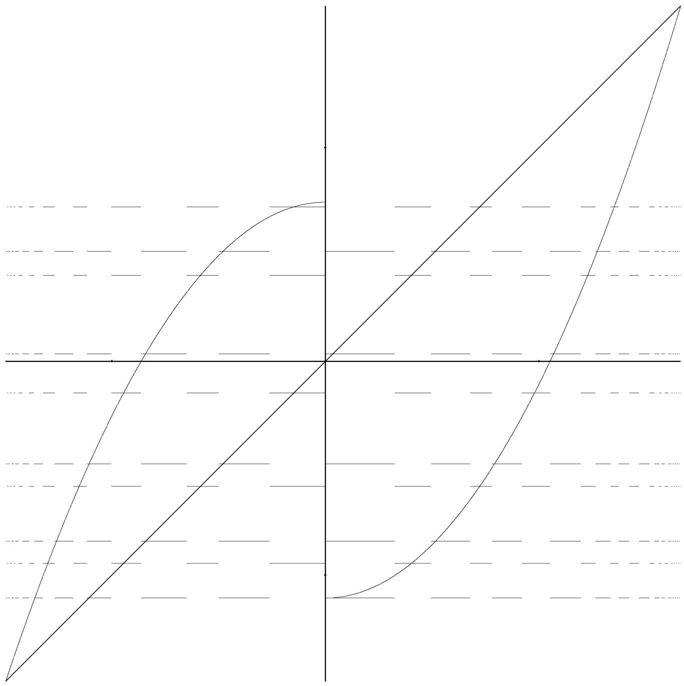


Fig. 3.5. A Lorenz map with two periodic attractors

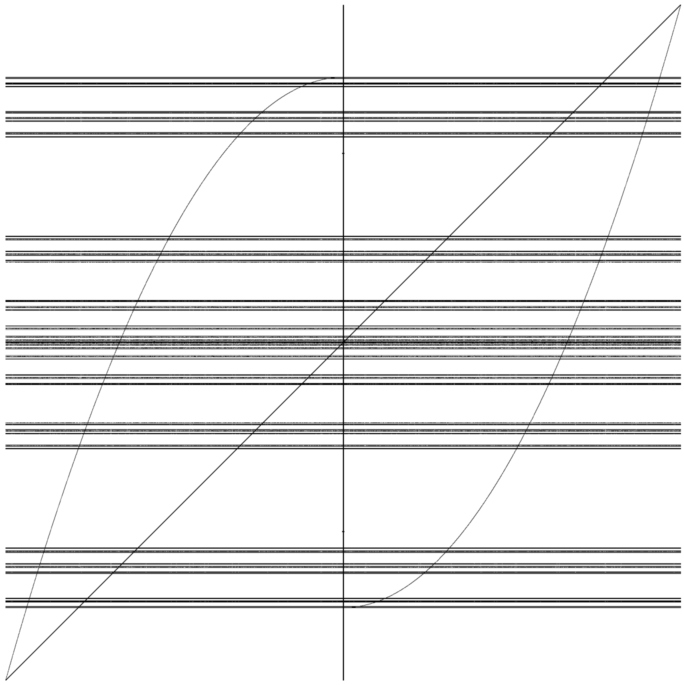


Fig. 3.6. A Lorenz map with a Cantor attractor

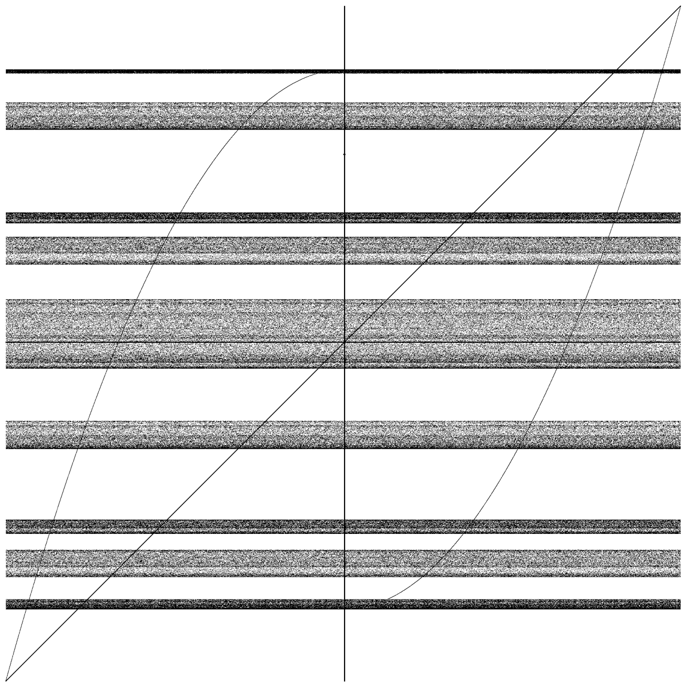


Fig. 3.7. A Lorenz map with an interval attractor

3.54. Proposition. *If f is finitely often renormalizable and if* $\mathrm{orb}(c_1^+)\cup\mathrm{orb}(c_1^-)$ *is not dense in* $X_{\mathrm{m}} := \pi(\widehat{X}_{\mathrm{m}})$ *then f is conservative on X_{m} if and only if $\widehat{f}$ is conservative on $\widehat{X}_{\mathrm{m}}$. In the conservative case $\omega(c_1^+)\cup\omega(c_1^-)$ has zero Lebesgue measure, and f has a σ-finite invariant measure $\mu = hm$ with density $h := \mathrm{P}_\pi\,\widehat{h}$.*

Proof. Just the "only if" part needs to be proved. So assume that $\widehat{f}$ is dissipative on $\widehat{X}_{\mathrm{m}}$. Choose an interval $I\subset\widetilde{X} := X_{\mathrm{m}}\setminus\mathrm{orb}(c_1^+)\cup\mathrm{orb}(c_1^-)$ and a cylinder set $Z\in\mathcal{Z}_n$ with $Z\subset I$. (If there is no such cylinder set then I contains a homterval, whence f is dissipative and there is nothing to prove.) Since $I\cap(\mathrm{orb}(c_1^+)\cup\mathrm{orb}(c_1^-))=\emptyset$ it follows that for every level $\widehat{D}\in\widehat{\mathcal{D}}$ the set $\pi^{-1}(Z)\cap\widehat{D}$ is either void or a complete n-cylinder set ([17]). Hence each of these sets is mapped to the same level $\widehat{E}\in\widehat{\mathcal{D}}$ in the tower by $\widehat{f}^n$. Let $\psi_n := \mathrm{S}_n\,1$ and $\widehat{\psi}_n := \widehat{\mathrm{S}}_n(1_{\widehat{D}_0^+\cup\widehat{D}_0^-})$. Obviously, $\psi_n = \mathrm{P}_\pi\,\widehat{\psi}_n$ on every component of $\widetilde{X}$. Since $\widehat{\mathrm{P}}$ is dissipative it follows that $\widehat{\psi}_\infty$ is finite on $\widehat{E}$ and from Lemma 3.29 it follows that $\psi_\infty(x) = \sum_{\widehat{x}\in\pi^{-1}(x)}\widehat{\psi}_\infty(x)$ is finite for every $x\in Z$. This proves that f is dissipative on X_{m}.

Now if f is conservative on X_{m} then it is also ergodic on X_{m} in view of Theorem 3.14 and Corollary 3.37(8). Consequently, $\mathrm{cl}(\mathrm{orb}(c_1^+))\cup\mathrm{cl}(\mathrm{orb}(c_1^-))$ has zero Lebesgue measure. Let $\widehat{h}$ be the invariant density for $\widehat{f}$ and $h := \mathrm{P}_\pi\,\widehat{h}$. Since $\widehat{h}$ is finite on $\widehat{E}$ it follows as before that h is finite on Z. Moreover, $h(x)/h(y)$ is bounded on Z, which implies that $\int_Z h\,dm<\infty$. Since $\bigcup_{n\in\mathbb{N}} f^{-n}(Z) = X \bmod m$ it follows that μ is σ-finite. ∎

3.55. Theorem. *If f is a Lorenz map with negative Schwarzian derivative then f has an absolutely continuous invariant measure μ of positive entropy if and only if $\widehat{f}$ is positively recurrent on $\widehat{X}_{\mathrm{m}}$. If this is the case then $\mu = hm$, where $h := \mathrm{P}_\pi\,\widehat{h}$.*

Proof. The "if" part and the last statement are obvious. From Corollary 3.37(8) it follows that the tail field for f has the same cardinality as the tail field for $\widehat{f}$, which is finite by Theorem 3.14, case 3(a), and Corollary 3.37(8). For the reverse implication, which relies on the fact that the topological entropy "at infinity" in the tower is zero, the reader is referred to Keller [39]. ∎

3.56. Remark. It also follows from the results in Keller [39] that f has an absolutely continuous invariant measure μ of positive entropy if and only if the upper Lyapunov exponent $\limsup_{n\to\infty} n^{-1}\log|(f^n)'(x)|$ is positive on a subset of positive Lebesgue measure (in which case it is positive almost everywhere). ◇

3.6. Shadowing the critical orbits. Throughout this section assume that f has no periodic attractors. We want to study to what extent the behaviour of the trajectory of the critical values c_1^+ and c_1^- determines the behaviour of the trajectory of typical points $x\in X$. Therefore we decompose the orbit of x into blocks $x_{\sigma_n+1},\ldots,x_{\sigma_{n+1}}$ which are alternately shadows of c_1^+ and c_1^- of maximal length. The natural candidates for the points of the critical orbit to be currently shadowed by x_n are the two endpoints of the interval $D_n[x] = f^n(Z_n[x])$, and the possible starting points for a new shadow are the moments just after a left or right cutting time occurred. Recall that

([17]) That is, $\pi(\pi^{-1}(Z)\cap\widehat{D}) = Z$ (cf. Remark 2.6).

$$D_n[x] = f^n(Z_n[x]) = (c^+_{n-s_n^<}, c^-_{n-s_n^>}) \quad \text{for } n > \max(S_0^<(x), S_0^>(x)),$$

where $s_n^{\lessgtr} := S^{\lessgtr}\langle n\rangle(x)$ denotes the last cutting$^{\lessgtr}$ time for x before time n (cf. Definition 2.17). Hence x_n shadows both critical trajectories simultaneously and the two shadows overlap.

To decide which of the current shadows is more suitable it is convenient to pass over to the Markov extension. Therefore, let $\widehat{x}$ be a lift of x to $\widehat{X}$. Under the action of $\widehat{f}$ the point $\widehat{x}$ climbs in the tower until some iterate $\widehat{x}_n$ reaches the "wrong" side of a critical level which acts like a trap-door and it falls on the other side of the tower. After that it climbs again on the other side, until it reaches the next trap-door, and so on. Let $\widehat{E}$ be the union of all trap-doors in the tower, i.e.,

$$\widehat{E} := \bigcup_{k\in\mathbb{N}} (\widehat{E}^+_{S_k^+} \cup \widehat{E}^-_{S_k^-}), \quad \text{where} \quad \widehat{E}^\pm_{S_k^\pm} := \widehat{D}^\pm_{S_k^\pm} \cap \pi^{-1}(Z^\pm).$$

The idea behind the decomposition is, roughly speaking, that x_n follows $c^+_{n-s_n^<}$ if $\widehat{x}_n$ is on a $\mathcal{D}^+$-level, and x_n follows $c^-_{n-s_n^>}$ if $\widehat{x}_n$ is on a $\mathcal{D}^-$-level. This is not quite correct, since a new shadow already starts before the jump, as we will see in a minute.

For $\widehat{x} \in \widehat{X}$ let $\tau_n := \tau_n(\widehat{x})$ be the nth instance where an iterate of $\widehat{x}$ visits $\widehat{E}$. The numbers $\tau_n(\widehat{x})$ can be defined recursively by

$$\tau_1(\widehat{x}) := \tau(\widehat{x}) \quad \text{and} \quad \tau_{n+1} := \tau(\widehat{f}^{\tau_n}(\widehat{x})),$$

where $\tau(\widehat{x}) := \inf\{n \in \mathbb{N} \mid \widehat{x}_n \in \widehat{E}\}$ is the first entry time to $\widehat{E}$. By Corollary 3.37, τ is finite except on the precritical points. Let $\varrho_n := \varrho_n(\widehat{x})$ denote the level of the corresponding trap-door, i.e.,

$$\widehat{x}_{\tau_n} \in \widehat{E}^+_{S^+_{\varrho_n}} \quad \text{if } n \text{ is even}, \quad \text{and} \quad \widehat{x}_{\tau_n} \in \widehat{E}^-_{S^-_{\varrho_n}} \quad \text{if } n \text{ is odd, or vice versa.}$$

We abbreviate this by $\widehat{x}_{\tau_n} \in \widehat{E}^\pm_{S^\pm_{\varrho_n}}$ ([18]). From here on $\widehat{x}_{\tau_n}$ follows the endpoint of $\widehat{D}^\pm_{S^\pm_{\varrho_n}}$ with the lower index, which is a lift of $c^\mp_{S^\mp_{Q^\pm(\varrho_n)}}$, but the actual shadow has already begun earlier, namely at time $\sigma_n + 1$, where $\sigma_n := \tau_n - (S^\pm_{\varrho_n} - S^\pm_{\varrho_n - 1}) = \tau_n - S^\mp_{Q^\pm(\varrho_n)}$. Figure 3.8 shows the orbit of $\widehat{x}$ in the tower during a shadow.

3.57. Remark. Because of Corollary 3.37(6), for any two points $\widehat{x}, \widehat{x}' \in \pi^{-1}(x)$ the construction yields the same numbers σ_n and τ_n, except for finitely many integers n. Hence the particular choice of the lift $\widehat{x} \in \pi^{-1}(x)$ is not important, since one is only interested in the long time behaviour of the shadowing. $\diamond$

3.58. Proposition. *Let f be a Lorenz map with negative Schwarzian derivative and without periodic attractors.*

1. *If $\widehat{f}$ is purely dissipative then $\sigma_{n+1}(\widehat{x}) - \sigma_{n-1}(\widehat{x}) \to \infty$ $\widehat{m}$-a.s.*
2. *If $\widehat{f}$ is essentially conservative then $\liminf_{n\to\infty}(\sigma_{n+1}(\widehat{x}) - \sigma_{n-1}(\widehat{x})) < \infty$ $\widehat{m}$-a.s. Moreover,*
   (a) *if $\int \widehat{h}\, d\widehat{m} = 1$ then $\sigma_n(\widehat{x})/n \to 1/\widehat{\mu}(\widehat{E})$ $\widehat{m}$-a.s.*
   (b) *if $\int \widehat{h}\, d\widehat{m} = \infty$ then $\sigma_n(\widehat{x})/n \to \infty$ $\widehat{m}$-a.s.*

([18]) Here the symbol "$\pm$" means "$(-)^n$" and "$\mp$" means "$(-)^{n+1}$", or vice versa. We apologize for the obvious violation of the agreement made in Remark 2.8.

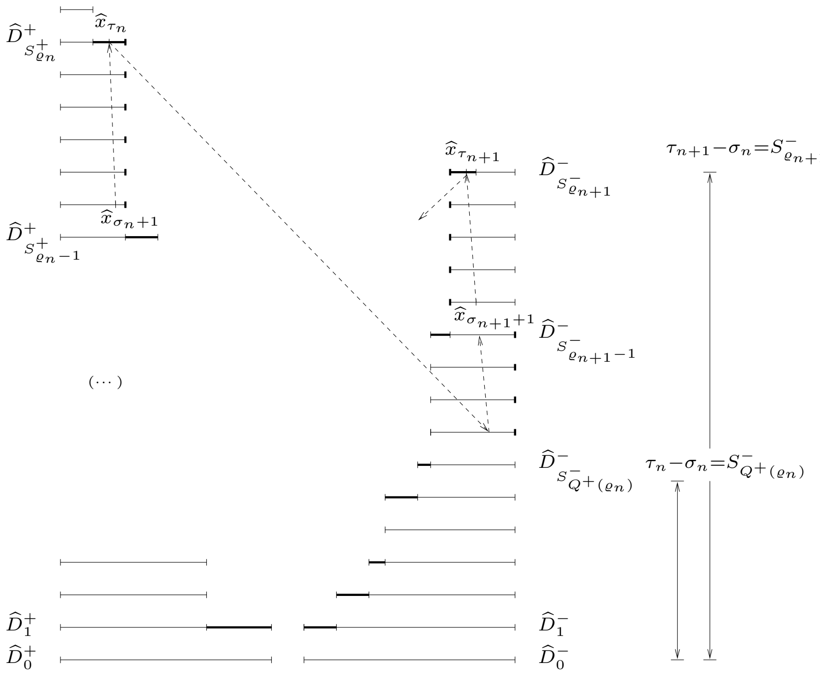


Fig. 3.8. A shadow viewed on the tower. A schematical view of a c_1^--shadow beginning at time $\sigma_n + 1$ and ending at time τ_{n+1} which overlaps a c_1^+-shadow beginning at time $\sigma_{n+1} + 1$. The currently shadowed endpoints are marked with thick marks and the trap-doors with thick lines.

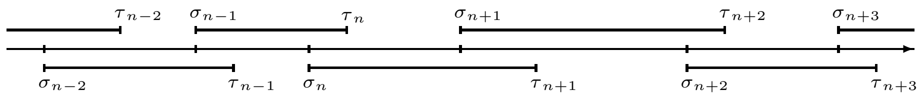


Fig. 3.9. The shadows of the orbit of c_1^+ and c_1^- alternate

Proof. Essentially, the proof follows the one of [39, Theorem 6], except for one detail: In [39] the inequality $S_k \leq 2S_{k-1}$ for the cutting times $(S_k)_{k\in\mathbb{N}}$ is used, which holds for unimodal maps without periodic attractor. This inequality has no counterpart for Lorenz maps (cf. Remark 2.11).

1. Fix $N \in \mathbb{N}$ and choose $M \geq 2N$ such that there are cutting times S_k^+ and S_l^- satisfying $2N < S_k^+, S_l^- \leq M$. Let $\widehat{X}_N := \bigcup_{n\leq N}(\widehat{D}_n^+ \cup \widehat{D}_n^-)$ be the union of the first N floors of the tower and let $\widehat{K}_N$ be the closure of the union of all complete M-cylinder sets contained in $\widehat{X}_N$. Because $\widehat{\mathrm{P}}$ is dissipative, a typical point visits $\widehat{K}_N$ only finitely often. Fix such a point $\widehat{x} \in \widehat{X}$ and let $\sigma_n := \sigma_n(\widehat{x})$ and $\tau_n := \tau_n(\widehat{x})$. For the following it may be useful to have Figures 3.8 and 3.9 in mind.

We want to show that $\sigma_{n+1} - \sigma_{n-1} \geq N$ for large n. If $\tau_n - \sigma_n \geq N$ then this is obviously the case. If $\tau_n - \sigma_n < N$ then $S^{\mp}_{Q^\pm(\varrho_n)} = \tau_n - \sigma_n < N$, which implies that $\widehat{x}_{\tau_n+1} \in \widehat{X}_N$ and hence $\widehat{x}_{\tau_n+1} \in \widehat{X}_N \setminus \widehat{K}_N$ for large n. By construction of $\widehat{X}_N \setminus \widehat{K}_N$ it follows that $S^{\mp}_{\varrho_{n+1}} \geq M$ and $S^{\mp}_{\varrho_{n+1}-1} \geq \min(S^+_k, S^-_l)$. Now

$$\sigma_{n+1} - \tau_n = S^{\mp}_{\varrho_{n+1}-1} - S^{\mp}_{Q^{\pm}(\varrho_n)} \geq \min(S^+_k, S^-_l) - N \geq N.$$

2. Choose N large enough such that $\widehat{E}_3 := \widehat{E} \cap \{\widehat{x} \mid \tau_3(\widehat{x}) \leq N\}$ has positive measure. Since $\widehat{f}$ is conservative and ergodic it follows that $\widehat{m}$-a.e. point $\widehat{x}$ visits $\widehat{E}_3$ infinitely often. But if $x_{\tau_{n-2}} \in \widehat{E}_3$ then $\tau_{n+1} - \tau_{n-2} \leq N$ and consequently $\sigma_{n+1} - \sigma_{n-1} \leq N$.

For the proof of cases (a) and (b) one can replace σ_n/n by τ_n/n, since $\sigma_{n-1} \leq \tau_{n-1} \leq \sigma_n$. In case (a) one can apply Birkhoff's Ergodic Theorem to $\widehat{\mu}$ and one obtains

$$\lim_{n\to\infty} \frac{n}{\tau_n} = \lim_{n\to\infty} \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} 1_{\widehat{E}} \circ \widehat{f}^{\,i} = \widehat{\mu}(\widehat{E}) > 0 \quad \widehat{m}\text{-a.s.}$$

In case (b) let $\widehat{K}_N$ be as before and $\widehat{L}_N := \widehat{f}^{-1}(\widehat{K}_N)$. Since $\widehat{\mu}(\widehat{L}_N) = \widehat{\mu}(\widehat{K}_N) < \infty$ by Proposition 3.7, Birkhoff's Ergodic Theorem now implies that

$$\lim_{n\to\infty} \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} 1_{\widehat{L}_N} \circ \widehat{f}^{\,i} = 0 \quad \widehat{m}\text{-a.s.},$$

which shows that

$$\limsup_{n\to\infty} \frac{n}{\tau_n} = \limsup_{n\to\infty} \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} 1_{\widehat{E}} \circ \widehat{f}^{i} = \limsup_{n\to\infty} \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} 1_{\widehat{E}\setminus\widehat{L}_N} \circ \widehat{f}^{i} \quad \widehat{m}\text{-a.s.}$$

By definition, $\widehat{x}_i \in \widehat{E}$ implies $i = \tau_n$ for some integer k. We claim that $\tau_{k+1} - \tau_{k-1} \geq N$ whenever $\widehat{x}_{\tau_k} \in \widehat{E} \setminus \widehat{L}_N$. Indeed, if $\tau_k - \tau_{k-1} < N$ then also $\tau_k - \sigma_k < N$ and it follows as in part 1 that $\widehat{x}_{\tau_k+1} \in \widehat{X}_N$. Even more, $\widehat{x}_{\tau_k+1} \in \widehat{X}_N \setminus \widehat{K}_N$, because $\widehat{x}_{\tau_k} \notin \widehat{L}_N$. But this implies that $\tau_{k+1} - \tau_k \geq M - N \geq N$. Now for any two indices $k < k'$ with the same parity the intervals (τ_{k-1}, τ_{k+1}) and $(\tau_{k'-1}, \tau_{k'+1})$ do not overlap. It follows that

$$\limsup_{n\to\infty} \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} 1_{\widehat{E}\setminus L_N}(\widehat{x}_i) = \limsup_{n\to\infty} \frac{1}{\tau_n} \operatorname{card}\{k \leq n \mid \widehat{x}_{\tau_k} \in \widehat{E} \setminus \widehat{L}_N\} \leq \frac{2}{N} \quad \widehat{m}\text{-a.s.}$$

This proves (b), since N was arbitrary. ∎

3.59. Remark. In a similar way to the proof of part 2 it can be shown that in the null recurrent case the average visit frequency of a typical point to every set $\widehat{X}_N$ tends to zero, even though $\widehat{X}_N$ does not necessarily have finite $\widehat{\mu}$-measure ([19]): Let $\widehat{X}_N$ and $\widehat{K}_N$ be as in Proposition 3.58. Whenever $\widehat{x}_{\tau_k+1} \in \widehat{X}_N \setminus \widehat{K}_N$, at most N of the next M iterates of $\widehat{x}$ including $\widehat{x}_{\tau_k+1}$ belong to $\widehat{X}_N$. Consequently,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\widehat{X}_N}(\widehat{x}_i) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\widehat{X}_N\setminus\widehat{K}_N}(\widehat{x}_i) \leq \frac{N}{M} \quad \widehat{m}\text{-a.s.}$$

This implies that $\lim_{n\to\infty} n^{-1} \operatorname{card}\{i \leq n \mid \widehat{x}_i \in \widehat{X}_N\} = 0$ $\widehat{m}$-a.s., since M can be chosen arbitrarily large. ◇

3.60. Definition. For a probability measure ν on X let $\omega^*(\nu)$ be the set of weak accumulation points of the sequence $(n^{-1}\sum_{k=0}^{n-1} \nu \circ f^{-k})_{n\in\mathbb{N}}$ and let $\omega^*(x) := \omega^*(\delta_x)$.

([19]) Compare Remark 2 after Theorem 2.7 in Bruin [4].

3.61. Theorem. *If f is a Lorenz map with negative Schwarzian derivative which has no absolutely continuous invariant probability measure of positive entropy then $\omega^*(x)$ is contained in the convex closure of $\omega^*(c_1^+) \cup \omega^*(c_1^-)$ for m-a.e. $x \in X$.*

Proof. Let $(\widehat{X}, \widehat{f})$ be the canonical Markov extension for f. By Theorem 3.55, the map $\widehat{f}$ is either purely dissipative or conservative without an integrable invariant density. If f has periodic attractors then the theorem holds trivially, because $\omega(x) = \omega(c^+)$ or $\omega(x) = \omega(c^-)$ m-a.s. by Theorem 3.44. Hence we can assume that f has no periodic attractors and it follows from parts 1 and 2(b) of Proposition 3.58 that $\sigma_n(\widehat{x})/n \to \infty$ m-almost surely, i.e., the orbit of a typical point $\widehat{x}$ spends most of its time in very long shadows of one of the critical orbits. Fix such a point and let $\sigma_k := \sigma_k(\widehat{x})$ and k_n be such that $\sigma_{k_n} \le n < \sigma_{k_n+1}$. Since $\sigma_n/n \to \infty$, it follows that $k_n/n \to 0$.

The Weak Shadowing Principle 3.42 implies that for every $\varepsilon > 0$ the average time where a shadow is not an ε-shadow becomes negligible when the length of the shadow tends to infinity. As in [39] one can now argue that

$$\left(\frac{1}{n}\sum_{j=1}^{n}\delta_{x_j} - \left(\sum_{k=1}^{k_n}\frac{\sigma_k - \sigma_{k-1}}{n}\,\mu_{\sigma_k-\sigma_{k-1}} + \frac{n-\sigma_{k_n}}{n}\,\mu_{n-\sigma_{k_n}}\right)\right)(\psi) \to 0$$

as $n \to \infty$ for every function $\psi \in \mathcal{C}(X)$ where $\mu_n^\pm := n^{-1}\sum_{i=1}^n \delta_{c_i^\pm}$, and using the weak compactness of $\omega^*(c^+)$ and $\omega^*(c^-)$ the claim follows. ∎

4. Kneading theory

The main goal of this chapter is to study the connection between the combinatorial properties of the kneading invariant and the structure of the Markov diagram.

4.1. The kneading invariant

4.1. Definition (Itinerary). The binary sequence $\zeta(x) := (\zeta_n(x))_{n\geq 0}$, defined by

$$\zeta_k(x) := 0 \text{ if } f^k(x) < c, \quad \text{and} \quad \zeta_k(x) := 1 \text{ if } f^k(x) > c,$$

is called the *itinerary* of x.

The itinerary of x is well defined unless x is a precritical point. As already mentioned in the introduction, for kneading theory it is convenient to double all the precritical points topologically. Let us describe this construction in more detail.

Let $\mathcal{K}$ denote the collection of precritical points. Since the map ζ from $[a,b]\setminus\mathcal{K}$ to the *shift space* $\Sigma := \{0,1\}^{\mathbb{N}}$ is monotonic, where Σ is endowed with the usual lexicographical order, the one-sided limits $\zeta(x^\pm) = \lim_{y\to x\pm}\zeta(y)$ exist everywhere and they are equal if and only if x is not a precritical point. For convenience we double all the precritical points topologically, i.e., replace every precritical point z by two points $z^- < z^+$. The resulting space is denoted by X, or X_f whenever necessary. It is totally ordered and a compact metrizable topological space w.r.t. its order topology.

4.2. Remark. The space X_f can be embedded naturally into the real line as a metric subspace: Just "blow up" each of the countably many precritical points z to an interval J_z such that the total length $\sum_{z\in\mathcal{K}}|J_z|$ is finite and take z^- and z^+ as the left and right endpoint of J_z, respectively. $\diamond$

The maps f and ζ can be extended to X_f by taking one-sided limits at z. The extensions will again be denoted by f and ζ. By definition of the itinerary ζ, the action of f on X_f is reflected by the one-sided shift σ on Σ, i.e., $\zeta\circ f = \sigma\circ\zeta$ on X_f.

4.3. Definition (Kneading invariant). Let $c_k^\pm := f^k(c^\pm)$ denote the forward iterates of the critical point $c^\pm$, and $\nu^\pm := \zeta(c^\pm)$ the itinerary of $c^\pm$. The pair $\nu := (\nu^+,\nu^-)$ is called the *kneading invariant* of f.

Since $\sigma\nu^+$ and $\sigma\nu^-$ are the itineraries of a and b, respectively, and since the map ζ is monotonic, the following condition is necessary for a sequence $\omega\in\Sigma$ to occur as the itinerary of some point $x\in[a,b]$:

$$\sigma\nu^+ \leq \sigma^n\omega \leq \sigma\nu^- \quad \forall n\geq 0. \tag{4.1}$$

It follows from Proposition 4.6 below that this condition is also sufficient. We call such sequences ν-*admissible*. The subshift of all ν-admissible sequences is denoted by Σ_ν. We will restrict ourselves to the interval $[a,b]$, since the itineraries of the remaining points $x \in [p,q] \setminus [a,b]$ are just of the form 0^∞, 1^∞, $0^n\omega$ or $1^n\omega$, where ω is the itinerary of some point in $[a,b]$.

4.4. Definition. A Lorenz map f is called (*topologically*) *expansive* if there exists some $\varepsilon > 0$ with the following property: For any two distinct points x and y there exists some n such that $d(x_n, y_n) \geq \varepsilon$.

An alternative way to define expansiveness is to require that f has no homtervals or that the set of precritical points is dense in $[a,b]$.

4.5. Definition. Two Lorenz maps f and $\widetilde{f}$ are (*semi*)*conjugate* if there exists a homeomorphism (resp. a monotonic continuous surjective map) $\phi : X_f \to X_{\widetilde{f}}$ satisfying $\phi \circ f = \widetilde{f} \circ \phi$.

4.6. Proposition. *The map $\zeta : X_f \to \Sigma_\nu$ is a semiconjugacy from (X_f, f) to (Σ_ν, σ). The preimage of each point $\omega \in \Sigma_\nu$ is a compact interval of X_f, possibly consisting of only one point. If f is expansive then ζ is a homeomorphism. In particular, any two expansive Lorenz maps are conjugate if and only if they have the same kneading invariant.*

Proof. The proof of the first part is standard. Note that the continuity of ζ is not magical at all, since all the discontinuities of the original map f were doubled. If f is expansive then ζ is invertible, and—since X_f is compact—closed sets are mapped to closed sets, whence ζ is a homeomorphism. ∎

4.7. Remark. The nontrivial intervals J which are collapsed by ζ to points are precisely the maximal homtervals (1) of f. The itinerary $\zeta := \zeta(J)$ determines the type of the homterval J: If ζ is not preperiodic then J is a wandering interval, and if ζ is preperiodic then J is a preperiodic homterval. More precisely, if $\zeta = vw^\infty$ for some finite words v and w then $f^{|v|}(J)$ is contained in a maximal homterval which is mapped into itself by $f^{|w|}$. It follows that every point $x \in J$ is attracted to a periodic attractor with itinerary w^∞. Moreover, the attractor can only be an essential (2) periodic attractor if $\sigma^k w^\infty = \nu^+$ or $\sigma^k w^\infty = \nu^-$ for some $k \geq 0$. Let us call J an *essential* preperiodic homterval if this is the case, and an *inessential* preperiodic homterval otherwise. ◇

Since ν^+ and ν^- are also subject to condition (4.1), any pair of binary sequences that occurs as the kneading invariant of some Lorenz map must necessarily satisfy the following *admissibility condition*.

4.8. Definition (Admissibility condition). A pair $\nu = (\nu^+, \nu^-)$ of binary sequences is called *admissible* if

$$\text{(AC)} \qquad \sigma\nu^+ \leq \sigma^n\nu^\pm \leq \sigma\nu^- \quad \text{for every } n \in \mathbb{N}.$$

Here and in the following we tacitly assume that $\nu = (10*, 01*)$.

(1) Cf. Definition 3.32.
(2) Cf. Remark 3.46.

Now it is interesting to ask whether this condition is also sufficient, i.e., whether any pair satisfying (AC) occurs as the kneading invariant of some Lorenz map. Before giving a first answer in this direction let us make a general observation:

4.9. Remark. A lot of the following combinatorial arguments become much clearer if one takes a geometrical viewpoint by thinking of points on the interval being iterated instead of infinite binary sequences being shifted. Indeed, it is much more intuitive to say "c_n^- lies to the left of c" than to say "$\sigma^n\nu^- \le \nu^-$". Although a priori it is not clear whether any subshift Σ_ν with admissible invariant ν really arises from a Lorenz map, this viewpoint is justified by the following construction:

Every element ω of the shift space Σ can be considered as the binary expansion of some point $\phi(\omega) := .\omega := \sum_{k=0}^{\infty}\omega_k/2^{k+1} \bmod 1$ on the circle $\mathbb{S}^1$ with angles measured in full turns. In this way one obtains a semiconjugacy between the shift σ on Σ and the angle doubling map $f_2 : x \mapsto 2x \bmod 1$ on $\mathbb{S}^1$. The map ϕ is one-to-one except for the usual ambiguity that for every finite word w the expansions $w01^\infty$ and $w10^\infty$ represent the same point. If one excludes the trivial case $\nu = (10^\infty, 01^\infty)$—which is the kneading invariant of a Lorenz map having two surjective branches, e.g. the angle doubling map itself—then ϕ is a homeomorphism of Σ_ν onto some compact f_2-invariant subset of the circle, which can be identified with Σ_ν. Indeed, let $\alpha := \sigma\nu^+$, $\beta := \sigma\nu^-$ and let $J := (.\beta, .\alpha) \subset \mathbb{S}^1$ be the open arc between $.\beta$ and $.\alpha$ containing 0. Then Σ_ν consists of exactly those points whose iterates never visit the interval J. The set Σ_ν is either a Cantor set of Lebesgue measure zero or a finite set, as in the case $\nu = ((10)^\infty, (01)^\infty)$ where $\Sigma_\nu = \{\,.\overline{01}, .\overline{10}\} = \{1/3, 2/3\}$. ◇

This observation already leads to a very simple answer to the above question for expansive Lorenz maps:

4.10. Theorem (Hubbard & Sparrow [35]). *A pair ν of binary sequences is realizable as the kneading invariant of an expansive Lorenz map f if and only if it satisfies the following expansive admissibility condition*:

$$\text{(EAC)} \qquad \sigma\nu^+ \le \sigma^n\nu^+ < \sigma\nu^- \quad \textit{and} \quad \sigma\nu^+ < \sigma^n\nu^- \le \sigma\nu^- \quad \textit{for every } n \in \mathbb{N}.$$

The map f is uniquely defined up to topological conjugacy.

Before proving the theorem we remark that condition (EAC) is equivalent to

$$\text{(EAC}'\text{)} \qquad \nu \text{ satisfies (AC)} \quad \text{and} \quad \sigma^n\nu^\pm \ne \nu^\mp \quad \text{for every } n \in \mathbb{N}.$$

It is obvious that (EAC′) is necessary for (EAC). Conversely, assume by contradiction that there is some $n \in \mathbb{N}$ such that $\sigma^n\nu^+ = \sigma\nu^-$ and choose n minimal. If $n = 0$ then condition (EAC′) is violated right away, whence $n > 0$. Now $\sigma^{n-1}\nu^+$ has to start with a 0, otherwise $\sigma^{n-1}\nu^+ = 1\sigma^n\nu^+ \ge \sigma^n\nu^+$ (3) and n was not chosen minimal. But $\sigma^{n-1}\nu^+ = 0*$ implies $\sigma^{n-1}\nu^+ = \nu^-$.

Proof of the theorem. (EAC′) is necessary: Since (AC) is necessary, it only remains to show that $\sigma^n\nu^\pm \ne \nu^\mp$ for all n. Indeed, assume $\sigma^n\nu^+ = \nu^-$ for some n. If $c_n^+ < c^-$ then (c_n^+, c^-) is a homterval, which contradicts expansiveness. So $c_n^+ = c^-$, which implies

(3) For the last inequality observe that $1\omega \ge \omega$ for all $\omega \in \Sigma$.

$\sigma^n\nu^+ = \nu^+ \neq \nu^-$, again a contradiction (recall that ν^+ and ν^- were defined using one-sided limits).

(EAC′) is sufficient: Again exclude the trivial case $\nu = (10^\infty, 01^\infty)$, let α, β and J be defined as in Remark 4.9 and identify Σ_ν with the corresponding subset of the circle. Then (EAC) assures that $\Sigma_\nu \cap [0, 1/2]$ and $\Sigma_\nu \cap [1/2, 1]$ are Cantor sets, i.e., have the cardinality of the continuum. Indeed, if J_{-k} and J_{-l} are connected components of $f_2^{-k}J$ and $f_2^{-l}J$, respectively, then (EAC) implies that either one of them is a subset of the other or they are disjoint and not adjacent. Removing J and collapsing all its preiterates (i.e., all components of $\bigcup_k f^{-k}J$) to points one gets the required Lorenz map f. The critical point of f corresponds to the component $(.0\beta, .1\alpha)$ of $f^{-1}J$ which contains $1/2$. The uniqueness is clear from Proposition 4.6. ■

4.11. Remark. The proof of Theorem 4.12 shows that nonexpansive kneading invariants imply the existence of homtervals on both sides of the critical point: If $\sigma^k\nu^+ = \nu^-$ then (c_k^+, c) is a homterval which can be pulled back monotonically to (c^+, z_k^+), where z_k^+ is the closest preimage of order S_k^+ on the right hand side of the critical point (cf. Definition 2.13). If ν^- is not periodic then these homtervals are either inessential preperiodic homtervals or wandering intervals (cf. Remark 4.7). ◇

Theorem 4.10 has two drawbacks: First, its proof cannot be modified to deal with nonexpansive admissible kneading invariants. The reason is that Σ_ν may be a very small set, even finite, in which case by collapsing all preiterates of J one ends up with a single point. Second, for every expansive kneading invariant the existence of a Lorenz map with the required combinatorial behaviour is established by the artificial construction of gluing together a Cantor set. In particular, the theorem does not answer the question whether in a given family of Lorenz maps every admissible kneading invariant occurs. Such families are called full families.

The following theorem fixes the first drawback but it is still based on an artificial construction. It is similar to one used by de Melo & van Strien [13] for l-modal maps. The question whether certain families of Lorenz maps are full families will be addressed in more detail in Chapter 5.

4.12. Theorem. *A pair $\nu = (\nu^+, \nu^-)$ of binary sequences is realizable as the kneading invariant of a Lorenz map f if and only if it satisfies the admissibility condition* (AC).

Proof. (i) Choose arbitrary points c, p and q on the real line such that $p < c < q$. The first step is to associate points $c_k^i \in [p, q]$ with all the shifted kneading invariants $\sigma^k\nu^i$, $k \in \mathbb{N}$, $i \in \{+, -\}$ (4). To begin, let $c_0^+ := c_0^- := c$. If $\sigma\nu^+ = 0^\infty$ let $c_1^+ := p$ and if this is not the case choose c_1^+ slightly larger than p. Proceed similarly for c_1^-. The points c_k^+ and c_l^- will be ordered essentially in the same way as the itineraries $\sigma^k\nu^+$ and $\sigma^l\nu^-$ in the shift space, i.e.

$$c_k^i \begin{smallmatrix}<\\=\\>\end{smallmatrix} c_l^j \quad \text{if} \quad \sigma^k\nu^i \begin{smallmatrix}<\\=\\>\end{smallmatrix} \sigma^l\nu^j \quad \text{for } k, l \in \mathbb{N} \text{ and } i, j \in \{+, -\},$$

(4) For the moment we treat $\sigma^k\nu^+$ and $\sigma^l\nu^-$ as distinct objects, even if the itineraries coincide. A more precise but clumsy notation would be to write $(+, \sigma^k\nu^+)$ and $(-, \sigma^l\nu^-)$ instead.

with the following exception: Given two itineraries ω_1 and ω_2 write $\omega_1 \sim \omega_2$ iff $\omega_1, \omega_2 \in \{w\nu^-, w\nu^+\}$ for some finite word w. Now

$$\text{if} \quad \sigma^k\nu^+ \sim \sigma^l\nu^- \quad \text{then choose} \quad \begin{cases} c_k^+ = c_l^- & \text{if } \sigma^k\nu^+ > \sigma^l\nu^-, \\ c_k^+ < c_l^- & \text{if } \sigma^k\nu^+ \leq \sigma^l\nu^-. \end{cases}$$

Note that the admissibility condition guarantees that no shift of the kneading invariants can fall between ν^- and ν^+. It can happen that two different itineraries belong to a single point on the line, and also the converse that there are two points on the line that have the same itinerary. In the former case the point is precritical and the two itineraries are its left and right hand side limits. The latter case occurs if the kneading invariant is not expansive (cf. Remark 4.11).

(ii) If there are infinitely many distinct points c_k^i then a little extra care is necessary to control where they accumulate: A point c_k^i which is associated with a single itinerary $\sigma^k\nu^i$ should be accumulated by other points from the left hand respectively right hand side if and only if the same holds for the itinerary $\sigma^k\nu^i$. The same holds analogously for points having a left hand and a right hand side itinerary associated and for intervals (c_k^+, c_l^-) with $\sigma^k\nu^+ = \sigma^l\nu^-$. This goal can easily be accomplished by inserting the points c_k^i successively with increasing index k into the interval $[c_1^+, c_1^-]$, leaving space on the appropriate side if necessary. We omit the lengthy details.

(iii) Let f_n be the piecewise linear map whose graph interpolates the points (p,p), (q,q), (c_k^+, c_{k+1}^+) and (c_k^-, c_{k+1}^-) for $k \leq n$. By construction f_n is a Lorenz map for every n and the kneading invariant of f_n coincides with the given one up to the nth digit.

The sequence f_n converges uniformly on compact sets to a Lorenz map f which is strictly monotonic on each side of c. This is trivial if there are only finitely many points c_k^i and otherwise guaranteed by step (ii). Now f is a Lorenz map with the required kneading invariant. ∎

4.13. Example. 1) From the kneading invariant $(\nu^+, \nu^-) = ((0110)^\infty, (1001)^\infty)$ one obtains seven distinct points:

$$\begin{array}{ccccccccccccccc} c_1^+ & \prec & c_3^- & < & c_2^+ & \prec & c^- & = & c^+ & \prec & c_2^- & < & c_3^+ & \prec & c_1^- \\ (0011) & & (0011) & & (0110) & & (0110) & & (1001) & & (1001) & & (1100) & & (1100) \end{array}$$

Here (w) is an abbreviation for $(w)^\infty$ and $c_1^+ \prec c_3^-$ means that the points $c_1^+ < c_3^-$ have the same itinerary. This is a good example of why it is necessary to associate two points with some of the itineraries. If one had not done that there would have been only three points

$$\begin{array}{ccccccc} c_1^+ & < & c^- & = & c^+ & < & c_1^- \\ (0011) & & (0110) & & (1001) & & (1100) \end{array}$$

and interpolating these three points one would have obtained only a Lorenz map with kneading invariant $((01)^\infty, (10)^\infty)$.

2) From the kneading invariant $(\nu^+, \nu^-) = ((100010)^\infty, (010100)^\infty)$ one obtains ten points:

$$c_1^+ \prec c_4^- < c_2^+ \prec c_5^- = c_5^+ \prec c_2^- < c_3^+ \prec c^- = c^+ \prec c_3^- < c_4^+ \prec c_1^- .$$

Note that the kneading invariant is both periodic and nonexpansive since (ν^+, ν^-) equals $((w_+w_-)^\infty, (w_-w_+)^\infty)$, where $w_- = 010$ and $w_+ = 100$ ([5]). The reason for the identification of c_5^- and c_5^+ is that the last digits of w_-w_+ and of w_+w_- are identical. $\diamond$

4.2. The splitting of itineraries. The cutting times can be calculated in a simple way from the kneading invariant: The first cutting$^+$ time is $S_0^+ = 1$, and if $m = S_{k-1}^+$ is a cutting$^+$ time, then $S_k^+ = n$ where $n > m$ is the minimal integer such that $c_n^+ < c < c_{n-m}^-$, i.e., the minimal integer at which the itineraries of c_{m+1}^+ and c_1^- differ:

$$\underbrace{\nu_{m+1}^+ \dots \nu_{n-1}^+}\, \nu_n^+ = \underbrace{\nu_1^- \dots \nu_{n-m-1}^-}\, \mathring{\nu}_{n-m}^-. \tag{4.2}$$

Here $\mathring{\nu}_k^- := \frac{0}{1}$ if $\nu_k^- = \frac{1}{0}$. Since we will be occupied a lot with comparing blocks of itineraries with others, let us introduce some useful shorthand:

4.14. Definition (Maximal shadows). For $\zeta \in \Sigma$ and $m \in \mathbb{N}$, $n \in \mathbb{N}_\infty$ let

$$\zeta(m,n] := \begin{cases} \zeta_{m+1}\zeta_{m+2}\dots\zeta_n & \text{if } n \text{ is finite,} \\ \zeta_{m+1}\zeta_{m+2}\dots & \text{if } n = \infty. \end{cases}$$

If the words $\zeta(n,m]$ and $\nu^\pm(0,n-m]$ differ precisely in the last digit then we say that $\zeta(m,n]$ is a *maximal shadow* of $\nu^\pm$ and write $\zeta(i,j] \mathrel{\mathring{=}} \nu^\pm(0,n-m]$.

Now the statement of (4.2) is simply that $\nu^+(m,n]$ is a maximal shadow of ν^-:

$$\nu^+(m,n] \mathrel{\mathring{=}} \nu^-(0,n-m]. \tag{4.2$'$}$$

The cutting$^-$ times are defined symmetrically. In this way one obtains a *splitting* of $\nu^\pm$ into *blocks* $\Delta_k^\mp$:

$$\nu^\pm = \tfrac{1}{0}\,\tfrac{0}{1}\,\Delta_1^\mp\,\Delta_2^\mp\,\Delta_3^\mp\,\dots, \tag{4.3}$$

where each block $\Delta_k^\mp := \nu^\pm(S_{k-1}^\pm, S_k^\pm] \mathrel{\mathring{=}} \nu^\mp(0, S_k^\pm - S_{k-1}^\pm]$ is a maximal shadow of $\nu^\mp$. The length $|\Delta_k^\mp|$ of each block equals $S_k^\pm - S_{k-1}^\pm$. It may happen that there are only finitely many blocks and the last block has infinite length. The cutting times satisfy the following recursion: $S_0^\pm = 1$ and

$$S_k^\pm = S_{k-1}^\pm + \#^\pm(S_{k-1}^\pm), \tag{4.4}$$

where $\#^\pm(m) := \min\{j > 0 \mid \nu_{m+j}^\pm \neq \nu_j^\mp\}$. Given any kneading invariant, the cutting times can be calculated in a purely combinatorial fashion by this splitting algorithm.

In a similar fashion one obtains the co-cutting times by a modified splitting of the kneading invariant starting at $n^\pm$, the so-called *co-splitting*:

$$\nu^\pm = \tfrac{1}{0}\,(\tfrac{0}{1})^{\widetilde{S}_0^\pm - 1}\,\tfrac{1}{0}\,\widetilde{\Delta}_1^\pm\,\widetilde{\Delta}_2^\pm\,\widetilde{\Delta}_3^\pm\,\dots, \tag{4.5}$$

i.e., starting from the second digit that equals $\frac{1}{0}$ one splits the itinerary $\nu^\pm$ into blocks $\widetilde{\Delta}_k^\pm := \nu^\pm(\widetilde{S}_{k-1}^\pm, \widetilde{S}_k^\pm] \mathrel{\mathring{=}} \nu^\pm(0, \widetilde{S}_k^\pm - \widetilde{S}_{k-1}^\pm]$ which are maximal shadows of $\nu^\pm$. The co-cutting times satisfy the following recursion: $\widetilde{S}_0^\pm = n^\pm$ is the minimal $j \geq 1$ such that

([5]) The kneading invariant of this example was obtained by period doubling from the kneading invariant $((100)^\infty, (010)^\infty)$ of a circle map with rotation number 1/3.

$\nu_j^{\pm} = \frac{1}{0}$ and

$$\widetilde{S}_k^{\pm} = \widetilde{S}_{k-1}^{\pm} + \widetilde{\#}^{\pm}(\widetilde{S}_{k-1}^{\pm}), \tag{4.6}$$

where $\widetilde{\#}^{\pm}(m) := \min\{j > 0 \mid \nu_{m+j}^{\pm} \neq \nu_j^{\pm}\}$. Note that the splitting of a kneading is obtained by comparing w.r.t. the other kneading, whereas the co-splitting is obtained by comparing w.r.t. itself.

4.15. Example. The following shows the splitting of a kneading invariant. Cutting times are marked with a "$\wedge$" below and co-cutting times with a "$\vee$" above the digit.

$$\nu^+ = 1\underset{\wedge}{0}\underset{\wedge}{0}\underset{\wedge}{0}\overset{\vee}{1}0\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}0\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}0\overset{\vee}{1}0\overset{\vee}{1}00\underset{\wedge}{0}1010\underset{\wedge}{0}1010\underset{\wedge}{0}10\overset{\vee}{1}0\overset{\vee}{1}00\underset{\wedge}{0}1010\underset{\wedge}{0}1010\overset{\vee}{1}0\ldots,$$

$$\nu^- = 0\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}0\underset{\wedge}{1}0\overset{\vee}{0}\underset{\wedge}{1}0\overset{\vee}{0}\underset{\wedge}{1}0\overset{\vee}{0}\overset{\vee}{0}1010\underset{\wedge}{1}00\underset{\wedge}{1}00\underset{\wedge}{1}0001010\overset{\vee}{0}1010\overset{\vee}{0}10\underset{\wedge}{1}0\overset{\vee}{0}\underset{\wedge}{1}0\overset{\vee}{0}\underset{\wedge}{1}0\underset{\wedge}{1}0\overset{\vee}{0}\underset{\wedge}{1}0\ldots$$

Note that cutting$^{\pm}$ and co-cutting$^{\pm}$ times never occur simultaneously. It will be shown in the next section that this property is an equivalent characterization of the admissible kneading invariants. $\diamond$

The splitting and the co-splitting are two variants of a general splitting technique: Let (ν^+, ν^-) be an admissible kneading invariant and $\zeta \in \Sigma$ an arbitrary binary sequence. The *left cutting times* $(S_k^<(\zeta))_{k\in\mathbb{N}}$ and the *right cutting times* $(S_k^>(\zeta))_{k\in\mathbb{N}}$ of ζ can be determined by the two splittings below. The left (resp. right) splitting starts at the first digit that equals 1 (resp. 0) and the blocks $\Delta_k^<(\zeta)$ (resp. $\Delta_k^>(\zeta)$) are maximal shadows of ν^+ (resp. ν^-).

$$\zeta = (\tfrac{0}{1})^{S_0^{\lessgtr}-1} 1\, \Delta_1^{\lessgtr}\, \Delta_2^{\lessgtr}\, \Delta_3^{\lessgtr}\, \ldots$$

Here we omitted the dependence on ζ.

4.16. Lemma. (i) *Let (ν^+, ν^-) be an admissible kneading invariant and $\zeta \in \Sigma_\nu$ be an admissible sequence. Then the difference between consecutive cutting$^{\lessgtr}$ times of ζ is a cutting$^{\pm}$ time, i.e., there exist maps $Q_\zeta^< : \mathbb{N} \to \mathbb{N}_\infty$ and $Q_\zeta^> : \mathbb{N} \to \mathbb{N}_\infty$ such that*

$$S_k^{\lessgtr}(\zeta) - S_{k-1}^{\lessgtr}(\zeta) = S^{\pm}_{Q_\zeta^{\lessgtr}(k)}.$$

(ii) *The map $\zeta \mapsto Q_\zeta^>$ is order preserving on Σ^- and the map $\zeta \mapsto Q_\zeta^<$ is order reversing on Σ^+, where $\Sigma^{\pm} := \{\zeta \in \Sigma \mid \zeta = \frac{1}{0}*\}$ and $(\mathbb{N}_\infty)^{\mathbb{N}}$ is endowed with the standard lexicographical order.*

Proof. We only prove the lemma for the left cutting times. Replacing ζ by $\sigma^{S_0^<(\zeta)-1}\zeta$ if necessary, we can assume w.l.o.g. that $\zeta \in \Sigma^+$. Furthermore we abbreviate $x_k := \sigma^k\zeta$ and $c_k^{\pm} := \sigma^k\nu^{\pm}$ in the spirit of Remark 4.9.

(i) Assume $l := |\Delta_1^<(\zeta)| < \infty$. If $\zeta_1 = 1$ then $l = 1 \in \mathcal{S}^+$, so assume $\zeta_1 = 0$. Now $x_1 \in [c_1^+, c^-]$ and by induction we get $x_n \in [c_n^+, c^-_{n-S^<\langle n\rangle}]$ for $1 < n \leq d$, where $S^<\langle n\rangle$ denotes the last cutting$^<$ time before n. Now $c \in [c_l^+, x_l] \subseteq [c_l^+, c^-_{l-S^<\langle l\rangle}]$ and it follows that l is a cutting$^+$ time. Because ζ is admissible, $\sigma\nu^+ \leq \sigma\zeta$ and it follows that the last digit of the first block, which is ζ_d, equals 1. Now $x_l = \sigma^l\zeta = 1*$ is again contained in Σ^+ and the assertion of the lemma follows by induction.

(ii) Assume $\zeta, \zeta' \in \Sigma^+$, $Q_\zeta^< > Q_{\zeta'}^<$ and choose k minimal such that $Q_\zeta^<(k) > Q_{\zeta'}^<(k)$. Then the first discrepancy between the strings ζ and ζ' occurs at the digit number $n = S_k^<(\zeta')$ where $\zeta'_n = 1$. Hence $\zeta < \zeta'$. ∎

As an immediate corollary to Lemma 4.16 we get the following necessary condition for the admissible kneading invariants (cf. (2.7) and (2.11)). In the next section it will turn out that this condition is in fact an equivalent characterization of admissibility.

4.17. Corollary. *If (ν^+, ν^-) satisfies the admissibility condition* (AC) *then the differences between consecutive cutting$^\pm$ times are cutting$^\mp$ times and the differences between consecutive co-cutting$^\pm$ times are cutting$^\pm$ times.*

Proof. The splitting and co-splitting of the kneading invariant coincide up to a shift with the left and right splittings of $\sigma\nu^+$ and $\sigma\nu^-$, more precisely $S_k^\pm = S_k^{\gtrless}(\sigma\nu^\pm) + 1$ and $\widetilde{S}_k^\pm = S_k^{\lessgtr}(\sigma\nu^\pm) + 1$. ∎

4.18. Remark. From Corollary 4.17 one finds in particular that $\nu^\pm_{S_k^\pm} = \frac{0}{1}$ for all k. That implies that if one moves all blocks one digit to the left, i.e., if one replaces every block $\nu^\pm(S_{k-1}^\pm, S_k^\pm]$ by the block $\nu^\pm[S_{k-1}^\pm, S_k^\pm)$ which includes $S_{k-1}^\pm$ and excludes $S_k^\pm$, then $\nu^\pm[S_{k-1}^\pm, S_k^\pm) = \nu^\mp[0, S_k^\pm - S_{k-1}^\pm)$, i.e., the blocks are identical copies. Similarly, $\nu^\pm_{\widetilde{S}_k^\pm} = \frac{1}{0}$, which implies that $\nu^\pm[\widetilde{S}_{k-1}^\pm, \widetilde{S}_k^\pm) = \nu^\pm[0, \widetilde{S}_k^\pm - \widetilde{S}_{k-1}^\pm)$. ◇

4.3. Admissibility conditions. If (ν^+, ν^-) is an arbitrary pair of binary sequences beginning with $10*$ and $01*$, respectively, then the splitting (4.3) and co-splitting (4.5) are well defined even if (ν^+, ν^-) does not satisfy the admissibility condition. These splittings determine the cutting and co-cutting times according to (4.4) and (4.6). The following theorem shows how the admissibility condition translates into the language of cutting and co-cutting times, and of the kneading map.

4.19. Theorem. *For (ν^+, ν^-) the following formulations of the admissibility condition are equivalent*:

(AC1) *For every $n \in \mathbb{N}$,*

$$\sigma\nu^+ \le \sigma^n\nu^\pm \le \sigma\nu^-. \tag{4.7}$$

(AC2) *For every $n \in \mathbb{N}$,*

$$\nu^+_{n-\widetilde{S}^+\langle n\rangle} \le \nu^+_n \le \nu^-_{n-S^+\langle n\rangle}, \tag{4.8}$$

$$\nu^+_{n-S^-\langle n\rangle} \le \nu^-_n \le \nu^-_{n-\widetilde{S}^-\langle n\rangle}, \tag{4.9}$$

(AC3) *All finite cutting$^\pm$ and co-cutting$^\pm$ times are distinct*:

$$\mathcal{S}^+ \cap \widetilde{\mathcal{S}}^+ \subseteq \{\infty\}, \tag{4.10}$$

$$\mathcal{S}^- \cap \widetilde{\mathcal{S}}^- \subseteq \{\infty\}. \tag{4.11}$$

(AC4) *The differences between consecutive cutting$^\pm$ times are cutting$^\mp$ times*:

$$S_k^\pm - S_{k-1}^\pm = S^\mp_{Q^\pm(k)}, \tag{4.12}$$

and the differences between consecutive co-cutting$^{\pm}$ *times are cutting*$^{\pm}$ *times*:

$$\widetilde{S}_k^{\pm} - \widetilde{S}_{k-1}^{\pm} = S^{\pm}_{\widetilde{Q}^{\pm}(k)}. \tag{4.13}$$

(AC5) *The differences between consecutive cutting*$^{\pm}$ *times are cutting*$^{\mp}$ *times*:

$$S_k^{\pm} - S_{k-1}^{\pm} = S^{\mp}_{Q^{\pm}(k)}, \tag{4.14}$$

and the kneading map Q satisfies the Hofbauer condition:

$$(Q^{\pm}(k+j))_{j\geq 1} \geq (Q^{\pm}(Q^{\mp}Q^{\pm}(k)+j))_{j\geq 1} \quad \forall k \geq 1, \tag{4.15}$$

where the ordering $\leq$ is just the lexicographical ordering on ${\mathbb{N}_\infty}^{\mathbb{N}}$.

(AC6) *For every $n \in \mathbb{N}$*,

$$\#^{\pm}(n) \in \mathcal{S}^{\pm}, \tag{4.16}$$

$$\widetilde{\#}^{\pm}(n) \in \mathcal{S}^{\pm}. \tag{4.17}$$

Before we prove the theorem we make some comments on these conditions first. Most of their unimodal analogues can already be found in the paper of Hofbauer & Keller [30]. Condition (AC2) is a little bit hidden, since co-cutting times were not explicitly introduced at that time yet. However, the sequences $a(n)$ and $b(n)$ defined in [30] are precisely the last cutting and co-cutting times before n, respectively. A more geometrical interpretation and the terminology of co-cutting times was introduced by Sands [55] and Bruin [4]. A compact survey on this matter can be found in Bruin [5].

Condition (AC2) tells us that the original admissibility condition (AC1) contains a lot of redundancies and that it is not really necessary to check the inequalities (4.7) for all integers n, but only for the "dynamically important" ones. It is also important from a practical viewpoint since it can be used to implement a fast algorithm to determine the finite prefixes of length n of all admissible kneading invariants by pruning the set of all pairs of binary words of length n using a backtracking method.

Condition (AC3) is rather amazing since it appears to be very weak. It shows nicely how strongly cutting and co-cutting times are tied together. The first one to point it out for unimodal maps was Thunberg [56]. Following the ideas of Hofbauer & Keller he showed that (AC3) implies the existence of a parameter value in the quadratic family with the given kneading invariant. We provide a simpler and purely combinatorial proof here.

While condition (AC2) is ideal to be programmed on a computer, condition (AC3) is the method of choice for human beings to check admissibility of finite prefixes: Just split the prefix marking each cutting and co-cutting time below respectively above the corresponding digit. The prefix is admissible if and only if no digit is marked twice (cf. the examples at the end of this section).

Conditions (AC4) and (AC5) are very similar, one formulated in the language of kneading maps and the other in the language of cutting and co-cutting times. Although rather strange looking, (AC5) is of great theoretical importance, since specifying a kneading map is the easiest way to define a Lorenz map whose Markov diagram has certain specific combinatorial properties. Note that any nondecreasing kneading map trivially satisfies the Hofbauer condition (4.15), because $Q^{\mp}Q^{\pm}(k) < k$ by Remark 2.11.

Condition (AC6) is a natural generalization of (AC4) which occurs when proving the implication (AC4)⇒(AC1).

Proof of Theorem 4.19. The statement of the theorem follows from the chains of implications (AC1)⇒(AC2)⇒(AC3)⇒(AC4)⇒(AC6)⇒(AC1) and (AC1)⇒(AC5)⇒(AC4). Additionally, we also include some redundant implications if there is a simple and short proof for them.

(AC1)⇒(AC2). Condition (AC1) implies the inequalities $\sigma\nu^+ \leq \sigma^{\widetilde{S}^+\langle n\rangle+1}\nu^+$, $\sigma^{S^+\langle n\rangle+1}\nu^+ \leq \sigma\nu^-$ and $\sigma\nu^+ \leq \sigma^{S^-\langle n\rangle+1}\nu^-$, $\sigma^{\widetilde{S}^-\langle n\rangle+1}\nu^- \leq \sigma\nu^-$. The first two inequalities imply (4.8), the second two imply (4.9).

(AC2)⇒(AC1). An equivalent reformulation of (AC2) is the following:

$$\sigma^{n-\widetilde{S}^+\langle n\rangle}\nu^+ \leq \sigma^n\nu^+ \leq \sigma^{n-S^+\langle n\rangle}\nu^-, \quad \forall n \in \mathbb{N},$$
$$\sigma^{n-S^-\langle n\rangle}\nu^+ \leq \sigma^n\nu^- \leq \sigma^{n-\widetilde{S}^-\langle n\rangle}\nu^-, \quad \forall n \in \mathbb{N}.$$

If we assume that (4.7) already holds for all integers less than n, in particular for $n-S^\pm\langle n\rangle$ and $n-\widetilde{S}^\pm\langle n\rangle$, then we get

$$\sigma\nu^+ \leq \sigma^{n-\widetilde{S}^+\langle n\rangle}\nu^+ \leq \sigma^n\nu^+ \leq \sigma^{n-S^+\langle n\rangle}\nu^- \leq \sigma\nu^-,$$
$$\sigma\nu^+ \leq \sigma^{n-S^-\langle n\rangle}\nu^+ \leq \sigma^n\nu^- \leq \sigma^{n-\widetilde{S}^-\langle n\rangle}\nu^- \leq \sigma\nu^+,$$

which implies that (4.7) also holds for n. Thus (AC1) follows by induction.

(AC2)⇒(AC3). If $n < \infty$ is a cutting$^+$ time and a co-cutting$^+$ time simultaneously, then $\nu^+_{n-\widetilde{S}^+\langle n\rangle} \neq \nu^+_n \neq \nu^-_{n-S^+\langle n\rangle}$, which contradicts (4.8), so this cannot occur. Likewise, $\mathcal{S}^- \cap \widetilde{\mathcal{S}}^- \not\subseteq \{\infty\}$ contradicts (4.9).

(AC3)⇒(AC4). Since cutting$^\pm$ and co-cutting$^\pm$ times are distinct, the set $\mathcal{S}^\pm \cup \widetilde{\mathcal{S}}^\pm$ can be grouped into chains of cutting$^\pm$ times with no co-cutting$^\pm$ times in between and chains of co-cutting$^\pm$ times with no cutting$^\pm$ times in between. The proof proceeds via induction from one chain to another. We distinguish two cases, namely (i) a chain of co-cutting$^\pm$ times between two cutting$^\pm$ times $S^\pm_{k-1}$ and $S^\pm_k$, and (ii) a chain of cutting$^\pm$ times between two co-cutting$^\pm$ times $\widetilde{S}^\pm_{k-1}$ and $\widetilde{S}^\pm_k$.

For the first chain of cutting$^\pm$ times, i.e., for $S^\pm_k < \widetilde{S}^\pm_0 = n^\pm$, the assertions of (AC4) are true, since $S^\pm_k - S^\pm_{k-1} = 1$. The induction continues in case (i) with $\widetilde{S}^\pm_l = n^\pm$ and $S^\pm_{k-1} = n^\pm - 1$ which satisfy the hypothesis (4.18) required below, since $\widetilde{S}^\pm_l - S^\pm_{k-1} = 1$.

Case (i). Let $\{\widetilde{S}^\pm_l, \ldots, \widetilde{S}^\pm_{l+d}\}$ be the set of co-cutting$^\pm$ times between $S^\pm_{k-1}$ and $S^\pm_k$ (d may be infinite if $S^\pm_k = \infty$). We use the additional induction assumption:

$$\widetilde{S}^\pm_l - S^\pm_{k-1} = S^\mp_m \quad \text{for some } m \in \mathbb{N}. \tag{4.18}$$

Since the strings $\nu^\pm(S^\pm_{k-1}, S^\pm_k]$ and $\nu^\mp(0, S^\pm_k - S^\pm_{k-1}]$ are identical except for the last digit, the co-splitting starting at $\nu^\pm_{\widetilde{S}^\pm_l}$ and the splitting starting at $\nu^\mp_{S^\mp_m}$ are the same (see Figure 4.1). This implies

$$\widetilde{S}^\pm_{l+j} = S^\pm_{k-1} + S^\mp_{m+j} \quad \text{for } 0 \leq j \leq d. \tag{4.19}$$

If $S^\pm_k = \infty$ the induction terminates. If this is not the case then the last block is $\nu^\pm(\widetilde{S}^\pm_{l+d}, S^\pm_k] \mathrel{\hat{=}} \nu^\mp(S^\mp_{m+d}, S^\pm_k - S^\pm_{k-1}]$ and since $S^\pm_k$ is not a co-cutting$^\pm$ time by (AC3),

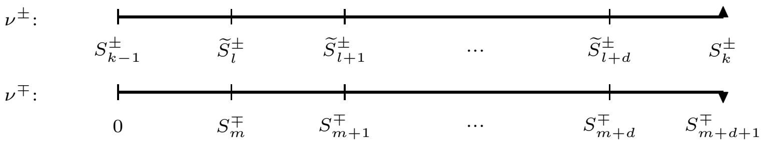


Fig. 4.1. A chain of co-cutting$^{\pm}$ times between consecutive cutting$^{\pm}$ times

because it already is a cutting$^{\pm}$ time, $S^{\pm}_k - S^{\pm}_{k-1}$ has to be a cutting$^{\mp}$ time, i.e., $S^{\pm}_k - S^{\pm}_{k-1} = S^{\mp}_{m+d+1}$, which proves (4.12) in this case.

Moreover, (4.19) implies $\widetilde{S}^{\pm}_{l+j} - \widetilde{S}^{\pm}_{l+j-1} = S^{\mp}_{m+j} - S^{\mp}_{m+j-1}$ for $1 \le j \le d$. By induction $S^{\mp}_{m+j} - S^{\mp}_{m+j-1} \in \mathcal{S}^{\pm}$, which proves (4.13) for $k = l+1, \dots, l+d$.

The induction continues in case (ii), where $\widetilde{S}^{\pm}_{l+d} < \widetilde{S}^{\pm}_{l+d+1}$ play the role of the two consecutive co-cutting$^{\pm}$ times. Note that $S^{\pm}_k$ is the first cutting$^{\pm}$ time following $\widetilde{S}^{\pm}_{l+d}$ and that $S^{\pm}_k - \widetilde{S}^{\pm}_{l+d} = S^{\mp}_{m+d+1} - S^{\mp}_{m+d}$. The latter is a cutting$^{\pm}$ time by induction, so assumption (4.20) is satisfied.

Case (ii). This case is completely analogous to case (i). Nevertheless, we provide the arguments in full detail since it is easy to get confused with all the $\pm$ and $\sim$ signs. Let $\{S^{\pm}_l, \dots, S^{\pm}_{l+d}\}$ be the set of cutting$^{\pm}$ times between $\widetilde{S}^{\pm}_{k-1}$ and $\widetilde{S}^{\pm}_k$ (d may be infinite if $\widetilde{S}^{\pm}_k = \infty$). We use the additional induction assumption:

$$S^{\pm}_l - \widetilde{S}^{\pm}_{k-1} = S^{\pm}_m \quad \text{for some } m \in \mathbb{N}. \tag{4.20}$$

Since the strings $\nu^{\pm}(\widetilde{S}^{\pm}_{k-1}, \widetilde{S}^{\pm}_k]$ and $\nu^{\pm}(0, \widetilde{S}^{\pm}_k - \widetilde{S}^{\pm}_{k-1}]$ are identical except for the last digit, the splitting starting at $\nu^{\pm}_{S^{\pm}_l}$ and the splitting starting at $\nu^{\pm}_{S^{\pm}_m}$ are the same (see Figure 4.2). This implies

$$S^{\pm}_{l+j} = \widetilde{S}^{\pm}_{k-1} + S^{\pm}_{m+j} \quad \text{for } 0 \le j \le d. \tag{4.21}$$

If $\widetilde{S}^{\pm}_k = \infty$ the induction terminates. If this is not the case then the last block is $\nu^{\pm}(S^{\pm}_{l+d}, \widetilde{S}^{\pm}_k] \circeq \nu^{\pm}(S^{\pm}_{m+d}, \widetilde{S}^{\pm}_k - \widetilde{S}^{\pm}_{k-1}]$ and since $\widetilde{S}^{\pm}_k$ is not a cutting$^{\pm}$ time by (AC3), $\widetilde{S}^{\pm}_k - \widetilde{S}^{\pm}_{k-1}$ has to be a cutting$^{\pm}$ time, i.e., $\widetilde{S}^{\pm}_k - \widetilde{S}^{\pm}_{k-1} = S^{\pm}_{m+d+1}$, which proves (4.13) in this case.

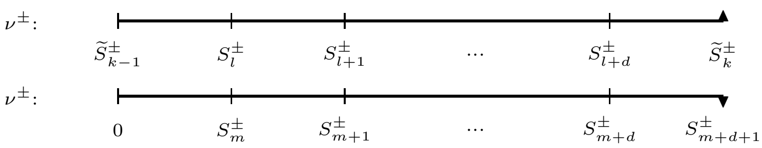


Fig. 4.2. A chain of cutting$^{\pm}$ times between consecutive co-cutting$^{\pm}$ times

Moreover, (4.21) implies $S^{\pm}_{l+j} - S^{\pm}_{l+j-1} = S^{\pm}_{m+j} - S^{\pm}_{m+j-1}$ for $1 \le j \le d$. By induction $S^{\pm}_{m+j} - S^{\pm}_{m+j-1} \in \mathcal{S}^{\mp}$, which proves (4.12) for $k = l+1, \dots, l+d$.

The induction continues in case (i), where $S^{\pm}_{l+d} < S^{\pm}_{l+d+1} < \infty$ play the role of the two consecutive cutting$^{\pm}$ times. Note that $\widetilde{S}^{\pm}_k$ is the first co-cutting$^{\pm}$ time following $S^{\pm}_{l+d}$ and that $\widetilde{S}^{\pm}_k - S^{\pm}_{l+d} = S^{\pm}_{m+d+1} - S^{\pm}_{m+d}$. The latter is a cutting$^{\pm}$ time by induction, so assumption (4.18) is satisfied.

(AC4)⇒(AC6). Obviously, (4.16) holds for all cutting$^\pm$ times and (4.17) holds for all co-cutting$^\pm$ times. So we can assume $S^\pm_{k-1} < n < S^\pm_k$ resp. $\widetilde{S}^\pm_{k-1} < n < \widetilde{S}^\pm_k$ for the proof of (4.16) resp. (4.17). First of all, from $\nu^\pm_1 = \frac{0}{1}$ and (4.12) one obtains $\nu^\pm_{S^\pm_k} = \frac{0}{1}$ for all k. The proof of the implication is by induction on n with the additional induction hypothesis:

$$n < \widetilde{S}^\pm_k < n + \#^\mp(n) \Rightarrow \widetilde{S}^\pm_k - n \in \mathcal{S}^\mp, \tag{4.22}$$

$$n < S^\pm_k < n + \widetilde{\#}^\pm(n) \Rightarrow S^\pm_k - n \in \mathcal{S}^\pm. \tag{4.23}$$

Fix some $n \in \mathbb{N}$ and assume that (4.16), (4.17), (4.22) and (4.23) hold for all $n' < n$.

Proof of (4.16). Throughout the proof abbreviate $n' := n - S^\pm_{k-1}$.

Case (i): $n+\#^\mp(n) < S^\pm_k$ (cf. Figure 4.3). Here $\nu^\pm(n, n+\#^\mp(n)] = \nu^\mp(n', n'+\#^\mp(n)]$, so $\#^\mp(n)$ equals $\widetilde{\#}^\mp(n')$ which is a cutting$^\mp$ time by induction.

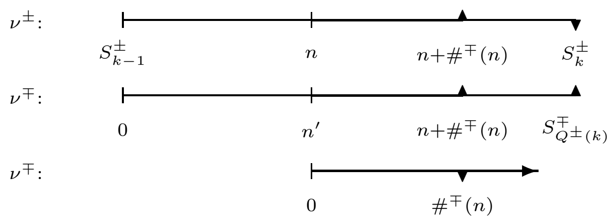


Fig. 4.3. Proof of (4.16), case (i)

Case (ii): $n+\#^\mp(n) > S^\pm_k$ (cf. Figure 4.4). Since $\nu^\pm(n, S^\pm_k] \circeq \nu^\mp(n', S^\mp_{Q^\pm(k)}]$, it follows that $S^\pm_k - n = \widetilde{\#}^\mp(n')$. The latter is a cutting$^\mp$ time by induction, so $S^\pm_k - n = S^\mp_m$ for some m. This yields the contradiction $\frac{1}{0} = \nu^\pm_{S^\pm_k} = \nu^\mp_{S^\mp_m} = \frac{0}{1}$, whence case (ii) cannot occur.

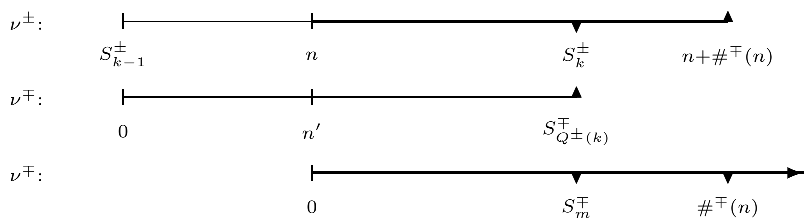


Fig. 4.4. Proof of (4.16), case (ii)

Case (iii): $n + \#^\mp(n) = S^\pm_k$ (cf. Figure 4.5). Since $\nu^\pm(n, S^\pm_k] \circeq \nu^\mp(n', S^\mp_{Q^\pm(k-1)}]$, we have $n' < S^\mp_{Q^\pm(k)} < n' + \widetilde{\#}^\mp(n')$. Using (4.23) for $n' = n - S^\pm_{k-1}$ instead of n, it follows that $\#^\mp(n) = S^\pm_k - n = S^\mp_{Q^\pm(k)} - n' = S^\mp_m$ for some m.

Now we verify the induction assumption (4.22) for n itself.

Proof of (4.22). Let $\{\widetilde{S}^\pm_l, \ldots, \widetilde{S}^\pm_{l+d}\} \ni S^\pm_k$ be the set of co-cutting$^\pm$ times between n and $n + \#^\mp(n)$ and abbreviate $\widetilde{n}' := n - \widetilde{S}^\pm_{l-1}$ (cf. Figure 4.6). From $\nu^\pm(n, \widetilde{S}^\pm_l] \circeq \nu^\pm(n - \widetilde{S}^\pm_{l-1}, S^\pm_{\widetilde{Q}^\pm(l)}]$ one gets $\widetilde{S}^\pm_l - n = \widetilde{\#}^\mp(\widetilde{n}') = S^\mp_m$ for some m by induction. Since $\nu^\pm(n, n + \#^\mp(n)] \circeq \nu^\mp(0, \#^\mp(n)]$, $\widetilde{S}^\pm_{l+j} = n + S^\mp_{m+j}$ for $0 \le j \le d$, in particular $\widetilde{S}^\pm_k - n =$

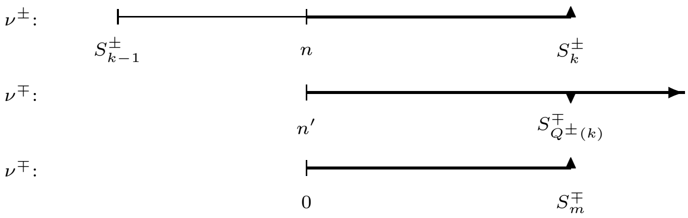


Fig. 4.5. Proof of (4.16), case (iii)

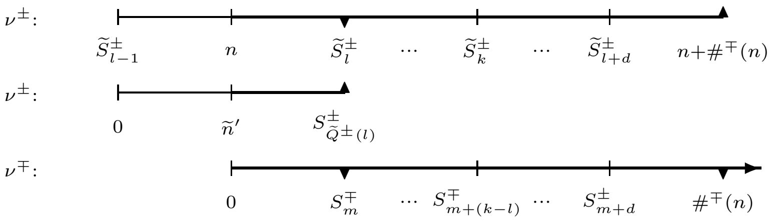


Fig. 4.6. Proof of (4.22)

$S^{\mp}_{m+(k-l)}$. This finishes the proofs of (4.16) and (4.22). The proofs of (4.17) and (4.23) are completely analogous.

(AC6)⇒(AC1). As in (AC4)⇒(AC6) one shows $\nu^{\pm}_{S^{\pm}_k} = \frac{0}{1}$ for all k. Now fix $n \in \mathbb{N}$ and show $\sigma^{n+1}\nu^+ \le \sigma\nu^-$: If $\#^-(n) = \infty$ then $\sigma^{n+1}\nu^+ = \sigma\nu^-$ so assume $\#^-(n) = S^-_k < \infty$. In this case the strings $\nu^+(n, n+\#^-(n)]$ and $\nu^-(0, S^-_k]$ differ only in the last digit where $\nu^-_{S^-_k} = 1$, which proves $\sigma^{n+1}\nu^+ < \nu^-$. The other three inequalities are treated similarly. This proves (AC1) except for $n = 0$. The inequalities $\nu^+ = 10* > 0* = \sigma\nu^+$ and $\nu^+ = 1\sigma\nu^+ \le 1\sigma^2\nu^- = \sigma\nu^-$ and their symmetric counterparts complete the proof.

(AC1)⇒(AC5). Since (AC1) implies (AC4), we immediately get (4.14). Let $n := S^+_k - S^+_{Q^-Q^+(k)}$. Applying (4.14) twice we see that the itineraries $\nu^+(n, S^+_k]$ and $\nu^+(0, S^+_{Q^-Q^+(k)}]$ coincide. Now Lemma 4.16(ii), applied to $\zeta := \sigma\nu^+$ and $\zeta' := \sigma^{n+1}\nu^+$, yields the Hofbauer condition (4.15) for Q^+, because $\zeta \le \zeta' = 0*$.

(AC5)⇒(AC4). Only (4.13) remains to be shown. The proof is almost identical to the proof of (AC3)⇒(AC4) with the following modifications:

Case (i). (AC3) was only used to guarantee that $S^{\pm}_k$ is not a co-cutting time, which proved (4.12). Here we turn this around and use (4.14) to prove that $S^{\pm}_k$ is not a co-cutting$^{\pm}$ time. Moreover,

$$S^{\pm}_k - \widetilde{S}^{\pm}_{l+d} = S^{\mp}_{Q^{\pm}(k)} - S^{\mp}_{Q^{\pm}(k-1)} = S^{\pm}_{Q^{\mp}Q^{\pm}(k)},$$

which implies the induction assumption (4.24) below for the next block.

Case (ii). Replace assumption (4.20) by

$$S^{\pm}_l - \widetilde{S}^{\pm}_{k-1} = S^{\pm}_{Q^{\mp}Q^{\pm}(l)}. \tag{4.24}$$

Then use (4.15) to show that $\widetilde{S}^{\pm}_k - \widetilde{S}^{\pm}_{k-1}$ satisfies (4.13) and $\widetilde{S}^{\pm}_k$ is not a cutting$^{\pm}$ time. ∎

4.20. Remark. The proof of the theorem sheds some light on the mysterious Hofbauer condition: Assume that $S_l^{\pm}$ is the minimal cutting$^{\pm}$ time between two consecutive co-cutting$^{\pm}$ times $\widetilde{S}_{k-1}^{\pm}$ and $\widetilde{S}_k^{\pm}$. Then $S_l^{\pm} - \widetilde{S}_{k-1}^{\pm} = S_{Q^{\mp}Q^{\pm}(l)}^{\pm}$ by (4.24) and it follows that $\widetilde{S}_k^{\pm} - \widetilde{S}_{k-1}^{\pm} = S_{Q^{\mp}(Q^{\pm}(l)+\beta(l))}^{\pm}$, where $\beta(l)$ denotes the smallest index where one has strict inequality in the Hofbauer condition, i.e., where

$$Q^{\pm}(l+\beta(l)) > Q^{\pm}(Q^{\mp}Q^{\pm}(l)+\beta(l))$$

(or $\beta(l) := \infty$ if there is no such index). It also shows that the Hofbauer condition $(Q^{\pm}(l+j))_{j\geq 1} \geq (Q^{\pm}(Q^{\mp}Q^{\pm}(l+j)))_{j\geq 1}$ only needs to be checked for those integers l for which there is a co-cutting$^{\pm}$ time between $S_{l-1}^{\pm}$ and $S_l^{\pm}$. $\diamond$

In the following we give some examples of kneading maps and apply our acquired knowledge in order to demonstrate that all cases in Theorem 3.44 are at least imaginable from a topological viewpoint. By Theorems 4.19 and 5.3 one can find for all of the kneading maps Q below a smooth Lorenz map whose Markov diagram is described by Q. We begin with a review of symmetric examples from a combinatorial viewpoint.

4.21. Example. If f is a symmetric Lorenz map then obviously its kneading invariant is symmetric, i.e., $\nu_n^+ + \nu_n^- = 1$ for all n. This implies that $S_k^+ = S_k^-$ and $\widetilde{S}_k^+ = \widetilde{S}_k^-$ for all k. In Example 3.52 it was shown how a symmetric Lorenz map f can be constructed from a unimodal map g. Let us see how the kneading invariants of f and g are related. Let $\mu = (\mu_n)_{n\geq 1}$ be the kneading invariant of g and let $(T_k)_{k\geq 0}$ and $(\widetilde{T}_k)_{k\geq 0}$ be the cutting and co-cutting times of μ (6). Let $\sigma_n := \sum_{k=1}^n \mu_k \bmod 1$, i.e., σ_n is the parity of ones in the block $\mu(0,n]$. From (3.15) it follows that $\mu_n = \nu_n^-$ if $\sigma_{n-1} = 0$, and $\mu_n = \nu_n^+$ if $\sigma_{n-1} = 1$. This can be used to show that $\sigma_n = \nu_n^-$ for all n. With this information it is a straightforward exercise to show by induction that $S_k^+ = S_k^- = T_k$ and $\widetilde{S}_k^+ = \widetilde{S}_k^- = \widetilde{T}_k$ for all $n \geq 0$. In other words, the construction of Example 3.52 respects cutting and co-cutting times. $\diamond$

4.22. Example (Feigenbaum map). For $Q^+(k) := Q^-(k) := \max(k-1,0)$ one obtains a map which is infinitely often properly $(2,2)$-renormalizable. The cutting times are $S_k^+ = S_k^- = 2^k$ and the kneading invariant begins as follows:

$$\nu^+ = 1001011100110100101101001100101100110100110010110100\ldots,$$

$$\nu^- = 0110100110010110100101100110100110010110011010010 11\ldots$$

Any Lorenz map with negative Schwarzian derivative and this kneading invariant has an attractor of type 2. $\diamond$

4.23. Example (Fibonacci map). For $Q^+(k) := Q^-(k) := \max(k-2,0)$ one obtains a map which is not renormalizable, because its Markov diagram is irreducible. The cutting times S_k^+ and S_k^- are precisely the Fibonacci numbers $1, 2, 3, 5, 8, 13, \ldots$ and it has the

(6) The cutting and co-cutting times of μ are determined by a splitting technique which is completely analogous to the one in Section 4.2 if one omits the "+" and "−" signs. The starting points of the splittings are $T_0 := 1$ and $\widetilde{T}_0 := \min\{n > 1 \mid n \text{ is not a cutting time}\}$, respectively. For more details, see Bruin [5].

following kneading invariant:

$$\nu^+ = 1000101101110011101000111010010001011101001000111000\ldots,$$
$$\nu^- = 0111010010001100010111000101101110100010110111001 11\ldots$$

The corresponding unimodal map has the same cutting times by Example 4.21. The interesting part about the unimodal Fibonacci map is that it is the optimal candidate for a nonrenormalizable map with a Cantor attractor, because the structure of its tower is as "close as possible" to the one of the Feigenbaum map. Indeed, it was shown in Bruin *et al.* [8] that if α is large enough then the map in the family $x \mapsto b - |x|^\alpha$ with the Fibonacci kneading invariant has a nonrenormalizable Cantor attractor. By Examples 3.52 and 4.21 the same statement holds for the Lorenz family (1.1). $\diamond$

Although the last example already provides examples of maps with attractors of type 3(b), we give another asymmetric example of a kneading invariant which is a good candidate for asymmetric maps with attractors of this type. From a combinatorial viewpoint it lies between the Feigenbaum and the Fibonacci Lorenz map.

4.24. Example. Let $Q^+(k) := \max(k-1, 0)$ and $Q^-(k) := \max(k-2, 0)$. Then one obtains an irreducible Markov diagram, which implies that $X_{\mathrm{m}} = (c_1^+, c_1^-)$. The kneading invariant starts as follows.

$$\nu^+ = 100101101110011101001011101001100101101110100110010\ldots,$$
$$\nu^- = 011101001100101110010110111010010110111001110100110\ldots$$

Since the cutting$^+$ and cutting$^-$ times both tend to infinity it follows that $c_1^+ \in \omega(c_1^-)$ and $c_1^- \in \omega(c_1^+)$. In order to show that $\omega(c_1^\pm) \subset X_{\mathrm{m}}$ we prove that the longest alternating sequence of zeros and ones (henceforth called an *alternating block*) occurring in $\nu^\pm$ is 1010 or 0101. This implies that there is a neighbourhood around the periodic orbit of period two ([7]) with itinerary $(01)^\infty \in \Sigma_\nu$, which is not visited by the iterates of $c_1^\pm$. We show this by induction. The main idea behind the proof is the fact that the block $\nu^\pm[S_{k-1}^\pm, S_k^\pm)$ is just an identical copy of a previous block $\nu^\mp[0, S_{Q^\pm(k)}^\mp)$ (cf. Remark 4.18). Hence $\nu^\pm[0, S_k^\pm)$ can only contain a longer alternating block than $\nu^\pm[0, S_{k-1}^\pm)$ and $\nu^\mp[0, S_{Q^\pm(k)}^\mp)$ together if that block starts before and ends after $S_{k-1}^\pm$. But this is impossible because for $k \geq 5$ all blocks $\nu^\pm[0, S_{k-1}^\pm)$ end with one of the words 1110, 1011 and 1001 and all blocks $\nu^\mp[0, S_{Q^\pm(k)}^\mp)$ start with 011, respectively 100. Since the longest alternating blocks contained in $\nu^+[0, S_4^+)$ and $\nu^-[0, S_4^-)$ are 1010 and 0101, the claim follows. $\diamond$

Now we are going to construct candidates for Lorenz maps with attractors of type 3(c) and 3(d). We do this by specifying a kneading invariant in such a way that every smooth Lorenz map with this kneading invariant is not renormalizable and has the following properties.

1. $\omega(c_1^+)$ and $\omega(c_1^-)$ are uncountable nowhere dense subsets of X_{m},

([7]) For simplicity we talk of a periodic orbit although it could also be a periodic homterval. However, if f has negative Schwarzian derivative then it is indeed a repelling periodic orbit.

2. $\omega(c_1^+) \supset \omega(c_1^-)$ and $c_1^- \in \omega(c_1^+)$,
3. $c_1^+ \in \omega(c_1^+)$ and $\mathrm{orb}(z) \notin \omega(c_1^+)$ (for type 3(c)), respectively
   $c_1^+ \notin \omega(c_1^+)$ and $\mathrm{orb}(z) \in \omega(c_1^+)$ (for type 3(d)),

where z is one of the points from the period two orbit. Note that the fact that $\omega(c_1^+)$ is nowhere dense in X_{m} follows from $\mathrm{orb}(z) \notin \omega(c_1^+)$ in the example for type 3(c), and from $c_1^+ \notin \omega(c_1^+)$ in the example for type 3(d).

The examples below will have $Q^+(1) = Q^+(2) = 0$, $Q^-(1) = 0$ and $Q^-(j) \in \{1, 2\}$ for $j > 1$, whence the Hofbauer condition for Q^- is satisfied and all cutting$^-$ and co-cutting$^-$ times can be determined without knowing the rest of Q^+. Fix a sequence $(Q^-(j))_{j\geq 2}$ from the subshift $\Omega := \{\omega \in \{1,2\}^{\mathbb{N}} \mid \omega_i = 1 \Rightarrow \omega_{i+1} = 2\}$ with the property that its orbit w.r.t. the shift is dense in Ω. Then $\omega(c_1^-)$ is uncountable, in particular c_1^- is not attracted to a periodic orbit and Corollary 3.37(2) can be applied. Obviously $S_k^- - S_{k-1}^- \leq 2$ for all k and it can easily be checked that also $\widetilde{S}_k^- - \widetilde{S}_{k-1}^- \leq 2$. It follows that $c_1^+ \notin \omega(c_1^-)$ and $c_1^- \notin \omega(c_1^-)$. Since $\nu^\pm[S_{k-1}^-, S_k^-) = 10$ if $Q^-(k) = 1$ and $\nu^\pm[S_{k-1}^-, S_k^-) = 100$ if $Q^-(k) = 2$ it follows that 01010 is the longest alternating block that occurs in ν^-, whence $\mathrm{orb}(c_1^-)$ does not accumulate at z.

For type 3(c) we do not specify the kneading map Q^+ but instead construct recursively a sequence $\widetilde{S}_{l_1}^+ < S_{k_1}^+ < \widetilde{S}_{l_2}^+ < S_{k_2}^+ < \dots$ of cutting$^+$ and co-cutting$^+$ times such that $S_{k_i-1}^+ < \widetilde{S}_{l_i}^+ < S_{k_i}^+$ and $\widetilde{S}_{l_{i+1}-1}^+ < S_{k_i}^+ < \widetilde{S}_{l_{i+1}}^+$ for all i and the differences $S_{k_i}^+ - S_{k_i-1}^+$ and $\widetilde{S}_{l_i}^+ - \widetilde{S}_{l_i-1}^+$ both tend to infinity as $i \to \infty$. One way to do this is the following: Let $\widetilde{S}_{l_1}^+ := 4$. If the sequence was defined up to $\widetilde{S}_{l_i}^+$ then let $S_{k_i-1}^+ := S^+\langle \widetilde{S}_{l_i}^+ \rangle$ and choose $Q^+(k_i) > Q^+(k_{i-1})$ large enough such that $S_{k_i}^+ > \widetilde{S}_{l_i}^+$. The choice of $S_{k_i}^+$ determines all co-cutting$^+$ times between $S_{k_i-1}^+$ and $S_{k_i}^+$. Number these co-cutting$^+$ times in increasing order and let $\widetilde{S}_{l_{i+1}-1}^+ := \widetilde{S}^+\langle S_{k_i}^+ \rangle$. Then choose $\widetilde{Q}^+(l_{i+1}) > \widetilde{Q}^+(l_i)$ large enough that there is a cutting$^+$ time between $\widetilde{S}_{l_{i+1}-1}^+$ and $\widetilde{S}_{l_{i+1}}^+$ and number the cutting$^+$ times in between.

Since Q^+ and $\widetilde{Q}^+$ are unbounded it follows that $c_1^- \in \omega(c_1^+)$ and $c_1^+ \in \omega(c_1^+)$. The proof that ν^+ does not have arbitrarily long alternating blocks is similar to the one in Example 4.24. It is done by induction, adding the overlapping blocks $\nu^+[S_{k_i-1}^+, S_{k_i}^+)$ and $\nu^+[\widetilde{S}_{l_i-1}^+, S_{l_i}^+)$ alternately. The first one is a copy of $\nu^-[0, S_{Q^+(k_i)}^-)$, where the length of alternating blocks is bounded, and the second one is a copy of $\nu^+[0, S_{\widetilde{Q}^+(l_i)}^+)$, where the length of alternating blocks can be assumed to be bounded by induction. Since the blocks start with 011 and 100, respectively, the length of alternating blocks cannot increase.

For type 3(d) we define the values $Q^+(k)$ recursively such that for certain integers $k_1 < k_2 < \dots$ the sequence $Q^+(k_i)$ tends to infinity, which assures $c_1^- \in \omega(c_1^+)$, and for the times in between Q^+ is chosen such that ν^+ has a long alternating block. Note that whenever $Q^+(k) \geq 1$, the Hofbauer condition for $Q^+(k-1)$ is satisfied since $Q^+(Q^-Q^+(k-1)+1) \leq 1$. Let $Q^+(3) = 1$ and fix a sequence $(k_i)_{i\in\mathbb{N}}$ such that $k_1 = 4$ and $k_{i+1} - k_i \to \infty$.

Now assume that $Q^+(k)$ has been defined for $k < k_i$ and assume by induction that $\widetilde{S}^+\langle S_{k_i-1}^+ \rangle = S_{k_i-1}^+ - 1$. This is true for k_1, since $S_3^+ = 5$ and $\widetilde{S}^+\langle S_3^+ \rangle = \widetilde{S}_0^+ = 4$.

The induction assumption implies that all integers $S^+_{k_i-1} + S^-_j$ are co-cutting times for $j < Q^+(k_i)$ (cf. Figure 4.1). In particular, the differences of all consecutive co-cutting$^+$ times between $S^+_{k_i-1}$ and $S^+_{k_i}$ is bounded by 3. By construction of Q^- there are arbitrarily large indices such that $Q^-(j) = 1$. Choose such an index j larger than $Q^+(k_{i-1})$ and let $Q^+(k_i) := j$. Then the last co-cutting$^+$ time $\widetilde{S}^+_{l_i} := \widetilde{S}^+\langle S^+_{k_i}\rangle$ before $S^+_{k_i}$ is $S^+_{k_i} - 2$. Next, let $Q^+(k_i + j) := 1$ for $j < k_{i+1} - k_j$. Then

$$\nu^+[\widetilde{S}^+_{l_i}, S^+_{k_{i+1}}) = \overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\dots\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1},$$

which shows that the differences of consecutive co-cutting$^+$ times in this block are bounded by 3 and that $S^+_{k_{i+1}-1}$ is a co-cutting$^+$ time, which yields the induction assumption for the next step. This finishes the construction of the kneading invariant. Now $c_1^+ \notin \omega(c_1^+)$ because the co-cutting$^+$ times are bounded, and $\mathrm{orb}(z) \in \omega(c_1^+)$ because there are arbitrarily long alternating blocks in ν^+.

4.25. Example. For the kneading map

$$\begin{aligned} Q^+ &= 0, 0, 1, 2, 1, 1, 4, 1, 1, 1, 6, 1, 1, 1, 1, 9, 1, 1, 1, 1, 1, \dots, \\ Q^- &= 0, 1, 2, 1, 2, 1, 2, 2, 1, 2, 1, 2, 2, 1, 2, 2, 2, 1, 2, 2, 1, \dots, \end{aligned}$$

one obtains the kneading invariant

$$\begin{aligned} \nu^+ &= 1\underset{\wedge}{0}\underset{\wedge}{0}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\overset{\vee}{1}0\overset{\vee}{1}00\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\overset{\vee}{1}0\overset{\vee}{1}00\overset{\vee}{1}0\overset{\vee}{1}00\overset{\vee}{1}0\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\underset{\wedge}{0}\overset{\vee}{1}\overset{\vee}{1}0\overset{\vee}{1}0\dots, \\ \nu^- &= 0\underset{\wedge}{1}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\underset{\wedge}{1}\overset{\vee}{0}\dots \end{aligned}$$

The cutting$^+$ times $S^+_{k_i} = S^+_4$, S^+_7, S^+_{11}, S^+_{16},… follow the large "gaps" in the splitting of ν^+. ◇

4.26. Remark. There is a lot of freedom of choice in the above construction, which makes it possible to prescribe a lot of additional properties for the omega limit sets of c_1^+ and c_1^- to hold. For example, it is possible to modify the construction in order to obtain examples of kneading invariants similar to the examples of Hofbauer & Keller [30, 31] for unimodal maps, which have the property that any Lorenz map with negative Schwarzian derivative and with such a kneading invariant has a strange SBR-measure like the point measure on the repelling period two orbit, or no SBR-measure at all. (The metric part of the arguments, i.e., the proof of [30, Propositions 1 and 2], carries over to the situation of Lorenz maps essentially without changes if one replaces the prerequisites from [39] and [37] used in the proof by the corresponding results from Chapter 3 and observes that the Weak Shadowing Principle 3.42 holds.) ◇

4.4. Renormalization from a combinatorial viewpoint. If a Lorenz map is renormalizable it is not only important to know how many iterates it takes to return to the renormalization interval on the left and right hand side, but also where the intermediate iterates of the left and right half are mapped in between. Since the intermediate iterates do not contain the critical point, one can describe their itineraries by finite words w_- and w_+. Consequently, the itinerary of any point in the renormalization interval, in particular the kneading invariant, can be written entirely in terms of those two words. This leads to the following combinatorial definition of renormalizability.

4.27. DEFINITION (Renormalizable). Let $(w_+, w_-) = (0*, 1*)$ be a pair of finite words such that the total length $|w_+| + |w_-|$ of the words is greater than or equal to three. A kneading invariant ν is called (w_+, w_-)-*renormalizable* if one can write it as

$$\nu^+ = w_+ w_-^{k_1} w_+^{k_2} \dots \quad \text{and} \quad \nu^- = w_- w_+^{l_1} w_-^{l_2} \dots$$

The renormalization is called *proper* if the itineraries w_-^∞ and w_+^∞ are ν-admissible, *minimal* if there is no other renormalization with shorter total length, and *trivial* if (w_+, w_-) equals $(10, 0)$ or $(1, 01)$. The pair (w_+, w_-) is called the *combinatorial type* of the renormalization.

A Lorenz map is called (w_+, w_-)-*renormalizable* if it is (m, n)-renormalizable and its kneading invariant is (w_+, w_-)-renormalizable, where $m = |w_+|$ and $n = |w_-|$.

4.28. REMARK. The last sentence needs a little explanation. It is possible that a Lorenz map is not (w_+, w_-)-renormalizable although its kneading invariant is (w_+, w_-)-renormalizable. This can happen in the situation when $\nu^- = w_- w_+^\infty$ and there is a periodic homterval with itinerary w_+^∞ (or in the symmetric situation). In this case it is possible that the right central branch of f^m has more than one fixed point and that c_n^- is mapped to the right of the innermost fixed point. Consequently, there is a fixed point in the dynamical interval (c_m^+, c_n^-) of the renormalized map, which was explicitly excluded in the definition of a Lorenz map. However, if $\nu^+ \neq w_+ w_-^\infty$ and $\nu^- \neq w_- w_+^\infty$ then the two notions are equivalent. $\diamond$

This combinatorial notion of renormalization was used by Glendinning & Sparrow [20] who studied renormalizations of kneading invariants for topologically expansive Lorenz maps. They showed among other things that if ν is a renormalizable admissible kneading invariant and if the minimal renormalization is not trivial then the sequences $(w_-)^\infty$ and $(w_+)^\infty$ are allowed and hence correspond to points on the interval. Stated in our language this means that every minimal renormalization is proper if it is nontrivial. This and more was already shown in Corollary 2.34.

It is an interesting combinatorial question to ask (i) which pairs (w_+, w_-) of binary words can appear as combinatorial type of a proper renormalization and (ii) how big the set of (w_+, w_-)-renormalizable kneading invariants is. It turns out that (i) these are exactly the nice pairs to be defined below and (ii) there is a one-to-one correspondence between admissible kneading invariants and (w_+, w_-)-renormalizable admissible kneading invariants given by the substitution $1 \to w_+$, $0 \to w_-$.

4.29. DEFINITION (Nice pair). A pair (ν^+, ν^-) of infinite binary strings is called *nice* if the orbits of ν^+ and ν^- w.r.t. the shift map never enter the interval (ν^-, ν^+), i.e., $\sigma^k \nu^\pm \notin (\nu^-, \nu^+)$ for all $k \in \mathbb{N}$. A pair (w_+, w_-) of finite aperiodic [8] binary strings is called *nice* if (w_+^∞, w_-^∞) is nice.

If (ν^+, ν^-) is a nice pair then $(\nu^+, \nu^-) = (10*, 01*)$, except for the case when $(\nu^+, \nu^-) = (1^\infty, 0^\infty)$. In the following we exclude this exceptional case.

[8] Aperiodic means that the length of $w_\pm$ coincides with the prime period of $w_\pm^\infty$.

4.30. PROPOSITION. *A pair* $(\nu^+,\nu^-)=(10*,01*)$ *of infinite binary strings is nice if and only if it is admissible. A pair* (w_+,w_-) *of finite aperiodic binary strings is nice if and only if* (w_+^∞,w_-^∞) *is admissible.*

PROOF. We only have to prove the first equivalence. Assume that (ν^+,ν^-) is not nice, i.e., there is an integer $k>0$ such that $\sigma^k\nu^\pm\in(\nu^-,\nu^+)$. If $\sigma^k\nu^\pm=0*$ then $\sigma\nu^-<\sigma^{k+1}\nu^\pm$, and if $\sigma^k\nu^\pm=1*$ then $\sigma^{k+1}\nu^\pm<\sigma\nu^+$. In both cases the admissibility condition is violated.

Conversely, let (ν^+,ν^-) be nice. We will show that $\sigma^k\nu^\pm\le\sigma\nu^-$ for all $k\ge 0$. The symmetric statement $\sigma^k\nu^\pm\ge\sigma\nu^+$ follows by analogy. Assume to the contrary that there exists some minimal integer $k\ge 0$ such that $\sigma\nu^-<\sigma^k\nu^\pm$. Since $\sigma\nu^-=1*$ it follows that $\nu^-\le\sigma\nu^-$, and because (ν^+,ν^-) is nice this in turn implies $\nu^+\le\sigma\nu^-$, whence $k>0$. If $\sigma^{k-1}\nu^\pm=1*$ then $\sigma^{k-1}\nu^\pm>\sigma^k\nu^\pm>\sigma\nu^-$, which contradicts the minimality of k. Hence $\sigma^{k-1}\nu^\pm=0*$ and $\nu^-<\sigma^{k-1}\nu^-=0*<\nu^+$. This is impossible because (ν^+,ν^-) is nice. ■

4.31. LEMMA. *Let* (w_+,w_-) *be a nice pair and let* $S:\Sigma\to\Sigma$ *be the map generated by the substitution* $1\mapsto w_+$, $0\mapsto w_-$. *Then the map* $(\nu^+,\nu^-)\mapsto(S\nu^+,S\nu^-)$ *is a bijection from the set of all admissible kneading invariants to the set of all* (w_+,w_-)*-renormalizable admissible kneading invariants.*

PROOF. Let $(\mu^+,\mu^-):=(S\nu^+,S\nu^-)$. By Proposition 4.30 it is equivalent to show that the pair (μ^+,μ^-) is nice if and only if (ν^+,ν^-) is nice.

(i) Assume (ν^+,ν^-) is not nice, i.e., $\sigma^l\nu^\pm\in(\nu^-,\nu^+)$ for some integer l. Let $m_\pm(l):=\sum_{i=0}^{l-1}\nu_i^\pm n_+ + (1-\nu_i^\pm)n_-$. Then

$$\sigma^{m_\pm}\mu^\pm=S(\sigma^l\nu^\pm)\in(S(\nu^-),S(\nu^+))=(\mu^-,\mu^+).$$

(ii) Conversely, assume that (μ^+,μ^-) is not nice, w.l.o.g. $\mu^-<\sigma^n\mu^\pm=0*$ for some integer n. If n equals $m_\pm(l)$ for some l then $\sigma^n\mu^\pm=S(\sigma^l\nu^\pm)$ and from

$$S(\nu^-)=\mu^-<\sigma^n\mu^\pm=S(\sigma^l\nu^\pm)$$

one concludes $\nu^-<\sigma^k\nu^\pm$, i.e., (ν^+,ν^-) is not nice. If n does not equal one of these numbers $m_\pm(l)$, then

$$\begin{aligned}\sigma^n\mu^\pm&=\sigma^k w_+*\quad\text{for some } 1<k<n_+ \text{ or}\\ \sigma^n\mu^\pm&=\sigma^k w_-*\quad\text{for some } 1<k<n_-,\end{aligned}$$

where the rest can be written entirely in terms of w_+ and w_-. In the first case, $\sigma^n\mu^\pm\le\sigma^k w_+^\infty\le w_-^\infty\le\mu^-$, since (w_+,w_-) is nice. This contradicts $\sigma^n\mu^\pm>\mu^-$.

In the second case, since $\sigma^n\mu^\pm>\sigma^k w_-^\infty$, we have $\sigma^n\mu^\pm=\sigma^k w_-^j w_+*$ for some j. Hence

$$\sigma^k w_-^\infty<w_-^\infty\le\mu^-<\sigma^n\mu^\pm=\sigma^k w_-^j w_+*\le\sigma^k w_-^j w_+^\infty.$$

Observe that the first strict inequality arises from the fact that w_- is aperiodic. Looking at the first $jn_- - k$ digits of the strict inequalities $\sigma^k w_-^\infty<w_-^\infty<\sigma^k w_-^j w_+^\infty$ we see that all three itineraries start with $\sigma^k w_-^j$. Cutting off this common initial part we obtain $w_-^\infty<\sigma^{n_- -k}w_-^\infty<w_+^\infty$, which contradicts the fact that (w_+,w_-) is nice. ■

4.5. Rotation numbers and rotation intervals. Every Lorenz map $f : [a,b] \to [a,b]$, where $[a,b]$ is the dynamical interval of f, can be considered in a natural way as a map from the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ into itself by rescaling the interval $[a,b]$ and gluing together the two endpoints a and b. For convenience we assume that $[a,b] = [0,1]$ and denote both the circle map and the interval map by f. The identification of the endpoints makes f continuous at the critical point c, whereas in general it is now discontinuous at the identification point. If $f(0) = f(1)$ then it is continuous and one obtains a *circle homeomorphism*. If $f(0) > f(1)$ then f is not surjective, i.e., there is a *gap* $J := (f(1), f(0)) = (c_2^-, c_2^+)$ in the image of f, and if $f(0) < f(1)$ then there is an *overlap* $J := (f(0), f(1)) = (c_2^+, c_2^-)$ of the images of the two branches of f . We call f briefly a *gap map* or an *overlap map*, accordingly. If f is not an overlap map, then we call it an *injective circle map*.

The combinatorial theory of circle maps has a long history which goes back to Poincaré, and it is a well established fact that there is a deep connection between the recurrence behaviour of the orbits of points on the circle and the continued fraction expansion of the rotation number of the map. For an expository introduction cf. de Melo & van Strien [13, Chapter I]. A detailed account on the symbolic dynamics of rotations can also be found in Gambaudo [15]. It is not our aim to reinvent this theory here but to demonstrate how it fits into the framework of the combinatorial theory of the Markov diagrams.

4.32. Definition (Rotation number). Let $F : \mathbb{R} \to \mathbb{R}$ be a lift of $f : \mathbb{S}^1 \to \mathbb{S}^1$ to the real line which is continuous on the interval $[0,1)$ and satisfies $F(x+1) = F(x)$ for all $x \in \mathbb{R}$. If f is an injective circle map then the limit

$$\varrho(f;\, x) := \lim_{n\to\infty} \varrho_n(F;\, x) := \lim_{n\to\infty} \frac{F^n(x) - x}{n} \bmod 1 = \lim_{n\to\infty} \frac{[F^n(x)]}{n} \bmod 1 \ (^9) \tag{4.25}$$

exists for all points $x \in \mathbb{R}$ and is independent of x and of the choice of the lift F. The number $\varrho(f) := \varrho(f;\, x)$ is called the *rotation number* of f.

The rotation number can be calculated in a purely combinatorial way from the itinerary $(\zeta_n(x))_{n\in\mathbb{N}}$ of a point x: Obviously, $[F^{n+1}(x)] = [F^n(x)] + \zeta_n(x)$, whence $\varrho(f)$ equals the asymptotic average of ones occurring in the itinerary of x, i.e.,

$$\varrho(f;\, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \zeta_k(x). \tag{4.26}$$

To reveal the dynamical information contained in the rotation number it is necessary to consider its continued fraction expansion.

4.33. Definition (Continued fraction expansion). Given countably many numbers $a_k \in \mathbb{N}_\infty \setminus \{0\}$ define

$$\frac{p_n}{q_n} := [1/a_1, a_2, \ldots, a_n] := \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

(^{9}) The notation $[y]$ is used for the integral part of $y \in \mathbb{R}$.

where $\frac{1}{\infty} := 0$, and p_n and q_n are coprime, and let

$$\alpha := \lim_{n\to\infty} \frac{p_n}{q_n} := [1/a_1, a_2, a_3, \ldots].$$

The representation $[1/a_1, a_2, a_3, \ldots]$ is called the *continued fraction expansion* of α. An expansion of the form $[1/a_1, a_2, \ldots, a_n, \infty, \ldots]$ is identified with the finite expansion $[1/a_1, a_2, \ldots, a_n]$.

Let $\gamma := \inf\{n \mid a_n = \infty\} - 1$. Then the following recursion holds for $n < \gamma$:

$$\begin{aligned} p_0 &= 0, & p_1 &= 1, & p_{n+1} &= a_{n+1}p_n + p_{n-1}, \\ q_0 &= 1, & q_1 &= a_1, & q_{n+1} &= a_{n+1}q_n + q_{n-1}. \end{aligned} \tag{4.27}$$

Every number $\alpha \in [0,1]$ has a continued fraction expansion, which is finite if and only if α is rational. The continued fraction expansion of an irrational number is unique whereas there is a little ambiguity in the continued fraction expansion of a rational number, namely $[1/a_1, \ldots, a_{n-1}, 1] = [1/a_1, \ldots, a_{n-1}+1]$ (e.g. $\frac{1}{1+1/1} = \frac{1}{2}$). For technical reasons it will be more convenient to use the first representation, so let us agree that the continued fraction expansion of a rational number ends with a one.

4.34. Proposition. *If f is a gap map and if D_2^+ is not critical, i.e., if $f(D_1^+) \subseteq D_1^-$, then it has a rotation number $\alpha = [1/1, a_2, a_3, \ldots] \geq 1/2$, and the tower has the following structure*:

1. *If $\alpha = [1/1, a_2, a_3, \ldots]$ is irrational, then for every $m \geq 2$ there are precisely a_m cutting times between q_{m-1} and q_m, including q_m , namely*

$$\begin{aligned} &q_{m-2} + 1 \cdot q_{m-1}, \\ &q_{m-2} + 2 \cdot q_{m-1}, \\ &\quad \cdots \\ &q_{m-2} + a_m q_{m-1} = q_m. \end{aligned}$$

They are cutting$^{\pm}$ times if m is odd/even, and the corresponding critical arrows are

$$D^{\pm}_{q_{m-2}+jq_{m-1}} \to D^{\mp}_{q_{m-1}+1}, \quad j = 1, \ldots, a_m.$$

In other words, the tower has alternately a_n critical levels on either side connected to the successor of the last critical level on the other side.

2. *If $\alpha = p_n/q_n = [1/1, a_2, a_3, \ldots, a_{n-1}, 1]$ is rational, then the critical levels with indices less than q_n are exactly the same as described in* (i). *q_n may be a cutting time or not*:

   (a) *If q_n is not a cutting time, then there are no cutting times $> q_n$.*
   (b) *If q_n is a cutting time—which is a cutting$^{\pm}$ time if n is even/odd—then there are no cutting$^{\pm}$ times but infinitely many cutting$^{\mp}$ times $> q_n$, namely $q_{n-1} + jq_n$, $j \geq 1$.*

If f is a gap map and if D_2^- is not critical, i.e., if $f(D_1^-) \subseteq D_1^+$, then it has a rotation number $\alpha = [1/1 + a_2, a_3, \ldots] \leq 1/2$, and the description in (i) *and* (ii) *holds if one interchanges + and − signs and replaces $[1/1, a_2, a_3, \ldots]$ by $\alpha = [1/1 + a_2, a_3, \ldots]$.*

Before going into the proof, we make a brief remark. If $f(D_1^+) \subseteq D_1^-$, then any zero in the itinerary of a point is followed by a one, and since α equals the asymptotic fraction of ones in the itinerary, one immediately obtains $\alpha \geq 1/2$. Similarly, in the case where $f(D_1^-) \subseteq D_1^+$ one obtains $\alpha \leq 1/2$. The latter case can be deduced from the former (and vice versa) by a symmetry argument, using the following property of continued fractions:

4.35. Lemma. *If $\alpha \in [0,1]$ has the continued fraction expansion $\alpha = [1/a_1, a_2, \ldots]$ then $\alpha \geq 1/2$ if $a_1 = 1$, and $\alpha \leq 1/2$ if $a_1 > 1$* [10]*. The continued fraction expansions for complementary angles $\alpha \leq 1/2 \leq \widetilde{\alpha}$, $\alpha + \widetilde{\alpha} = 1$, can be determined from each other using the following rule*:

$$[1/1, a_2, a_3, \ldots] + [1/1 + a_2, a_3, \ldots] = 1. \tag{4.28}$$

Proof. The first part follows from $1/(a_1+1) \leq \alpha \leq 1/a_1$. Now assume that $\alpha \geq 1/2$ and let $\alpha' := [1/a_2, a_3, \ldots]$ be such that $\alpha = 1/(1+\alpha')$. If $\alpha' = 0 = [1/\infty]$ then the equality above obviously holds, and if $\alpha' > 0$ then the equality $1/(1+\alpha') + 1/(1+1/\alpha') = 1$ implies $\widetilde{\alpha} = 1/(1+1/\alpha')$. ∎

Now if f is a gap map with rotation number α and τ an orientation reversing homeomorphism of the circle (e.g., $x \mapsto -x \bmod 1$), then the *mirror image* $\widetilde{f} = \tau^{-1} f \tau$ of f has the complementary rotation number $\widetilde{\alpha} = 1 - \alpha$.

The main part of the proof of Proposition 4.34 is the following lemma.

4.36. Lemma. *Let f be a gap map with rotation number $\alpha = [1/1, a_2, a_3, \ldots] \geq 1/2$. Then $f(D_1^+) \subseteq D_1^-$ and $a_2 = \sup\{k \mid D_{1+1}^+, \ldots, D_{1+k}^+ \subseteq D_1^-\}$. Moreover*:

1. *If $a_2 = \infty$, then all levels $\widehat{D}_{1+1}^+, \widehat{D}_{1+2}^+, \widehat{D}_{1+3}^+, \ldots$ are noncritical and all levels $\widehat{D}_{1+1}^-, \widehat{D}_{1+2}^-, \widehat{D}_{1+3}^-, \ldots$ are critical. The entire interval D_1^+ is attracted by a fixed point in the interval $(c, 1)$.*

2. *If $a_2 < \infty$, then the levels $\widehat{D}_{1+1}^+, \widehat{D}_{1+2}^+, \ldots, \widehat{D}_{1+a_2}^+$ are noncritical and the levels $\widehat{D}_{1+1}^-, \widehat{D}_{1+2}^-, \ldots, \widehat{D}_{1+(a_2-1)}^-$ are critical. The interval $D_{1+(a_2+1)}^+$ is the first one that intersects D_1^+ again, and $\widehat{D}_{1+(a_2+1)}^+$ is critical if and only if $\widehat{D}_{1+a_2}^-$ is.*
   - (a) *If the level $\widehat{D}_{1+a_2}^-$ is not critical then there are no critical levels above $\widehat{D}_{1+a_2}^-$ and $\widehat{D}_{1+(a_2+1)}^+$.*
   - (b) *If the level $\widehat{D}_{1+a_2}^-$ is critical then f is renormalizable on the renormalization interval $D_{1+a_2}^- = (c_1^+, c_{1+a_2}^-)$ and the renormalized map $\mathcal{R}f$ is the mirror image of a gap map $\widetilde{\mathcal{R}}f$ which maps its left half into its right half and has the rotation number $[1, a_3, a_4, \ldots]$.*

Proof. Let $a := \sup\{k \mid D_{1+1}^+, \ldots, D_{1+k}^+ \subseteq D_1^-\}$. We will show later on that $a = a_2$.

For $1 \leq m \leq a$, the level $D_{1+m}^+ = (c_{1+m}^+, c_m^-)$ is mapped monotonically to the level $D_{1+(m+1)}^+$ which lies to the left of D_{1+m}^+ and is separated from it by the gap (c_{1+m}^-, c_{1+m}^+). Since D_{1+m}^+ and $D_{1+(m-1)}^-$ share their right endpoint, it follows by induction that $D_{1+m}^- = (c_1^+, c_{1+m}^-)$ is critical for $m = 1, \ldots, a$.

If $a = \infty$, then the sequence D_{1+m}^+ of intervals converges monotonically to some fixed point in $(c, 1)$, so $\alpha = 1$ and $a = a_2$ and we are finished in this case.

[10] Recall that $1/2$ has two representations $1/2 = [1/2] = [1/1, 1]$ and that we use the second one.

Now assume $a < \infty$. From the above we see that $D^+_{1+1}, D^+_{1+2}, \ldots, D^+_{1+a}$ are noncritical and the levels $D^-_{1+1}, D^-_{1+2}, \ldots, D^-_{1+(a_2-1)}$ are critical. By definition, the interval $D^+_{1+(a+1)}$ is the first one that intersects D^+_1 again. Since $D^+_{1+(a+1)}$ and D^-_{1+a} share their right endpoint, D^-_{1+a} is critical iff $D^+_{1+(a+1)}$ is not contained in D^+_1 iff $D^+_{1+(a+1)}$ is critical.

(a) D^-_{1+a} is not critical: In this case D^+_1 is mapped monotonically by f^{a+1} into itself and f has a periodic orbit with itinerary 01^a. Hence the rotation number of f equals $a/(a+1) = [1, a]$ and $a = a_2$. Clearly, there is no cutting$^-$ time greater than $1 + a$. Additionally, since $D^+_{1+(a+1)} \subseteq D^+_1$, there is no cutting$^-$ time greater than $1 + (a + 1)$.

b) D^-_{1+a} is critical: It is easy to check that the first return map

$$\mathcal{R}f : x \mapsto \begin{cases} f^{1+a}(x) & \text{if } x \in (c^+_1, c^-), \\ f(x) & \text{if } x \in (c^+, c^-_{1+a}), \end{cases}$$

is a gap map on D^-_{1+a} which maps its right half into its left half. In fact, the gap of $\mathcal{R}f$ is precisely $(c^-_{2+a}, c^+_{2+a}) \subseteq D^+_1$. Let $\widetilde{f}$ be the mirror image of $\mathcal{R}f$ and denote its left and right halves by $\widetilde{D}^+_1$ and $\widetilde{D}^-_1$, respectively. By what we have shown, $\widetilde{f}$ has a unique rotation number $\beta = [1, b_2, b_3, \ldots] \geq 1/2$. Let us show that $\alpha = [1, a, b_2, b_3, \ldots]$, thus proving $a = a_2$ and the last statement of (b): Every point $\widetilde{x} \in \widetilde{D}^-_1$ corresponds to a point $x \in D^+_1$ and the itinerary of x can be obtained from the itinerary of $\widetilde{x}$ by the substitution $1 \to 01^a$, $0 \to 1$. Let $\widetilde{w}_n$ be the word consisting of the first n digits of the itinerary of $\widetilde{x}$ and w_n be the word obtained from $\widetilde{w}_n$ by substitution. Let z_n and o_n denote the cardinalities of zeros resp. ones of w_n, and let $\widetilde{z}_n$ and $\widetilde{o}_n$ denote the corresponding cardinalities for $\widetilde{w}_n$. Then

$$\frac{o_n}{o_n + z_n} = \frac{a\widetilde{o}_n + \widetilde{z}_n}{a\widetilde{o}_n + (\widetilde{o}_n + \widetilde{z}_n)} = 1 - \frac{\widetilde{o}_n}{a\widetilde{o}_n + (\widetilde{o}_n + \widetilde{z}_n)}$$

for all n. Taking limits on both sides we get

$$\alpha = 1 - \frac{\beta}{a\beta + 1} = 1 - \frac{1}{a + 1/\beta} = 1 - [a + 1, b_2, b_3, \ldots] = [1, a, b_2, b_3, \ldots]. \blacksquare$$

Proof of Proposition 4.34. The proof is carried out by induction using Lemma 4.36 and the following induction assumption for $2 \leq m \leq \gamma$:

> The map $\widetilde{\mathcal{R}}^{m-2}f$ is a gap map with rotation number $[1, a_m, a_{m+1}, \ldots]$. If n is even it coincides with the first return map $\mathcal{R}^{m-2}f$ to the interval $(c^+_{q_{m-1}}, c^-_{q_{m-2}})$:
>
> $$\mathcal{R}^{m-2}f : x \mapsto \begin{cases} f^{q_{m-2}}(x) & \text{if } x \in (c^+_{q_{m-1}}, c^-), \\ f^{q_{m-1}}(x) & \text{if } x \in (c^+, c^-_{q_{m-2}}), \end{cases}$$
>
> and if n is odd it is the mirror image of the first return map $\mathcal{R}^{m-2}f$ to the interval $(c^+_{q_{m-2}}, c^-_{q_{m-1}})$:
>
> $$\mathcal{R}^{m-2}f : x \mapsto \begin{cases} f^{q_{m-1}}(x) & \text{if } x \in (c^+_{q_{m-2}}, c^-), \\ f^{q_{m-2}}(x) & \text{if } x \in (c^+, c^-_{q_{m-1}}). \end{cases}$$

For $m = 2$ this assumption is trivially satisfied. Now assume that it holds for some even $m < \gamma$ and apply Lemma 4.36 to $\widetilde{\mathcal{R}}^{m-2}f = \mathcal{R}^{m-2}f$ instead of f. This can be easily done by replacing each occurrence of $D^{\pm}_{1+j}$ in the lemma by $D^{\pm}_{q_{m-2}+jq_{m-1}}$. It follows that the levels $D^{-}_{q_{m-2}+jq_{m-1}}$, $1 \le j \le a_m - 1$, are critical and the levels $D^{+}_{q_{m-2}+jq_{m-1}}$, $1 \le j \le a_m$, are noncritical. The induction stops if $a_m = \infty$ (case 1 of the lemma) or if $a_m < \infty$ and $D^{-}_{q_{m-2}+a_m q_{m-1}} = D^{-}_{q_m}$ is noncritical (case 2(a) of the lemma). In the remaining case $\widetilde{\mathcal{R}}^{m-1} f$ is defined and it is the mirror image of the first return map to $D^{-}_{q_{m-2}+a_m q_{m-1}} = D^{-}_{q_m} = (c^{-}_{q_m}, c^{+}_{q_{m-1}})$. The proof for m odd is similar.

To prove the case where the rotation number is less than $1/2$ one just has to replace f by its mirror image. ∎

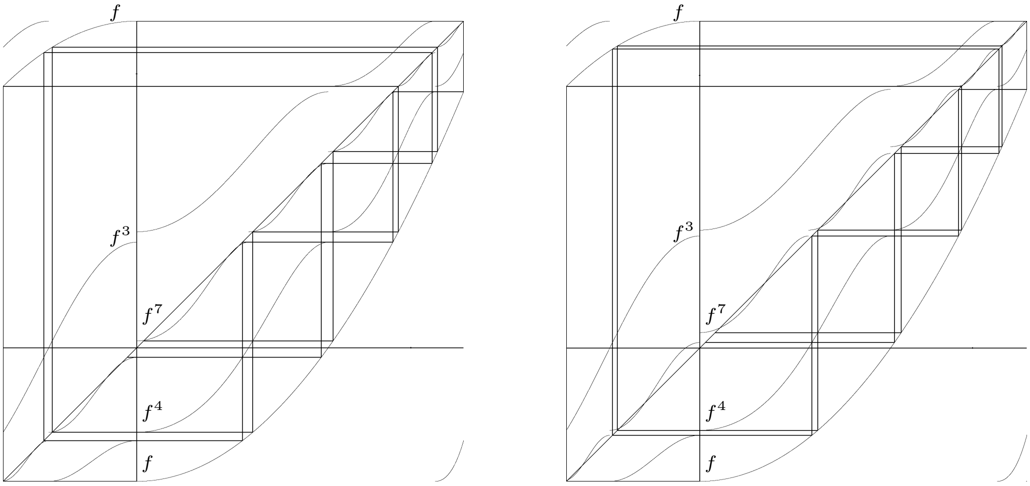


Fig. 4.7. Two gap maps with rotation number $5/7 = [1/1, 2, 1, 1]$

4.37. Example. The difference between cases 2(a) and 2(b) in Proposition 4.34 is illustrated in Figure 4.7. It shows two Lorenz maps with rotation number $5/7 = [1/1, 2, 1, 1]$ but with different itineraries, plotted over their dynamical interval. The kneading invariant on the l.h.s. equals (w_+^∞, w_-^∞) with $w_+ := 1011011$ and $w_- := 0111011$. It has the following splitting (cutting times are marked with a "$\wedge$" below).

$$\nu^+ = 1\underset{\wedge}{0}11\underset{\wedge}{0}11101101110110111011011\ldots,$$
$$\nu^- = 0\underset{\wedge}{1}\underset{\wedge}{1}\underset{\wedge}{1}011011101101110110111011\ldots$$

The kneading invariant on the r.h.s. equals $((1011011)^\infty, 0111011(1011011)^\infty)$, which splits as follows.

$$\nu^+ = 1\underset{\wedge}{0}11\underset{\wedge}{0}111011\underset{\wedge}{0}111011\underset{\wedge}{0}111011\underset{\wedge}{0}11\ldots,$$
$$\nu^- = 0\underset{\wedge}{1}\underset{\wedge}{1}\underset{\wedge}{1}011\underset{\wedge}{1}01101110110111011011\ldots$$

The approximating denominators for $5/7$ are $q_1 = 1$, $q_2 = 3$, $q_3 = 4$, and $q_4 = 7$. The corresponding iterates f^{q_i} are plotted in the same figure. ◇

For overlap maps the limit (4.25) may not exist, or depend on the point x. Instead of using rotation *numbers* one defines the rotation *set* $\varrho(f)$ of f as

$$\varrho(f) := \bigcup_{x \in [0,1]} \varrho(f;x), \quad \text{where} \quad \varrho(f;,x) := \omega(\varrho_n(f;x))_{n \in \mathbb{N}}$$

denotes the set of limit points of the sequence $(\varrho_n(f;\, x))_{n\in\mathbb{N}}$. It can be shown that $\varrho(f)$ is a compact interval ([11]). If $\varrho(f) = \{\alpha\}$ consists of a single point only then f is called *frequency locked* with rotation number α. If f is an overlap map then in general it has a nondegenerate rotation interval, but the following proposition shows that there are two cases where f is frequency locked. In the first case the iterates of the overlap interval (c_2^+, c_2^-) never hit the critical point and in the second case a proper renormalization interval $(\widetilde{p}, \widetilde{q})$ exists whose intermediate iterates are rotated on the circle according to a rational angle. The second case corresponds to the situation in Malkin [46, Theorem 3].

4.38. Proposition. *A Lorenz map f is frequency locked on its dynamical interval if and only if the decomposition of its Markov diagram* (cf. Theorem 2.22) *has one of the following properties.*

1. *The entire diagram consists of a single transient chain $\widehat{\mathcal{T}}_1$.*
2. *The first irreducible component $\widehat{\mathcal{X}}_1$ consists of a single loop only.*

In the second case the map has a rational rotation number $\alpha = r/s$ and the kneading invariant of f is properly (w_+, w_-)-renormalizable, where (w_+^∞, w_-^∞) is the kneading invariant of the (*Lorenz*) *map $x \mapsto x + \alpha \bmod 1$* ([12]).

Proof. Note that by Remark 2.26 the tower is entirely transient if and only if D_n^+ and D_n^- are never critical simultaneously. Since $D_2^+ \cap D_2^- = (c_2^+, c_2^-)$ it follows that $D_n^+ \cap D_n^- = (c_n^+, c_n^-)$ as long as $\{2, \ldots, n-1\} \cap \mathcal{S}^- \cap \mathcal{S}^+ = \emptyset$, whence the tower is entirely transient if and only if the overlap interval (c_2^+, c_2^-) is a homterval.

If (c_2^+, c_2^-) is a homterval then the *upper map* $F_u : x \mapsto \sup_{y \le x} F(y)$ and the *lower map* $F_l : x \mapsto \inf_{y \ge x} F(y)$ have the same rotation number, which shows that f is frequency locked. Moreover, the Markov diagram of f has precisely the same structure as described in Proposition 4.34. The proof is practically the same. The only difference is that in the proof of Lemma 4.36 there is no gap between the intervals $D^+_{1+(m+1)}$ and D^+_{1+m} any more but instead they overlap on the interval (c^+_{1+m}, c^-_{1+m}). However, this does not matter because the overlap interval never hits the critical point.

If (c_2^+, c_2^-) is not a homterval then let $s \ge 2$ be the minimal integer such that (c_s^+, c_s^-) contains the critical point. By Remark 2.26 one has $s = S_k^+ = S_l^- = S_{k-1}^+ + S_{l-1}^-$, where k and l are the minimal indices such that $Q^+(k) = l - 1$ and $Q^-(l) = k - 1$ (in other words, $\widehat{C}_k^+ = \widehat{D}_s^+$ and $\widehat{C}_l^- = \widehat{D}_s^-$ are the lowest critical levels contained in $\widehat{\mathcal{X}}_1$ on either side). Since $Q^-(Q^+(k) + 1) = k - 1 < k$ and $Q^-(Q^+(l) + 1) = l - 1 < l$, it follows as in the proof of Theorem 2.32 that both central branches of f^s have fixed points $\widetilde{p}$ and $\widetilde{q}$ with itineraries w_-^∞ and w_+^∞, respectively. Assume for simplicity that $\widetilde{p}$ and $\widetilde{q}$ are not contained in a periodic homterval. The general case is similar. Since $|w_+| = |w_-| = s$ and since the Markov diagram below the levels D_s^+ and D_s^- is entirely transient it follows that (w_+^∞, w_-^∞) is the kneading invariant of a rigid rotation $g : x \mapsto x + \alpha \bmod 1$ with

([11]) See e.g. Malkin [46], where rotation sets for topologically expansive Lorenz maps are studied.

([12]) It is assumed that r and s are coprime and that $|w_+| = |w_-| = s$.

rational angle $\alpha = r/s$, where α equals the fraction of ones in the words w_+ and w_-. In the following we distinguish two cases, namely $(c_s^+, c_s^-) \subseteq (\widetilde{p}, \widetilde{q})$ and $(c_s^+, c_s^-) \not\subseteq (\widetilde{p}, \widetilde{q})$.

If $(c_s^+, c_s^-) \subseteq (\widetilde{p}, \widetilde{q})$ then the map is properly (s, s)-renormalizable, which implies that $\widehat{\mathcal{D}}_{kl}^{\wedge}$ is invariant, whence $\widehat{\mathcal{X}}_1$ consists only of the closed loop $\widehat{C}_k^+ \rightsquigarrow \widehat{C}_l^- \rightsquigarrow \widehat{C}_k^+$. The combinatorial type of the renormalization is (w_+, w_-). A characteristic property of the kneading invariants of rational rotations is that they can be renormalized to the kneading invariant $((10)^\infty, (01)^\infty)$ by finitely many trivial renormalizations. From that it follows that $\sigma^t w_+^\infty = w_-^\infty$ for some $t < s$, i.e., $\widetilde{p}$ and $\widetilde{q}$ belong to the same periodic orbit.

In particular, the words w_+ and w_- contain the same number of ones, which we denote by r. This implies that every point $x \in [\widetilde{p}, \widetilde{q}]$ has rotation number r/s, because its itinerary consists entirely of the words w_+ and w_-. Since *every* point in the interval $[0, 1] = [c_1^+, c_1^-]$ enters the interval $[\widetilde{p}, \widetilde{q}]$ after finitely many iterations by Remark 2.36 it follows that $\varrho(f) = \{r/s\}$.

If $(c_s^+, c_s^-) \not\subseteq (\widetilde{p}, \widetilde{q})$ then either $f^s(\widetilde{p}, c^-) \supset (\widetilde{p}, \widetilde{q})$ or $f^s(c^+, \widetilde{q}) \supset (\widetilde{p}, \widetilde{q})$. Assume that we are in the former situation. It follows that there exists a small one-sided neighbourhood $(\widetilde{q}, \widetilde{q} + \varepsilon) \subset (\widetilde{q}, c_s^-)$ of $\widetilde{q}$ which is mapped by f^t homeomorphically onto a one-sided neighbourhood $(\widetilde{p}, \widetilde{p} + \delta)$ of $\widetilde{p}$ which can be chosen such that $f^{ms}(\widetilde{p}, \widetilde{p} + \delta) = (\widetilde{p}, c_s^-)$ for some m. It follows that $(\widetilde{q}, \widetilde{q} + \delta)$ is mapped onto itself homeomorphically by f^{ms+t}, whence there is a periodic orbit whose period divides $ms+t$. This implies that the rotation set does not consist of a single point. ■

4.39. Remark. In case 1 the overlap is a wandering interval if and only if the rotation number is irrational. More on wandering intervals can be found in Section 3.4. From the results in that section it follows that there are no $\mathcal{C}^2$-Lorenz maps with overlapping branches which are frequency locked with irrational rotation number. ◇

5. Families of Lorenz maps

The aim of this chapter is to study parametrized families of Lorenz maps and to understand how the dynamics in such a family depends on the parameters. The families we consider are assumed to have the following properties.

5.1. Definition ($\mathcal{C}^1$-Lorenz family). Let Λ be an open simply connected subset of $\mathbb{R}^2$. A family $\mathcal{F} := (f_{a,b})_{(a,b)\in\Lambda}$ of maps $f_{a,b} : \mathbb{R} \to \mathbb{R}$ is called a *$\mathcal{C}^1$-Lorenz family* if the following properties hold.

(L1) The map $(a, b, x) \mapsto f_{a,b}(x)$ is continuously differentiable for $x \neq c := 0$ and the one-sided limits $\mathrm{D}_a f_{a,b}(0^\pm) = \mathrm{D}_a c_1^\pm(a, b)$ and $\mathrm{D}_b f_{a,b}(0^\pm) = \mathrm{D}_b c_1^\pm(a, b)$ exist for all parameters $(a, b) \in \Lambda$.

(L2) $\mathrm{D}_x f_{a,b}(x) > 0$ and $\mathrm{D}_x f_{a,b}(x) \to 0$ as $x \to 0$.

(L3) The map $f_{a,b}$ has precisely two fixed points $p^-(a, b) < 0$ and $p^+(a, b) > 0$ which are hyperbolic repellers.

(L4) There are no neutral or repelling fixed points in the interval $(p^-(a, b), p^+(a, b))$.

For the study of parameter dependence we will make some additional assumptions for convenience, namely that $\mathbb{R}^2_+ \subseteq \Lambda$ and that the following properties hold.

(L5) $c_1^+(0, b) = 0$ and $c_1^-(a, 0) = 0$ for all $a, b \geq 0$.

(L6) $\mathrm{D}_a c_1^+(a, b) < 0$ and $\mathrm{D}_b c_1^-(a, b) > 0$.

(L7) $\mathrm{D}_a f_{a,b}(x) \leq 0$ and $\mathrm{D}_b f_{a,b}(x) \geq 0$ for $x \in (p^-(a, b), p^+(a, b))$.

(L8) There exists a parameter such that $f_{a,b}$ has full branches, i.e., $c_1^+(a, b) = p^-(a, b)$ and $c_1^-(a, b) = p^+(a, b)$.

Prototypes of such $\mathcal{C}^1$-Lorenz families are the families

$$f_{a,b} : x \mapsto \begin{cases} -a + |x|^\alpha & \text{if } x > 0, \\ b - |x|^\alpha & \text{if } x < 0, \end{cases} \tag{5.1}$$

with a fixed constant $\alpha > 1$, in particular the *quadratic Lorenz family*:

$$f_{a,b} : x \mapsto \begin{cases} -a + x^2 & \text{if } x > 0, \\ b - x^2 & \text{if } x < 0. \end{cases} \tag{5.2}$$

To satisfy all conditions (L1)–(L8) choose $\Lambda = (-\varepsilon, \infty) \times (-\varepsilon, \infty)$ for a small $\varepsilon > 0$.

5.2. Remark. The third assumption implies that the location of the fixed points $p^-(a, b)$ and $p^+(a, b)$ depends continuously on the parameters. On the contrary, nonhyperbolic fixed points can bifurcate into arbitrarily many fixed points or even disappear. The assumptions (L3) and (L4) were made in order to avoid such silly problems. ◇

Not all maps from the $\mathcal{C}^1$-Lorenz family $\mathcal{F}$ are really Lorenz maps in the sense of Definition 2.1, since we neither assume that $c_1^+(a,b) < 0 < c_1^-(a,b)$ nor that the interval $[p^-(a,b), p^+(a,b)]$ is invariant under $f_{a,b}$. However, we will show in Lemma 5.21 that under the assumptions (L1)–(L8) of Definition 5.1 the following holds:

(L9) There is an open subset J of Λ such that

$$p^-(a,b) < c^+(a,b) < 0 < c^-(a,b) < p^+(a,b) \quad \text{for all } (a,b) \in J$$

and the closure of J is the homeomorphic image of the unit square $[0,1] \times [0,1]$ in such a way that the left, right, bottom, and top edge of the square are mapped into the sets $\{c^+ = 0\}$, $\{c^+ = p^-\}$, $\{c^- = 0\}$, and $\{c^- = p^+\}$, respectively.

These four arcs constitute the edges of the "square" J.

One of the main goals is to answer the question whether $\mathcal{C}^1$-families are full families, i.e., whether every admissible kneading invariant actually occurs in this family. We prove the following theorem.

5.3. Theorem. *Let $(f_{a,b})_{(a,b)\in\Lambda}$ be a $\mathcal{C}^1$-Lorenz family and let $\nu = (\nu^+, \nu^-)$ be an admissible kneading invariant such that either*

(i) *ν is expansive, or*

(ii) *at least one of ν^+ and ν^- is periodic.*

Then there exists a parameter $(a,b) \in \operatorname{cl} J$ where $f_{a,b}$ has the kneading invariant ν.

As already mentioned in the introduction, we will not treat the question of full families in isolation, but rather as part of a detailed analysis of the parameter dependence of the kneading invariant. Nevertheless, we begin in Section 5.1 with a presentation of the classical approach to the proof of the Full Family Theorem using the Thurston map, which is well known for continuous multimodal maps, and show that it can easily be applied without major modifications to discontinuous maps like the Lorenz maps. Here we follow essentially the arguments of de Melo & van Strien [13] with some modifications due to Martens & de Melo [47].

Although Section 5.1 is a little bit independent of the others, there are two reasons why we included it. First of all, for its interesting practical aspect, since it provides an efficient algorithm to find Lorenz maps with specific combinatorial properties, and second, because the Thurston map will be needed for the discussion of monotonicity of the kneading invariant in the quadratic family in Section 5.7.

5.1. The Thurston algorithm. In this section we study the question whether every admissible invariant ν can be realized within certain smooth families of Lorenz maps for *finite*, i.e., preperiodic, admissible kneading invariants. It is even possible to find maps where the critical orbits $\operatorname{orb}(c^+)$ and $\operatorname{orb}(c^-)$ are also finite. Such maps are called *post-critically finite*.

5.4. THEOREM. *Let $(f_{a,b})_{(a,b)\in\Lambda}$ be a $\mathcal{C}^1$-Lorenz family. Then for any finite admissible kneading invariant ν which is periodic or expansive* (1) *there is a parameter $(a,b)\in\Delta$ where $f_{a,b}$ is post-critically finite and has the kneading invariant ν.*

Strictly speaking, we do not need to assume (L1)–(L8) in order to prove the theorem, but only (L1)–(L4) and the following condition.

(L9′) There exist constants $P<0<Q$ and a subset Δ of Λ which is homeomorphic to $V:=[P,0]\times[0,Q]$ such that $P\le p^-(a,b)<p^+(a,b)\le Q$ for all $(a,b)\in\Delta$ and the map
$$F:\Delta\to V,\quad (a,b)\mapsto (c_1^+(a,b),c_1^-(a,b)),$$
is continuous, $F(\operatorname{int}\Delta)\subseteq\operatorname{int}V$, $F(\operatorname{bd}\Delta)\subseteq\operatorname{bd}V$ and $\deg F|_{\operatorname{bd}\Delta}\neq 0$.

This condition is satisfied by every $\mathcal{C}^1$-Lorenz family after a coordinate change: As already mentioned, condition (L9) holds for every $\mathcal{C}^1$-Lorenz family (for a proof cf. Lemma 5.21). Now let J be the region in parameter space defined in (L9) and let $\widetilde{f}_{a,b}$ be the Lorenz map obtained from $f_{a,b}$ by conjugating it with a Möbius transformation mapping $p^\pm(a,b)$ to ±1 and 0 to 0. Then the normalized family $(\widetilde{f}_{a,b})_{(a,b)\in\operatorname{cl}J}$ satisfies condition (L9′) with $\Delta=\operatorname{cl}J$ and $V=[-1,0]\times[0,1]$, since $\deg\widetilde{F}|_{\operatorname{bd}J}=-1$, where $\widetilde{F}$ denotes the analogue of F for the family $(\widetilde{f}_{a,b})_{(a,b)\in\Lambda}$.

5.5. REMARK. For the prototype families (5.1) there is a simpler way to verify assumption (L9′): Just choose $P=-2$, $Q=2$, and $\Delta=[0,2]\times[0,2]$. ◇

Let us briefly introduce the notion of the degree of a map which appears in condition (L9′) and state the Brouwer Fixed Point Theorem—or rather a slight generalization of it—which will be needed in the following. In our notation we closely follow Massey [49, Chapters VI–VIII].

Denote by $\mathbb{B}^n$ the n-dimensional unit ball and by $\mathbb{S}^{n-1}$ its bounding sphere.

5.6. DEFINITION (Degree of a map). Let $f:(\mathbb{B}^n,\mathbb{S}^{n-1})\to(\mathbb{B}^n,\mathbb{S}^{n-1})$ be a *map of pairs*, i.e., a continuous map $f:\mathbb{B}^n\to\mathbb{B}^n$ such that $f(\mathbb{S}^{n-1})\subseteq\mathbb{S}^{n-1}$, and let $g:\mathbb{S}^{n-1}\to\mathbb{S}^{n-1}$ be a continuous map.

Then the homomorphism $f_*:H_n(\mathbb{B}^n,\mathbb{S}^{n-1})\to H_n(\mathbb{B}^n,\mathbb{S}^{n-1})$ induced by f on the relative homology group $H_n(\mathbb{B}^n,\mathbb{S}^{n-1})=\mathbb{Z}$ is of the form $z\mapsto(\deg f)\cdot z$ for some integer $\deg f$, which is called the *degree* of f.

Similarly, the homomorphism $g_*:H_{n-1}(\mathbb{S}^{n-1})\to H_{n-1}(\mathbb{S}^{n-1})$ induced by g on the (reduced) homology group $H_{n-1}(\mathbb{S}^{n-1})=\mathbb{Z}$ is of the form $z\mapsto(\deg g)\cdot z$ for some integer $\deg g$, called the *degree* of g.

5.7. PROPOSITION. $\deg f=\deg f|_{\mathbb{S}^{n-1}}$ *for every map* $f:(\mathbb{B}^n,\mathbb{S}^{n-1})\to(\mathbb{B}^n,\mathbb{S}^{n-1})$.

PROOF. See Massey [49]. ■

5.8. THEOREM (Brouwer). *Let $g:(\mathbb{B}^n,\mathbb{S}^{n-1})\to(\mathbb{B}^n,\mathbb{S}^{n-1})$ be a map of pairs with nonzero degree and let $f:\mathbb{B}^n\to\mathbb{B}^n$ be an arbitrary continuous map. Then there exists a point $x\in\mathbb{B}^n$ such that $f(x)=g(x)$.*

(1) Recall that a kneading invariant is called expansive if it satisfies condition (EAC) on page 72.

PROOF. Assume to the contrary that $f(x) \neq g(x)$ for all $x \in \mathbb{B}^n$. Then the ray starting at $f(x)$ with direction $g(x) - f(x)$ intersects $\mathbb{S}^{n-1}$ at a unique point $v(x)$ which depends continuously on x. Since $g(\mathbb{S}^{n-1}) \subseteq \mathbb{S}^{n-1}$ one has $v(x) = g(x)$ for x in $\mathbb{S}^{n-1}$. Moving every point $g(x)$ with constant speed along a straight line towards $v(x)$ one obtains a homotopy $g_t = (1-t)g + tv : (\mathbb{B}^n, \mathbb{S}^{n-1}) \to (\mathbb{B}^n, \mathbb{S}^{n-1})$ from $g_0 = g$ to $g_1 = v$ which maps $\mathbb{B}^n$ into $\mathbb{S}^{n-1}$. Consequently, the induced homomorphism $g_* = v_* : H_n(\mathbb{B}^n, \mathbb{S}^{n-1}) \to H_n(\mathbb{B}^n, \mathbb{S}^{n-1})$ is trivial, whence $\deg g = 0$, a contradiction. ■

To obtain the classical Brouwer Fixed Point Theorem choose $F := \mathrm{id} : \mathbb{B}^n \to \mathbb{B}^n$.

Proof of Theorem 5.4. The proof divides into three steps. In the first step the combinatorial information contained in the kneading invariant is translated into a form which is more suitable to be used as "input" for the definition of the so-called Thurston map. In step two we define the Thurston map and prove the theorem in the case where F is a homeomorphism, and in step three we do the same for general F. The last two parts of the proof are only sketched, since the arguments are essentially the same as the ones in the corresponding proof for multimodal maps by de Melo & van Strien [13, Chapter II, Theorem 4.1] (with minor modifications by Martens & de Melo [47]).

(i) Assume w.l.o.g. that the family $\mathcal{F}$ satisfies (L9$'$) with $P = -1$ and $Q = 1$. Using the algorithm introduced in the proof of Theorem 4.12, one constructs a post-critically finite Lorenz map f on $[-1, 1]$ with the required kneading invariant ν. Denote the points of the post-critical set by

$$\mathrm{orb}(c^+) \cup \mathrm{orb}(c^-) = \{-1 \leq z_1 < z_2 < \ldots < z_l = 0 = z_{l+1} < \ldots < z_{l+r} \leq 1\}$$

and let $\mathcal{I} := \{1, \ldots, l+r\}$. The action of f on the post-critical set is determined by a map $\pi : \mathcal{I} \to \mathcal{I}$ via $f(z_i) = z_{\pi(i)}$. The requirement that the kneading invariant is periodic or expansive implies that the *post-critical map* π has the following additional property:

(5.3) For every $i \in \mathcal{I} \setminus \{l+r\}$ there is some $n \geq 0$ such that $\pi^n(i) \leq l \leq \pi^n(i+1)$.

Indeed, since $f^n[z_i, z_{i+1}] = [z_{\pi^n(i)}, z_{\pi^n(i+1)}]$ as long as $0 \notin f^j(z_i, z_{i+1})$ for $j < n$, the above property simply means that none of the intervals $[z_i, z_{i+1}]$ is an inessential preperiodic homterval (cf. Remark 4.7).

(ii) To introduce the Thurston map, first assume that F is a homeomorphism from Δ to V. Although it would not involve much additional effort to treat the general case right away, we refrain from doing it, because we do not want the technical details of the proof to obscure the striking simplicity of the algorithm. The modifications for the general case will be treated in step (iii). In the case where F is a homeomorphism the Thurston map T can be defined as a continuous map on the $(l+r-2)$-dimensional simplex

(5.4) $W := \{(x_1, \ldots, x_{l+r}) \mid -1 \leq x_1 \leq \ldots \leq x_l = 0 = x_{l+1} \leq \ldots \leq x_{l+r} \leq 1\}.$

The map T assigns to every vector $(x_1, \ldots, x_{l+r}) \in W$ a new vector $(y_1, \ldots, y_{l+r}) \in W$ which is obtained by the following procedure:

1. Choose $(a, b) \in \Delta$ such that $f_{a,b}(0^+) = x_1$ and $f_{a,b}(0^-) = x_{l+r}$. Since F has a continuous inverse, there is a unique such parameter which depends continuously on (x_1, x_{l+r}).

2. Pull back the points $x_1, \dots, x_{l+r}$ by taking preimages along the appropriate branches as prescribed by the map π, i.e., choose $y_1, \dots, y_{l+r}$ such that

$$y_1 \le \dots \le y_l = 0 = y_{l+1} \le \dots \le y_{l+r} \quad \text{and} \quad f_{a,b}(y_i) = x_{\pi(i)} \tag{5.5}$$

(cf. Figure 5.1). This is possible, because by the first step both branches of f are long enough. The condition $-1 \le p^-(a,b) < p^+(a,b) \le 1$ guarantees that $|y_i| \le 1$ for all i, whence $T(x_1, \dots, x_l) := (y_1, \dots, y_l)$ is again contained in W.

The continuous map $T : W \to W$ obtained in this way is called the *Thurston map* associated with π. For an arbitrary constellation of points $(y_1, \dots, y_l)$ the map T has no dynamical meaning. Yet, by construction it has the property that if $(z_1, \dots, z_l)$ is a fixed point of the Thurston map then $(a,b) := F^{-1}(z_1, z_l)$ is the desired parameter, since $f_{a,b}(z_i) = z_{\pi(i)}$ for all i by (5.5).

Since W is a simplex and T is continuous, the Brouwer Fixed Point Theorem guarantees the existence of a fixed point for T in the closure of W. However, the fixed point may lie in the boundary of W and hence be of no use, since this means that some of the points x_i are identical. The solution to this problem lies at hand if one takes a closer look at how the proof of the Brouwer Theorem 5.8 works: Let $\widetilde{T} := h \circ T \circ h^{-1}$ where $h : W \to \mathbb{B} := \mathbb{B}^{l+r-2}$ is a lippeomorphism ([2]). If T has no fixed points in W then $\widetilde{T}$ does not have any in $\mathbb{B}$ and every point $\widetilde{x} = h(x) \in \mathbb{B}$ can be moved along the ray starting at $\widetilde{x}$ and pointing in the direction $-(\widetilde{T}(\widetilde{x}) - \widetilde{x})$ towards the intersection point $\widetilde{v}(\widetilde{x})$ of the ray with the boundary $\operatorname{bd}\mathbb{B}$. In this way one obtains a deformation retraction of the ball onto its boundary, which is impossible. The idea how to fix the proof is to show that near the boundary the vector $\widetilde{T}(\widetilde{x}) - \widetilde{x}$ points "inwards". More precisely, one proves that

$$\frac{\operatorname{dist}(T(x), x)}{\operatorname{dist}(x, \operatorname{bd} W)} \to \infty \quad \text{as } x \to \operatorname{bd} W. \tag{5.6}$$

It is here where condition (5.3) enters: If one moves x towards the boundary of W then one of the intervals (x_k, x_{k+1}) must also collapse. Assuming to the contrary that the ratio (5.6) is bounded and using (5.3) one can show that one of the neighbourhoods (x_{l-1}, x_l) and (x_{l+1}, x_{l+2}) of the critical point must also collapse. But the smaller these neighbourhoods are, the more they are expanded in the pullback, because the derivative at the critical point vanishes. This effect causes $\operatorname{dist}(Tx, \operatorname{bd} W)/\operatorname{dist}(x, \operatorname{bd} W)$ to diverge as $x \to \operatorname{bd} W$, whence the ratio (5.6) cannot be bounded, a contradiction. For the details of the proof of (5.6) the reader is referred to [13, Chap. II, Lemma 4.1].

Since h is a lippeomorphism, it follows that

$$\frac{\operatorname{dist}(\widetilde{T}(\widetilde{x}), \widetilde{x})}{\operatorname{dist}(\widetilde{x}, \operatorname{bd} \mathbb{B})} \to \infty \quad \text{as } \widetilde{x} \to \operatorname{bd} \mathbb{B}. \tag{5.7}$$

Now if T has no fixed point inside W then $\widetilde{v}(\widetilde{x})$ can be defined as before for $\widetilde{x} \in \operatorname{int}\mathbb{B}$ and condition (5.7) ensures that $\widetilde{v}$ extends continuously to a map of $\mathbb{B}$ which fixes the boundary points. As above, one obtains a deformation retraction of $\mathbb{B}$ onto its boundary.

(iii) In the case where F is not a homeomorphism the simplex W defined in (5.4) is not suitable as domain for the Thurston map, because for a given point $x \in W$ there is

([2]) Shorthand for: Lipschitz-continuous homeomorphism with Lipschitz-continuous inverse.

in general more than one choice for the pullback map $f_{a,b}$, whence T is not single-valued on W. We will fix this problem more or less by replacing the coordinates x_1 and x_{l+r} by the coordinates a and b.

Let $\widehat{\Delta} := \Delta \times W'$ where $W' := \{x \in W \mid x_1 = -1,\ x_{l+r} = 1\}$ and let $\phi_{u,v} : [-1,1] \to \mathbb{R}$ be the Möbius transformation ([3]) satisfying $\phi_{u,v}(-1) = u$, $\phi_{u,v}(0) = 0$ and $\phi_{u,v}(1) = v$. To every parameter $(a,b,x') \in \widehat{\Delta}$ corresponds a point $x = \widehat{F}(a,b,x') \in W$ given by

$$\widehat{F} : \widehat{\Delta} \to W, \quad (a,b,x') \mapsto \phi_{F(a,b)}(x').$$

Intuitively, the information contained in the coordinates $(a,b,x') \in \widehat{\Delta}$ consists of two parts, namely (i) the parameter (a,b) of the pullback map, which uniquely determines the critical values x_1 and x_{l+r}, and (ii) the relative coordinates $x'_2, \dots, x'_{l+r-1}$ of the remaining points inside the interval $(c_1^+(a,b), c_1^-(a,b))$. Now the Thurston map can be well defined as a map $\widehat{T} : \widehat{\Delta} \to W$ in the following way: For every $(a,b,x') \in \widehat{\Delta}$ let $T(a,b,x') := y$ be the pullback of $x = \widehat{F}(a,b,x')$ with respect to $f_{a,b}$, i.e., choose y such that (5.5) holds. Instead of looking for a fixed point of the Thurston map one now seeks a parameter $(a,b,x') \in \widehat{\Delta}$ such that $\widehat{T}(a,b,x') = \widehat{F}(a,b,x')$.

Because $\widehat{F}$ maps $\operatorname{bd}\widehat{\Delta}$ into $\operatorname{bd}\widehat{W}$ and both $\widehat{\Delta}$ and W are homeomorphic to $\mathbb{B}^{l+r-2}$, the degree of the map $\widehat{F}|_{\operatorname{bd}\widehat{\Delta}}$ is defined. Even more, $\widehat{F}$ maps every leaf of the continuous foliation $((a,b) \times W')_{(a,b)\in \operatorname{int}\Delta}$ of $\widehat{\Delta}$ homeomorphically onto the leaf $W_{F(a,b)}$ of the continuous foliation $(W_{u,v})_{(u,v)\in \operatorname{int} V}$ of W, where $W_{u,v} := \{x \in W \mid x_1 = u,\ x_{l+r} = v\}$, which implies that $\deg \widehat{F}|_{\operatorname{bd}\widehat{\Delta}} = \deg F|_{\operatorname{bd}\Delta}$.

Now Theorem 5.8 can be applied to the maps $\widehat{F}$ and $\widehat{T}$ to obtain a solution for the equation $\widehat{F}(a,b,x') = \widehat{T}(a,b,x')$ in $\widehat{\Delta}$. As in step (ii) one can modify the argument by showing that $\widehat{T}(a,b,x') - \widehat{F}(a,b,x')$ points "inwards" when $\widehat{F}(a,b,x')$ is close to the boundary of W in order to obtain a solution where $\widehat{F}(a,b,x') \in \operatorname{int} W$. ∎

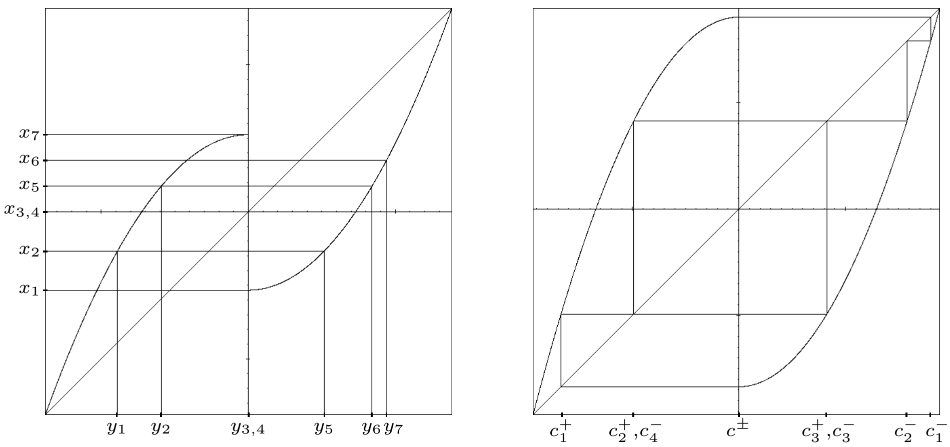


Fig. 5.1. The Thurston map for Example 5.9 and its fixed point

([3]) Any continuous family of homeomorphisms interpolating these three values will serve the purpose.

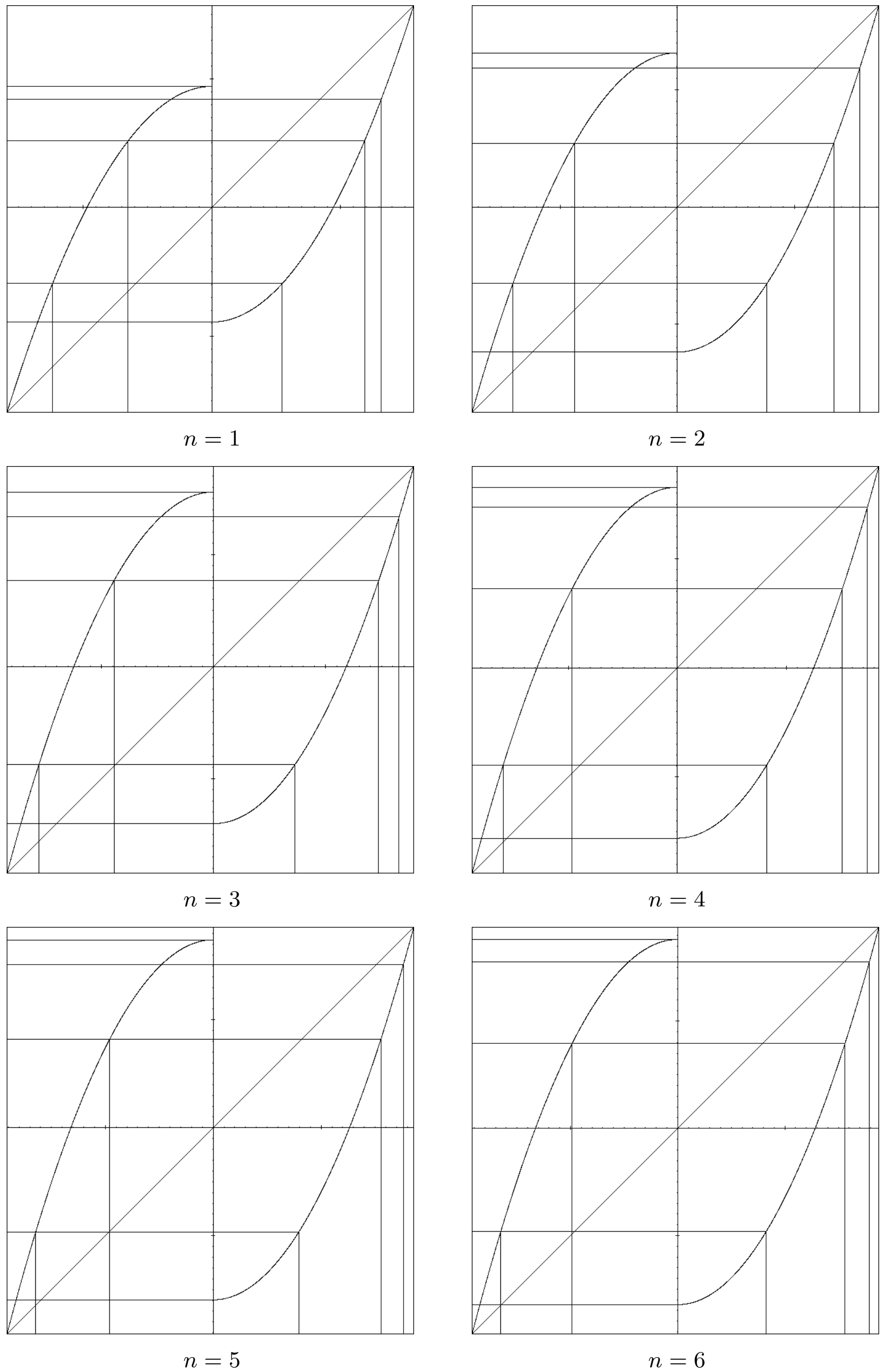


Fig. 5.2. The Thurston iteration for Example 5.9

5.9. Example. Suppose we are searching for a map in the quadratic Lorenz family (5.2) with the kneading invariant $\nu = (10(01)^\infty, 011(10)^\infty)$. From the kneading invariant we obtain the post-critical map

$$\pi = \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7 \\ 2\,5\,7\,1\,2\,5\,6 \end{pmatrix}, \quad \text{i.e.,} \quad \begin{array}{l} c^- = z_3 \mapsto z_7 \mapsto z_6 \mapsto z_5 \leftrightarrow z_2, \\ c^+ = z_4 \mapsto z_1 \mapsto z_2 \leftrightarrow z_5. \end{array}$$

The left hand side of Figure 5.1 shows how the Thurston map acts on a collection $(x_1, \ldots, x_l)$ of points by pulling each point back along the correct branch. The right hand side shows the fixed point of T (at the parameter $(a, b) \approx (1.670763, 1.803022)$).

As we already mentioned, the algorithm used to make the figures in this thesis is based on the assumption that the Thurston map is a contraction and the fixed point is found by iteration. Figure 5.2 shows the next few steps of the Thurston iteration started in Figure 5.1. $\diamond$

5.10. Remark. Hubbard & Schleicher [34] use a similar algorithm for the complex quadratic family $z \mapsto az(1 - z)$. They call the map T the *spider map* instead of the Thurston map, because strictly speaking it is only a "toy version" of the true Thurston map which is defined on a Teichmüller space. For an explanation of the name *spider* map the reader is referred to [34, Figure 1]. $\diamond$

In principle, Theorem 5.4—which is the finite version of Theorem 5.3—could be used to prove Theorem 5.3 by approximation: Choose a sequence $\nu^{(n)}$ of finite kneading invariants approximating the given kneading invariant and parameters $(a^{(n)}, b^{(n)})$ corresponding to these finite kneading invariants. Finally, let (a, b) be a limit point of this sequence. If the kneading invariant $\nu(a, b)$ depended continuously on the parameters (a, b) then one would be finished, but of course it does not. In fact, we already mentioned the possibility that some of the infinite kneading invariants do not exist. So they must get lost somehow in the limit. It seems that in order to understand which kneading invariants can get lost and which not, one needs a more detailed knowledge of the parameter dependence of the kneading invariant. Since the study of the parameter dependence provides us with an independent proof of Theorem 5.3, we do not follow this approach any further.

5.2. Parameter dependence of the kneading invariant. Now we are going to define the bifurcation diagram of the parameter space mentioned earlier. Because of assumptions (L5) and (L6), the kneading invariant of $f_{a,b}$ is very simple for $a < 0$ or $b < 0$, namely one of $(0^\infty, 1^\infty)$, $(01^\infty, 1^\infty)$ and $(0^\infty, 10^\infty)$. We will concentrate on the most interesting part of the parameter space, namely the region J where the interval $[p^-(a, b), p^+(a, b)]$ is invariant by $f_{a,b}$.

5.11. Definition (Bifurcation diagram). Let $\nu = (\nu^+, \nu^-)$ be an admissible kneading invariant. For $n \in \mathbb{N}$ let $J_n(\nu)$ be the set of parameters $(a, b) \in J$ such that

1. $c_i^+(a, b) \neq c$ and $c_i^-(a, b) \neq c$ for $i = 1, \ldots, n$.
2. $\nu_i^+(a, b) = \nu_i^+$ and $\nu_i^-(a, b) = \nu_i^-$ for $i = 0, \ldots, n$.

The collection of partitions $\mathcal{J}_n := \{J_n(\nu) \mid \nu \text{ admissible}\}$ is called the *bifurcation diagram* of the Lorenz family $\mathcal{F}$.

5.12. Remark. Here and in the following $c_i^\pm(a,b)$ and $\nu_i^\pm(a,b)$ denote the critical iterates respectively the digits of the kneading invariant for a fixed map $f_{a,b}$, whereas $c_i^\pm$ and $\nu_i^\pm$ are considered as functions of the parameters, i.e., $c_i^\pm : (a,b) \mapsto c_i^\pm(a,b)$ and $\nu_i^\pm : (a,b) \mapsto \nu_i^\pm(a,b)$. $\diamond$

The partitions $\mathcal{J}_n$ of the parameter space J for the quadratic Lorenz family (5.2) are plotted in Figure 1.2 for $n = 2, \dots, 7$; see also the front cover ([4]). The region shaped like an almond is the set J, which is contained in the square $[0,2] \times [0,2]$. For fixed n, the family $\mathcal{J}_n$ is a finite partition of J modulo the boundaries. Let us have a closer look at the border lines separating two adjacent pieces of the partition $\mathcal{J}_n := \{J_n(\nu) \mid \nu \text{ admissible}\}$. Since the kneading invariant changes there, one or both of the critical points has to be periodic of some period less than or equal to n. In the figure the lines going from left to right respectively from bottom to top correspond to Lorenz maps where the critical point c^- respectively c^+ is periodic.

5.13. Definition (Homoclinic orbit, hom-point). A periodic orbit containing one of the critical points c^+ and c^- is called a *homoclinic orbit*. Let $w_\pm$ denote a finite binary word and let $n^\pm$ denote the length of $w_\pm$.

1. $c^\pm(a,b)$ is called *homoclinic* of type $w_\pm$ if
   (a) $c^\pm_{n^\pm}(a,b) = 0$ and $c^\pm_k(a,b) \neq 0$ for $k < n^\pm$,
   (b) $\nu^\pm(a,b) = w_\pm^\infty$.

The set of parameters where $c^\pm$ is homoclinic of type $w_\pm$ is denoted by $C_{w_\pm}$.

2. $f_{(a,b)}$ is called *homoclinic* of type (w_+, w_-) if $(a,b) \in C_{w_+} \cap C_{w_-}$. If this is the case then (a,b) is called a *hom-point* of type (w_+, w_-).

5.14. Remark. The term “homoclinic orbit” is motivated by the origin of Lorenz maps as first return maps for the flow on the geometric Lorenz attractor. There the cross section Σ is taken transverse to the stable manifold of the singularity. The critical point c corresponds to the intersection of the stable manifold with Σ, and the iterates c_n^- and c_n^+ correspond to the successive intersections of the left and right half of the unstable manifold with Σ, whence periodic critical orbits of the Lorenz map correspond to homoclinic loops on the geometric Lorenz attractor (cf. Figure 1.1). $\diamond$

In order to avoid later confusion we emphasize that the sets of homoclinic parameters are denoted by $C_{w_\pm}$ and not by $C^\pm_{w_\pm}$. This is done for efficiency, since the upper “±” sign is redundant: The first digit of $w_\pm$ indicates whether it is c^+ or c^- which is periodic, e.g., $C_{10} = C^+_{10}$ and $C_{01} = C^-_{01}$. Also note that the sets C_0 and C_1 are closed, but in general the sets $C_{w_\pm}$ are not closed for $n^\pm > 0$, because of the second part of condition 1a. For example, in Figure 1.2 the origin belongs to $C_1 \cap C_0$ but not to any of the other curves emerging from the origin (cf. Section 5.3).

A first simple observation is that the gradient $\mathrm{D}c^\pm_{n^\pm} = (\mathrm{D}_a c^\pm_{n^\pm}, \mathrm{D}_b c^\pm_{n^\pm})$ is nonzero at every point $(a,b) \in C_{w_\pm} \subseteq \{c^\pm_{n^\pm} = 0\}$ and points “northwest”. This follows from

([4]) A similar figure can be found in Gambaudo *et al.* [16].

conditions (L6) and (L7) of Definition 5.1, since by the chain rule

$$\mathrm{D}_a c_n^{\pm} = \sum_{i=0}^{n-1} \mathrm{D}_a f(c_i^{\pm}) \cdot \prod_{k=i+1}^{n-1} \mathrm{D}_x f(c_k^{\pm}) < 0 \quad \text{and}$$

$$\mathrm{D}_b c_n^{\pm} = \sum_{i=0}^{n-1} \mathrm{D}_b f(c_i^{\pm}) \cdot \prod_{k=i+1}^{n-1} \mathrm{D}_x f(c_k^{\pm}) > 0 \quad \text{for all } n \in \mathbb{N}.$$

Consequently, the sets C_{w_+} and C_{w_-} are $\mathcal{C}^1$-smooth one-dimensional manifolds. In general, they can have more than one connected component.

5.15. Definition (Hom-line). A connected component of $C_{w_\pm}$ is called a *hom*$^{\pm}$*-line* of type $w_\pm$, or just a *hom-line*, if the distinction does not matter.

For parameter values contained in C_{w_+} respectively C_{w_-} one gets the following local information about the bifurcation diagram:

5.16. Lemma. *Choose $(a_0, b_0) \in C_{w_\pm} \subseteq \{c^{\pm}_{n^\pm} = 0\}$. Then there is a small neighbourhood U around (a_0, b_0) such that the gradient $\mathrm{D}c^{\pm}_{n^\pm}$ does not vanish on U and $C_{w_\pm} \cap U$ is a smooth curve cutting U in two halves. Let U^+ and U^- be the two components of $U \setminus C_{w_\pm}$ such that the gradient $\mathrm{D}c^{\pm}_{n^\pm}(a_0, b_0)$ points into U^+. Then one has the following local bifurcation diagram.*

1. *Moving from $C_{w_\pm}$ into $U^\pm$ the homoclinic orbit of type $w_\pm$ turns into a periodic attractor of type $w_\pm$ which contains $c^\pm$ in its immediate basin. The kneading invariant remains the same*: $\nu^\pm(a,b) = w_\pm^\infty = \nu^\pm(a_0, b_0)$.
2. *Moving from $C_{w_\pm}$ into $U^\mp$ the homoclinic orbit of type $w_\pm$ disappears and the kneading invariant changes to $\nu^\pm(a,b) = w_\pm \frac{0}{1} * \neq \nu^\pm(a_0, b_0)$. If the other critical point $c^\mp$ has the itinerary $\nu^\mp = w_\mp *$ and is not periodic of any period less than $|w_\mp|$ at (a_0, b_0) then $U^\mp$ can be chosen small enough such that $\nu^\pm(a,b) = w_\pm w_\mp *$ for $(a,b) \in U^\mp$.*

Proof. This is just a simple application of the Implicit Function Theorem. Observe that the existence of a periodic attractor in statement 1 follows from the fact that the derivative of $f^{n^\pm}$ vanishes at $c^\pm$ which guarantees that the central branch of $f^{n^\pm}$ has a fixed point for small perturbations into $U^\pm$ ([5]). The second part of statement 2 follows from a continuity argument. ∎

5.17. Remark. Observe that the half neighbourhood $U^\pm$ of a point $(a_0, b_0) \in C_{w_\pm}$ intersects no other set $C_{\widetilde{w}_\pm}$ of whatever period, since $c^\pm$ is attracted to a periodic attractor. This means that every hom$^\pm$-line is isolated on one side from all other hom$^\pm$-lines, except for its endpoints (which do not belong to it). ◇

The fact that the gradient points northwest everywhere on $C_{w_\pm}$ has an important consequence: Every line $M_t := \{(a,b) \mid a + b = t\}$ intersects $C_{w_\pm}$ at most once. In other words, after changing to (x,t)-coordinates, where $t := a+b$ and $x := b-a$, every connected component I of $C_{w_\pm}$ can be parametrized as the graph $\phi : t \mapsto (t, \varphi(t))$ of a $\mathcal{C}^1$-function φ defined on some interval $(t_*, t_\dagger)$. The interval $(t_*, t_\dagger)$ could be unbounded on one side

([5]) This is in contrast to Lorenz maps with exponent $\alpha < 1$ where the derivative diverges at c.

or both. The corresponding parametrization $\phi : t \mapsto (a(t), b(t))$ in (a, b)-coordinates is called the *natural parametrization* of the hom-line I.

5.18. Definition. Let I be a hom-line with natural parametrization $\phi : (t_*, t_\dagger) \to I$. If $t_* > -\infty$ and if the point $(a_*, b_*) := \lim_{t\to t_*} \phi(t)$ is contained in Λ then we say that I is *created* in (a_*, b_*). If $t_\dagger < \infty$ then we say that I is *annihilated* at the point $(a_\dagger, b_\dagger) := \lim_{t\to t_\dagger} \phi(t)$ (6). The interval $(t_*, t_\dagger)$ is called the *life-span* of I.

Because of Lemma 5.16 the creation resp. annihilation point of a hom-line I never belongs to I itself, only to its closure. Also note that if I and $\widetilde{I}$ are two different hom-lines of the same type $w_\pm$ then their life-spans are disjoint, because otherwise there would be a line M_t intersecting $C_{w_\pm}$ twice. The origin of the terms "creation", "annihilation" and "life-span" becomes evident if one interprets the parameter t as time. We will frequently make use of this interpretation when utilizing formulations like "some curve I enters (or leaves) some set A", etc.

In the following it will be useful to extend the definition of the sets C_{w_+} and C_{w_-} by defining $C_{10^\infty} := \{c_1^+ = p^-\}$ and $C_{01^\infty} := \{c_1^- = p^+\}$. This is another reason for the requirement that the fixed points $p^-(a, b)$ and $p^+(a, b)$ are hyperbolic repellers. It guarantees that the sets C_{10^∞} and C_{01^∞} behave very similarly to the sets C_{w_+} and C_{w_-}, namely they are smooth manifolds with a natural parametrization and the following analogue to Lemma 5.16 holds.

5.19. Lemma. *Choose* $(a_0, b_0) \in C_{10^\infty} \subseteq \{c_2^+ - c_1^+ = 0\}$. *Then there is a small neighbourhood* U *around* (a_0, b_0) *such that the gradient* $\mathrm{D}(c_2^+ - c_1^+)$ *does not vanish on* U *and* $C_{10^\infty} \cap U$ *is a smooth curve cutting* U *in two halves. Let* U^+ *and* U^- *be the two components of* $U \setminus C_{10^\infty}$ *such that the gradient* $\mathrm{D}(c_2^+ - c_1^+)(a_0, b_0)$ *points into* U^+. *Then one has the following local bifurcation diagram.*

1. *Moving from* C_{10^∞} *into* U^+ *the itinerary* $\nu^+(a, b)$ *changes from* 10^∞ *to* $10^{n-1}1*$, *where* $n = \widetilde{S}_0^+(a, b)$. *The lower bound* $\min\{\widetilde{S}_0^+(a, b) \mid (a, b) \in U^+\}$ *for the first co-cutting*$^+$ *time in* U^+ *can be chosen arbitrarily large by shrinking* U *appropriately.*

2. *Moving from* C_{10^∞} *into* U^- *the itinerary* $\nu^+(a, b)$ *does not change, but the iterates of* c^+ *now escape to* $-\infty$ *instead of ending up at the repelling fixed point* $p^-(a, b)$.

A similar statement holds for $(a_0, b_0) \in C_{01^\infty} \subseteq \{c_2^- - c_1^- = 0\}$.

Proof. At the point $(a_0, b_0) \in C_{10^\infty} \subseteq \{c_2^+ - c_1^+ = 0\}$ one has

$$\begin{aligned} &\mathrm{D}_a(c_2^+ - c_1^+) = \mathrm{D}_a f(c_1^+) + (f'(p^-) - 1) \cdot \mathrm{D}_a f(0^+) < 0 \quad \text{and} \\ &\mathrm{D}_b(c_2^+ - c_1^+) = \mathrm{D}_b f(c_1^+) + (f'(p^-) - 1) \cdot \mathrm{D}_b f(0^+) \geq 0, \end{aligned} \tag{5.8}$$

which follows from assumptions (L6) and (L7) and the fact that $f'(p^-) > 1$ by assumption (L3). Hence the gradient $\mathrm{D}(c_2^+ - c_1^+)$ does not vanish at (a_0, b_0) and points northwest and the Implicit Function Theorem can be applied again. The first statement now follows using a continuity argument, and the second one from the assumption that there are no additional fixed points outside the interval $[p^-, p^+]$. ■

(6) Observe that both limits exist in $\mathbb{R}^2$ because $\|\varphi'\| \leq 1$, but (a_*, b_*) may lie outside Λ.

5.20. Remark. Our aim is to treat the sets C_{10^∞} and C_{01^∞} in just the same way as the sets C_{w_+} and C_{w_-} as far as possible. For that reason let us make the agreement that a hom-line can also be a component of C_{10^∞} or C_{01^∞}, although strictly speaking these lines correspond to heteroclinic and not to homoclinic orbits. $\diamond$

Now we are ready to show that (L9) follows from assumptions (L1)–(L8) of Definition 5.1. Before doing that we make a last comment: Since the sets C_1, C_0, C_{10^∞} and C_{01^∞} are closed relative to Λ, they cannot have creation or annihilation points, in particular they must be simply connected.

5.21. Lemma. *Let $(f_{a,b})_{a,b\in\Lambda}$ be a $\mathcal{C}^1$-Lorenz family. Then there is an open subset J of Λ such that*

$$p^-(a,b) < c^+(a,b) < 0 < c^-(a,b) < p^+(a,b) \quad \textit{for all } (a,b)\in J$$

and the closure of J is the homeomorphic image of the unit square $[0,1]\times[0,1]$ in such a way that the left, right, bottom and top edge of the square are mapped into the hom-lines $C_1=\{c_1^+=0\}$, $C_{10^\infty}=\{c_1^+=p^-\}$, $C_0=\{c_1^-=0\}$, and $C_{01^\infty}=\{c_1^-=p^+\}$, respectively.

Proof. By condition (L8) of Definition 5.1 there is a parameter (a_0,b_0) in the sector $\mathbb{R}^2_+$ which is contained in $C_{10^\infty}\cap C_{01^\infty}$, whence C_{10^∞} and C_{01^∞} are nonvoid. Since C_{10^∞} has no creation point and since it is not allowed to intersect the b-axis (this would imply $C_1\cap C_{10^\infty}\neq\emptyset$), it must have entered $\mathbb{R}^2_+$ through the positive a-axis at some point $(a^+,0)$ before arriving at (a_0,b_0). This point is unique by (5.8). Similarly, C_{01^∞} enters $\mathbb{R}^2_+$ through the positive ob-axis at a unique point $(0,b^+)$. This shows that the connected component of the set $\{c^+<0\}\cap\{c_2^+>c_1^+\}\cap\{c^->0\}\cap\{c_2^-<c_1^-\}$ which contains the origin in its boundary has the required property. The corners of the "square" J are the points $(0,0)$, $(a^+,0)$, (a_1,b_1) and $(0,b^+)$, where (a_1,b_1) is the first intersection point of C_{10^∞} and C_{01^∞} after entering $\mathbb{R}^2_+$ (7). ∎

5.3. The gluing bifurcation. Looking at Figure 1.2 one gets the impression that all hom-lines separating the sets $J_n(\nu)$ are "created" at some point which is the intersection of two other hom-lines and then extend northeast until they reach the boundary of the set J. (Actually, they cross the boundary and continue in the positive cone $\mathbb{R}^2_+$ but this is not shown in the figure). At every such creation point a whole bunch of new curves is created. The new curves always appear in pairs forming the boundary of a "bubble" and the bubbles lie next to each other touching only at the creation point. In the following we try to understand the mechanism that creates all these curves and the meaning of the bubbles. We start investigating the creation of hom-lines at the origin, which is the intersection of C_1 and C_0, where this phenomenon has a simple and well known explanation. Later on we will see that the creation of hom-lines at all the other intersections of hom$^+$- and hom$^-$-lines can be understood in a similar way using renormalization.

(7) Formally: $(a_1,b_1):=\phi_{10^\infty}(t_1)=\phi_{10^\infty}(t_1)$, where ϕ_{10^∞} and ϕ_{10^∞} are the respective natural parametrizations and $t_1:=\inf\{t>\max(a^+,b^+)\mid \phi_{10^\infty}(t)=\phi_{01^\infty}(t)\}$.

5.22. Proposition. *There is a neighbourhood U of the origin such that for all parameters $(a,b) \in U^{++} := U \cap \mathbb{R}^2_+$ the restriction of $f_{a,b}$ to the interval $[c_1^+(a,b), c_1^-(a,b)]$ is a gap map* (8). *Moreover,*

1. *There are infinitely many bubbles emerging from the origin into U^{++} corresponding to the frequency locked regions where there is a unique periodic attractor with rational rotation number that attracts both critical points. The boundary of each bubble consists of two hom-lines where c^+ respectively c^- becomes part of the periodic orbit.*

2. *The bubbles are ordered in the same way as the Farey tree*, *i.e.*, *for fixed n the rotation numbers of the bubbles visible in $\mathcal{J}_n$ are precisely the numbers from the nth Farey sequence $\mathcal{F}_n := \{p/q \in \mathbb{Q} \mid 0 \le p \le q \le n\}$* (9) *in increasing order*, *where the rotation number of the bubbles increases in positive mathematical sense with respect to their location in the parameter plane.*

Injective circle maps and families thereof have been studied by various authors, e.g. Keener [36], Tresser [57] and Gambaudo [15]. We do not want to reinvent the wheel here and go through all the details of the proof of Proposition 5.22. Instead, we would like to indicate a simple alternative method how the proposition can be shown using induction along the Farey tree. As a byproduct one obtains a combinatorial rule to obtain all kneading invariants corresponding to rational rotation numbers through concatenation going down the Farey tree (cf. Figure 5.4). (A similar rule was used by Gambaudo [15] to obtain the shift-maximum of the periodic itineraries produced by rational rotations.) The induction step consists in showing the following claim.

Claim. *Two bubbles in $\mathcal{J}_n$ emerging from the origin are adjacent iff their rotation numbers p/q and p'/q' are Farey neighbours in $\mathcal{F}_n$. If n is increased then the first bubble to appear between those two is the bubble with rotation number $(p+p')/(q+q')$, called the* mediant *of p/q and p'/q'. The kneading invariant of the new bubble can be obtained from the ones of its parents by concatenation in the way specified in Figure* 5.3.

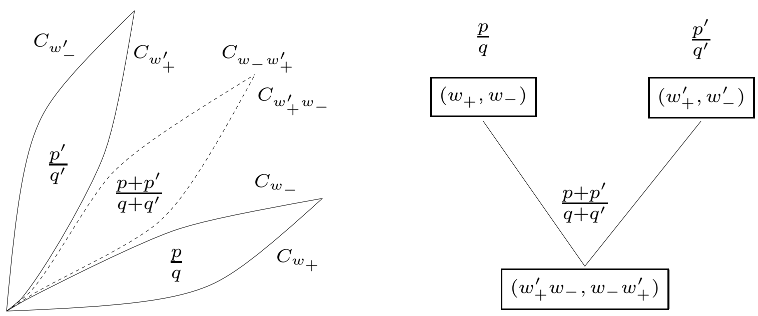


Fig. 5.3. The concatenation rule for the frequency locked regions. The kneading invariant corresponding to the mediant can be obtained by concatenating the itineraries w'_+ and w_- of its parents in the two possible ways. Note that the other itineraries w'_- and w_+ are not used.

(8) That is, a nonsurjective circle map (cf. Section 4.5).
(9) All fractions are assumed to be in reduced form.

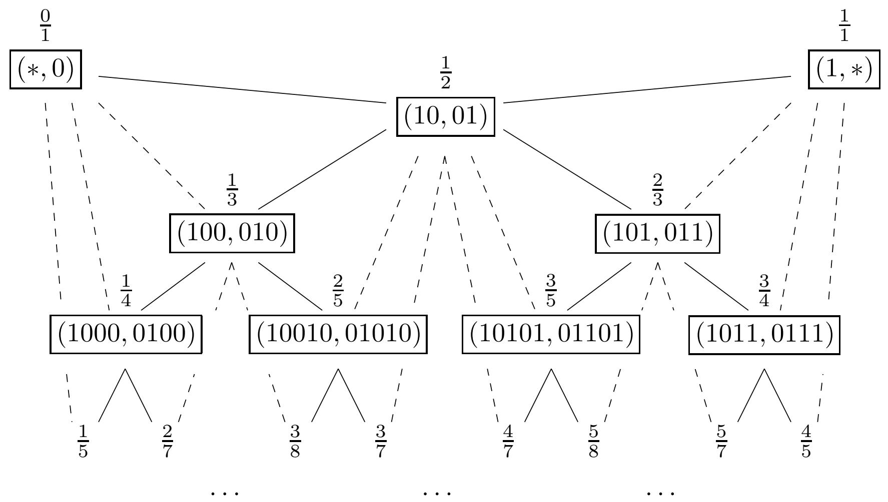


Fig. 5.4. The Farey tree for the kneading invariants of rational rotations. The kneading invariants and the corresponding rotation numbers are ordered in a Farey tree. The notation $(10, 01)$ is an abbreviation for $((10)^\infty, (01)^\infty)$, and so on. The kneading invariant of every node can be obtained by concatenating the kneading invariants of its parents according to Figure 5.3.

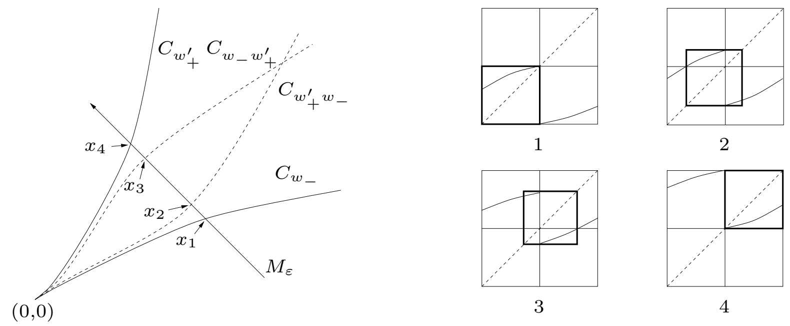


Fig. 5.5. The creation of new bubbles. Figures 1–4 show the left central branch of f^q, $q = |w_-|$, and the right central branch of $f^{q'}$, $q' = |w'_+|$, while moving from (x_1, ε) to (x_4, ε) along M_ε. The small squares indicate the renormalization interval $[c^+_{q'}, c^-_q]$.

Sketch of the proof. Since $c_1^+(0,0) = c_1^-(0,0) = 0$ and the derivative of the Lorenz maps vanishes at the origin, one can choose $\varepsilon > 0$ small enough such that $f'_{a,b}(x) \le 1/2$ for all (a, b) in the line segment $M_\varepsilon \cap \mathbb{R}^2_+$ ([10]). In particular, f is a gap map. Let (w_+, w_-) and (w'_+, w'_-) be the kneading invariants, corresponding to the Farey neighbours $p/q < p'/q'$. Here we adopted the abbreviation (w_+, w_-) for (w_+^∞, w_-^∞), which we are going to keep in the following. By induction we assume that the line M_ε intersects both $C_{w'_+}$ and C_{w_-} at points with (x, t)-coordinates (x_1, ε) and (x_4, ε), where $x_1 < x_4$ (compare Figure 5.3 with Figure 5.5). This is obviously the case for C_1 and C_0, the starting point of the induction. In Figure 5.5 we watch the left central branch of f^q (corresponding to the

([10]) For the definition of M_t, see page 105.

word w_-) and the right central branch of $f^{q'}$ (corresponding to the word w'_+) while moving along M_ε from (x_1, ε) to (x_4, ε). Along this straight line both branches move up and the map becomes (w'_+, w_-)-renormalizable ([11]) to an injective circle map $\mathcal{R}f$ on the interval $[c^+_{q'}, c^-_q]$ with a gap of positive length. This follows from the fact that the derivative of $\mathcal{R}f$ is less than $1/2$ on the entire interval $[c^+_1, c^-_1]$. The little pictures 1 and 4 of Figure 5.5 show the constellation of the two central branches at the parameters (x_1, ε) and (x_4, ε), respectively. By continuity, there exist $x_2,\ x_3$ in between such that c^+ respectively c^- is homoclinic of period two for the renormalized map $\mathcal{R}f$. This implies that $(x_2, \varepsilon) \in C_{w'_+ w_-}$ and $(x_3, \varepsilon) \in C_{w_- w'_+}$. The presence of the gap implies $x_2 < x_3$ and between those two parameters the renormalized map has a periodic attractor of period two that attracts the entire interval $[c^+_{q'}, c^-_q]$, and the kneading invariant of $\mathcal{R}f$ equals $(10, 01)$. Accordingly, the original map has kneading invariant $(w'_+ w_-, w_- w'_+)$ and is frequency locked. The rotation number equals the proportion of ones in the kneading invariant, which is $(p + p')/(q + q')$.

It remains to show that this frequency locked region is indeed a bubble, i.e., bounded only by hom-lines of type $w'_+ w_-$ and $w_- w'_+$ as indicated in the figure. For this one shows as above that for every parameter t such that the set M_t intersects $C_{w'_+ w_-}$ below $C_{w_- w'_+}$, the entire line segment in between belongs to the frequency locked region of rotation number $(p + p')/(q + q')$. In particular, this segment does not intersect any other hom-lines. We skip the remaining details. ∎

5.23. Remark. For the flows $\Phi_{a,b}$ on the geometric Lorenz attractor corresponding to the Lorenz maps $f_{a,b}$ the previous proposition yields the following bifurcation scenario: For $a < 0$ and $b < 0$ the flow $\Phi_{a,b}$ has two attracting cycles which turn into homoclinic loops in the "butterfly" constellation at $(a, b) = (0, 0)$. By arbitrarily small perturbations one obtains flows with attracting cycles that follow the two former homoclinic loops within a small ε-strip but may wind more than once around one of the two loops. More precisely, if one encodes the periodic cycle in a binary sequence writing down a 0 or 1 whenever it follows the left or right loop then for every itinerary corresponding to a rational rotation number there is a flow close to $\Phi_{0,0}$ that has an attracting cycle with the prescribed itinerary. It appears that all these new periodic cycles are created from the two original cycles (with itineraries 1^∞ and 0^∞, respectively) by gluing them together in various ways at the codimension 2 homoclinic bifurcation point $(0, 0)$. For that reason this bifurcation is called the *gluing bifurcation* (french: *collage de cycles*). This type of bifurcation has been introduced by Coullet, Gambaudo and Tresser [10] and was studied in great detail in Gambaudo [15]. For similar types of bifurcations see also Homburg [33]. ◇

5.4. Homoclinic bifurcation points. Let us have a closer look at the neighbourhood of a hom-point (a_0, b_0) where two sets C_{w_-} and C_{w_+} intersect.

5.24. Definition (Orientation of transverse hom-points). If (a_0, b_0) is a hom-point of type (w_+, w_-) then we say that it is a *transverse* hom-point if C_{w_-} and C_{w_+} intersect

([11]) These renormalizations are nonproper (cf. Section 4.5).

transversely at (a_0, b_0), i.e., if the gradients $\mathrm{D}c^-_{n^-}(a_0, b_0)$ and $\mathrm{D}c^+_{n^+}(a_0, b_0)$ are linearly independent, where $n^\pm := |w_\pm|$. If this is the case then the hom-point is called *positively* (resp. *negatively*) *oriented* if the determinant

$$\det \begin{pmatrix} \mathrm{D}_a c^-_{n^-} & \mathrm{D}_b c^-_{n^-} \\ \mathrm{D}_a c^+_{n^+} & \mathrm{D}_b c^+_{n^+} \end{pmatrix} (a_0, b_0) \tag{5.9}$$

has positive (resp. negative) sign. In other words, a hom-point is positively oriented if C_{w_+} crosses C_{w_-} from bottom to top and it is negatively oriented if C_{w_+} crosses C_{w_-} from left to right.

For simplicity assume for the moment that all intersections of hom-lines are transverse. We will return to the general case later.

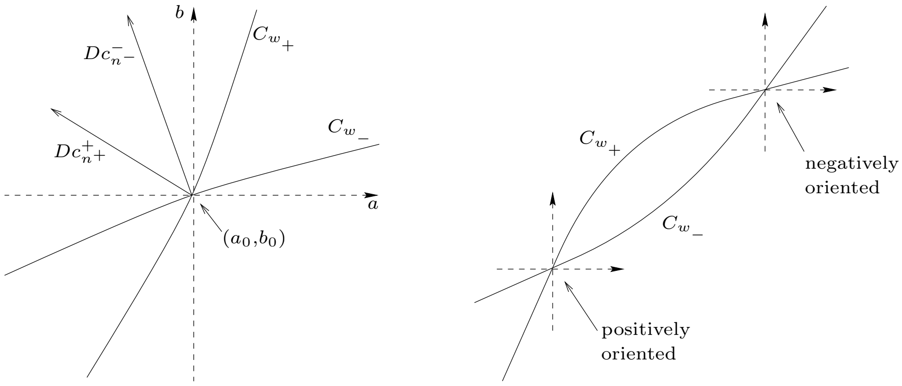


Fig. 5.6. Orientation of transverse hom-points. The left figure shows a close-up of a positively oriented intersection of C_{w_+} and C_{w_-}. Two consecutive transverse intersections of connected components of C_{w_+} and C_{w_-} must have different orientations (right figure).

5.25. Lemma. *Let (a_0, b_0) be a positively oriented hom-point of type (w_+, w_-). Then there is an orientation preserving local reparametrization $\psi : (\alpha, \beta) \mapsto (a, b)$ mapping a small disk V around $(0, 0)$ to a small disk U around (a_0, b_0) such that*

$$c^+_{n^+} \circ \psi(\alpha, \beta) = -\alpha \quad \text{and} \quad c^-_{n^-} \circ \psi(\alpha, \beta) = \beta.$$

If the hom-point is negatively oriented then there is an orientation preserving local reparametrization $\psi : (\alpha, \beta) \mapsto (a, b)$ mapping a small disk V around $(0, 0)$ to a small disk U around (a_0, b_0) such that

$$c^+_{n^+} \circ \psi(\alpha, \beta) = -\beta \quad \text{and} \quad c^-_{n^-} \circ \psi(\alpha, \beta) = \alpha.$$

Proof. For the positively oriented case apply the Inverse Function Theorem to the map $(a, b) \mapsto (-c^+_{n^+}(a, b), c^-_{n^-}(a, b))$ and observe that the determinant of its linearization at (a_0, b_0) equals the determinant (5.9), whence it is positive. The determinant of the linearization of $(a, b) \mapsto (-c^+_{n^+}(a, b), c^-_{n^-}(a, b))$ at (a_0, b_0) is just the opposite of determinant (5.9), whence it is positive in the orientation reversing case. ■

If (a_0, b_0) is a transverse hom-point then Lemma 5.25 tells us that the two homoclinic orbits of f_{a_0,b_0} unfold completely when the two parameters are perturbed, i.e., one can

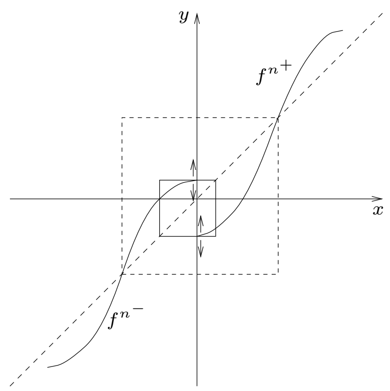


Fig. 5.7. Two homoclinic orbits unfold

move $c^-_{n^-}$ and $c^+_{n^+}$ up and down independently by slightly perturbing the parameter (a_0, b_0) as shown in Figure 5.7. As long as the perturbation is small enough the two central branches of f^{n^-} respectively f^{n^+} both have fixed points q^- and q^+ and the restriction of these two branches to the invariant interval $[q^-, q^+]$ is again a proper Lorenz map (dashed box). In the positively oriented case the bifurcation diagram near (a_0, b_0) looks just the same as the one at the origin. Northeast of the hom-point the map is renormalizable to a gap map (solid box). In the negatively oriented case the local picture is the mirror image with respect to the second diagonal of the one at the origin. In particular, the regime of circle renormalizations lies southwest of the hom-point. These observations are collected in the proposition below (cf. Figure 5.8).

5.26. Remark. If all maps of the Lorenz family have negative Schwarzian derivative then the fixed points q^- and q^+ of the two central branches are again hyperbolic repellers, since they obviously do not attract any of the two critical points. Hence the renormalization interval $[q^-, q^+]$ depends continuously on the parameters. $\diamond$

5.27. Proposition. *Assume that* (a_0, b_0) *is a transverse hom-point of type* (w_+, w_-) *and let* $U^{st} := \psi(V \cap (\mathbb{R}_s \times \mathbb{R}_t))$ *for* $s, t \in \{-, +\}$, *where* ψ *is the reparametrization of Lemma* 5.25. *If the hom-point is positively oriented then one has the following bifurcation diagram.*

1. *In* U^{--} *the map* f *has two periodic orbits of type* w_- *resp.* w_+ *containing* c^- *resp.* c^+ *in their immediate basin. The kneading invariant equals* (w_+^∞, w_-^∞).
2. *In* U^{-+} *the map* f *has a periodic orbit of type* w_+ *containing* c^+ *and* $c^-_{n^-}$ *in its immediate basin. The kneading invariant equals* $(w_+^\infty, w_- w_+^\infty)$.
3. *In* U^{+-} *the map* f *has a periodic orbit of type* w_- *containing* c^- *and* $c^+_{n^+}$ *in its immediate basin. The kneading invariant equals* $(w_+ w_-^\infty, w_-^\infty)$.
4. *In* U^{++} *the map* f *is* (w_+, w_-)*-renormalizable to an injective circle map. There are infinitely many hom-lines emerging from* (a_0, b_0) *into* U^{++} *corresponding to hom-lines of the renormalized circle map.*

If the orientation is negative then the same holds with the roles of U^{++} *and* U^{--} *interchanged.*

Proof. This is an immediate consequence of Proposition 5.22 and Lemma 5.25. ∎

Just as in Proposition 5.22, the hom-lines in the sector U^{++} form bubbles which are ordered in the way of the Farey tree, where the place of a bubble in the tree is determined by the rotation number of the renormalized map. The kneading invariants of the bubbles can be obtained by the same concatenation rule traversing the Farey tree as in Figure 5.4. One just has to replace $(*, 0)$ and $(1, *)$ by $(*, w_-)$ and $(w_+, *)$ at the top of the tree, respectively.

5.28. Definition. The bifurcation that occurs at a positively oriented hom-point is called a *gluing bifurcation* and the bifurcation that occurs at a negatively oriented hom-point is called a *dissolving bifurcation*.

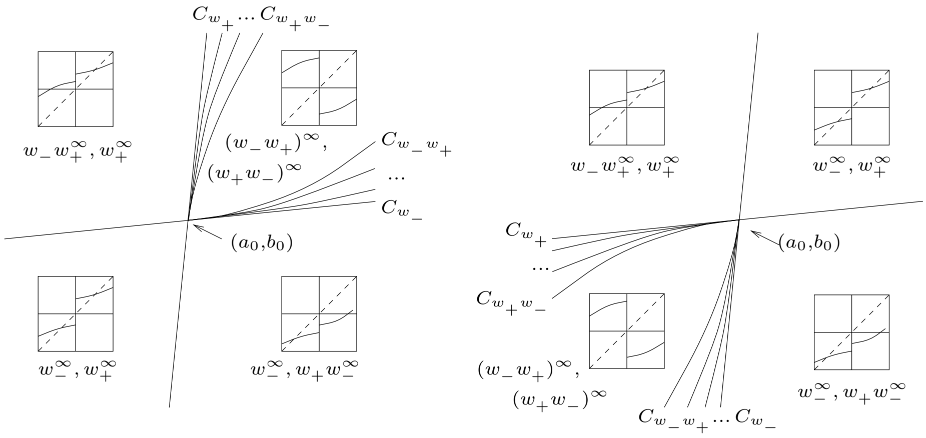


Fig. 5.8. The gluing and dissolving bifurcations. If the hom-point is positively oriented (left figure) then northeast of the homoclinic value (a_0, b_0) the map f becomes renormalizable to an injective circle map and infinitely many new hom-lines are created. Similarly, infinitely many curves are annihilated in a negatively oriented hom-point (right figure).

The following proposition shows that gluing and dissolving bifurcations are the only possible ways to create or annihilate hom-lines.

5.29. Proposition. *Let $(f_{a,b})_{a,b}$ be a $\mathcal{C}^1$-Lorenz family such that all hom-points are transverse and let I be a hom-line.*

1. *If I is created in (a_*, b_*) then (a_*, b_*) is a positively oriented hom-point and I is one of the lines created in the gluing bifurcation at (a_*, b_*).*
2. *If I is annihilated in $(a_\dagger, b_\dagger)$ then $(a_\dagger, b_\dagger)$ is a negatively oriented hom-point and I is one of the lines annihilated in the dissolving bifurcation at $(a_\dagger, b_\dagger)$.*

If I is created in (a_*, b_*) and annihilated in $(a_\dagger, b_\dagger)$ then the hom-lines intersecting at (a_*, b_*) and $(a_\dagger, b_\dagger)$ are the same and the creation and annihilation points of I are two successive intersections of the same two hom-lines.

Proof. We only prove the first part, the second one is completely analogous. Assume for definiteness that I is of type w_+ and $|w_+| = n$. Let $\phi : (t_*, t_* + \varepsilon) \to I$ be the natural parametrization of I near (a_*, b_*), where $t_* := a_* + b_*$. We reverse time and approach the creation point (a_*, b_*) of I along the curve $\phi(t)$ for $t \searrow t_*$. In the limit, some lower iterate c_k^+, $k < n$, also has to become homoclinic. If this were not the case then (a_*, b_*) would actually belong to I and by Lemma 5.16 the curve I could be extended beyond (a_*, b_*), yielding a contradiction.

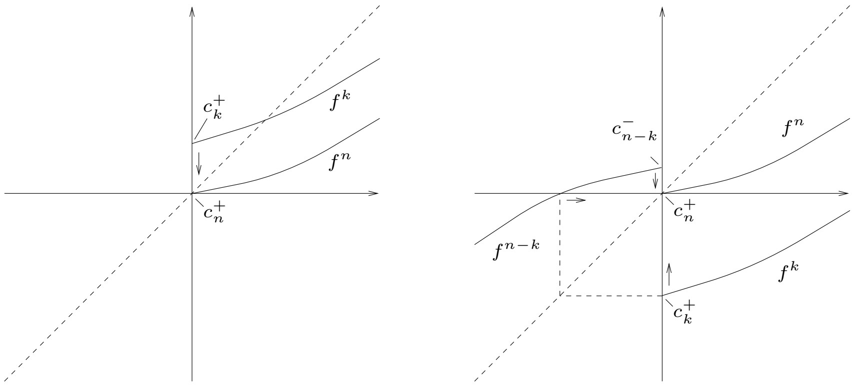


Fig. 5.9. The creation of a hom-line. Moving towards the creation point (a_*, b_*) of C_{w_+}, some lower iterate c_k^+ has to become homoclinic. It cannot approach c from above (left figure), only from below and c_{n-k}^- approaches c from above simultaneously (right figure). The arrows indicate the movement of the branches after "reversing time".

We claim that together with c_k^+ also c_{n-k}^- becomes homoclinic as $t \to t_*$. First of all, as $t \to t_*$ the point c_k^+ has to approach the discontinuity c from below. If this were not the case, then one could choose ε small enough such that for all $t \in (t_*, t_* + \varepsilon)$ the right central branch of f^k would have an attracting fixed point containing c^+ in its immediate basin (cf. Figure 5.9, l.h.s.) But then c^+ could not be periodic of any period, contradicting the fact that $c_n^+ = c$. Choose ε small enough such that c_k^+ is contained in the left central branch of f^{n-k} for $t \in (t_*, t_* + \varepsilon)$. Then $\mathrm{dist}(c, c_{n-k}^-) = (f^{n-k})'(\zeta) \cdot \mathrm{dist}(c_k^+, c)$ for some $\zeta \in (c_k^+, c^-)$, so $\mathrm{dist}(c, c_{n-k}^-)$ tends to 0 as $t \to t_*$ and the claim follows (cf. Figure 5.9, r.h.s.) We proved that at the parameter (a_*, b_*) both critical points are periodic. If one chooses minimal periods for c^+ and c^- then Proposition 5.27 applies.

To prove the last statement one shows that every line M_t that intersects I must also intersect C_{w_+} above I and C_{w_-} below I, whence the hom-lines of type w_+ and w_- that intersect at (a_*, b_*) exist throughout the entire interval between t_* and $t_\dagger$. Now the only possibility how I can be annihilated at $(a_\dagger, b_\dagger)$ is when the two hom-lines intersect a second time with negative orientation. ∎

Now we want to drop the condition that all intersections of hom-lines are transverse. Without this condition, the local bifurcation diagram can look a lot messier than it did in the transverse case. The reason for that is that two hom-lines I^+ and I^- that intersect nontransversely do not have to cross each other, instead they may only touch and separate again. Even worse, their intersection does not have to be a single point, it could be an arc or a Cantor set. Still one has the following analogue to Proposition 5.27, except that now the four different types of dynamical behaviour near the hom-point are not separated neatly into four sectors any more.

5.30. Proposition. *Let (a_0, b_0) be a hom-point of type (w_+, w_-). Then there is a neighbourhood U of (a_0, b_0) such that for every parameter in $(a, b) \in U$ one of the following holds*:

1. *The map $f_{a,b}$ has two periodic orbits of type w_- resp. w_+ containing c^- resp. c^+ in their immediate basin. The kneading invariant equals (w_+^∞, w_-^∞).*
2. *The map $f_{a,b}$ has a periodic orbit of type w_+ containing c^+ and $c_{n^-}^-$ in its immediate basin. The kneading invariant equals $(w_+^\infty, w_- w_+^\infty)$.*
3. *The map $f_{a,b}$ has a periodic orbit of type w_- containing c^- and $c_{n^+}^+$ in its immediate basin. The kneading invariant equals $(w_+ w_-^\infty, w_-^\infty)$.*
4. *The map $f_{a,b}$ is (w_+, w_-)-renormalizable to an injective circle map.*

Proof. Apply Lemma 5.16 twice, once for C_{w_+} and once for C_{w_-}. ∎

Although it has now become senseless to talk about the orientation of a hom-point, the notion of gluing and dissolving points can be adapted to the general setting. For this it is convenient to change to (x, t)-coordinates where the local picture is less nasty: Every line M_t intersects at most one of the regimes 1 and 4, depending on whether C_{w_+} lies below or above C_{w_-} at time t.

5.31. Definition. Let $I^\pm$ be a hom$^\pm$-line and let $\phi^\pm$ denote the natural parametrization of $I^\pm$ in (x, t)-coordinates. Furthermore, let $T = (t_g, t_u)$ be a connected component of the open set $\{t \mid \phi^+(t) > \phi^-(t)\}$. If $t_g > t_*^+ \vee t_*^-$ then $(a_g, b_g) := \phi^+(t_g) = \phi^-(t_g)$ is called a *gluing point* of I^+ and I^-. If $t_u < t_\dagger^+ \wedge t_\dagger^-$ then $(a_d, b_d) := \phi^+(t_d) = \phi^-(t_d)$ is called a *dissolving point* of I^+ and I^-.

Here we used the abbreviation $u \vee v := \max(u, v)$ and $u \wedge v := \min(u, v)$. With the arguments used in the proof of Proposition 5.22 one can show that a gluing (resp. dissolving) point is just what its name suggests: Northeast (resp. southwest) of this point the map is renormalizable to a gap map and infinitely many hom-lines are created (resp. annihilated) in (a_0, b_0). Instead of using the Implicit Function Theorem which is not applicable here, one proves this directly by considering lines M_t which pass arbitrarily close above (resp. below) the homoclinic bifurcation point. Note that a nontransverse hom-point can be a gluing and a dissolving point simultaneously, or neither of the two.

Now Proposition 5.29 holds without the assumption of transversality if one replaces the term "positively (negatively) oriented hom-point" by "gluing (dissolving) point".

5.5. Monotonic Lorenz families. As we already saw in the previous subsection, the presence of nontransverse homoclinic bifurcation points complicates the local picture of the bifurcation diagram. But even if all intersections are transverse, the global bifurcation diagram may still be complicated since the hom-lines can get tangled and the components of the partition $\mathcal{J}_n$ may get a very complicated shape. For that reason we start proving Theorem 5.3 in the special case of monotonic Lorenz families and treat the general case later.

5.32. Definition (Monotonic Lorenz family). A $\mathcal{C}^1$-Lorenz family is *monotonic* if all intersections of curves C_{w_-} and C_{w_+} are transverse and positively oriented.

5.33. Proposition. *Let* $(f_{a,b})_{a,b}$ *be a monotonic Lorenz family. Then all sets* C_{w_+} *and* C_{w_-} *are simply connected. All but two hom-lines, namely* C_0 *and* C_1, *are created in a gluing bifurcation and no hom-line is annihilated* ([12]). *Two sets* C_{w_+} *and* C_{w_-} *intersect at most once.*

Proof. Proposition 5.29 shows that monotonic Lorenz families have no annihilation points. Now $C_{w_\pm}$ has to be simply connected, because if I and I' were two different components of $C_{w_\pm}$ then at least one of them would have an annihilation point. If C_{w_+} and C_{w_-} intersect more than once then—since they are both connected—not all intersection points can be positively oriented, whence this is impossible. ∎

5.34. Theorem. *Let* $(f_{a,b})_{(a,b)\in\Lambda}$ *be a monotonic Lorenz family and let* $\nu = (\nu^+, \nu^-)$ *be an admissible kneading invariant. Then the sequence* $(J_n(\nu))_{n\in\mathbb{N}}$ *is a decreasing sequence of simply connected open regions. The boundary of each set* $J_n(\nu)$ *consists of two or four differentiable arcs where one of the critical points is periodic of some period less than or equal to* n ([13]) *or it is mapped onto one of the two repelling fixed points in the boundary. For every parameter* (a, b) *in* $J_\infty(\nu) := \bigcap_{n\in\mathbb{N}} J_n(\nu)$ *the critical points are not periodic and the Lorenz map* $f_{a,b}$ *has the kneading invariant* ν. *If* ν *is expansive or periodic then* J_∞ *is nonvoid.*

This implies in particular that Theorem 5.3 holds for monotonic Lorenz families. Another corollary is the conclusion that the kneading invariant of a monotonic Lorenz family depends monotonically on the parameters in the sense of the definition made in the introduction, which we repeat here.

5.35. Definition (Monotonicity of the kneading invariant). Let $\mathcal{F} = (f_{a,b})_{(a,b)\in J}$ be a $\mathcal{C}^1$-Lorenz family. We say that the *kneading invariant* of $\mathcal{F}$ *depends monotonically on the parameters* if for every admissible kneading invariant ν and every n the set $J_n(\nu)$ is simply connected.

5.36. Corollary. *The kneading invariant of a monotonic Lorenz family depends monotonically on the parameters.*

Proof of Theorem 5.34. We fix some admissible invariant (ν^+, ν^-) and look at the decreasing sequence of sets $J_n(\nu)$. As already mentioned in the discussion of Theorem 4.19 on page 78, the admissible prefixes can be ordered hierarchically in the form of a tree. The nodes of this tree correspond to the admissible prefixes and the subnodes of every admissible $(n-1)$-prefix are determined by the possible choices for ν_n^+ and ν_n^- in the inequalities of the admissibility condition (AC2):

$$\nu^+_{n-\widetilde{S}^+\langle n\rangle} \le \nu^+_n \le \nu^-_{n-S^+\langle n\rangle}, \quad \nu^+_{n-S^-\langle n\rangle} \le \nu^-_n \le \nu^-_{n-\widetilde{S}^-\langle n\rangle}. \tag{5.10}$$

This hierarchical ordering of the admissible prefixes is reflected by the ordering of the sets $J_n(\nu)$ in parameter space where the hierarchy is just given by inclusion (cf. Figure 1.2). We will use it to prove Theorem 5.34 by induction on n.

([12]) Instead, the hom$^+$-lines depart through the upper boundary C_{01^∞} and the hom$^-$-lines depart through the right boundary C_{10^∞} and continue beyond J, which is not shown in Figure 1.2.

([13]) More precisely, the period for $c^\pm$ equals $S^\pm\langle n+1\rangle$ or $\widetilde{S}^\pm\langle n+1\rangle$.

Fix a set $J_{n-1}(\nu)$ and assume by induction that it has the shape of a "square", i.e., it is simply connected and its boundary consists of four points, the *vertices* or *corners* of $J_{n-1}(\nu)$, which are connected by four arcs, the *edges* of $J_{n-1}(\nu)$, such that the following holds (cf. Figure 5.10):

1. The edges of $J_{n-1}(\nu)$ are parts of hom-lines where c^+ or c^- is periodic of some period less than n, namely:

- The left edge is contained in C_{w_+} and the right edge is contained in $C_{\widetilde{w}_+}$.
- The bottom edge is contained in C_{w_-} and the top edge is contained in $C_{\widetilde{w}_-}$.

Here $w_\pm$ is the $S^\pm\langle n\rangle$-prefix and $\widetilde{w}_\pm$ is the $\widetilde{S}^\pm\langle n\rangle$-prefix of $\nu^\pm$. Recall that $S^\pm\langle n\rangle$ and $\widetilde{S}^\pm\langle n\rangle$ denote the last cutting respectively co-cutting time before n.

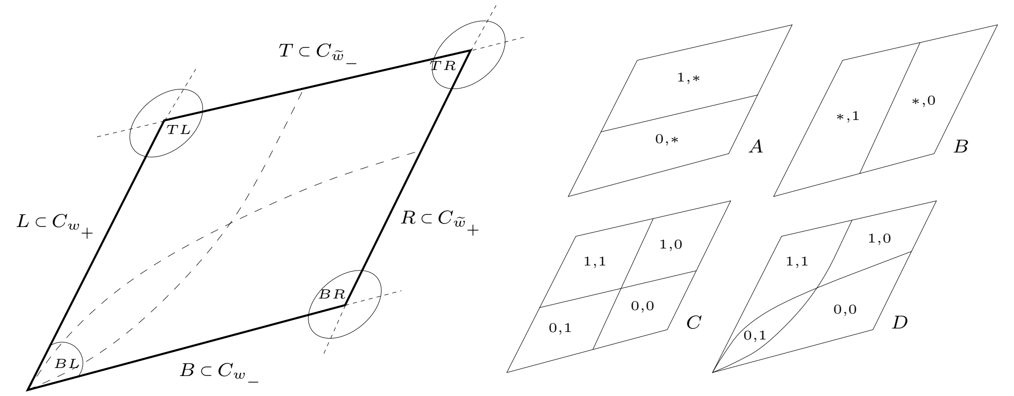


Fig. 5.10. The splitting of $J_{n-1}(\nu)$ for monotonic families. The edges of $J_{n-1}(\nu)$ are parts of C_{w_+}, $C_{\widetilde{w}_+}$, C_{w_-} and $C_{\widetilde{w}_-}$. If $J_{n-1} \notin \mathcal{J}_n$ then it splits into two or four fragments of $\mathcal{J}_n$ as in one of the figures A–D. The fragments are labelled by the digits that are appended to the $(n-1)$-prefix of (ν^+, ν^-). One of these fragments is $J_n(\nu)$.

2. The corners of $J_{n-1}(\nu)$ are hom-points. Except for corner 4 they are the transverse intersections of the hom-lines containing the adjacent edges. The left and bottom hom-line do not necessarily intersect in corner 4. Instead, corner 4 could be a hom-point of a shorter combinatorial type (w'_-, w'_+) where one or both hom-lines are created. In this case w_- and w_+ are both (w'_-, w'_+)-renormalizable and the renormalized words correspond to rotation numbers which are Farey neighbours.

To be precise, assumption 1 only holds after the first co-cutting times have occurred. For $n \le \widetilde{S}_0^\pm$ one has to replace $C_{\widetilde{w}_+}$ and $C_{\widetilde{w}_-}$ by C_{10^∞} and C_{01^∞}, respectively. The starting point of the induction is $n = 2$ and $J_1 = J$, which is bounded by C_1, C_{10^∞}, C_0 and C_{01^∞}. To avoid tedious repetitions we only treat the case were $n > \max(\widetilde{S}_0^+, \widetilde{S}_0^-)$. The necessary modifications for the remaining case are obvious.

We apply Proposition 5.27 to small neighbourhoods of the corners of $J_{n-1}(\nu)$ to determine the kneading invariant inside $J_{n-1}(\nu)$ near these corners. These regions are indicated by disks in Figure 5.10 and the sectors lying inside $J_n(\nu)$ are going to be called the TR-, BR-, TL-, and BL-corner [14].

[14] That is, the "top right, "bottom right", "top left", and "bottom left" corner, respectively.

In the BL-corner only a part of the disk is drawn to emphasize that the curves C_{w_+} and C_{w_-} may not intersect transversely there. Instead, they could be the boundaries of two adjacent bubbles created in a gluing bifurcation at that corner (as in Figure 5.5, where $C_{w'_+}$ plays the role of C_{w_+}). In this case Proposition 5.27 has to be applied to the ancestors of C_{w_+} and C_{w_+} that initiated the gluing bifurcation.

1. TR-corner: $\nu = (\widetilde{w}_+^\infty, \widetilde{w}_-^\infty) \Rightarrow \nu_n^+ = \nu^+_{n-\widetilde{S}^+\langle n\rangle}$, $\nu_n^- = \nu^-_{n-\widetilde{S}^-\langle n\rangle} \Rightarrow n$ is neither a co-cutting$^+$ time nor a co-cutting$^-$ time.

2. BR-corner: $\nu = (\widetilde{w}_+^\infty, w_-\widetilde{w}_+^\infty) \Rightarrow \nu_n^+ = \nu^+_{n-\widetilde{S}^+\langle n\rangle}$, $\nu_n^- = \nu^+_{n-S^-\langle n\rangle} \Rightarrow n$ is neither a co-cutting$^+$ time nor a cutting$^-$ time.

3. TL-corner: $\nu = (w_+\widetilde{w}_-^\infty, \widetilde{w}_-^\infty) \Rightarrow \nu_n^+ = \nu^-_{n-S^+\langle n\rangle}$, $\nu_n^- = \nu^-_{n-\widetilde{S}^-\langle n\rangle} \Rightarrow n$ is neither a cutting$^+$ time nor a co-cutting$^-$ time.

4. BL-corner: $\nu = (w_+w_-*, w_-w_+*) \Rightarrow \nu_n^+ = \nu^-_{n-S^+\langle n\rangle}$, $\nu_n^- = \nu^-_{n-S^-\langle n\rangle} \Rightarrow n$ is neither a cutting$^+$ time nor a cutting$^-$ time as long as $n < S^+\langle n\rangle + S^-\langle n\rangle$.

If $n = S^+\langle n\rangle + S^-\langle n\rangle$ then a bubble appears in the bottom left corner, bounded by two hom-lines of type w_+w_- and w_-w_+. Inside the bubble the kneading invariant is (w_+^∞, w_-^∞), so n is a co-cutting$^+$- and a co-cutting$^-$ time and there are no further cutting or co-cutting times.

The numbers of the items are chosen such that they match the respective case of Proposition 5.27 which applies. The implication arrows can easily be verified if one recalls how the splitting and co-splitting are obtained (cf. Section 4.2). As an example, in the bottom left corner n is not a cutting$^+$ time for $n < S^+\langle n\rangle + S^-\langle n\rangle$, because the last cutting time was $S^+\langle n\rangle$ and the itineraries $\nu^+[S^+\langle n\rangle, \dots) = w_-*$ and $\nu^-[0, \dots) = w_-w_+*$ coincide at least in $|w_-| = S^-\langle n\rangle$ digits. All other implications are proven similarly.

For the last statement about the bottom left corner we remark in advance that if $(\nu^+, \nu^-) = (w_+w_-w_+*, w_-w_+w_-*)$ then the "last chance to cut" (cf. Remark 2.11) is missed, since $n = S^+\langle n\rangle + S^-\langle n\rangle$ is neither a cutting$^+$ nor a cutting$^-$ time. Consequently, $(\nu^+, \nu^-) = ((w_+w_-)^\infty, (w_+w_-)^\infty)$ and there are no further cutting$^+$ or cutting$^-$ times.

Comparing the kneading invariants in the four corners with the inequalities (5.10), we see that no matter how many choices we have for ν_n^+ and ν_n^- to satisfy these inequalities there will always be at least one corner where this choice is realized, in particular $J_n(\nu)$ is nonvoid. To finish the induction we have to verify that $J_n(\nu)$ again satisfies the induction assumption. In order to do this we first show that $J_{n-1}(\nu)$ can only be split into pieces in one of the four ways depicted in Figures 5.10(A)–(D).

Assume that $\nu^+_{n-\widetilde{S}^+\langle n\rangle} < \nu^-_{n-S^+\langle n\rangle}$. Then the digit ν_n^+ differs in the TR- and TL-corner, whence it has to change somewhere along the upper edge in between, which means that there is a hom-point on the upper edge. By Proposition 5.33 it is the unique point of this combinatorial type. This means that there is a hom$^+$-line V (= "vertical") leaving $J_{n-1}(\nu)$ through the upper boundary where c_n^+ is periodic. Where did this curve come from? We claim that it either entered $J_{n-1}(\nu)$ through the lower edge C_{w_-} or that it was created in the bottom left corner by gluing together C_{w_+} and C_{w_-}, i.e., $V = C_{w_+w_-}$. We prove this by excluding the remaining possibilities.

First of all, V could not have been created somewhere in the interior of $J_{n-1}(\nu)$ since this would imply the existence of homoclinic parameter values of period less than n, which is excluded by the definition of $J_{n-1}(\nu)$. It did not enter $J_{n-1}(\nu)$ through the upper edge by Proposition 5.33, since there exists already such an intersection. The right edge is forbidden because of Remark 5.17 and by the same argument it could not have come from the left edge unless it was created in the BL-corner. This proves the claim for ν_n^+ and the analogous claim for ν_n^- is proved similarly.

If V is created in the BL-corner then $n = S^+\langle n\rangle + S^-\langle n\rangle$ and $V = C_{w_+w_-}$. In this case there is also a hom$^-$-line $H = C_{w_-w_+}$ (= "horizontal") emerging from the corner which forms a bubble together with V. The hom-line H has to leave $J_{n-1}(\nu)$ through the right edge, as one can see by similar arguments to those above.

If $\nu^+_{n-\widetilde{S}^+\langle n\rangle} = \nu^-_{n-S^+\langle n\rangle}$ then the digit ν_n^+ does not change between the TR- and the TL-corner, in fact it does not change in $J_{n-1}(\nu)$ at all, so there is no vertical splitting of $J_{n-1}(\nu)$. Repeating all the arguments for the lower inequality (5.10) we finally see that either $J_n(\nu)$ equals $J_{n-1}(\nu)$, or $J_{n-1}(\nu)$ splits into pieces in one of the ways shown in Figures 5.10(A)–(D) and $J_n(\nu)$ is one these parts. If case D applies and if $J_n(\nu)$ coincides with the bubble, then the induction terminates, because then n is neither a cutting$^+$ nor a cutting$^-$ time, whence the "last chance to cut" is missed and the bubble will not undergo any further splitting. In the other cases the set $J_n(\nu)$ has two edges which are parts of edges of $J_{n-1}(\nu)$ and the remaining two are parts of V and H where c^+ respectively c^- is periodic of period n. It can easily be checked that n equals $S^-\langle n+1\rangle$ resp. $\widetilde{S}^-\langle n+1\rangle$ if H is the bottom resp. top edge of $J_n(\nu)$ and n equals $S^+\langle n+1\rangle$ resp. $\widetilde{S}^+\langle n+1\rangle$ if V is the left resp. top right edge of $J_n(\nu)$.

It follows that $J_n(\nu)$ again satisfies conditions 1 and 2 of the induction assumption. This completes the induction and proves the theorem except for the last sentence. If ν is periodic then $J_\infty(\nu)$ is not void because there is an open set G such that $G \subseteq J_n(\nu)$ for all n. If ν is expansive then there are infinitely many cutting and co-cutting times for ν^+ and ν^-. At every cutting$^+$ time $J_n(\nu)$ lies in the left half of $J_{n-1}(\nu)$ and at every co-cutting$^+$ time it lies in the right half of $J_{n-1}(\nu)$. Similarly, at every cutting$^-$ time $J_n(\nu)$ lies in the lower half and at every co-cutting$^-$ time it lies in the upper half. Since all four possibilities occur infinitely often, it follows that $\bigcap_{n\in\mathbb{N}} J_n(\nu) = \bigcap_{n\in\mathbb{N}} \operatorname{cl} J_n(\nu)$, whence $J_\infty(\nu)$ is nonvoid by compactness. ∎

5.37. Remark. (i) If $\nu = (w_+^\infty, w_-^\infty)$ is periodic then there are two possibilities: If w_-^∞ and w_+^∞ are shifts of the same periodic itinerary then the sequence $(J_n(\nu))$ is eventually constant, namely $J_n(\nu)$ coincides with one of the frequency locked bubbles for large n and does not split any further. If this is not the case, then one can choose the open set G from the end of the proof as follows: Apply Proposition 5.27 to the hom-point of type (w_+, w_-) and let $G := U^{--}$. Inside G every Lorenz map has two periodic attractors of type w_- and w_+, respectively, and each one contains one of the critical points in its immediate basin, whence the kneading invariant equals (w_+^∞, w_-^∞) inside G.

(ii) In the expansive case we cheated a little bit when we said that $J_\infty(\nu)$ equals $\bigcap_{n\in\mathbb{N}} \operatorname{cl} J_n(\nu)$. There is one exception which is a little bit technical. It occurs precisely when the kneading invariant is properly (w_+, w_-)-renormalizable to a kneading invariant of a circle map with irrational rotation number, i.e., when $J_\infty(\nu)$ is one of the irrational “rays” between the rational “bubbles”. In this case for eventually all n either $J_n(\nu) = J_{n-1}(\nu)$ or $J_{n-1}(\nu)$ splits according to type D in Figure 5.10 and $J_n(\nu)$ is the top left or the bottom right piece. Thus the two sets $J_\infty(\nu)$ and $\bigcap_{n\in\mathbb{N}} \operatorname{cl} J_n(\nu)$ differ by one point, namely the hom-point of type (w_+, w_-) which is the common bottom left edge of eventually all sets $J_n(\nu)$. But this does not cause any problems here, because the local analysis of the gluing bifurcation shows that the sets $\operatorname{cl} J_n(\nu)$ cannot shrink down to that single point, whence $J_\infty(\nu) \neq \emptyset$. $\diamond$

5.6. Proof of the Full Family Theorem. As mentioned earlier, the main complications in the nonmonotonic case arise from the fact that the hom-lines can get tangled, which gives rise to complicated shapes and arrangements of the partition elements: It is possible that pieces $J_n(\nu) \in \mathcal{J}_n$ consist of several connected components and the connected components can have various shapes. Figure 5.11 shows an example how starting from the simple situation described in the last section—where the region $J_{n-1}(\nu)$ is a “square”—the situation can become very involved already within two induction steps. In order to adopt the proof from the last section one has to choose one of the connected components $G_n(\nu)$ of $J_n(\nu)$ instead of $J_n(\nu)$ itself. Here it turns out to be crucial to make the right choice: Some of the pieces are incomplete in the sense that they do not contain all possible kneading invariants which have the same n-prefix as ν. Whether or not the region is complete can be determined by having a look at the boundary of the region. For example, on the left hand side of Figure 5.11 there are two regions labeled 1 and $1'$ which belong to the same set $J_n(\nu)$ for some ν. But the boundary of region 1 has edges of four different types and vertices of four different types, whereas region $1'$ has only edges of two different types and its two vertices even have the same type. The right hand side of Figure 5.11, which shows one further step of the splitting, contains many more examples. The numbers of all bad regions are marked with primes.

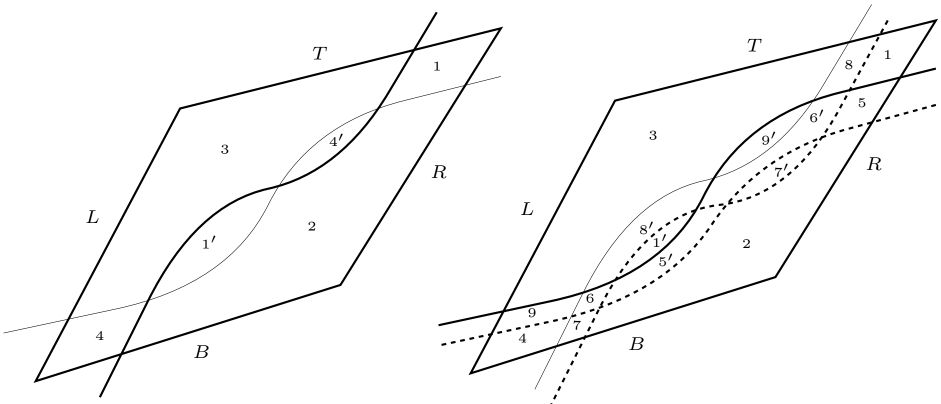


Fig. 5.11. The splitting of $J_{n-1}(\nu)$ in the general case. A region $J_{n-1}(\nu)$ splits in two steps into pieces of $\mathcal{J}_n$ (l.h.s.) and $\mathcal{J}_{n+1}$ (r.h.s.). Pieces with the same numbers are of the same type, but not all pieces are suited for the induction. The numbers of the “bad” pieces are marked with primes.

The following theorem generalizes Theorem 5.34 and finishes the proof of the Full Family Theorem 5.3 in the case of nonmonotonic $\mathcal{C}^1$-Lorenz families.

5.38. THEOREM. *Let $(f_{a,b})_{(a,b)\in\Lambda}$ be a $\mathcal{C}^1$-Lorenz family and let $\nu = (\nu^+, \nu^-)$ be an admissible kneading invariant. Then there is a decreasing sequence $(G_n(\nu))_{n\in\mathbb{N}}$ of nonvoid simply connected open regions $G_n(\nu) \subseteq J_n(\nu)$. The boundary of each $G_n(\nu)$ is contained in the union of four differentiable arcs where one of the critical points is either periodic of some period less than or equal to n or is mapped onto one of the repelling fixed points in the boundary.*

For every parameter (a,b) in $G_\infty(\nu) := \bigcap_{n\in\mathbb{N}} G_n(\nu)$ the critical points are not periodic and the Lorenz map $f_{a,b}$ has the kneading invariant ν. If ν is expansive or periodic then $G_\infty(\nu)$ is nonvoid.

The main work will be to formalize the intuitive notion of what are the good pieces and to prove that in every induction step there exists at least one such piece. Before doing that, we describe the possible shapes of an arbitrary component of $J_n(\nu)$.

5.39. LEMMA. *Let $G_n(\nu)$ be an arbitrary connected component of $J_n(\nu) \in \mathcal{J}_n$. Then $G_n(\nu)$ can be described in (x,t)-coordinates as follows*:

There exist parameters $t_0 < t_1$ and four differentiable arcs which are the graphs of functions ϕ_B, ϕ_T, ϕ_L, $\phi_R : [t_0, t_1] \to \mathbb{R}$ satisfying $\phi_B \vee \phi_R < \phi_L \wedge \phi_T$ on the open interval (t_0, t_1) and $\phi_B \vee \phi_R = \phi_L \wedge \phi_T$ on the boundary $\{t_0, t_1\}$ such that

1. *$G_n(\nu)$ is the set between the graphs of $\phi_B \vee \phi_R$ and $\phi_L \wedge \phi_T$:*

$$G_n(\nu) = \{(x,t) \mid t \in (t_0, t_1),\ (\phi_B \vee \phi_R)(t) < x < (\phi_L \wedge \phi_T)(t)\}.$$

2. *For $X \in \{L, R, B, T\}$ let $t_0^X < t_1^X$ be the first and last moment when X intersects $\operatorname{cl} G_n(\nu)$ and denote the graph of ϕ_X between t_0^X and t_1^X by X. Then X is part of a hom-line*, *namely*

- *c^+ is homoclinic of type w_+ on L and homoclinic of type $\widetilde{w}_+$ on R.*
- *c^- is homoclinic of type w_- on L and homoclinic of type $\widetilde{w}_-$ on T.*

Here $w_\pm$ is the $S^\pm\langle n\rangle$-prefix and $\widetilde{w}_\pm$ is the $\widetilde{S}^\pm\langle n\rangle$-prefix of $\nu^\pm$.

In particular, the two vertices $v_i := (\phi_B \vee \phi_R)(t_i) = (\phi_L \wedge \phi_T)(t_i)$, $i = 0, 1$, are hom-points. Since $\operatorname{dist}(B,T) > 0$ and $\operatorname{dist}(L,R) > 0$ it follows that either $v_i = \phi_B(t_i) = \phi_L(t_i)$ or $v_i = \phi_T(t_i) = \phi_R(t_i)$. In the former case we call v_i a BL-vertex, and in the latter case we call it a TR-vertex. These two vertices are the only two points where $L \cup T$ and $B \cup R$ intersect. The intersection of T- and L-lines (resp. B- and R-lines) may be more complicated, but the first and last point of intersection—if they intersect at all—are well defined hom-points which are distinct from v_0 and v_1. We call them the TL_1- and TL_2-vertex (resp. BR_1- and BR_2-vertex) (cf. left hand side of Figure 5.12).

Proof of Lemma 5.39. The proof proceeds by induction and the arguments are very similar to the ones used for the proof of Theorem 5.34. There are four cases to distinguish according to the possible types of v_0 and v_1 and in each case there are four possible types of splittings corresponding to the cases A–D in the previous subsection. Since we do not

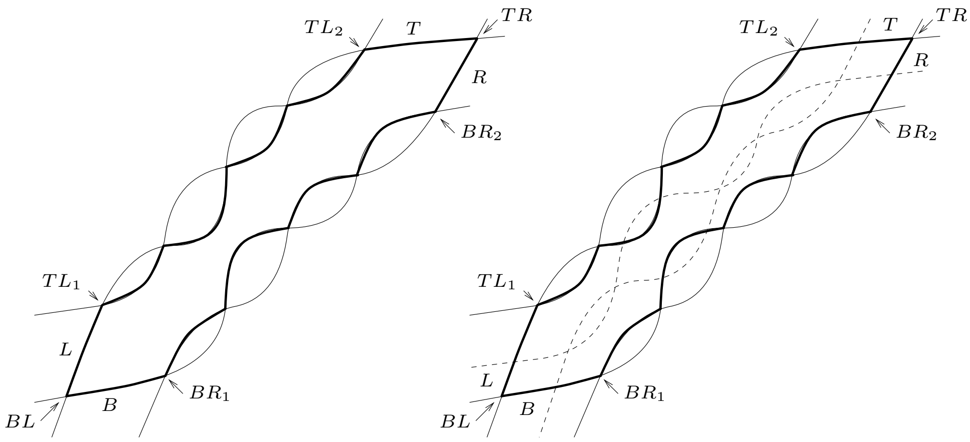


Fig. 5.12. The splitting of a set $G_{n-1}(\nu)$ with winding number one

want to bore the reader with repetitions, we discuss only the case where v_0 is a BL- and v_1 is a TR-vertex which is the one we are mainly interested in.

Assume that $G_{n-1}(\nu)$ splits according to type C (cf. right hand side of Figure 5.12), the other cases are analogous. As before one shows—using (5.10) and the local analysis at the corresponding vertices—that there is a hom$^+$-curve entering $G_{n-1}(\nu)$ through the B-line between the vertices BL and BR_1 and a hom$^+$-curve of the same type leaving $G_{n-1}(\nu)$ through the T-line between the vertices TL_2 and TR.

An important observation is that these two hom$^+$-curves are in fact the same, let's call it the V-line (= "vertical"). This V-line must initially enter $G_{n-1}(\nu)$ through the B-line between the vertices BL and BR_1 and eventually leave $G_{n-1}(\nu)$ through the T-line between the vertices TL_2 and TR. In between, it may leave $G_{n-1}(\nu)$ temporarily through the B- or T-line but it is not allowed to intersect the L- and R-line, because they are also hom$^+$-curves. So even if the V-line leaves $G_{n-1}(\nu)$ it is still caught between the L- and R-lines, which prevents it from being annihilated and recreated outside $G_{n-1}(\nu)$ before reentering. An analogous statement holds for the hom$^-$-curve, which we call the H-line (= "horizontal"). ∎

If the vertices v_0 and v_1 of $G_{n-1}(\nu)$ are of different type then $G_{n-1}(\nu)$ has vertices of all four types. In particular, $G_{n-1}(\nu)$ contains a component of $J_n(\nu)$. So the good pieces $G_{n-1}(\nu)$ are the ones where v_0 and v_1 are of different type. What remains to show is that a good piece $G_{n-1}(\nu)$ always contains at least one good piece $G_n(\nu)$. In order to prove this we introduce the notion of a winding number for the sets $G_n(\nu)$.

5.40. Definition (Winding number). Let $U := (0,1)\times(0,1)$ be the unit square and denote the left, right, bottom and top edge of its boundary by l, r, b and t, respectively. Furthermore, let $G_n(\nu)$ be a connected component of $J_n(\nu)$ and let $\phi : \operatorname{cl} G_n(\nu) \to \operatorname{cl} U$ be a continuous map which maps $G_n(\nu)$ into U and $\operatorname{bd} G_n(\nu)$ into $\operatorname{bd} U$ in such a way that $\phi(X \cap \operatorname{bd} G_n(\nu)) \subseteq x$ for $(X,x) \in \{(B,b),(L,l),(T,t),(R,r)\}$. Then ϕ is called a *winding map* for $G_n(\nu)$ and $\deg \phi|_{\operatorname{bd} G_n(\nu)}$ is called the *winding number* of $G_n(\nu)$.

5.41. Remark. Such a winding map exists, since $\operatorname{dist}(B,T) > 0$ and $\operatorname{dist}(L,R) > 0$. One possible choice is

$$\phi(\lambda) := \left(\frac{\operatorname{dist}(\lambda,L)}{\operatorname{dist}(\lambda,L)+\operatorname{dist}(\lambda,R)}, \frac{\operatorname{dist}(\lambda,B)}{\operatorname{dist}(\lambda,B)+\operatorname{dist}(\lambda,T)}\right) \qquad \text{for } \lambda \in \operatorname{cl} G_n(\nu).$$

But the explicit form is not important, because if ϕ and ϕ' are two winding maps then $\phi|_{\operatorname{bd} G}$ and $\phi'|_{\operatorname{bd} G}$ are homotopic. In particular, the winding number of G is well defined. $\diamond$

5.42. Lemma. *The winding number of $G_n(\nu)$ equals either* 1, 0 *or* -1. *It equals* 1 *if v_0 is a BL-vertex and v_1 is a TR-vertex,* -1 *if v_0 is a TR-vertex and v_1 is a BL-vertex, and* 0 *if v_0 and v_1 are vertices of the same type.*

Proof. The points $(0,0)$ and $(1,1)$ have at most one preimage, namely one of the two vertices v_0 and v_1. This implies $|\deg\phi| \le 1$. If both vertices are mapped to the same point then ϕ is obviously not surjective and the degree is zero. If the two vertices are mapped to different points then the degree equals 1 if $\phi(v_0) = (0,0)$ and $\phi(v_1) = (1,1)$ and it equals -1 if $\phi(v_0) = (1,1)$ and $\phi(v_1) = (0,0)$. ∎

5.43. Lemma. *Let ν be an admissible kneading invariant and assume that $G_{n-1}(\nu)$ is a component of $J_{n-1}(\nu)$ with nonzero winding number. Then there exists a component $G_n(\nu)$ of $J_n(\nu)$ with nonzero winding number which is contained in $G_{n-1}(\nu)$.*

Proof. Let $\phi : \operatorname{cl} G_{n-1}(\nu) \to \operatorname{cl} U$ be a winding map for $G_{n-1}(\nu)$ with the additional property that the preimages of $\{1/2\} \times (0,1)$ and $(0,1) \times \{1/2\}$ are precisely the points of $G_{n-1}(\nu)$ which lie on the V-line and H-line, respectively. If there is no H- or V-line then the corresponding requirement is void, yet for definiteness assume that both lines exist. Then ϕ maps every connected component $G_n(\nu)$ of $J_n(\nu) \cap G_{n-1}(\nu)$ into the same subsquare U' of U of side length $1/2$ (which depends only on ν_n^+ and ν_n^-) and the map $\phi|_{\operatorname{cl} G_n(\nu)}$ is a winding map for $G_n(\nu)$ up to a rescaling of the image. For every component $G_n(\nu)$ which has winding number 0 there is a homotopy $\psi'_t : \operatorname{cl} G_n(\nu) \to \operatorname{cl} U'$ fixing the points in $\operatorname{bd} G_n(\nu)$ such that $\psi'_0 = \phi|_{\operatorname{cl} G_n(\nu)}$ and $\psi'_1(G_n(\nu)) \subseteq \operatorname{bd} U'$.

Now if all components of $J_n(\nu) \cap G_{n-1}(\nu)$ have winding number zero, then one can choose such a homotopy for every component and by trivial extension one obtains a homotopy $\psi_t : \operatorname{cl} G_{n-1}(\nu) \to \operatorname{cl} U$ fixing the points in $\operatorname{bd} G_{n-1}(\nu)$ such that $\psi_0 = \phi|_{\operatorname{cl} G_{n-1}(\nu)}$ and $\psi_1(G_{n-1}(\nu)) \subseteq U \setminus U'$. This contradicts the fact that $\deg\phi \neq 0$. ∎

5.7. The quadratic Lorenz family. We end this chapter with some remarks on the quadratic Lorenz family

$$f_{a,b} : x \mapsto \begin{cases} -a + x^2 & \text{if } x > 0, \\ b - x^2 & \text{if } x < 0. \end{cases} \tag{5.2}$$

Recall Figure 1.2 which shows the bifurcation diagram of this family. Looking at the figure one gets the strong impression that the quadratic Lorenz family is a monotonic family in the sense of Definition 5.32. In view of Corollary 5.36 this yields the following conjecture.

5.44. Conjecture. *The kneading invariant of the quadratic Lorenz family* (5.2) *depends monotonically on the parameters.*

In the following we want to demonstrate that there is a close connection between this conjecture and another one made earlier, namely the conjecture that the Thurston map for the quadratic family is a contraction.

Before being more precise about that, let us recall the situation in the case of unimodal maps. For the quadratic unimodal family $f_a : x \mapsto a - x^2$ it was shown by Douady, Hubbard & Sullivan ([15]) that the kneading invariant depends monotonically on the parameter. The main part of the proof consists in showing that any two real quadratic unimodal maps with periodic critical points have the same kneading invariant if and only if they are linearly conjugate. This is done by defining the complex Thurston map on an appropriate Teichmüller space which is related to the combinatorial type of this periodic orbit and to show that this Thurston map contracts the Teichmüller distance, whence it has a unique fixed point. From this fixed point one obtains a quadratic unimodal map with the required dynamics which is unique up to conjugacy by (affine) automorphisms of the complex plane. Now if the kneading invariant did not depend monotonically on the parameters then one could find two distinct maps in this family with periodic critical points of the same combinatorial type, a contradiction since different maps from $(f_a)_{a\in\mathbb{R}}$ are not linearly conjugate.

Recently, Tsujii [58, 59] gave a simplified proof for the monotonicity in the quadratic family based on the following two observations.

1. The question whether the kneading invariant depends monotonically on the parameters can be reduced to a local problem, namely to showing the following: If the critical point is periodic of period $n+1$ then

$$\frac{\mathrm{D}_a c_{n+1}(a)}{(f^n)'(c_1(a))} > 0. \tag{5.11}$$

The interpretation of this property is that whenever the critical point is periodic and the parameter a is increased, $c_{n+1}(a)$ moves towards the side of c corresponding to the larger kneading invariant.

2. Local monotonicity at a periodic parameter follows if the Thurston map corresponding to the periodic orbit is locally contracting: If $\mathrm{D}T(c_0,\dots,c_n)$ denotes the linearization of the Thurston map at the fixed point $(c_0,\dots,c_n)$ corresponding to the periodic orbit, then

$$\frac{\mathrm{D}_a c_{n+1}(a)}{(f^n)'(c_1(a))} = \det(\mathrm{I} - \mathrm{D}T(c_0,\dots,c_n)). \tag{5.12}$$

Since the characteristic polynomial $\det(\lambda\,\mathrm{I} - \mathrm{D}T(c_0,\dots,c_n))$ is a real polynomial in λ which diverges to $+\infty$ as $\lambda \to +\infty$, inequality (5.11) follows from (5.12) if the linearized Thurston map is a contraction.

In the following we are going to derive an equation analogous to (5.12) for a homoclinic Lorenz map f_{a_0,b_0} where both critical points are homoclinic (= periodic). It has the consequence that if the Thurston map is locally a contraction (near the fixed point corresponding to the post-critical orbit of f_{a_0,b_0}) then (a_0,b_0) is a positively oriented

([15]) Unpublished. For a proof see the monograph of Milnor & Thurston [51].

hom-point (cf. Definition 5.24). In view of Lemma 5.25 this conclusion has a similar dynamical interpretation to the one we gave for (5.11): If both critical points of a map f_{a_0,b_0} are homoclinic then these two homoclinic orbits can be unfolded in an orientation preserving way by a small perturbation of the parameter (a_0, b_0).

Now let us be more precise and give a definition of the complex Thurston map. For practical reasons the notation will be slightly different from the one in Section 5.1. First of all, we use a different parametrization of the quadratic Lorenz family, namely

$$f_{a,b} : x \mapsto \begin{cases} a + f_+(x) & \text{if } x > 0, \\ b + f_-(x) & \text{if } x < 0, \end{cases} \tag{5.13}$$

with $f_\pm : \mathbb{C} \to \mathbb{C}$, $z \mapsto \pm z^2$. We do this in order to simplify the notation in the proofs, although in general we prefer the original definition (5.2) for reasons of symmetry.

Fix some parameter (a_0, b_0) where the Lorenz map f_{a_0,b_0} is homoclinic of some type $w := (w_+, w_-)$ with $|w_+| = n + 1$, $|w_-| = m + 1$ ([16]) and let $(\nu^+, \nu^-) = (w_+^\infty, w_-^\infty)$ be the kneading invariant of f_{a_0,b_0}.

5.45. Definition (Thurston map). Let $g_\pm$ be the one of the two inverse branches of $f_\pm$ for which $g_\pm \circ f_\pm = \mathrm{id}$ on $\mathbb{R}_\pm$. Then the *Thurston map* $T : U \to \mathbb{C}^{n+m}$ is defined on a neighbourhood $U \subset \mathbb{C}^{n+m}$ of $\widetilde{c}_0 := \big(c_1^+, c_2^+, \ldots, c_n^+, c_1^-, c_2^-, \ldots, c_m^-\big)(a_0, b_0)$ as follows.

$$\begin{aligned} T(z) &:= T(z_1^+, z_2^+, \ldots, z_n^+, z_1^-, z_2^-, \ldots, z_m^-) \\ &:= (g_{\nu_1^+}(z_2^+ - z_1^{\nu_1^+}),\ g_{\nu_2^+}(z_3^+ - z_1^{\nu_2^+}),\ \ldots,\ g_{\nu_n^+}(z_n^+ - z_1^{\nu_n^+}), \\ &\qquad g_{\nu_1^-}(z_2^- - z_1^{\nu_1^-}),\ g_{\nu_2^-}(z_3^- - z_1^{\nu_2^-}),\ \ldots,\ g_{\nu_m^-}(z_m^- - z_1^{\nu_m^-})). \end{aligned}$$

Here the kneading invariants are considered as sequences over the alphabet $\{-, +\}$ instead of $\{0, 1\}$.

By construction the point $\widetilde{c}_0$ is fixed by T. Except for the ordering of the coordinates, T is just the complex extension of the Thurston map defined in Section 5.1 to a neighbourhood of the fixed point $\widetilde{c}_0$.

We look at the linearization $\mathrm{D}T(\widetilde{c}_0)$ of T at the fixed point and relate its characteristic polynomial $p_{n,m}(\lambda) := \det(\lambda \mathrm{I} - \mathrm{D}T(\widetilde{c}_0))$ to the orientation of the hom-point (a_0, b_0). The orientation of (a_0, b_0) is determined by the sign of the determinant (5.9), which now has changed due to the re-parametrization $-a \to a$. We compensate this change by swapping the two rows of the matrix (5.9). Moreover, we divide the rows by $\mathrm{D}_x f^n(c_1^+)$ and $\mathrm{D}_x f^m(c_1^-)$, respectively, which is possible because $\mathrm{D}_x f^n(c_1^+) > 0$ and $\mathrm{D}_x f^m(c_1^-) > 0$ at the parameter value (a_0, b_0) by minimality of m and n. With all the changes introduced above a hom-point now is positively oriented if

$$V_{n,m}(a_0, b_0) > 0 \quad \text{where} \quad V_{n,m} := \det \begin{pmatrix} \dfrac{\mathrm{D}_a c_{n+1}^+}{\mathrm{D}_x f^n(c_1^+)} & \dfrac{\mathrm{D}_b c_{n+1}^+}{\mathrm{D}_x f^n(c_1^+)} \\ \dfrac{\mathrm{D}_a c_{m+1}^-}{\mathrm{D}_x f^m(c_1^-)} & \dfrac{\mathrm{D}_b c_{m+1}^-}{\mathrm{D}_x f^m(c_1^-)} \end{pmatrix}. \tag{5.14}$$

Now we are able to state the analogue to (5.12).

([16]) Note that the periods of c^+ and c^- are $n+1$ and $m+1$ instead of n and m, respectively.

5.46. Proposition. *Let $V_{n,m}$ and $p_{n,m}(\lambda)$ be defined as above. Then*

$$V_{n,m} = p_{n,m}(1) = \det(\mathrm{I} - \mathrm{D}T(\widetilde{c}_0)).$$

Proof. In Lemmas 5.47 and 5.48 explicit expressions for $V_{n,m}$ and $p_{n,m}(\lambda)$ are derived. Comparing (5.15) and (5.16) one immediately verifies the claim. ∎

5.47. Lemma.

$$V_{n,m} = \sum_{i=0}^{n}\sum_{j=0}^{m}(\delta_i^{++}\delta_j^{--} - \delta_i^{+-}\delta_j^{-+})\frac{1}{\mathrm{D}_x f^i(c_1^+)} \cdot \frac{1}{\mathrm{D}_x f^j(c_1^-)}, \tag{5.15}$$

where $\delta_i^{st} := 1_{\{\nu_i^s = t\}}$ for $s,t \in \{+,-\}$.

Proof. By the chain rule $\mathrm{D}_a c_{n+1}^+ = \mathrm{D}_a f(c_n^+) + \mathrm{D}_x f(c_n^+) \cdot \mathrm{D}_a c_n^+$, and induction yields

$$\mathrm{D}_a c_{n+1}^+ = \sum_{i=0}^{n} \mathrm{D}_a f(c_i^+) \prod_{k=i+1}^{n} \mathrm{D}_x f(c_k^+) = \sum_{i=0}^{n} \delta_i^{++} \prod_{k=i+1}^{n} \mathrm{D}_x f(c_k^+).$$

For the last equation we used the fact that $\mathrm{D}_a f(x) = 1_{\{x>0\}}$. Now

$$\frac{\mathrm{D}_a c_{n+1}^+}{\mathrm{D}_x f^n(c_1^+)} = \sum_{i=0}^{n} \delta_i^{++} \prod_{k=1}^{i} \mathrm{D}_x f(c_k^+) = \sum_{i=0}^{n} \delta_i^{++} \mathrm{D}_x f^i(c_1^+).$$

The calculations for the other entries of the matrix are similar and we obtain

$$V_{n,m} = \det \begin{pmatrix} \sum_{i=0}^{n} \delta_i^{++} \frac{1}{\mathrm{D}_x f^i(c_1^+)} & \sum_{i=0}^{n} \delta_i^{+-} \frac{1}{\mathrm{D}_x f^i(c_1^+)} \\ \sum_{j=0}^{m} \delta_j^{-+} \frac{1}{\mathrm{D}_x f^j(c_1^-)} & \sum_{j=0}^{m} \delta_j^{--} \frac{1}{\mathrm{D}_x f^j(c_1^-)} \end{pmatrix}. \quad ∎$$

5.48. Lemma.

$$p_{n,m}(\lambda) = \sum_{i=0}^{n}\sum_{j=0}^{m}(\delta_i^{++}\delta_j^{++} - \delta_i^{+-}\delta_j^{-+})\frac{\lambda^{n-i}}{\mathrm{D}_x f^i(c_1^+)} \cdot \frac{\lambda^{m-j}}{\mathrm{D}_x f^j(c_1^-)}, \tag{5.16}$$

where $\delta_i^{st} := 1_{\{\nu_i^s = t\}}$ for $s,t \in \{+,-\}$.

Proof. First let us see what the matrix representation of $\mathrm{D}T$ looks like. Therefore, let $N^+ \cup M^- := \{1^+ < \ldots < n^+ < 1^- < \ldots < m^-\}$ be the index set for the canonical base of $\mathbb{C}^{n+m}$ and let δ_{st} denote Dirac's delta function for $s,t \in N^+ \cup M^-$. Then the partial derivatives are

$$\mathrm{D}_{j^\pm} T_{i^+}(z) = \mathrm{D}_x g_{\nu_i^+}(z_{i+1}^+ - z_1^{\nu_i^+}) \cdot (\delta_{j^\pm,(i+1)^+} - \delta_{j^\pm,1^{\nu_i^+}}),$$

$$\mathrm{D}_{j^\pm} T_{i^-}(z) = \mathrm{D}_x g_{\nu_i^-}(z_{i+1}^- - z_1^{\nu_i^-}) \cdot (\delta_{j^\pm,(i+1)^-} - \delta_{j^\pm,1^{\nu_i^-}}).$$

Using the fact that $\mathrm{D}_x g_{\nu_i^\pm}(c_{i+1}^\pm - c_1^{\nu_i^\pm}) = \mathrm{D}_x g_{\nu_i^\pm}(f_{\nu_i^\pm}(c_i^\pm)) = 1/\mathrm{D}_x f(c_i^\pm)$ we get the matrix representation

$$\lambda I - \mathrm{D}T(\widetilde{c}_0) = \left(\begin{array}{ccccc|ccccc}
\lambda & -t_1^+ & & & & t_1^+ & & & & \\
*_2^+ & \lambda & -t_2^+ & & & *_2^+ & & & & \\
\cdot & & \lambda & \ddots & & \cdot & & & & \\
\cdot & & & \lambda & -t_{n-1}^+ & \cdot & & & & \\
*_n^+ & & & & \lambda & *_n^+ & & & & \\
\hline
t_1^- & & & & & \lambda & -t_1^- & & & \\
*_2^- & & & & & *_2^- & \lambda & -t_2^- & & \\
\cdot & & & & & \cdot & & \lambda & \ddots & \\
\cdot & & & & & \cdot & & & \lambda & -t_{m-1}^- \\
*_m^- & & & & & *_m^- & & & & \lambda
\end{array}\right)$$

where we abbreviated $t_i^\pm := \frac{1}{\mathrm{D}_x f(c_i^\pm)}(a_0, b_0)$. Precisely one of the two entries $*_i^\pm$ equals $t_i^\pm$, the other one is zero, depending on the kneading invariant.

In order to calculate the determinant of this matrix we develop it with respect to the columns 1^+ and 1^- simultaneously. For this we need to introduce some notation. Let $A := \lambda I - \mathrm{D}T(\widetilde{c}_0)$ and for $R, C \subseteq N^+ \cup M^-$ let $A_{R,C}$ (resp. $\widetilde{A}_{R,C}$) be the submatrix obtained from A using only the rows and columns contained (resp. not contained) in R and C. Define $C := \{1^+, 1^-\}$ and let $\mathcal{R}$ be the collection of all subsets of $N^+ \cup M^-$ of cardinality two. Finally, let $B_{ij} := A_{\{i^+,j^-\},C}$ and $\widetilde{B}_{ij} := \widetilde{A}_{\{i^+,j^-\},C}$. Then

$$\begin{aligned}
\det A &= \operatorname{sgn} C \cdot \sum_{R\in\mathcal{R}} \operatorname{sgn} R \cdot \det A_{R,C} \cdot \det \widetilde{A}_{R,C} \ (^{17}) \qquad (5.17)\\
&= \operatorname{sgn} C \cdot \sum_{i=1}^{n}\sum_{j=1}^{m} \operatorname{sgn}\{i^+, j^-\} \cdot \det A_{\{i^+,j^-\},C} \cdot \det \widetilde{A}_{\{i^+,j^-\},C}\\
&= (-1)^{n-1} \sum_{i=1}^{n}\sum_{j=1}^{m} (-1)^{(i-1)+(n+j-2)} \det B_{ij} \cdot \det \widetilde{B}_{ij}\\
&= (-1)^{i+j} \sum_{i=1}^{n}\sum_{j=1}^{m} \det B_{ij} \cdot \det \widetilde{B}_{ij}.
\end{aligned}$$

For the second equality observe that $\det \widetilde{A}_{R,C}$ vanishes if $R \subseteq N^+$ or $R \subseteq M^-$. The remaining subdeterminants are

$$\begin{aligned}
\det B_{ij} &= \det \begin{pmatrix} \lambda\delta_{1,i} + t_i^+\delta_i^{++} & t_i^+\delta_i^{+-} \\ t_j^-\delta_j^{-+} & \lambda\delta_{1,j} + t_j^-\delta_j^{--} \end{pmatrix}\\
&= t_i^+ t_j^- (\delta_i^{++}\delta_j^{--} - \delta_j^{-+}\delta_i^{+-})\\
&\quad + \lambda\delta_{1,i} t_j^- \delta_j^{--} + \lambda\delta_{1,j} t_i^+ \delta_i^{++} + \lambda^2 \delta_{1,i}\delta_{1,j}
\end{aligned}$$

([17]) The sign of a subset $R \subset N^+ \cup M^-$ is the sign of the permutation of $N^+ \cup M^-$ which moves the elements of R all the way to the left, preserving the respective orders of R and its complement.

and

$$\begin{aligned}\det \widetilde{B}_{ij} &= \prod_{k=1}^{i-1}(-t_k^+)\cdot\lambda^{n-i}\prod_{k=1}^{j-1}(-t_k^-)\cdot\lambda^{m-j}\\ &= (-1)^{i+j-2}\prod_{k=1}^{i-1}t_k^+\cdot\prod_{k=1}^{j-1}t_k^-\cdot\lambda^{m+n-i-j}.\end{aligned}$$

Inserting the expressions for $\det B_{ij}$ and $\det \widetilde{B}_{ij}$ into (5.17) yields

$$\begin{aligned}\det A = &\sum_{i=1}^{n}\sum_{j=1}^{m}\prod_{k=1}^{i}t_k^+\cdot\prod_{k=1}^{j}t_k^-\cdot\lambda^{n-i}\lambda^{m-j}(\delta_i^{++}\delta_j^{--}-\delta_j^{-+}\delta_i^{+-})\\ &+\sum_{j=1}^{m}\prod_{k=1}^{0}t_k^+\cdot\prod_{k=1}^{j}t_k^-\cdot\lambda^{n-0}\lambda^{m-j}\delta_j^{--}\\ &+\sum_{i=1}^{n}\prod_{k=1}^{i}t_k^+\cdot\prod_{k=1}^{0}t_k^-\cdot\lambda^{n-i}\lambda^{m-0}\delta_i^{++}\\ &+\prod_{k=1}^{0}t_k^+\cdot\prod_{k=1}^{0}t_k^-\cdot\lambda^{n-0}\lambda^{m-0}\\ =&\sum_{i=0}^{n}\sum_{j=0}^{m}\prod_{k=1}^{i}t_k^+\cdot\prod_{k=1}^{j}t_k^-\cdot\lambda^{n-i}\lambda^{m-j}(\delta_i^{++}\delta_j^{--}-\delta_j^{-+}\delta_i^{+-}). \blacksquare\end{aligned}$$

Since $p_{n,m}(\lambda)$ is a real polynomial with positive leading coefficient, the Monotonicity Conjecture 5.44 now can be reduced to the following conjecture:

5.49. CONJECTURE. *At every hom-point (a_0, b_0) the spectrum of the linearized Thurston map at $\widetilde{c}_0$ is strictly contained in the unit disk, where T and $\widetilde{c}_0$ are defined as before.*

Originally, it was our hope to be able to prove the local contraction of the Thurston map with similar methods to the ones used by Tsujii [58, 59] (see also Hubbard & Schleicher [34]), but unfortunately this turned out to be not so easy. Although the above question is of local nature, its solution turns into a global problem as soon as one tries to apply the complex arguments due to their inherent global nature. Here it now becomes a fundamental problem that the Thurston map is defined using two different holomorphic maps for the pullback of points. Nevertheless, we think that the above considerations are not completely useless, because they link two conjectures together, the conjecture about the contraction of the Thurston map and the conjecture about the monotonicity of the kneading invariant for the quadratic family. Since we have numerical evidence for both conjectures independently, we believe very strongly that they are true.

References

[1] D. Berry and B. D. Mestel, *Wandering intervals for Lorenz maps with bounded variation*, Bull. London Math. Soc. 23 (1991), 183–189.

[2] A. M. Blokh and M. Y. Lyubich, *Attractors of maps of the interval*, in: Banach Center Publ. 23, PWN, Warszawa, 1989, 427–442.

[3] —, —, *On decomposition of one-dimensional dynamical systems into ergodic components. The case of negative Schwarzian*, Leningrad Math. J. 1 (1990), no. 1, 137–155.

[4] H. Bruin, *Invariant measures of interval maps*, PhD thesis, Technische Universiteit Delft, 1994.

[5] —, *Combinatorics of the kneading map*, Internat. J. Bifurcation Chaos 5 (1995), 1339–1349.

[6] —, *Induced maps, Markov extensions and invariant measures in one-dimensional dynamics*, Comm. Math. Phys. 168 (1995), 571–580.

[7] —, *Topological conditions for the existence of absorbing Cantor sets*, Trans. Amer. Math. Soc. 350 (1998), 2229–2263.

[8] H. Bruin, G. Keller, T. Nowicki, and S. van Strien, *Wild Cantor attractors exist*, Ann. of Math. 143 (1996), 97–130.

[9] P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhäuser, Boston, 1980.

[10] P. Coullet, J.-M. Gambaudo et C. Tresser, *Une nouvelle bifurcation de codimension 2: le collage de cycles*, C. R. Acad. Sci. Paris Sér. I 299 (1984), 253–256.

[11] P. Coullet et C. Tresser, *Itérations d'endomorphismes et groupe de renormalisation*, ibid. 287 (1978), 577–580.

[12] —, —, *Itérations d'endomorphismes et groupe de renormalisation*, J. Phys. C5 (1978), 25–28.

[13] W. de Melo and S. van Strien, *One-Dimensional Dynamics*, Ergeb. Math. Grenzgeb. 25, Springer, Berlin, 1993.

[14] A. Douady and J. H. Hubbard, *A proof of Thurston's topological characterization of rational functions*, Acta Math. 171 (1993), 263–297.

[15] J.-M. Gambaudo, *Ordre, désordre, et frontière des systèmes Morse–Smale*, thèse d'Etat, l'Université de Nice, 1987.

[16] J.-M. Gambaudo *et al.*, *New universal scenarios for the onset of chaos in Lorenz-type flows*, Phys. Rev. Lett. 57 (1986), 925–928.

[17] J.-M. Gambaudo, P. A. Glendinning, and C. Tresser, *Stable cycles with complicated structure*, J. Phys. Lett. 46 (1985), 653–657.

[18] —, —, —, *Collage de cycles et suites de Farey*, C. R. Acad. Sci. Paris Sér. I 299 (1984), 711–714.

[19] J.-M. Gambaudo, O. Lanford III et C. Tresser, *Dynamique symbolique des rotations*, C. R. Acad. Sci. Paris Sér. I 299 (1984), 823–826.

[20] P. Glendinning and C. Sparrow, *Prime and renormalisable kneading invariants and the dynamics of expanding Lorenz maps*, Phys. D 62 (1993), 22–50.

[21] J. Guckenheimer, *A strange, strange attractor*, in: J. E. Marsden and M. McCracken (eds.), The Hopf Bifurcation Theorem and its Applications, Springer, 1976, 368–381.

[22] J. Guckenheimer and R. F. Williams, *Structural stability of Lorenz attractors*, Publ. Math. I.H.E.S. 50 (1979), 307–320.

[23] F. Hofbauer, *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy*, Israel J. Math. 34 (1979), 213–237.

[24] —, *The topological entropy of the transformation* $x \mapsto ax(1-x)$, Monatsh. Math. 90 (1980), 117–141.

[25] —, *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy II*, Israel J. Math. 38 (1981), 107–115.

[26] —, *The structure of piecewise monotonic transformations*, Ergodic Theory Dynam. Systems 1 (1981), 159–178.

[27] —, *Monotonic mod one transformations*, Studia Math. 80 (1984), 17–40.

[28] —, *Piecewise invertible dynamical systems*, Probab. Theory Related Fields 72 (1986), 359–386.

[29] F. Hofbauer and G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. 180 (1982), 119–140.

[30] —, —, *Quadratic maps without asymptotic measure*, Comm. Math. Phys. 127 (1990), 319–337.

[31] —, —, *Some remarks on recent results about S-unimodal maps*, Ann. Inst. H. Poincaré 53 (1990), 413–425.

[32] F. Hofbauer and P. Raith, *Topologically transitive subsets of piecewise monotonic maps, which contain no periodic points*, Monatsh. Math. 107 (1989), 217–239.

[33] A. Homburg, *Some global aspects of homoclinic bifurcations of vector fields*, PhD thesis, Rijksuniversiteit Groningen, 1993.

[34] J. H. Hubbard and D. Schleicher, *The spider algorithm*, in: Proc. Sympos. Appl. Math. 49, Amer. Math. Soc., 1994, 155–180.

[35] J. H. Hubbard and C. T. Sparrow, *The classification of topologically expansive Lorenz maps*, Comm. Pure Appl. Math. 43 (1990), 431–443.

[36] J. P. Keener, *Chaotic behaviour in piecewise continuous difference equations*, Trans. Amer. Math. Soc. 261 (1980), 589–604.

[37] G. Keller, *Lifting measures to Markov extensions*, Monatsh. Math. 108 (1989), 183–200.

[38] —, *Markov extensions, zeta functions, and Fredholm theory for piecewise invertible dynamical systems*, Trans. Amer. Math. Soc. 314 (1989), 433–497.

[39] —, *Exponents, attractors and Hopf decompositions for interval maps*, Ergodic Theory Dynam. Systems 10 (1990), 717–744.

[40] —, *Zeta functions and transfer operators for piecewise monotone transformations*, Comm. Math. Phys. 127 (1990), 459–477.

[41] G. Keller and T. Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet–Eckmann maps*, Comm. Math. Phys. 149 (1992), 31–69.

[42] U. Krengel, *Ergodic Theorems*, de Gruyter Stud. Math. 6, de Gruyter, Berlin, 1985.

[43] M. Lin, *Mixing for Markov operators*, Z. Wahrsch. Verw. Gebiete 19 (1971), 231–242.

[44] E. N. Lorenz, *Deterministic non-periodic flow*, J. Atmospheric Sci. 20 (1963), 130–141.

[45] M. I. Malkin, *On continuity of entropy of discontinuous mappings of the interval*, Selecta Math. Sov. 8 (1989), 131–139.

[46] M. I. Malkin, *Rotation intervals and the dynamics of Lorenz type mappings*, ibid. 10 (1991), 265–275.

[47] M. Martens and W. de Melo, *Universal models for Lorenz maps*, preprint, Oct. 1996.

[48] M. Martens, W. de Melo, P. Mendez, and S. van Strien, *Cherry flows on the torus: towards a classification*, Ergodic Theory Dynam. Systems 10 (1990), 531–553.

[49] W. S. Massey, *A Basic Course in Algebraic Topology*, Grad. Texts in Math. 127, Springer, Berlin, 1991.

[50] J. Milnor, *On the concept of attractor*, Comm. Math. Phys. 99 (1985), 177–195.

[51] J. Milnor and W. Thurston, *On Iterated Maps of the Interval*, Lecture Notes in Math. 1342, Springer, Berlin, 1988.

[52] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Publ. Math. I.H.E.S. 53 (1981).

[53] W. Parry, *The Lorenz attractor and a related population model*, in: Lecture Notes in Math. 729, Springer, Berlin, 1979, 169–187.

[54] D. Rand, *The topological classification of Lorenz attractors*, Math. Proc. Cambridge Philos. Soc. 83 (1978), 451–460.

[55] D. Sands, *Topological conditions for positive Lyapunov exponent in unimodal maps*, PhD thesis, St. John's College, Cambridge, 1994.

[56] H. Thunberg, *Some problems in unimodal dynamics*, PhD thesis, KTH, Stockholm, 1996.

[57] C. Tresser, *Nouveaux types de transitions vers une entropie topologique positive*, C. R. Acad. Sci. Paris Sér. I 296 (1983), 729–732.

[58] M. Tsujii, *A note on Milnor and Thurston's monotonicity theorem*, in: Geometry and Analysis in Dynamical Systems (Kyoto, 1993), Adv. Ser. Dynam. Systems 14, World Sci., 1994, 60–62.

[59] —, *A simple proof of monotonicity of entropy in quadratic family using Ruelle operator*, preprint, October 1997.

[60] R. F. Williams, *The structure of Lorenz attractors*, Publ. Math. I.H.E.S. 50 (1979), 321–347.

Index