Invent. math. 101, 335–372 (1990)

Inventiones mathematicae

The mapping class group of a generic quadratic rational map and automorphisms of the 2-shift

Lisa R. Goldberg[1,*] and Linda Keen[2,**]

[1] School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA and Department of Mathematics, Brooklyn College, CUNY, Brooklyn, NY 11210, USA

[2] Department of Mathematics, Lehman College, CUNY, Bronx, NY 10468, USA

Introduction

A rich class of dynamical systems is obtained by iterating rational maps of the Riemann sphere. An outline for classifying these dynamical systems was developed by Sullivan. He exploits the intimate connection between this classification problem and the theory of moduli for Riemann surfaces. In particular, he defines *mapping class groups* for rational maps, and shows how the mapping class group of a generic rational map can be built from subgroups of the mapping class groups of punctured tori.

An open question is to find presentations for these subgroups. We attack this problem for a class of rational maps that is of interest in symbolic dynamics.

Definition. A rational map is *generic* if it is hyperbolic, and its critical points have independent orbits containing no periodic cycles.

The well-known *Generic Hyperbolicity Conjecture* states that generic rational maps form an open dense subset of the space of all rational maps (with the coefficient topology).

In this article, we focus on an important class of rational maps with strong ties to abstract symbolic dynamics.

It is not difficult to show that the Julia set of a shift-like map R is a Cantor set; furthermore, the dynamics of R restricted to its Julia set are conjugate to a one-sided shift. The mapping class group (MCG) of a generic, shift-like rational map is infinitely generated [GK], and in Sect. 4 we give a complete presentation in case R is quadratic.

There is a representation of MCG(R) as a subgroup of automorphisms of the shift. The groups Aut_d of the one-sided d-shift were originally studied by Hedlund [H], and recent work of Boyle, Franks and Kitchens [BFK] and Ashley [A] show that these groups have a complicated, but not intractible

* Research Supported in part by the Alfred P. Sloan Foundation, The National Science Foundation, and the PSC-CUNY Fellowship Award Program.

** Research Supported in part by The National Science Foundation and the Max-Planck-Institut, Bonn, FRG

structure. Furthermore, shift automorphisms can be realized as homeomorphisms of the Riemann sphere which commute with polynomials [BDK].

One focus of this article is the introduction of methods to study the representation of $\mathrm{MCG}(R)$ as a group of automorphisms of the shift. In Sect. 5, we describe explicity this representation in the quadratic case. It is an open problem to extend these results to higher degrees.

All of our constructions depend on an ability to analyze the parameter spaces of rational maps. At a conference in Paris in 1988, J.H. Hubbard outlined a technique to analyze the parameter space for cubic polynomials. This "critical point surgery" technique models pieces of the cubic connectivity locus on the dynamical planes of quadratic polynomials. The technique first appears in the thesis of Wittner [Wit].

In this paper, we develop an analogous technique to analyze the parameter space of quadratic rationals. A careful description of this technique comprises most of Sect. 3.

In our context, it is natural to identify two rational maps if they are conjugate by a rational homeomorphism of the Riemann sphere $\hat{\mathbb{C}}$. The group of all such homeomorphisms is the Moebius group $\mathrm{PSL}(2, \mathbb{C})$. For a rational map R, we denote by $M(R)$, the space of $\mathrm{PSL}(2, \mathbb{C})$ conjugacy classes of rational maps which are quasiconformally (qc) conjugate to R.

Examples. i) If $R(z)=z^2+c$ for $c \neq 0$ in the cardioid (fixed point) region of the Mandelbröt set, $M(R)$ is parametrized by the cardioid with the origin deleted.

ii) If R is a quadratic polynomial with a cycle of Siegel disks, then $M(R)$ is a point.

From [S], it follows that for a generic rational map R, $M(R)$ is a $K(\pi, 1)$ space which has the structure of a complex analytic manifold with singularities. We define a contractible covering space for $M(R)$ in the following way:

Let $Q(R)$ be the space of qc homeomorphisms $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ for which fRf^{-1} is again a rational map. Homeomorphisms f_0, f_1 are *equivalent* if there is a Moebius transformation h and an isotopy between hf_0 and f_1 through elements of $Q(R)$. The quotient space of $Q(R)$ by this equivalence relation is the *Teichmüller space* of R and we denote it $\mathrm{Teich}(R)$.

Equivalent homeomorphisms in $Q(R)$ conjugate R to rational maps that are $\mathrm{PSL}(2, \mathbb{C})$ conjugate to one another, so there is a projection:

$$\mathbf{P}\colon \mathrm{Teich}(R) \to M(R)$$

given by

$$\mathbf{P}([f]) = [fRf^{-1}].$$

Let $Q_0(R) \subset Q(R)$ be the subgroup of qc homeomorphisms which commute with R. The mapping class group $\mathrm{MCG}(R)$ is the quotient of $Q_0(R)$ obtained by identifying homeomorphisms which are isotopic to the identity through elements of $Q_0(R)$.

There is an action of $\mathrm{MCG}(R)$ on $\mathrm{Teich}(R)$ given by

$$[g] \cdot [f] = [f \circ g]$$

and the orbit of any $[f] \in \mathrm{Teich}(R)$ is precisely a fibre of $\mathbf{P}$. This gives

Proposition 0.1. $M(R) \cong \mathrm{Teich}(R)/\mathrm{MCG}(R)$. □

Examples. i) If $R(z)=z^2+c$ for $c\neq 0$ in the cardioid (fixed point) region of the Mandelbröt set, $\mathrm{MCG}(R)=\pi_1(M(R))\cong\mathbb{Z}$.

ii) If $R(z)=z^d$, then $\mathrm{MCG}(R)$ is the dihedral group generated by the conformal homeomorphisms $z\to 1/z$ and $z\to\zeta z$, $\zeta^{d-1}=1$.

The above definitions are modelled on the classical theory of Riemann surfaces which we now recall.

Let S be a Riemann surface and define, as above

$$Q(S)=\{\text{qc homeomorphisms } f\colon S\to f(S)\}.$$

The equivalence relation here is:

$$f_0\approx f_1, \quad \text{if there is a conformal homeomorphism}$$

$$M\colon f_0(S)\to f_1(S) \text{ and an isotopy between } Mf_0 \text{ and } f_1.$$

The *Teichmüller space* $\mathrm{Teich}(S)$ is $Q(S)$ modulo this equivalence relation. The *mapping class group* $\mathrm{MCG}(S)$ is the group of isotopy classes of qc homeomorphisms $g\colon S\to S$.

As for rational maps, $\mathrm{MCG}(S)$ acts on $\mathrm{Teich}(S)$ via

$$[g]\cdot[f]=[f\circ g]$$

and the analog of $M(R)$ is the moduli space

$$M(S)=\mathrm{Teich}(S)/\mathrm{MCG}(S).$$

$M(S)$ parametrizes the ways a fixed topological surface can be made into a Riemann surface.

Mapping class groups of finite type Riemann surfaces are well understood (see for example [Bi] and Sect. 4 of this article). They are generated by a finite set of mapping classes which are represented by homeomorphisms of a particularly simple type.

By contrast, little detailed information is available for the mapping class groups of very basic "generic" rational maps. In [G], we generalize example ii) above, where we study the mapping class group of a special class of cubic polynomials – those with a single attracting fixed point which attracts all finite critical points. These groups are infinitely generated subgroups of the mapping class group of a twice punctured torus.

In this article, we give a complete description of $M(R)$ in case R is a generic, shift like rational map of degree 2. A rough statement of our result from Sect. 3 is:

Theorem. *The moduli space $M(R)$ is a fibre space over the punctured disk whose fibre is modelled on the filled Julia set of a quadratic polynomial.*

Using this, we prove

Theorem. i) MCG(R) *is an infinitely generated subgroup of the mapping class group of a twice punctured torus* MCG(T^2).

ii) *There is a distinguished set of generators for* MCG(R) *consisting of Dehn twists and spins (see Sect. 4 for definitions).*

As mentioned above, the restriction of R to its Julia set induces a representation ρ of MCG(R) as a group of automorphisms of the 2 shift. As an application of the previous theorem, we prove:

Theorem. i) *The representation* ρ: MCG$(R) \to \mathrm{Aut}_2$ *is surjective.*

ii) *There is a distinguished set of generators for the kernel of* ρ *consisting of Dehn twists and spins.*

This result is similar to the corresponding result in [BDK] but the methods are quite different.

The organization of this paper is as follows:

Section 1 introduces the two-parameter family of quadratic rational maps which contains a model for our moduli space $M(R)$.

Section 2 describes the dynamical planes of these quadratic rational maps.

Section 3 describes the parameter space of these maps. This section contains the details of the surgery technique mentioned above.

Section 4 contains a presentation for MCG(R).

Finally in Section 5, we analyze the representation ρ: MCG$(R) \to \mathrm{Aut}_2$.

Section 1

Consider the two dimensional family of quadratic rational maps

$$R_{\lambda,b}(z) = 1/\lambda(z+b+1/z)$$

parametrized by pairs $(\lambda, b) \in D_0 \times \mathbb{C}$. Each $R_{\lambda,b}$ has an attracting fixed point at ∞ with derivative λ, a pole at 0, and critical points at ± 1. Label the critical values $v_+ = R_{\lambda,b}(+1)$ and $v_- = R_{\lambda,b}(-1)$, respectively. For a fixed λ, v_+ and v_- depend linearly on b (and on each other), so that the family is essentially parametrized by $R'(\infty)$ and either v_+ or v_-.

The attractive basin of $R_{\lambda,b}$ at ∞, denoted $A_{\lambda,b}$, is a connected open set that is completely invariant under iteration and contains at least one critical point. The complement of $A_{\lambda,b}$ is the *filled Julia set*, $K_{\lambda,b}$; as in the case of quadratic polynomials, the filled Julia set must be one of two types:

Lemma 1.1. *Either*

i) $K_{\lambda,b}$ *is a connected set containing one critical point*

or

ii) $K_{\lambda,b}$ *is a Cantor set disjoint from the critical points.*

Sketch of Proof. Let Δ be a neighborhood of ∞ which is homeomorphic to a disk such that

i) Δ contains one critical value and no critical points

and

ii) $R_{\lambda,b}^{-1}(\Delta) \supset \Delta$.

If $K_{\lambda,b}$ contains one critical point, then the family $R_{\lambda,b}^{-n}(\Delta)$ forms a nested increasing sequence of topological disks whose union is homeomorphic to a disk and whose complement is $K_{\lambda,b}$.

If, on the other hand, $K_{\lambda,b}$ contains no critical points, then there is a smallest integer N such that $R_{\lambda,b}^{-N}(\Delta)$ contains both critical points and the complement of $R_{\lambda,b}^{-N}(\Delta)$ consists of two topological disks. It follows that for $m>0$, the complement of $R_{\lambda,b}^{-N+m}(\Delta)$ consists of 2^m disks and that the diameter of these disks tends to 0 as m tends to ∞. The complement, $K_{\lambda,b}$, of the union of the regions $R_{\lambda,b}^{-N+m}(\Delta)$ is therefore a Cantor set. □

Denote by M, the subset of $D_0 \times \mathbb{C}$ for which $K_{\lambda,b}$ contains a critical point. The set M is a disjoint union of subsets M_+ and M_- for which the critical values v_+ and v_- respectively are in $K_{\lambda,b}$.

If (λ, b) is not in M, both critical points are in $A_{\lambda,b}$ and $R_{\lambda,b}$ is a shift-like map. Generically, the orbits of the critical points are disjoint from each other and from ∞; we enumerate below the non-generic cases.

i) *Poles.* The subset of (λ, b)'s such that

$$R_{\lambda,b}^{j}(v_+)=\infty \quad \text{or} \quad R_{\lambda,b}^{j}(v_-)=\infty.$$

is the pole set P.

ii) *Orbit relations.* The subset of (λ, b)'s such that

$$R_{\lambda,b}^{i}(v_+)=R_{\lambda,b}^{j}(v_-).$$

is the set O of orbit relations.

Let Rat_2 be the space of rational maps parametrized by

$$D_0 \times \mathbb{C} - M \cup O \cup P.$$

A standard argument from [MSS] shows that any two maps in Rat_2 are conjugate by a quasiconformal homeomorphism of $\hat{\mathbb{C}}$. Hence, for any $R \in \mathrm{Rat}_2$, there is a projection $\mathbb{P}_R$ from Rat_2 to the moduli space $M(R)$ that sends $R_{\lambda,b}$ to its $\mathrm{PSL}(2, \mathbb{C})$ conjugacy class.

Proposition 1.2. *For any $R \in \mathrm{Rat}_2$, $\mathbb{P}_R$ is a degree 2 covering of $M(R)$ by* Rat_2 *ramified along the curve $D_0 \times 0$.*

Proof. $\mathbb{P}_R$ is surjective, since any rational map that is quasiconformally conjugate to an element of Rat_2 is actually $\mathrm{PSL}(2, \mathbb{C})$ conjugate to *some* $R_{\lambda,b} \in \mathrm{Rat}_2$.

The degree of $\mathbb{P}_R$ is 2 since $R_{\lambda,b}$ is $\mathrm{PSL}(2, \mathbb{C})$ conjugate to $R_{\lambda',b'}$ if and only if $\lambda=\lambda'$ and $b=\pm b'$. It follows that along the curve $D_0 \times 0$, the functions $R_{\lambda,0}$ are unique in their $\mathrm{PSL}(2, \mathbb{C})$ conjugacy classes.

Since each $R_{\lambda,0}$ is an odd function, the orbits of the critical points are disjoint and neither critical point is a preimage of ∞. We conclude that $R_{\lambda,0} \in \mathrm{Rat}_2$ for all λ. □

An immediate corollary which will be used in Sect. 3 is

Corollary 1.3. *If b and b' are in the right half plane and $R_{\lambda,b}$ is conformally conjugate to $R_{\lambda,b'}$, then $b=b'$.* □

We are interested in the action on fundamental groups induced by the $\mathbb{P}_R$'s. To write this action precisely we must pick a basepoint in Rat_2. We choose $R=R_{1/2,0}$ here, and wherever else in this paper we need a basepoint. For the sake of coherence, we work with the moduli space $M(R_{1/2,0})=M(R)$ which has the canonical choice of basepoint, R.

As we will see in Sect. 2, it is often easier to study rational maps if the critical points are labelled. Analogously, in the theory of Riemann surfaces with punctures, one often labels the punctures.

Since a map $f\in Q_0(R)$ or $Q_0(S)$ permutes the labelled critical points or punctures, we define the pure moduli spaces $M_*(R)$ and $M_*(S)$, and pure mapping class groups $\mathrm{MCG}_*(R)$ and $\mathrm{MCG}_*(S)$ by restricting to maps which take each labelled point to itself. If n is the number of labelled points,

$$M_* \to M$$

is an $n!$ to 1 projection and MCG_* is a subgroup of MCG of index $n!$.

Lemma 1.4. $\mathrm{MCG}_* \cong$ *image* $\mathbb{P}_*$.

Proof. This is immediate since the critical points of Rat_2 are labelled $+1$ and -1. □

Section 2

The Poincaré linearization theorem [P] guarantees that the restriction of any $R_{\lambda,b}$ to a suitable neighborhood of ∞ is conjugate to a linear map. We construct these conjugacies in an organized fashion.

Lemma 2.1. *For each $\lambda\in D_0$ and $b\in\mathbb{C}$, the sequence of functions*

$$\{\lambda^n R^n_{\lambda,b}\}$$

converge uniformly on compact subsets of $A_{\lambda,b}$ to a nonconstant, analytic surjection

$$\varphi_{\lambda,b}\colon A_b \to \hat{\mathbb{C}}-0$$

which fixes ∞, and satisfies the functional equation

$$\varphi_{\lambda,b}\circ R_{\lambda,b}=\lambda^{-1}\cdot\varphi_{\lambda,b}. \tag{1}$$

Remarks. i) The derivative $\varphi'_{\lambda,b}(\infty)$ is always equal to 1, so that $\varphi_{\lambda,b}$ is a conjugacy in a neighborhood of ∞. Of course, $\varphi_{\lambda,b}$ cannot be a homeomorphism on all of $A_{\lambda,b}$, and below, we characterize sets of points in $\hat{\mathbb{C}}$ identified by $\varphi_{\lambda,b}$.

ii) The maps $\varphi_{\lambda,b}$ depend analytically on the parameters λ and b.

Sketch of Proof. Set $\varphi_{\lambda,b,n}=\lambda^n R^n_{\lambda,b}$. The heart of the argument is to show that the $\varphi_{\lambda,b,n}$'s form a normal family on a neighborhood of ∞.

By definition $\varphi_{\lambda,b,n}(\infty)=\infty$, and it is not difficult to show that on a sufficiently small neighborhood N_∞ of ∞, there is a $K=K(b)$ such that

$$|R_{\lambda,b}(z)|>1/\lambda\,|z|-K.$$

It follows immediately that

$$|\varphi_{\lambda,b,n}(z)|=\lambda^n\,|R^n_{\lambda,b}(z)|>|z|-K$$

for $z\in N_\infty$ so that the $\varphi_{\lambda,b,n}$'s form a normal family on N_∞ as claimed.

No limit function of the $\varphi_{\lambda,b,n}$'s can be constant since $\varphi'_{\lambda,b,n}(\infty)=1$ for all b and n, and by construction, any limit of the $\varphi_{\lambda,b,n}$'s must satisfy the functional equation (1). Therefore, the $\varphi_{\lambda,b,n}$'s converge to a unique, nonconstant limit satisfying (1).

Since, under iteration of $R_{\lambda,b}$, every point in $A_{\lambda,b}$ eventually falls in N_∞, (1) defines $\varphi_{\lambda,b}$ on the rest of $A_{\lambda,b}$. □

To see which points have the same image under $\varphi_{\lambda,b}$ we recall from [S] that there are two important equivalence relations induced on $\hat{\mathbb{C}}$ by any rational map R:

Definitions. Points z and w are *grand orbit equivalent*, if there are integers n, $m\geqq 0$ such that $R^n(z)=R^m(w)$. If the stronger relation $R^n(z)=R^n(w)$ holds, z and w are *small orbit equivalent.*

Now $\varphi_{\lambda,b}(z)=\infty$ if and only if z is in the grand orbit of ∞, and we denote by $A^0_{\lambda,b}$, the complement in $A_{\lambda,b}$ of the grand orbit of ∞.

Lemma 2.2. *The restriction of $\varphi_{\lambda,b}$ to $A^0_{\lambda,b}$ identifies each small orbit equivalence class to a unique point in $\mathbb{C}-0$.*

Proof. If z and w are in $A_{\lambda,b}$, then there is an $M>0$ such that $R^M_{\lambda,b}(z)$ and $R^M_{\lambda,b}(w)$ are in N_∞. If z and w are small orbit equivalent, then (taking M larger if necessary) $R^M_{\lambda,b}(z)=R^M_{\lambda,b}(w)$ so that

$$\varphi_{\lambda,b}(z)=\lim_{n\to\infty}\lambda^n R^n_{\lambda,b}=\varphi_{\lambda,b}(w).\quad\square$$

Let G_λ be the cyclic group generated by the map $z\to\lambda z$. G_λ acts discontinuously on $\mathbb{C}-0$ and the map

$$p_\lambda\colon \mathbb{C}-0\to(\mathbb{C}-0)\,G_\lambda=T_\lambda$$

is a holomorphic projection onto a torus T_λ of modulus λ. Let $\Phi_{\lambda,b}$ be the map obtained by composing p_λ with the projection $\varphi_{\lambda,b}$.

$$\Phi_{\lambda,b}\colon A^0_{\lambda,b}\xrightarrow{\varphi_{\lambda,b}}\mathbb{C}-0\xrightarrow{p_\lambda}T_\lambda$$

From Lemma 2.2, it follows that $\Phi_{\lambda,b}$ maps the dynamical space of $R_{\lambda,b}$ to the torus T_λ by identifying each grand orbit equivalence class in $A^0_{\lambda,b}$ with a unique point in T_λ. Note that the image of $\Phi_{\lambda,b}$ is the same torus for all b and fixed λ.

If $R_{\lambda,b} \in \mathrm{Rat}_2$, then both $+1$ and -1 are in $A^0_{\lambda,b}$, and their images under $\Phi_{\lambda,b}$ are distinct on T_λ. Call $T^2_{\lambda,b}$ the twice punctured torus obtained by deleting from T_λ the special points $\Phi_{\lambda,b}(+1)$ and $\Phi_{\lambda,b}(-1)$. The set of critical points of $\Phi_{\lambda,b}$ is comprised of the backward orbits of $+1$ and -1.

Let $A^*_{\lambda,b} = \Phi^{-1}_{\lambda,b}(T^2_{\lambda,b})$; $A^*_{\lambda,b} \subset A_{\lambda,b}$ is the complement of the grand orbits of ∞ and of the critical points. It is proved in [S] that the restricted map

$$\Phi_{\lambda,b}: A^*_{\lambda,b} \to T^2_{\lambda,b}$$

is a covering projection.

The transformation Ψ: $R_{\lambda,b} \to T^2_{\lambda,b}$ gives a correspondence between Rat_2 and the pure moduli space of a twice punctured torus. We choose $T^2 = T^2_{1/2,0}$ as the basepoint for moduli space since it is the image of our basepoint $R \in \mathrm{Rat}_2$.

Again from [S] it follows that Ψ is a covering map, and so induces an injection on fundamental groups:

$$\Psi_*: \pi_1(\mathrm{Rat}_2) \to \pi_1(M_*(T^2))$$

which can be rewritten

$$\Psi_*: \mathrm{MCG}_*(R) \to \mathrm{MCG}_*(T^2).$$

Therefore we can view $\mathrm{MCG}_*(R)$ as a subgroup of $\mathrm{MCG}_*(T^2)$, and in Sect. 4, we give a presentation for this subgroup. This presentation depends on an analysis of the parameter space $D_0 \times \mathbb{C}$, which is the focus of the next section.

The remainder of this section contains a topological description of the projection from the dynamical plane of $R_{\lambda,b}$ onto the torus T_λ. To simplify the exposition, we surpress subscripts.

Notation. Oriented simple closed curves on T are called δ. The collection of components of $\Phi^{-1}(\delta)$ is called $\Lambda(\delta)$, and an element of $\Lambda(\delta)$ is called $\tilde{\delta}$. Because, in later sections, our curves are equipped with cylindrical neighborhoods, we make the following convention: if δ (or $\tilde{\delta}$) bounds a disk, we define $\Delta = \Delta(\delta)$ (or $\tilde{\Delta} = \tilde{\Delta}(\tilde{\delta})$) to be an open disk containing the bounded disk as well as the cylindrical neighborhood.

Under the map Φ: $A^0 \to T$, the curve $|\varphi(z)| = \text{constant}$ is identified with the curve $|\varphi(z)| = 2 \cdot \text{constant}$. Fix any constant and let α be the image of the level curve on T; the homology class of α on T is independent of the choice of constant. Let β be any curve on T which, together with α, generates the homology of T.

Let S be the collection of oriented simple closed curves δ on T, and let

$$i: S \times S \to \mathbb{Z}$$

be the intersection pairing.

Lemma 2.3. i) *If $i(\alpha, \delta) = 0$, then each $\tilde{\delta} \in \Lambda(\delta)$ is compact.*

ii) *If $i(\alpha, \delta) \neq 0$, then every element of $\Lambda(\delta)$ is noncompact.*

Proof. In a deleted neighborhood $N(\infty)$ of ∞, the map Φ is analytically equivalent (by φ) to the standard projection

$$p_{1/2}\colon \mathbb{C}-0 \to T.$$

The components of $p_{1/2}^{-1}(\delta)$ are either compact or noncompact, for $i(\alpha,\delta)=0$ or $i(\alpha,\delta)\neq 0$ respectively. Choosing $N(\infty)$ smaller, if it is necessary to remove "fragments", the same holds for any $\tilde{\delta}\cap N(\infty)$.

Since each $\tilde{\delta}\in\Lambda(\delta)$ can be mapped by an iterate of R which is a finite to 1 branched covering, so that it intersects $N(\infty)$, the lemma follows. $\square$

Remark. If δ is contained entirely in the punctured torus T, the restricted map

$$\Phi\colon \tilde{\delta}\to\delta$$

is a covering projection, so that the elements of $\Lambda(\delta)$ are all homeomorphic to $\mathbb{R}$ if $i(\alpha,\delta)\neq 0$; otherwise they are all homeomorphic to S^1.

In the next lemma, assume δ is properly contained in the punctured torus T^2 and that $i(\alpha,\delta)=i(\beta,\delta)=0$. The curve δ divides T^2 into two components; one of which is either a disk, a punctured disk, or a twice punctured disk.

Lemma 2.4. *If $\Delta(\delta)$ is a twice punctured disk, then either*

i) *Some component $\tilde{\Delta}$ of $\Phi^{-1}(\Delta)$ contains both v_+ and v_-, and $\hat{\mathbb{C}}-\Phi^{-1}(\Delta)$ is disconnected, or*

ii) *$\hat{\mathbb{C}}-\Phi^{-1}(\Delta)$ is connected.*

Proof. If some component $\tilde{\Delta}$ contains v_+ and v_-, then $\tilde{\Delta}$ is homeomorphic to a disk since it contains no critical points of Φ. The preimage $R^{-1}(\tilde{\Delta})$ contains both $+1$ and -1 and by the Riemann Hurwitz formula, $R^{-1}(\tilde{\Delta})$ is a doubly connected region which disconnects the sphere.

If no component contains both critical values, the discussion breaks into two steps.

First, there are the special components $\tilde{\Delta}_+$ and $\tilde{\Delta}_-$ which each contain one critical value. Depending on whether or not $\tilde{\Delta}_+$ and $\tilde{\Delta}_-$ are in the same grand orbit, the restriction of Φ to $\tilde{\Delta}_+$ and $\tilde{\Delta}_-$ is a degree 2 map with one branch point, or a degree 4 map with 3 branch points. Either way, by the Riemann Hurwitz formula they are both topological disks.

Second, $\tilde{\Delta}$ is an arbitrary component. Figure 2.1 illustrates the four possibilities for $\tilde{\Delta}$:

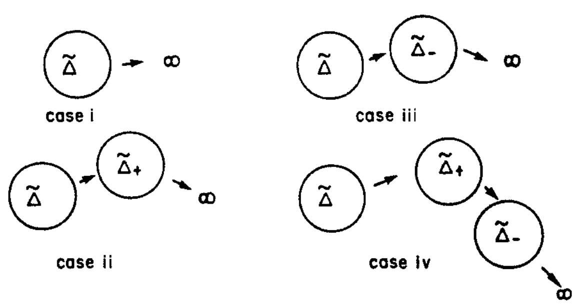

Fig. 2.1

In i) $\Phi|\tilde{\Delta}$ is a homeomorphism.

In ii) $\Phi|\tilde{\Delta}$ is a 2:1 branched covering with one simple branch point.

Case iii) is just case ii) with the roles of $\tilde{\Delta}_-$ and $\tilde{\Delta}_+$ reversed.

In case iv), $\Phi|\tilde{\Delta}$ is a degree 4 branched covering with three branch points.

In all four cases, the Riemann Hurwitz formula implies that $\tilde{\Delta}$ is a topological disk. □

For the following lemma, δ is a non-dividing, oriented simple closed curve on T^2 such that $i(\alpha, \delta)=0$, $i(\beta, \delta)=1$.

Lemma 2.5. *There are distinguished sequences of lifts* $\{\tilde{\delta}_n|0\leqq n\}\subset\Lambda(\delta)$ *and disks* $\{\tilde{\Delta}_n=\tilde{\Delta}_n(\tilde{\delta}_n)\}$, *and a non-negative integer* $K=K(\delta)$ *satisfying:*

i) *Each* $\tilde{\delta}_n$ *is an oriented simple closed curve.*

ii) R *maps* $\tilde{\Delta}_n$ *onto* $\tilde{\Delta}_{n+1}$, $\tilde{\Delta}_n\supset\tilde{\Delta}_{n+1}$ *and* $\bigcap\tilde{\Delta}_n=\infty$.

iii) R *is a homeomorphism for* $n\geqq K$ *and is a* 2:1 *branched covering for* $n<K$.

Proof. Suppose that δ passes through $\Phi(-1)$.

As in the proof of Lemma 2.3, there is a neighborhood of ∞ containing an infinite sequence of Φ-lifts of δ which pass through the forward iterates of -1 and bound nested disks containing ∞. This sequence contains a first element which we temporaily call $\tilde{\delta}_0$. It is defined by the property that the disk it bounds contains all the other curves in the sequence. Label the other curves $\tilde{\delta}_n=R(\tilde{\delta}_{n-1})$ and let $\tilde{\Delta}_n=\tilde{\Delta}(\tilde{\delta}_n)$.

Now proceed according to the following algorithm:

i) $j=0$.

ii) If $\tilde{\Delta}_j$ contains both critical values, or if $\tilde{\delta}_j$ contains one critical value, renumber the sequence by $n=n-j$ and stop.

iii) Else, let $\tilde{\delta}_{j-1}$ be the component of $R^{-1}(\tilde{\delta}_j)$ that has winding number 1 about ∞. Set $j=j-1$ and go to ii).

The process terminates in a finite number of steps, and produces the required sequence $\{\tilde{\delta}_n\}$. It is no problem to choose the $\tilde{\Delta}_n$ so that $R(\tilde{\Delta}_n)=\tilde{\Delta}_{n+1}$.

To complete the proof, let $K\geqq 0$ be the smallest integer such that $\tilde{\Delta}_{K-1}$ contains at most one critical point. □

Lemma 2.6. *If* $R\in\mathrm{Rat}_2$, *there is a simple closed curve* α_0 *on* T *passing through both* $\Phi(+1)$ *and* $\Phi(-1)$ *such that some* $\tilde{\alpha}_0\in\Lambda(\alpha_0)$ *passes through both critical values* v_+ *and* v_-.

Proof. See the beginning of Sect. 4 for the definition of spins.

Let δ be a simple closed curve on T passing through $\Phi(+1)$ and $\Phi(-1)$ and satisfying $i(\alpha, \delta)=1$ and $i(\alpha, \beta)=0$. There is a unique element $\tilde{\delta}\in\Lambda(\delta)$ which passes through v_-. Now $\tilde{\delta}$ must pass through at least one point in the grand orbit of v_+ which we can assume is of the form $R^m(v_+)$ for some $m\geqq 0$. (Otherwise, just reverse the roles of "+" and "−".) It follows that δ contains no other points in the grand orbits of the critical points.

If $m=0$, there is nothing to prove so assume $m>0$.

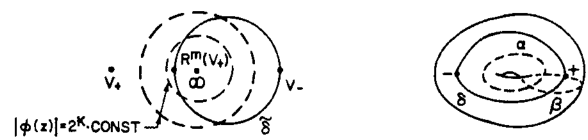

Fig. 2.2

We can assume, up to homological equivalence in T that β passes through $\Phi(+1)$ and not $\Phi(-1)$. Then either $S_\beta^m(\delta)$ or $S_{-\beta}^m(\delta)$ is the curve α_0 we desire, where S_β is the spin about β defined in Sect. 4. □

Remark. The curve α_0 is *not* unique.

Section 3

In this section, we analyze the parameter space for the $R_{\lambda,b}$'s. It is instructive to consider one dimensional slices of $D_0 \times \mathbb{C}$ obtained by fixing λ and varying b. Figure 3.1a shows the b-plane for $\lambda=1/2$.

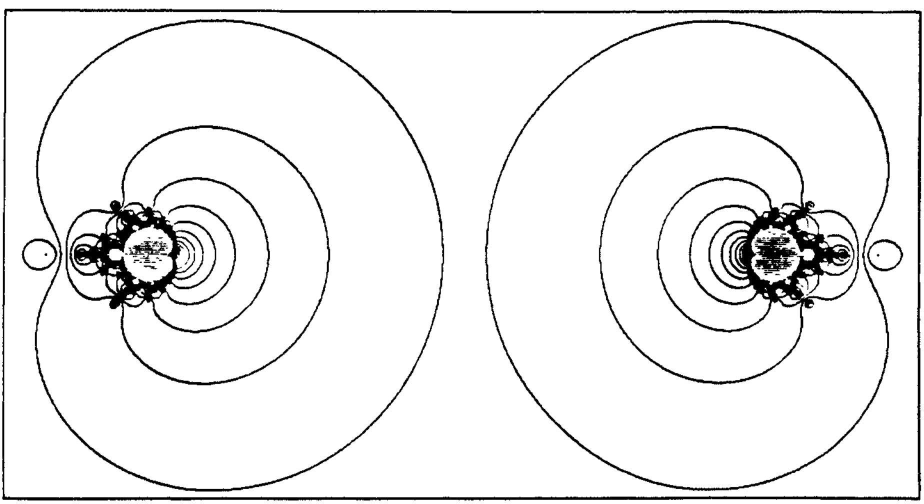

Fig. 3.1a

The algorithm that draws Fig. 3.1a counts the minimum number of iterates it takes for both critical points to leave a ball of radius 10. The grey region contains parameter values for which one of the critical points remains inside the ball after 100 iterates; it consists of two connected components, each of

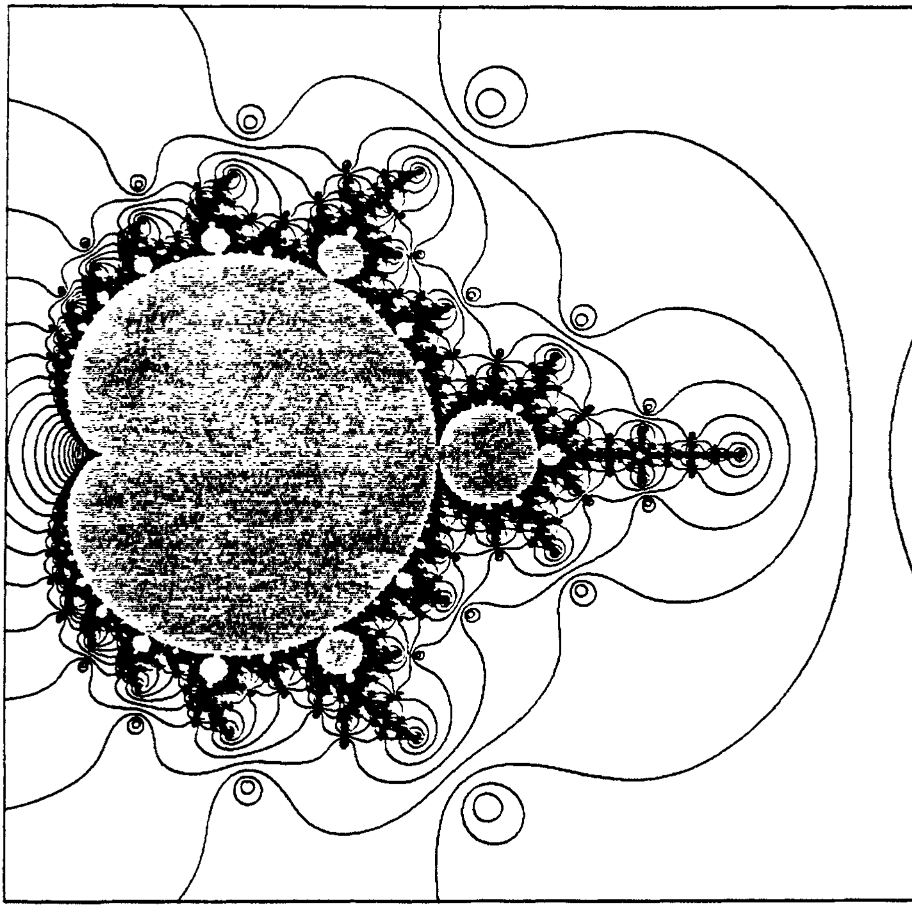

Fig. 3.1 b

which is homeomorphic to a Mandelbröt set. This is illustrated in Fig. 3.1 b which shows a blow-up of the right half plane.

The projection π: $\mathrm{Rat}_2 \to D_0$ that sends $R_{\lambda,b}$ to λ is a locally trivial fibration. The fibre $F_\lambda = \pi^{-1}(\lambda)$ is the complement in the b-plane (λ = constant) of the set $M \cup O \cup P$. With respect to the basepoint $R = R_{1/2,0} \in F_{1/2}$, the long exact homotopy sequence of π yields a short exact sequence:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \pi_1(F_{1/2}) & \longrightarrow & \pi_1(\mathrm{Rat}_2) & \xrightarrow{\pi_*} & \pi_1(D_0) & \longrightarrow & 1 \\ & & & & \| \wr & & & & \\ & & & & \mathrm{MCG}_*(R) & & & & \end{array}$$

so that $\mathrm{MCG}_*(R)$ is a twisted product of $\pi_1(F_{1/2})$ and $\mathbb{Z}$.

For the remainder of this section, we fix $\lambda = 1/2$ and simplify our notation by writing R_b for $R_{1/2,b}$, A_b for $A_{1/2,b}$, etc.

The affine map $z \to -z$ conjugates R_b to R_{-b}, so that the b-plane is symmetric about the origin. To understand this symmetry, we return to the dynamical plane of a particular R_b. The level curves of $|\varphi_b|$ are the leaves of a singular foliation of A_b^0 whose singularities occur precisely along the backward orbits of $+1$ and -1. There are two possibilities for the structure of this foliation.

i) Both critical points lie on the same leaf of the foliation. This occurs exactly when b is on the y axis.

ii) R_b has a *preferred* critical point. In this case, there is a distinguished leaf of the foliation which is a figure eight curve and has the preferred critical point as its cut point; this distinguished leaf will be called γ_0. The *principal loop* of γ_0 encloses the fixed point ∞, and the *secondary loop* encloses the pole 0. The region bounded by the principal loop is the largest dynamically defined region on which φ_b is a conjugacy.

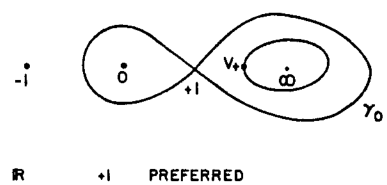

Fig. 3.2

The critical points $+1$ and -1 are preferred in the right and left half b-planes (**R** and **L**) respectively; the affine conjugacy $z \to -z$ interchanges **R** and **L**, and therefore changes the label of the preferred critical point.

If $b \in \mathbf{R}$ then any pole has the form

$$R_b^n(v_-) = \infty,$$

any orbit relation has the form

$$R_b^n(v_+) = R_b^m(v_-) \qquad m \geqq n,$$

and if $b \in \mathbf{R} \cap M$, the iterates $R_b^n(v_-)$ stay bounded.

It is a common phenomenon to find similarities between the dynamical plane of a rational map and the parameter space of a family of rational maps. Following ideas of Hubbard, we will use as a model for $\mathbf{R} - M$, a subset of the filled Julia set of a quadratic polynomial.

Let $Q(z) = z^2 + 1/2\, z$. The map Q has a single attracting fixed point at 0 with multiplier 1/2, a critical point at $-1/4$ and a critical value at $-1/8$. In anticipation of the ensuing construction, we label the critical point and critical value as c_+ and w_+. The filled Julia set K is a quasidisk whose interior $\mathring{K}$ constitutes the entire attractive basin of 0.

The map Q can be linearized in a neighborhood of its fixed point. As in the proof of Lemma 2.1, there is an analytic map

$$\varphi : \mathring{K} \to \mathbb{C}$$

defined by

$$\varphi(z) = \lim_{n \to \infty} 2^n Q^n(z).$$

The map φ fixes 0 and conjugates Q to $z \to 1/2z$ in a neighborhood of the fixed point.

Note that $\varphi(z)=0$ if and only if $Q^n(z)=0$ for some $n \geqq 0$; it follows that

$$\varphi(c_+) = re^{i\theta} \qquad \text{for some } r \neq 0.$$

As for the R_b's, we denote by γ_0, the preferred component of the level set $|\varphi|^{-1}(r)$ containing c_+. The curve γ_0 is a figure eight whose *principal loop* encloses 0, and whose *secondary loop* encloses the preimage $-1/2$ of 0.

To continue our description of K, we need the following:

Definition. A *generalized figure eight curve* in $\hat{\mathbb{C}}$ is a union of simple closed curves consisting of:

i) a primary loop which separates the sphere into two disks.

ii) $n \geqq 1$ secondary loops, each attached to the primary loop at a unique point and all of which are contained in the same disk.

A generalized figure eight divides the plane into three types of regions: the disk not containing the secondary loops is type A, the disks bounded by the secondary loops are type B, and the remaining "queer" component is type C. We denote by $A(\gamma)$, $B(\gamma)$ and $C(\gamma)$ the different regions bounded by a generalized figure eight γ.

Let $\gamma_n = Q^{-n}(\gamma_0)$. Then γ_n is a generalized figure eight curve with 2^n secondary loops and $\gamma_n \subset A(\gamma_{n+1})$ for all n. Each $B(\gamma_n)$ contains a single point in the set $Q^{-(n+1)}(0)$.

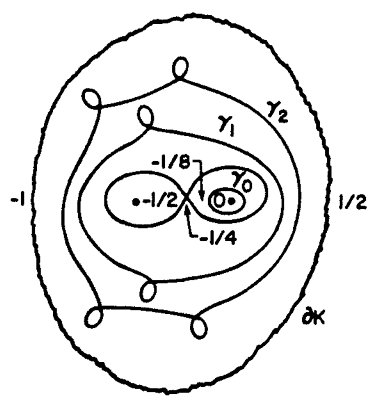

Fig. 3.3

The image $Q(\gamma_0)$ of the figure eight γ_0 is a simple closed curve contained in $A(\gamma_0)$. Let Δ be the closed disk bounded by $Q(\gamma_0)$; the remaining region $\mathring{K}-\Delta$, is a topological annulus.

Lemma 3.1. *There is an injective holomorphic map*

$$E\colon \mathbf{R}-M_{-}\to \mathring{K}-\Delta$$

which has the following properties:

i) *E maps each pole in* $\mathbf{R}-M_{-}$ *to a preimage of* 0; *that is, if b satisfies*

$$R_b^n(v_{-})=\infty$$

then $w=E(b)$ *satisfies* $Q^n(w)=0$.

ii) *E maps each point b satisfying an orbit relation in* $\mathbf{R}-M_{-}$ *to a point in the grand orbit of the critical value* w_{+}; *that is, if*

$$R_b^n(v_{+})=R_b^m(v_{-}),$$

then $w=E(b)$ *satisfies*

$$Q^n(w_{+})=Q^m(w_{-}).$$

iii) *As b tends to* $\partial\mathbf{R}$, $w=E(b)$ *tends to* $\partial\Delta$.

Proof. The map E from $\mathbf{R}-M_{-}$ to $\mathring{K}$ is defined in the following way. Since $R_b'(\infty)=Q'(0)=1/2$, there is a unique conformal homeomorphism ξ_b from a neighborhood of ∞ in A_b to a neighborhood of 0 in K which conjugates R_b to Q and is normalized so that $\xi_b(+1)=c_{+}$. Since b is in $\mathbf{R}-M_{-}$, the map ξ_b can be analytically continued in a unique fashion to v_{-} and we define

$$E(b)=\xi_b(v_{-}).$$

Since the critical point $+1$ is preferred for $b\in\mathbf{R}$, $\xi_b(v_{-})$ lies outside Δ. The conjugacies ξ_b depend holomorphically on b, so that E is holomorphic. By construction, E satisfies properties i) and ii).

Furthermore, E is injective: if $\xi_{b'}(v_{-})=\xi_b(v_{-})$, the map

$$\xi_{b'}^{-1}\circ\xi_b$$

maps a neighborhood of ∞ in A_b to a neighborhood of ∞ in $A_{b'}$; it extends to a conjugacy of R_b to $R_{b'}$ on the stable sets. It follows from [MSS] that this map extends further to a conformal conjugacy between R_b and $R_{b'}$ on all of $\hat{\mathbb{C}}$. Since b and $b'\in\mathbf{R}$, Corollary 1.3 implies $b=b'$.

Property iii) follows easily. As $b\to\partial\mathbf{R}$, the critical value v_{-} tends toward the leaf of the dynamically defined foliation containing v_{+} in the dynamical plane of R_b, and therefore $w=E(b)$ tends to $\partial\Delta\ni w_{+}$. □

Remark. The map ξ_b ties together domains in the dynamical spaces of R_b and Q with the parameter space of the R_b's.

The rest of this section contains a proof that E is a homeomorphism onto $\mathring{K}-\Delta$. The following technical lemma will be used repeatedly:

Lemma 3.2. *Suppose S is a Riemann surface homeomorphic to $\hat{\mathbb{C}}-n$ disks and let $w_+, w_- \in S$. There is exactly one isomorphism class of degree 2 ramified covering projections of S by planar Riemann surfaces which ramify simply over w_+ and w_-. These projections are normal, and the total space is homeomorphic to $\hat{\mathbb{C}}-2n$ disks.*

Proof. To show existence, fix an embedding $e: S \to \hat{\mathbb{C}}$ that maps w_+ and w_- to 0 and ∞ respectively. Consider the map $\Pi(z)=z^2$. The region $S'=\Pi^{-1}(e(S))$ is planar, therefore the Riemann Hurwitz formula implies that S' is homeomorphic to $\hat{\mathbb{C}}-2n$ disks. The required covering is

$$e^{-1} \circ \Pi: S' \to S.$$

To show uniqueness, let $q: S' \to S$ be a covering satisfying the hypotheses. As above, the degree 2 assumption implies S' is homeomorphic to $\hat{\mathbb{C}}-2n$ disks. Since q is a ramified covering, it is injective in a neighborhood of each end. Therefore, the classifying homomorphism

$$q_*: \pi_1(S-\{w_+, w_-\}) \to \mathbb{Z}_2$$

maps generating loops of $\pi_1(S-\{w_+, w_-\})$ about w_+ and w_- to the generator of $\mathbb{Z}_2$ and maps all other generating loops to the identity. Since q_* determines the isomorphism class of q, the result follows [Ma, p. 159]. □

Theorem 3.3. *The map E: $\mathbf{R}-M_- \to \mathring{K}-\Delta$ is a homeomorphism.*

Proof. We construct an inverse to the map E.

Outline of the construction:

i) We think of the points in $\mathring{K}-\Delta$ as potential places to add a new critical value. For any choice of $w_- \in \mathring{K}-\Delta$, the stable region of a rational map depending only on w_- is built up by an inductive procedure. At the n^{th} stage, we obtain a Riemann surface Ω_n which is topologically $\hat{\mathbb{C}}-2^n$ disks, and a holomorphic degree 2 endomorphism $Q_n: \Omega_n \to \Omega_n$. The direct limit is a Riemann surface Ω_∞ of infinite type, with a holomorphic, degree 2 endomorphism

$$Q_\infty: \Omega_\infty \to \Omega_\infty.$$

The map Q_∞ has two branch values and an attracting fixed point with multiplier 1/2.

ii) The measurable Riemann mapping theorem is used to construct a conformal embedding $e: \Omega_\infty \to \hat{\mathbb{C}}$ such that

$$e \circ Q_\infty = R_b \circ e$$

for a unique $b \in \mathbf{R}-M_-$ satisfying $\xi_b(v_-)=w_-$.

Details of the construction

Step 0. Fix $w_- \in \mathring{K} - \Delta$. There are two distinct elements of $Q^{-1}(w_-)$; choose one and label it c_-. Let N be the smallest integer such that $w_- \in A(\gamma_n) \cup B(\gamma_n)$. The curve γ_n is a component of the level set $|\varphi|^{-1}(2^N \cdot r)$. For all sufficiently small $\varepsilon > 0$, each level set $|\varphi|^{-1}(2^N \cdot r + \varepsilon)$ has a component which is a simple analytic curve bounding a disk which contains γ_N, w_+, c_+, and w_- but not c_-. Fix such an ε and call the disk Ω_0.

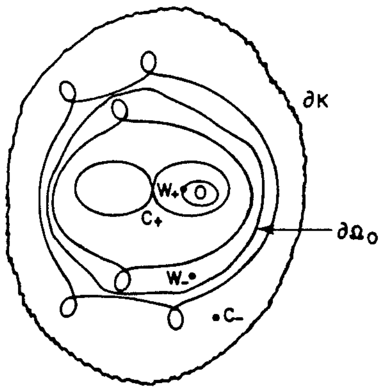

Fig. 3.4

Lemma 3.2 implies that there is a holomorphic degree 2 covering

$$\Pi_1 : \Omega_1 \to \Omega_0$$

ramified over w_+ and w_- with Ω_1 doubly connected.

Since the topological disk Ω_0 contains one critical point c_+, $Q(\Omega_0)$ is a topological disk and

$$Q : \Omega_0 \to Q(\Omega_0)$$

is a 2:1 covering ramified over w_+; (w_- is not in $Q(\Omega_0)$).

Set $\Omega_1' = \Pi_1^{-1}(Q(\Omega_0))$; then

$$\Pi_1 : \Omega_1' \to Q(\Omega_0)$$

is also a 2:1 covering, ramified over w_+.

Consider the diagram

$$\begin{tikzcd}
\Omega_1' \subset \Omega_1 \arrow[d, "\Pi_1"] & \\
Q(\Omega_0) \subset \Omega_0 & \Omega_0 \subset \hat{\mathbb{C}} \arrow[l, "Q"'] \arrow[ul, dashed]
\end{tikzcd}$$

We construct a homeomorphism $i_1: \Omega_0 \to \Omega_1'$ satisfying

$$\Pi_1 \circ i_1 = Q$$

by lifting locally near the critical point and extending by analytic continuation. Since Π_1 is a normal covering, there are two choices for i_1. However the rational map we construct does not depend on this choice.

Define $Q_1 = i_1 \circ \pi_1 : \Omega_1 \to \Omega_1$; Q_1 is a holomorphic endomorphism which is a 2:1 branched covering of its image ramified over $i_1(w_+)$ and $i_1(w_-)$. The point $i_1(0)$ is a fixed point of Q_1 with multiplier 1/2.

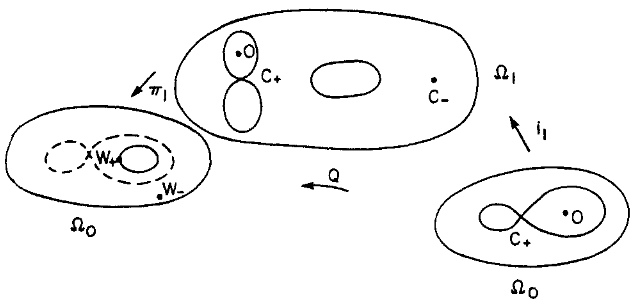

Fig. 3.5

We proceed inductively. The hypotheses are:

i) For $0 \leqq j \leqq n$, there is a planar Riemann surface Ω_j which is homeomorphic to the complement in $\hat{\mathbb{C}}$ of 2^j topological disks.

ii) For $0 \leqq j \leqq n$, there is a holomorphic endomorphism

$$Q_j: \Omega_j \to \Omega_j$$

which is a 2:1 branched covering of its image. Each Q_j has a fixed point with multiplier 1/2.

iii) For $0 \leqq j \leqq n-1$, there is a holomorphic embedding

$$i_{j+1}: \Omega_j \to \Omega_{j+1}$$

satisfying $Q_{j+1} \circ i_{j+1} = i_{j+1} \circ Q_j$.

Inductive Step. Lemma 3.2 gives a planar Riemann surface $\Omega_{n+1} \approx \hat{\mathbb{C}} - 2^{n+1}$ disks and a 2:1 normal covering

$$\Pi_{n+1}: \Omega_{n+1} \to \Omega_n$$

ramified over w_+ and w_-.

The Riemann Hurwitz formula implies that $Q_n(\Omega_n)$ is homeomorphic to the Riemann sphere with 2^{n-1} disks deleted. Set $\Omega'_{n+1} = \Pi_{n+1}^{-1}(Q_n(\Omega_n))$. Both

$$\Pi_{n+1}: \Omega'_{n+1} \to Q_n(\Omega_n)$$

and

$$Q_n: \Omega_n \to Q_n(\Omega_n)$$

are 2:1 covering projections ramified over w_+ and w_-. Therefore Lemma 3.2 implies that there is a holomorphic isomorphism

$$i_{n+1}: \Omega_n \to \Omega'_{n+1}$$

satisfying

$$\Pi_{n+1} \circ i_{n+1} = Q_n \quad \text{on } \Omega_n.$$

Once again, there are two possibilities for i_{n+1}, but the choice does not matter.

Define $Q_{n+1}: \Omega_{n+1} \to \Omega_{n+1}$ by

$$Q_{n+1} = i_{n+1} \circ \Pi_{n+1}.$$

Then

$$\begin{aligned} Q_{n+1} \circ i_{n+1} &= i_{n+1} \circ \Pi_{n+1} \circ i_{n+1} \\ &= i_{n+1} \circ Q_n \end{aligned}$$

and i_{n+1} conjugates Q_n to Q_{n+1}. Therefore both induction hypotheses ii) and iii) are satisfied. This completes the inductive step.

The direct limit Ω_∞ of the system (Ω_n, i_n) is the quotient of the union

$$\bigcup_n \Omega_n$$

by the all identifications of the form $z \approx i_n(z)$.

The resulting space is a Riemann surface Ω_∞ of infinite type whose fundamental group is given by:

$$\Pi_1(\Omega_\infty) = \lim_{n \to \infty} i_{n*}: \Pi_1(\Omega_n) \to \Pi_1(\Omega_{n+1})$$

A twofold self-covering $Q_\infty: \Omega_\infty \to \Omega_\infty$ is defined by

$$Q_\infty([z]) = [Q_n(z)] \qquad z \in \Omega_n.$$

The map Q_∞ is holomorphic and ramifies over $[w_+]$ and $[w_-]$ and has an attracting fixed point at $[0]$ with multiplier $1/2$.

Topologically, Ω_∞ is homeomorphic to the complement of a binary Cantor set in $\hat{\mathbb{C}}$. We claim that there is a unique conformal embedding

$$e: \Omega_\infty \to \hat{\mathbb{C}},$$

that conjugates Q_∞ to an R_b for a unique $b \in \mathbf{R}$.

Consider first the generic case for which the grand orbits of $[0]$, $[w_+]$ and $[w_-]$ are distinct. We construct a topological conjugacy between $(Q_\infty, \Omega_\infty)$ and the restriction of the base map $R \in \mathrm{Rat}_2$ to its stable set A.

Let Ω_∞^* be the region obtained by deleting from Ω_∞ the grand orbits of $[0]$, $[c_+]$ and $[c_-]$. The quotient of Ω_∞^* by the grand orbit equivalence relation is a twice punctured torus T_∞^2 of modulus $1/2$, and the projection is denoted

$$\Phi_\infty : \Omega_\infty^* \to T_\infty^2 .$$

It is analogous to the projection

$$\Phi_0 : A^* \to T^2,$$

of Sect. 2. From Lemma 2.6, there is a simple curve α_0 on T^2 passing through both punctures and a lift $\tilde{\alpha}_0$ which passes through v_+ and v_-.

Similarly, there is a curve α_∞ on T_∞^2 which passes through both punctures, and a lift $\tilde{\alpha}_\infty$ passing through both $[w_+]$ and $[w_-]$.

By inductive application of Lemma 3.2, any orientation preserving homeomorphism $\tilde{f}$ from T^2 to T_∞^2 which preserves the labelling on the punctures and maps α to α_∞ lifts to a topological conjugacy f between $R|A_0$ and Q_∞. We use this topological conjugacy to find the conformal embedding.

Without loss of generality, assume that $\tilde{f}$ is K-quasiconformal with Beltrami coefficient $\tilde{\mu}$. Then $\tilde{\mu}$ lifts to a Beltrami differential μ on A_0^* which is compatible with R, that is

$$\mu(R(z))\,\overline{R'(z)}/R(z) = \mu(z) \qquad z \in A_b^* .$$

Extend μ to be 0 on $\hat{\mathbb{C}} - A_0^*$.

By the measurable Riemann mapping theorem [AB], there is a unique K-quasiconformal homeomorphism $g: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ fixing 0, 1, and ∞, such that $g \circ R \circ g^{-1}$ is analytic on $\hat{\mathbb{C}}$. From Corollary 1.3, it follows that $g \circ R \circ g^{-1}$ is of the form R_b for a unique $b(w_-) \in \mathbf{R} \cap F_{1/2}$.

The map $e = g \circ f^{-1} : \Omega_\infty \to A_{b(w_-)}$ is the required conformal embedding.

To show that the correspondence $w_- \to b(w_-)$ gives an inverse to E, consider the commutative diagram, in which i_∞ is the direct limit of the maps i_j:

$$\begin{tikzcd} K & & \\ \cup & & \\ \Omega_0 \arrow[r, "i_\infty"] & \Omega_\infty \arrow[r, "e"] & A_{b(w_-)} \subset \hat{\mathbb{C}} \arrow[uull, "\xi_{b(w_-)}"'] \end{tikzcd}$$

By construction $e \circ i_\infty(w_-) = b(w_-)$, $v_- = R_b(-1)$ and $\xi_b(v_-) = w_-$.

In the non-generic cases, the critical value w_- is added along the orbit of 0 or of w_+ in $\mathring{K} - \Delta$ and the quotient of the region Ω_∞^* by the grand orbit equivalence relation of Q_∞ is a once punctured torus. A construction similar to, but easier than, the one above defines a conformal conjugacy e between Q_∞ and a unique R_b, $b \in O \cup P \cap \mathbf{R}$.

This completes the proof of Theorem 3.1. $\square$

Remarks. i) By placing w_- outside of Δ in $\mathring{K}$, we are permitted to think of w_+ as the preferred critical point. It is possible to imagine adding w_- along $\partial\Delta$, in which case there is no prefered critical point. By taking limits we extend E to a homeomorphism

$$\bar{E}: \partial\mathbf{R} \to \partial\Delta$$

where $\partial\mathbf{R} = y$ axis $\bigcup\{\infty\}$. The point at ∞ is the image of $w_- = w_+$ since there is no quadratic rational map with a single critical value.

ii) In the left half plane, -1 is preferred; we could just as easily have mapped $\mathbf{L} - M_+$ onto $\mathring{K} - \Delta$, by the map E' defined as

$$E'(b) = \xi_b(v_+)$$

iii) The curves $\gamma_n \subset \mathring{K}$ accumulate on the quasicircle boundary component of $\mathring{K}$ as $n \to \infty$; in fact $\bigcap A(\gamma_n)$ is the complement of $\mathring{K}$. The corresponding regions $A(\Gamma_n)$ also nest and we see that the corresponding boundary component of $F_{1/2}$ is the set $M_-(1/2)$. Moreover, we have:

Corollary 3.4. $M \cap b$-plane consists of two bounded, connected sets

$$M_- = M \cap \mathbf{R} \quad \text{and} \quad M_+ = M \cap \mathbf{L}. \quad \square$$

Remark. A straightforward application of the straightening theorem of [DH2] shows that M_- and M_+ are each homeomorphic to the Mandelbröt set.

Section 4

Material from Sects. 2 and 3 is used to give a presentation for the subgroup $\mathrm{MCG}(R) \subset \mathrm{MCG}(T^2)$. We begin by reviewing some basics about mapping class groups of surfaces.

Let δ be an oriented simple closed curve on any Riemann surface S. Let $N(\delta) \subset S$ be a cylindrical neighborhood of δ with coordinates t and θ, $-1 \leqq t \leqq 1$, $0 \leqq \theta \leqq 2\pi$; δ is given by the equation $t=0$. We assume that, at any point $p \in N(\delta)$, the tangent vectors $(dt(p), d\theta(p))$ form a basis for the tangent space $TS\,(p)$ compatible with the complex structure of S.

Suppose that r is a real number.

Definition. An r-Dehn twist D_δ^r: $N(\delta) \to N(\delta)$ is the homeomorphism defined by

$$D_\delta^r(t, \theta) = (t, \theta + \pi \cdot r(1+t))$$

If r is an integer, then D_δ^r is the identity on $\partial N(\delta)$ and so canonically extends to a homeomorphism on all of S. We use the notation D_δ^r to represent both the local and the global homeomorphism, and we surpress the superscript in case $r=1$.

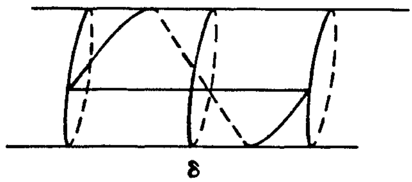

Fig. 4.1

Now suppose that S is a punctured Riemann surface, and δ is an oriented simple closed curve on S that passes through one puncture X. Choose simple closed curves $\delta_+, \delta_- \subset S$ which bound a cylindrical neighborhood of δ in $S \cup X$. Let $N(\delta_+)$ and $N(\delta_-)$ be chosen with δ as a common boundary and let $N(\delta) = N(\delta_+) \cup N(\delta_-)$.

Definition. The spin S_δ: $N(\delta) \to N(\delta)$ is the homeomorphism $D_{\delta_+}^{-1} \circ D_{\delta_-}$. As for Dehn twists, S_δ has a canonical extension to S which is also denoted S_δ.

The spin S_δ "drags" the puncture X along the curve δ as drawn below:

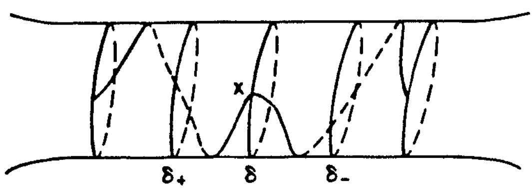

Fig. 4.2

The isotopy class of D_δ or S_δ depends only on the class of δ in $H_1(S)$ or $\pi_1(S \cup X, X)$, respectively.

The following elementary fact about twists and spins will be useful later in this section.

Lemma 4.1. [Bi]. *If f: $S \to S$ is a homeomorphism, then*

i) $f \circ D_\delta \circ f^{-1} = D_{f(\delta)}$

ii) $f \circ S_\delta \circ f^{-1} = S_{f(\delta)}$.

Namely, conjugates of twists and spins are twists and spins. $\square$

In Sect. 2, we constructed a map

$$\Phi: A^0 \to T$$

from the attractive basin of ∞ with the grand orbit of ∞ removed, onto a torus T of modulus 1/2. Recall that we defined the twice punctured torus T^2 as T minus the points $\Phi(+1)$ and $\Phi(-1)$. We also described a dynamically defined simple closed curve $\alpha \subset T$ and a complementary curve $\beta \subset T$ which together, generate the homology $H_1(T)$.

It is well known [Bi] that the mapping class group of T is the modular group $\mathrm{PSL}(2, \mathbb{Z})$, and that the Dehn twists D_α and D_β generate a subgroup of index 6.

The map that assigns to any twice punctured torus, the unique torus obtained by filling in the punctures defines a map $\rho: M_*(T^2) \to M(T)$. Since the fundamental group of a moduli space is a mapping class group, the induced map on fundamental groups can be expressed as

$$\rho_*: \mathrm{MCG}_*(T^2) \to \mathrm{MCG}(T).$$

Consider next the once punctured torus $T^1 = T - \Phi(+1)$ and choose $\Phi(-1)$ as basepoint on T^1. Let i: $T^1 \to M_*(T^2)$ be the map that assigns to any point $p \in T^1$ the twice punctured torus $T^1 - \{p\}$. This choice is asymmetric; we make it because we have chosen to work in the right half b-plane.

The following lemma is proved in [Bi]:

Lemma 4.2. i) *The sequence of maps*

$$T^1 \xrightarrow{i} M_*(T^2) \xrightarrow{\rho} M(T)$$

is a fibration whose long exact homotopy sequence yields the short exact sequence:

$$1 \longrightarrow \pi_1(T^1, \Phi(-1)) \xrightarrow{i_*} \mathrm{MCG}_*(T^2) \xrightarrow{\rho_*} \mathrm{MCG}(T) \longrightarrow 1.$$

ii) *The map*

$$i_*: \pi_1(T^1, \Phi(-1)) \to \pi_1(M_*(T^2) \cong \mathrm{MCG}_*(T^2)$$

is an injection that sends an oriented based loop δ on T^1 to the spin S_δ on T^2. □

The fundamental group of the parameter space $F_{1/2}$ and the mapping class group $\mathrm{MCG}_*(R)$ satisfy a similar relation. We next find the analog for R of the short exact sequence of Lemma 4.2. In Sect. 2, we introduced the fibration

$$F_{1/2} \to M_*(R) \cong \mathrm{Rat}_2 \xrightarrow{\pi} D_0$$

given by $\pi(R_{\lambda,b}) = \lambda$.

We also defined a covering projection

$$\Psi: \mathrm{Rat}_2 \to M_*(T^2)$$

which assigns to $R_{\lambda,b}$, the twice punctured torus $T^2_{\lambda,b}$, (the quotient of $A^*_{\lambda,b}$ by the grand orbit equivalence relation).

Since the Riemann surface obtained by filling in the punctures on $T^2_{\lambda,b}$ is a torus T_λ of modulus λ, the map Ψ induces the "modulus" map

$$\mu: D_0 \to M(T)$$

which sends λ to the torus T_λ. Consequently, Φ induces a covering map of fibres

$$\upsilon: F_{1/2} \to T^1$$

and there is commutative diagram of spaces:

$$\begin{tikzcd} F_{1/2} \arrow[r, "i"] \arrow[d, "\upsilon"'] & \mathrm{Rat}_2 \arrow[r, "\pi"] \arrow[d, "\Psi"'] & D_0 \arrow[d, "\mu"'] \\ T^1 \arrow[r, "i"] & M_*(T^2) \arrow[r, "\rho"] & M(T) \end{tikzcd}$$

in which the horizontal sequences are fibrations and the vertical maps are covering projections. This implies

Lemma 4.3. *There is a commutative diagram of fundamental groups*

$$\begin{tikzcd} 1 \arrow[r] & \pi_1(F_{1/2}) \arrow[r, "i_*"] \arrow[d, "\upsilon_*"'] & M_*(R) \arrow[r, "\pi_*"] \arrow[d, "\Phi_*"] & \pi_1(D_0) \arrow[r] \arrow[d, "\mu_*"] & 1 \\ 1 \arrow[r] & \pi_1(T^1) \arrow[r] & \mathrm{MCG}_*(T^2) \arrow[r] & \mathrm{MCG}(T) \arrow[r] & 1 \end{tikzcd}$$

in which the horizontal sequences are exact and the vertical arrows are injective. □

Since $F_{1/2}$ is contained in Rat_2 the implicit choice of basepoint for $\pi_1(F_{1/2})$ is $b=0$. Taking $q=\upsilon(0)$ as basepoint on T^1 and identifying $\pi_1(F_{1/2})$ with its image in $\pi_1(T^1)$ under υ_*, we obtain

Corollary 4.4. *$\pi_1(F_{1/2})$ is a subgroup of the free group $\pi_1(T^1)$.* □

The map $\upsilon: F_{1/2} \to T^1$ was defined abstractly, but it can be understood concretely in terms of the dynamics of our quadratic polynomial Q and the homeomorphism $E: \mathbf{R} \cap M_- \to \mathring{K} - \Delta$. We will use this explicit formula in Sect. 5 to relate loops in the parameter space $F_{1/2}$ with loops in the dynamical plane of the rational map R.

In analogy with prior notation, let K^* be the region obtained by removing from $\mathring{K}$, the grand orbits of 0 and the critical value $w_+ = -1/8$. As above, K^* projects via the grand orbit equivalence relation onto a once punctured torus T_Q' of modulus 1/2: The puncture corresponds to the grand orbit of w_+. Let

$$\Phi_Q: K^* \to T_Q'$$

denote the projection.

The conformal homeomorphism ξ_0 that conjugates R to Q induces a conformal isomorphism between T^1 and T_Q'. We treat this isomorphism as an identification and prove:

Lemma 4.5. *The restricted map*

$$\upsilon: F_{1/2} \cap \mathbf{R} \to T^1$$

equals the composite $\Phi_Q \circ E$.

Proof. We must show that the following diagram commutes:

$$\begin{tikzcd} K^* - \Delta \arrow[drr, "\Phi_Q"'] & & F_{1/2} \cap \mathbf{R} \arrow[ll, "E"'] \arrow[r] \arrow[d, "\upsilon"] & \mathrm{Rat}_2 \arrow[d, "\Psi"] \\ & & T^1 \arrow[r] & M_*(T^2) \end{tikzcd}$$

Choose $b \in F_{1/2} \cap \mathbf{R}$ and let $w = E(b)$. As in the proof of Lemma 3.1, there is a conformal isomorphism ξ_b from a neighborhood of ∞ in A_b to a neighborhood of 0 in K that maps υ_+ to w_+ and υ_- to w_-. Therefore, ξ_b descends to a conformal isomorphism from $T^2_{1/2,b}$ to the twice punctured torus $T_Q - \{\Phi_Q(w_-)\}$. □

To understand the action of υ on $F_{1/2} \cap \mathbf{L}$, recall that the linear map

$$l: z \to -z$$

conjugates R_b to R_{-b} and interchanges $+1$ and -1. On the level of Riemann surfaces, l descends to a conformal isomorphism between the twice punctured tori T_b^2 and T_{-b}^2 which interchanges the marked points. There is an induced map

$$L: M_*(T^2) \to M_*(T^2)$$

of degree 2 and the quotient of $M_*(T^2)$ by the relation $z \approx L(z)$ is $M(T^2)$. Each fibre of the map $\rho: M_*(T^2) \to M(T)$ is preserved so there is an induced map

$$\varepsilon: T^1 \to T^1$$

obtained by restricting L to the fibre T^1. The map ε is the *elliptic involution* of the torus T^1. An explicit construction of ε is given at the end of this section.

We conclude

Lemma 4.6. *The map* υ *satisfies the functional equation*

$$\upsilon(-b) = \varepsilon \circ \upsilon(b). \quad \square$$

Recall from Sect. 3 the structure on $\mathring{K}$ given by the generalized figure eight curves γ_n. The region $B(\gamma_0)$ contains the preimage, $-1/2$, of the fixed point of Q. It also contains a sequence of points w_m, $m \geqq 1$, which accumulate on $-1/2$ and satisfy

$$Q(w_m) = Q^m(w_+)$$

Each $B(\gamma_n)$ is mapped onto $B(\gamma_0)$ by Q^n in a 1:1 manner. It follows that $B(\gamma_n)$ contains a single element of the set $Q^{-(n+1)}(0)$ and a sequence of points $w_{n,m}$ satisfying

$$Q^{n+1}(w_{n,m}) = Q^m(w_+).$$

The union of the cut points w of the γ_n's comprise the backward orbit of the critical value w_+; they satisfy equations of the form

$$Q^n(w) = w_+$$

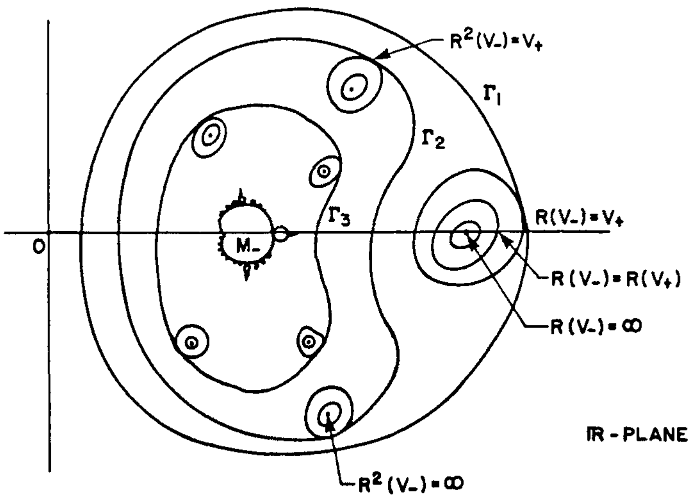

Fig. 4.3

Under the homeomorphism E, $\mathbf{R} - M_-$ has an analogous structure:

i) Each $B(\Gamma_n)$ contains a single pole b satisfying

$$R_b^{(n+1)}(v_-) = \infty$$

and a sequence of b's satisfying the orbit relations

$$R_b^{(n+1)}(v_-) = R^m(v_+) \qquad m \geqq 1$$

ii) The cut points b of each Γ_n satisfy an orbit relation of the form

$$R_b^n(v_-) = v_+$$

From the symmetry of the b-plane, it follows that $\mathbf{L} \cap F_{1/2}$ has the same structure but the roles of v_+ and v_- are interchanged.

We conclude that the b's satisfying orbit relations are isolated points in $M \cup O \cup P$ and accumulate on the poles, while the poles accumulate on M. Therefore $\pi_1(F_{1/2})$ is an infinitely generated free group, which is generated by:

i) One (based) loop enclosing each orbit relation.
ii) One loop enclosing each pole.
iii) Two loops, one enclosing M_+, one enclosing M_-.

It is certainly possible to choose these loops so that each is contained either in the closed left or right half plane. If $\{\gamma_j\}$ is a set of generators for $\pi_1(\overline{\mathbf{R}}\cap F_{1/2})$ then $\{\gamma_j, -\gamma_j\}$ are generators for $\pi_1(F_{1/2})$. Lemma 3.6 implies

$$\upsilon_*[-\gamma_j]=\varepsilon_*\circ\upsilon_*[\gamma_j].$$

We will construct a family of generators in the right half plane of types i)–iii) above. We use the maps υ_* and i_* to interpret their action as elements of $\mathrm{MCG}_*(T^2)$. We thus obtain a presentation for the subgroup $\mathrm{MCG}_*(R)$ of $\mathrm{MCG}_*(T^2)$.

Since the map υ is an ∞ to 1 covering projection, the preimages in $F_{1/2}$ of the basepoint $q=\upsilon(0)\in T^1$ form an infinite discrete subset. If ω is a directed path in $F_{1/2}$ from 0 to $x\in\upsilon^{-1}(q)$, then $\upsilon(\omega)$ is a based (but not necessarily simple) loop in T^1. The element $[\upsilon(\omega)]\in\pi_1(T^1)$ therefore corresponds to the homeomorphism

$$h_\omega=i_*[\upsilon(\omega)]$$

in $\mathrm{MCG}_*(T^2)$.

Let $N_0\subset F_{1/2}$ be a small deleted neighborhood of an orbit relation $b\in O\subset\mathbf{R}$. For every such b, the image of the restriction of υ to N_0 is a deleted neighborhood of the puncture in T^1.

Let k be a homotopically nontrivial loop in N_0 parametrized by θ, $0\leqq\theta<2\pi$. Since υ is a covering projection, there is a directed path $\zeta\in F_{1/2}$ connecting *some* element x of $\upsilon^{-1}(q)$ to $k(0)$, so that $\upsilon(\zeta)$ is a simple path on T^1. Now

$$\gamma=\zeta^{-1}k\zeta$$

is a simple loop based at x which encloses the orbit relation b, and no other points in $M\cup O\cup P$.

There are two cases to consider

i) The point b is a cut point of a Γ_n.
ii) Otherwise.

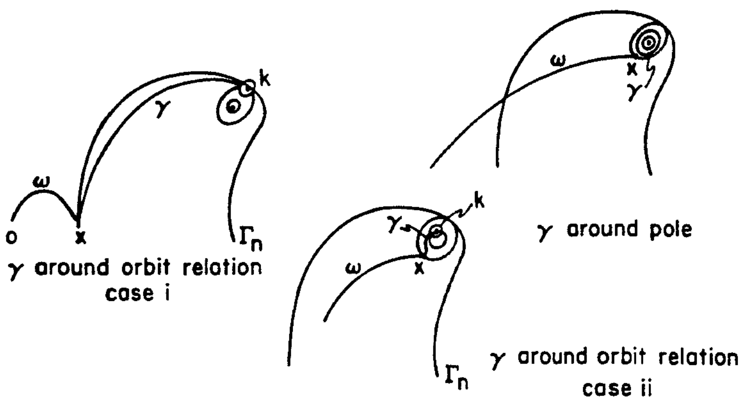

Fig. 4.4

Case ii) is easier since the restriction of υ to D_0 is a homeomorphism, and so, taking the circle k smaller if necessary, the restricted map

$$\upsilon: \gamma \to \upsilon(\gamma)$$

is a homeomorphism as well. The based loop $\upsilon(\gamma)$ on T^1 is ε-approximated by a simple loop δ for any ε, so that γ corresponds to the spin S_δ in $\mathrm{MCG}_*(T^2)$.

Finally, if we choose a directed path ω from 0 to x in $F_{1/2}$, then the based loop

$$\omega^{-1}\gamma\omega$$

in $F_{1/2}$ encloses the orbit relation b, and, by Lemma 4.1, corresponds to the spin

$$h_\omega^{-1} S_\delta h_\omega$$

in $\mathrm{MCG}_*(T^2)$.

In case i), the restriction of υ to N_0 is 2:1, so that k projects to a curve that winds twice about the puncture in T^1, therefore, the projection of the loop

$$\gamma = \zeta^{-1} k \zeta$$

based at x in $F_{1/2}$ is a based loop $\upsilon(\gamma)$ on T^1 that winds twice about the puncture. Consequently, there is a simple based loop δ that winds once around the punctures such that $\upsilon(\gamma)$ is homotopic to δ^2. It follows that the image of $[\upsilon(\gamma)]$ in $\mathrm{MCG}_*(T^2)$ is the spin-square S_δ^2. By Lemma 4.1, the based loop

$$\omega^{-1}\gamma\omega$$

in $F_{1/2}$ corresponds to the spin-square

$$h_\omega^{-1} S_\delta^2 h_\omega$$

in $\mathrm{MCG}_*(T^2)$.

Suppose next that $b \in P \subset \mathbf{R}$ is a pole. Then b lies at the "center" of some $B(\Gamma_n)$ and the restricted map

$$\upsilon: B(\gamma_n) \cap F_{1/2} \to T^1$$

is analytically equivalent to the projection

$$p_{1/2}: \mathbb{C} - \{0, 2^n \mid n \in \mathbb{Z}\} \to T^1,$$

defined in Sect. 2.

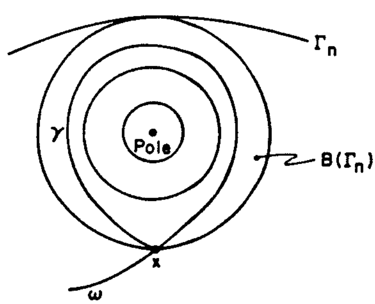

Fig. 4.5

Therefore, $B(\Gamma_n) \cap F_{1/2}$ is tiled by a nested sequence of fundamental domains for υ, each of which is homeomorphic to a punctured annulus. Fix such a fundamental domain D, and let x be the unique element of $\upsilon^{-1}(q)$ in D. Let γ be a simple loop in D based at x which generates the fundamental group of the unpunctured annulus. The image $\upsilon(\gamma)$ is a based simple closed curve on T^1 which corresponds (under i_*) to the spin $S_{\upsilon(\gamma)}$. If ω is a directed path connecting 0 to x in $F_{1/2}$, then

$$\omega^{-1}\gamma\omega$$

is a based loop in $F_{1/2}$ that encloses the pole b, and corresponds to the spin

$$h_\omega^{-1} S_{\upsilon(\gamma)} h_\omega$$

in $\mathrm{MCG}_*(T^2)$.

The last part is the easiest. The y axis and the primary loop of Γ_0 in $\mathbf{R}$ bound an annulus A which is a fundamental domain for υ. Let $\gamma \subset A \subset \mathbf{R}$ be a simple closed curve based at 0 generating $\pi_1(A)$. The curve γ separates M_- (and all the poles and orbit relations in $\mathbf{R}$) from M_+. As above, γ corresponds to the spin $S_{\upsilon(\gamma)}$ in $\mathrm{MCG}_*(T^2)$.

The relation $\upsilon_*[-\gamma_j] = \varepsilon_* \circ \upsilon_*[\gamma_j]$ is used to interpret the action of the generators $\{-\gamma_j\}$ as elements of $\mathrm{MCG}_*(T^2)$. If S_{γ_j} is the spin corresponding to γ_j then $S_{\varepsilon(\gamma_j)}$ is the spin corresponding $-\gamma_j$.

To complete the analysis of $\mathrm{MCG}_*(R)$, recall the commutative diagram in Lemma 4.3. We can choose a fundamental domain for the projection Φ_0 which is an annulus $\mathbf{A}$ containing a lift $\tilde{\alpha}$ of the curve α on T. The disk D_0 is the moduli space of $\mathbf{A}$; and the group $\pi_1(D_0) = \mathrm{MCG}(\mathbf{A})$ is an infinite cyclic group generated by the Dehn twist $D_{\tilde{\alpha}}$. Therefore the group $\mathrm{MCG}(\mathbf{A})$ injects into the subgroup of $\mathrm{MCG}(T)$ generated by D_α.

We summarize this analysis in

Theorem 4.7. The pure mapping class group $\mathrm{MCG}_*(R)$ is an infinitely generated subgroup of $\mathrm{MCG}_*(T^2)$ which can be expressed as a twisted product

$$\mathbb{Z} * \pi_1(F_{1/2}).$$

The factor $\mathbb{Z} = \pi_1(D_0)$ is generated by a Dehn twist about a dynamically defined simple closed curve α in T^2. The factor $\pi_1(F_{1/2})$ is an infinitely generated subgroup of $\pi_1(T^1) \subseteq \mathrm{MCG}_*(T^2)$ with three types of generators:

ia) Loops about orbit relations of the type $R^n(\pm v) = \pm v$ in $F_{1/2}$ which are spin squares about simple closed curves in T^2. These curves satisfy $i(\alpha, \delta) = i(\beta, \delta) = 0$.

ib) Loops about orbit relations in $F_{1/2}$ of the type $R^n(\pm v) = R^m(\pm v)$ for $m \geqq 1$ which are spins about simple closed curves in T^2. These curves satisfy $i(\alpha, \delta) = i(\beta, \delta) = 0$.

ii) Loops about poles which are spins about simple closed curves δ in T^2. These curves satisfy $i(\alpha, \delta) = 0$, $i(\beta, \delta) = 1$.

iii) One loop about each Mandelbröt set M_+ and M_-, these are spins about simple closed curves δ in T^2 that satisfy $i(\alpha, \delta)=0$, $i(\beta, \delta)=1$.

Moreover we conclude

Theorem 4.8. *The mapping class group* $\mathrm{MCG}(R)$ *is a degree* 2 *extension of* $\mathrm{MCG}_*(R)$. *It is generated by* $\mathrm{MCG}_*(R)$ *together with the unique involution of* T^2 *that interchanges the punctures and induces the identity on* $H_*(T)$. □

Remark. As b varies in $F_{1/2}$, the quotient $\varphi_b(+1)/\varphi_b(-1)$ takes on all complex values except 0 and the integral powers of 2. Therefore the map

$$b \to \varphi_b(+1)/\varphi_b(-1) \bmod 2^n$$

projects $F_{1/2}$ onto a once punctured torus, and it is not difficult to show that this map is just υ. The projection

$$p_{1/2}\colon \mathbb{C}-\{0, 2^n \mid n \in \mathbb{Z}\} \to T^1$$

commutes with the map $\tilde{\varepsilon}$: $z \to 1/z$ on $\mathbb{C}-\{0, 2^n \mid n \in \mathbb{Z}\}$. Therefore $\tilde{\varepsilon}$ descends to a degree two map

$$\varepsilon\colon T^1 \to T^1$$

which is once again, the elliptic involution.

Section 5

One of the original motivations for this article was to study the relationship between the dynamics of shift-like rational maps and automorphisms of the shift. We review the basic definitions [D]:

Let Σ_2 be the space of semi-infinite sequences whose entries are the integers 0 or 1. A neighborhood basis for a topology on Σ_2 is defined at the point $s=(s_0, s_1, s_2, \ldots)$ by

$$U_j=\{t=(t_0, t_1, t_2, \ldots) \mid t_i=s_i, 0 \leqq i \leqq j\}.$$

In this topology Σ_2 is a Cantor set.

The one-sided shift map σ: $\Sigma_2 \to \Sigma_2$ defined by

$$\sigma(s_0, s_1, s_2, \ldots)=(s_1, s_2, s_3, \ldots)$$

is a continuous, 2:1 endomorphism.

The automorphism group Aut_2 is the group of self-homeomorphisms of Σ_2 that commute with the shift.

A well-known result of Hedlund is:

Theorem 5.1. [H] Aut_2 *is generated by the homeomorphism that interchanges the symbols* 0 *and* 1; *in particular*, $\mathrm{Aut}_2 \cong \mathbb{Z}_2$. □

The relationship with the rational maps we have been studying is given by:

Proposition 5.2. *If* $R \in \mathrm{Rat}_2$ *there is a homeomorphism from the Julia set of* R, J_R *to* Σ_2 *which conjugates* $R \mid J_R$ *to* σ.

Sketch of Proof. For convenience, suppose that R has $+1$ as its preferred critical point. Define the curve δ on $T^2=\Phi(A_0)$ by

$$\delta=\Phi(\{z \,|\, |\varphi(z)|=|\varphi(v_-)|\}).$$

Since $i(\alpha,\delta)=0$ and $i(\beta,\delta)=1$, Lemma 2.5 implies that there is a sequence of lifts $\{\tilde{\delta}_n \,|\, n\geqq 0\}$ and nested disks $\{\tilde{\Delta}(\tilde{\delta}_n)\}$ which contain ∞.

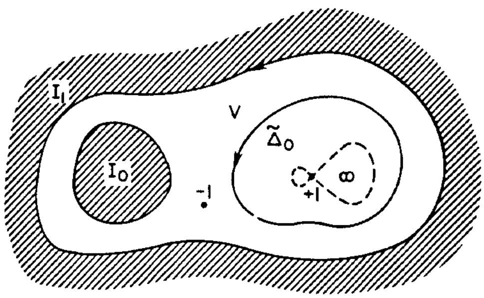

Fig. 5.1

The disk $\tilde{\Delta}_0$ contains both critical values so that its preimage $V=R^{-1}(\tilde{\Delta}_0)$ is doubly connected. The complement of V consists of two disjoint topological disks which we label I_0 and I_1.

The disks I_0 and I_1 generate a Markov partition for the Julia set: a map from J_R to Σ_2 is obtained by labelling points of J_R using their itineraries relative to this partition. That is, set

$$\begin{aligned}\rho(z)&=0 \quad \text{if } R^j(z)\in I_0\\ &=1 \quad \text{if } R^j(z)\in I_1.\end{aligned}$$

By definition $\sigma\circ\rho=\rho\circ R$. The proof that p is a homeomorphism is standard and can be found, for example in [D]. □

Remarks. i) Hedlund's theorem implies that ρ is unique up to postcomposition by the homeomorphism that interchanges the symbols 0 and 1.

ii) A proof virtually identical to the proof of 5.2 shows that if R is a shift-like rational map of degree d, there is a homeomorphism from the Julia set of R to Σ_d which conjugates $R|J_R$ to σ.

The map ρ induces a representation

$$\rho_*\colon \mathrm{MCG}_*(R)\to \mathrm{Aut}_2$$

in the obvious way: an element f of $\mathrm{MCG}_*(R)$ restricted to $J_R\approx\Sigma_2$ induces an element $\rho_* f$ of Σ_2.

In Sect. 3 we defined a set of generators for $\mathrm{MCG}_*(R)$ which fall into four categories:

Generators for $\pi_1(F_{1/2})$

i) Spins generated by loops about orbit relations.
ii) Spins generated by loops about poles.
iii) Spins generated by a loop about a Mandelbröt set.

Generator for $\pi_1(D_0)$

iv) A Dehn twist of the annulus.

The remainder of this section contains a proof of:

Theorem 5.3. Under the map ρ_*: $\mathrm{MCG}_*(R) \to \mathrm{Aut}_2$, the generators of type iii) map to the generator of Aut_2, all others map to the identity.

Outline of Proof. i) Lemmas 2.4 and 2.5 are used to describe lifts of curves on T^2 to the dynamical plane of R.

ii) The Julia set is located relative to these lifts (as in the proof of Proposition 5.2).

iii) Lemmas 5.4, 5.5, and 5.6 (below) are used to describe how the spins and twists lift to the stable set of R.

iv) Lemma 5.7 (below) is used to extend this action to the Julia set of R.

We first review how a Dehn twist on a Riemann surface S lifts locally to a branched covering S'.

Recall notation from Sect. 4; suppose δ is an oriented simple closed curve on a Riemann surface, and $N(\delta)$ is a cylindrical neighborhood with coordinates (t, θ), $-1 \leqq t \leqq 1$, $0 \leqq \theta < 2\pi$. Suppose

$$\Pi: S' \to S$$

is a map of Riemann surfaces, $\tilde{\delta}$ a component of $\Pi^{-1}(\delta)$, and $N(\tilde{\delta})$ the component of $\Pi^{-1}(N(\tilde{\delta}))$ containing $\tilde{\delta}$. If the restricted map

$$\Pi: N(\tilde{\delta}) \to N(\delta)$$

is a degree q unbranched covering, then $\tilde{\delta}$ is a simple closed curve. Moreover, there are q different ways to give $N(\tilde{\delta})$ "Π-preferred coordinates" (t, θ), $-1 \leqq t \leqq 1$, $0 \leqq \theta < 2\pi$, so that

$$\Pi(t, \theta) = (t, q\theta \bmod 2\pi).$$

Lemma 5.4. *There are precisely q lifts of the p^{th} power Dehn twist*

$$D_\delta^p: N(\delta) \to N(\delta)$$

to $N(\tilde{\delta})$, each of which can be expressed in any Π-preferred coordinate system for $N(\tilde{\delta})$ as a p/q-Dehn twist followed by a rotation by $2\pi n/q$, $1 \leqq n \leqq q$. That is,

$$\mathrm{Rot}_{2\pi n/q} \circ D_{\tilde{\delta}}^{p/q}. \quad \square$$

Recall that a spin S_δ about a curve δ through a puncture is a composition of Dehn twists. If

$$\Pi: N(\tilde{\delta}) \to N(\delta)$$

is a degree q regular covering. Then $\tilde{\delta}$ is a simple closed curve with q punctures. These facts, together with Lemma 5.4, immediately imply the following lemma.

Lemma 5.5. *There are q distinct lifts f of the p^{th} power spin to $N(\tilde{\delta})$. Among these is a distinghuished lift which restricts to the identity on both components of $\partial N(\tilde{\delta})$. All the other lifts restrict to nontrivial rotations on both components of $\partial N(\tilde{\delta})$.* $\square$

If the map

$$\Pi: N(\tilde{\delta}) \to N(\delta)$$

is a degree 2 cover branched over the puncture in δ, the curve $\tilde{\delta}$ is a figure eight curve and $\partial N(\tilde{\delta})$ has three components labelled so that, $\tilde{\delta}_-$ double covers δ_- and $\tilde{\delta}_{+0}$ and $\tilde{\delta}_{+1}$ each single cover δ_+.

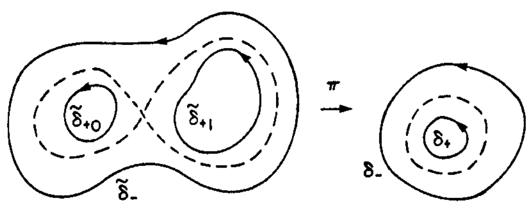

Fig. 5.2

In this case, Lemma 5.4 implies:

Lemma 5.6. *There are two possible lifts of S_δ to $N(\tilde{\delta})$: the first is a rotation of π radians about $\tilde{\delta}_-$ and extends to the identity on $\tilde{\delta}_{+0}$ and $\tilde{\delta}_{+1}$, the second is the identity on $\tilde{\delta}_-$ and interchanges $\tilde{\delta}_{+0}$ and $\tilde{\delta}_{+1}$.* $\square$

For the last technical lemma, define $W = A - \Phi^{-1}(N(\delta))$, where A is the stable region of R. In analogy with notation of Sect. 2, set $W^* = \Phi^{-1}(T^2 - N(\delta))$. The set $W - W^*$ consists of the grand orbits of the critical points and ∞.

Let U be a component of W.

Lemma 5.7. *If f fixes some point in U, then $f|U$ is the identity.*

Proof. If U contains one point, there is nothing to show. Each component of W which is not a point is the closure in W of a component U^* of W^*.

The restricted map

$$\Phi\colon U^* \to T^2 - N(\delta)$$

is a covering projection of its image (which may not be all of $T^2 - N(\delta)$). Since $N(\delta)$ contains the support of f, the restriction of $f \mid U^*$ is a covering translation.

If $z \in U$ and $f(z) = z$, either $z \in U$ or $z \in \partial U$, and in either case, f maps a deleted neighborhood V^* of z in U^* to a deleted neighborhood of z. Since any covering translation that fixes a point is the identity, the restriction of f to V^* is the identity. Since f is analytic, f is the identity on all of U^*, and its closure U. □

Proof of Theorem 5.3

The spin generator cases. Let γ be a generator of $\pi_1(F_{1/2})$ of type i)–iii) which we assume, without loss of generality, is contained in the closed right half plane. From Theorem 4.7, we know that the corresponding generator $i_* \upsilon_*(\gamma) \in \mathrm{MCG}_*(T^2)$ is either a spin or spin square about an oriented simple closed curve δ on T^2 which passes through the puncture corresponding to $\Phi(-1)$.

Case i) The loop γ encloses an orbit relation b such that

$$R_b^n(v_-) = R_b^m(v_+)$$

where $n \geq 0$, $m \geq 1$ are the smallest integers for which the relation holds. Theorem 4.7 implies

$$\begin{aligned} i_* \upsilon_* [\gamma] &= S_\delta \qquad m > 1 \\ &= S_\delta^2 \qquad m = 1. \end{aligned}$$

Since $i(\alpha, \delta) = i(\beta, \delta) = 0$, the curve δ divides T^2 into two parts, one of which is topologically a punctured disk Δ. We claim that the region $\hat{\mathbb{C}} - \Phi^{-1}(\Delta)$ is connected.

To see this, slightly enlarge Δ to a disk Δ' which contains both punctures. The grand orbits of both critical points are contained in Φ-lifts of Δ', but no single component of $\Phi^{-1}(\Delta')$ contains both v_+ and v_- since there is no orbit relation $v_+ = v_-$. Then Lemma 2.4 implies that $\hat{\mathbb{C}} - \Phi^{-1}(\Delta')$ is connected.

Shrinking Δ back to its original size, we conclude that $\hat{\mathbb{C}} - \Phi^{-1}(\Delta)$ is connected as well. Therefore,

$$\begin{aligned} U &= A - \Phi^{-1}(\Delta) \\ &= \hat{\mathbb{C}} - \{\Phi^{-1}(\Delta) \cup \text{Julia set}\} \end{aligned}$$

is also connected. The region U contains ∞ and $f(\infty) = \infty$, so by Lemma 5.7 $f \mid U$ is the identity. Since the Julia set of R is part of ∂U, it follows that f is the identity on the Julia set.

If γ is a generator of type ii) or iii), the curve δ on T^2 for which $i_* \upsilon_*(\gamma) = S_\delta$ satisfies $i(\alpha, \delta) = 0$ and $i(\beta, \delta) = 1$.

The analysis of cases ii) and iii) begins as in the proof of Proposition 5.2. From Lemma 2.5, there is a distinguished set of lifts $\{\tilde{\delta}_n\}$, $n \geqq 0$ and nested disks $\{\tilde{\Delta}_n\}$ containing ∞.

The disk $\tilde{\Delta}_0$ contains both critical values so its preimage $V = R^{-1}(\tilde{\Delta}_0)$ is doubly connected. The complement of V in $\hat{\mathbb{C}}$ is a pair of disks I_0 and I_1 which generate a Markov partition for the Julia set.

There is a distinguished lift $\tilde{\delta}_v \in \Lambda(\delta)$ that contains the critcal value v_-. The curve $\tilde{\delta}_v$ is contained in $\tilde{\Delta}_0$ and its preimage

$$\tilde{\delta}_c = R^{-1}(\tilde{\delta}_v)$$

is a figure eight curve contained in V.

In case ii), $\tilde{\delta}_v$ bounds a disk containing a single pole, while in case iii), the curve $\tilde{\delta}_0 = \tilde{\delta}_v$ so the disk bounded by $\tilde{\delta}_v$ contains ∞. Figure 5.3 illustrates the different locations of $\tilde{\delta}_c$ in cases ii) and iii).

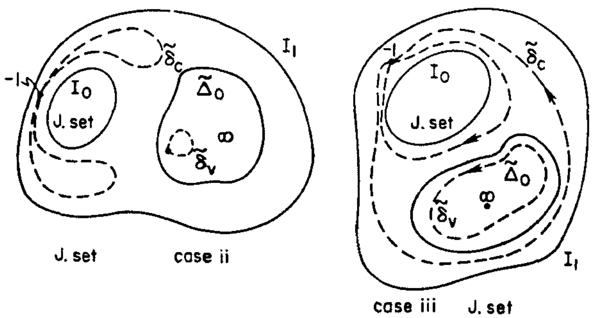

Fig. 5.3

Case ii) the loop γ encloses a pole of the form $R^n(v_-) = \infty$.

We will show that in this case, the lift f of S_δ to the dynamical plane of R restricts to the identity map on $\partial V = \partial I_0 \cup \partial I_1$.

The map f fixes each forward iterate of -1, and therefore maps each $N(\tilde{\delta}_n)$ homeomorphically to itself. Let $K = K(\delta)$ be the integer from Lemma 2.5. For $n \geqq K$, the restricted map

$$\Phi\colon N(\tilde{\delta}_n) \to N(\delta)$$

is a homeomorphism so that $f \mid N(\tilde{\delta}_n)$ is the spin $S_{\tilde{\delta}_n}$ for $n \geqq K$.

If $0 \leqq n < K$, the map

$$\Phi\colon N(\tilde{\delta}_n) \to N(\delta)$$

is a 2^n to 1 covering. Lemma 5.5 and induction imply the restricted map $f \mid N(\tilde{\delta}_n)$ is the unique Φ-lift of the spin S_δ that restricts to the identity on $N(\tilde{\delta}_n)$. We conclude that

$$\begin{aligned} f \mid \partial N(\tilde{\delta}_0) &= f \mid \partial \tilde{\Delta}_0 \\ &= \text{identity}. \end{aligned} \tag{1}$$

Refer to Fig. 5.3; there is a component U of $\hat{\mathbb{C}}-\Phi^{-1}(N(\delta))$ whose boundary consists of $\partial\tilde{\Delta}_0 \cup \partial I_0 \cup \partial I_1 \cup \tilde{\delta}_{c_-}$. From (1) and Lemma 5.7, the restricted map

$$f|\{\partial I_0 \cup \partial I_1\} = \text{identity}$$

so that f does *not* interchange the generators of the Markov partition for J_R. Consequently, Hedlund's theorem implies that $S_* f = \text{identity}$.

Case iii) the loop γ encloses the Mandelbröt set and all poles and orbit relations.

Let U be the component of $\hat{\mathbb{C}}-\Phi^{-1}(N(\delta))$ containing $+1$. The map f fixes $+1$, hence Lemma 2.7 implies that $f|U = \text{identity}$. Since ∂U contains $\tilde{\delta}_{c_-}$, the restriction of f to $\tilde{\delta}_{c_-}$ is the identity as well. Lemma 5.6 implies that f interchanges the simple boundary curves $\tilde{\delta}_{c_{+0}}$ and $\tilde{\delta}_{c_{+1}}$ and thus interchanges I_0 and I_1 as well. It follows from Theorem 5.1 that $S_* f$ generates $\mathrm{Aut}_2 \cong \mathbb{Z}_2$.

Case iv) The Dehn twist about the curve α on T^2 is the easiest to analyze. Recall from Sect. 2 that we defined the curve α on the torus T as the image under $\Phi\colon A^0 \to T$ of a level curve of the form

$$\tilde{\alpha}_0 = \{z \,|\, |\varphi(z)| = \text{constant}\}.$$

By choosing the constant appropriately, we can assume α does not pass through the branch values $\Phi(+1)$ or $\Phi(-1)$.

Lemma 2.3 implies that each $\tilde{\alpha} \in \Lambda(\alpha)$ is compact so the restriction

$$\Phi\colon \tilde{\alpha} \to \alpha$$

is a finite to one covering. We claim that Φ is 1:1 for every $\tilde{\alpha}$. To verify this, we refer to the diagram below.

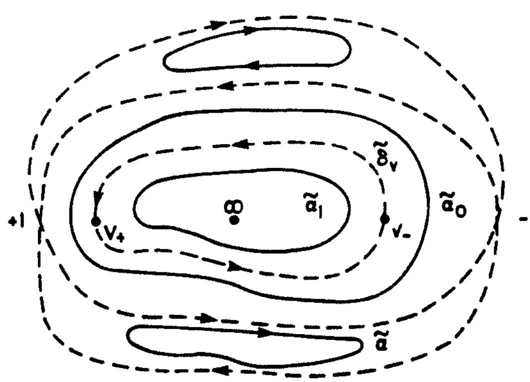

Fig. 5.4

The curve α satisfies the hypotheses of Lemma 2.5, so there is a family of distinguished lifts $\{\tilde{\alpha}_n | n \geqq 0\}$. Since our basepoint map R is symmetric, the disk $\Delta(\tilde{\alpha}_0)$ contains both critical values, the disks $\Delta(\tilde{\alpha}_n)$ contain neither critical value if $n > 0$.

Every curve in $\Lambda(\alpha)$ is mapped onto some $\tilde{\alpha}_n$ by an iterate of R. Suppose that R^j maps $\tilde{\alpha}$ onto $\tilde{\alpha}_n$; then either

i) $n > 0$: The region $R^{-j}(\tilde{\Delta}_n)$ consists of 2^j topological disks, each of which is mapped homeomorphically onto $\tilde{\Delta}_n$ under R^j. One such component is bounded by $\tilde{\alpha}$, therefore $R^j | \tilde{\alpha}$ is a homeomorphism.

ii) $n = 0$: The region $U = R^{-j}(\tilde{\Delta}_0)$ is connected and the restricted map

$$R^j \colon U \to \tilde{\Delta}_0$$

is a degree 2^j covering branched over v_+ and v_-. The Riemann Hurwitz formula implies that U has 2^j boundary components, each mapped homeomorphically onto $\partial \tilde{\Delta}_0 = \tilde{\alpha}_0$ by R^j. Since $\tilde{\alpha} \subset \partial U$, $R^j | \tilde{\alpha}$ is a homeomorphism.

In either case, $\Phi | \tilde{\alpha}$ is a homeomorphism. We conclude that the lift f of the Dehn twist D_α to the dynamical plane of R is given by

$$\begin{aligned} f(z) &= D_{\tilde{\alpha}}(z) \qquad && z \in N(\tilde{\alpha}),\ \tilde{\alpha} \in \Lambda(\alpha) \\ &= \text{identity} && \text{otherwise.} \end{aligned}$$

Therefore, f restricts to the identity on the Julia set as claimed.

Acknowledgements. It is a pleasure to thank the Institute for Advanced Study, the Max-Planck-Institut and the IBM Research Laboratories for their hospitality during the preparation of this article.

Special thanks to John Milnor for his many contributions to this article. The authors also wish to thank John Hubbard who introduced us to critical point surgery.

References

[AB] Ahlfors, L., Bers, L.: The Riemann mapping theorem for variable metrics. Ann. Math. **72**, 385–404 (1960)

[A] Ashley, J.: Marker automorphisms of the one-sided d-shift (to appear) J. Erg. Theory Dyn. Syst.

[Bi] Birman, J.S.: The algebraic structure of surface mapping class groups. In: Discrete Groups and Automorphic Functions, Harvey, W.S. (ed.). Academic Press 1977

[BDK] Blanchard, P., Devaney, R., Keen, L.: Geometric realizations of automorphisms of the one-sided d-shift (manuscript in preparation)

[BFK] Boyle, M., Franks, J., Kitchens, B.: Automorphisms of the one-sided shift and subshifts of finite type (to appear) J. Erg. Theory Dyn. Syst.

[D] Devaney, R.L.: An Introduction to Chaotic Dynamical Systems. Addison Wesley 1987

[DH1] Douady, A., Hubbard, J.H.: Etude Dynamique Des Polynomes Complexes I and II. Publ. Math. Orsay (unpublished notes)

[DH2] Douady, A., Hubbard, J.H.: On the dynamics of polynomial-like mappings. Ann. Sci. Ec. Norm. Super. N. Ser. **18**, 287–343 (1985)

[G] Goldberg, L.R.: On the mapping class group of a degree 3 Blaschke product. (manuscript in preparation)

[GK] Goldberg, L.R., Keen, L.: (unpublished)

[H] Hedlund, G.A.: Endomorphisms and automorphisms of the shift dynamical system. (Math. Syst. Theory **3**, 320–375 (1969)

[Ma] Massey, W.S.: Algebraic Topology: An Introduction. Harcourt, Brace and World, Inc., 1967
[Mc] McMullen, C.T.: Automorphisms of rational maps. In: Holomorphic Functions and Moduli. Drasin, D. (ed.). Berlin-Heidelberg-New York: Springer 1988
[MSS] Mane, R., Sad, P., Sullivan, D.: On the dynamics of rational maps. Ann. Sci. Ec. Norm. Sup., **16**, 193–217 (1983)
[R] Rees, M.: Hyperbolic rational maps of degree two. I. Preprint
[S] Sullivan, D.: Quasiconformal Homeomorphisms and Dynamics. III. Preprint
[WW] Whittaker, E.T., Watson, G.N.: A Course in Modern Analysis. Cambridge University Press, 1973
[Wit] Wittner, B.: Ph.D. Thesis. Cornell University, 1987 (unpublished)

Oblatum 17 – I – 1989 & 12-VI-1989