Astérisque

Mary Rees

Views of parameter space : topographer and resident

Astérisque, tome 288 (2003)

<http://www.numdam.org/item?id=AST_2003__288__R1_0>

ASTÉRISQUE 288

VIEWS OF PARAMETER SPACE: TOPOGRAPHER AND RESIDENT

Mary Rees

Société Mathématique de France 2003
Publié avec le concours du Centre National de la Recherche Scientifique

M. Rees
University of Liverpool, Division of Pure Mathematics,
Department of Mathematical Sciences, M & O Building, Liverpool, L693BX, UK.
E-mail : maryrees@liverpool.ac.uk

2000 *Mathematics Subject Classification.* — 37J40, 37J50, 32H50, 57NXX, 57N16, 30F60.

Key words and phrases. — Rational maps, branched coverings, critical point, Teichmüller space.

VIEWS OF PARAMETER SPACE: TOPOGRAPHER AND RESIDENT

Mary Rees

Abstract. — In this work, we investigate the structure of certain parameter spaces. The aim is to understand the variation of dynamics — in particular, of hyperbolic dynamics — in certain parameter spaces of rational maps. In order to do this, we examine the topological and geometric structure of larger parameter spaces, of branched coverings of the Riemann sphere $\overline{\mathbf{C}}$, where some of the critical points are constrained to have finite forward orbits.
We obtain a complete topological description of the spaces under consideration, from two points of view, which we call the *Topographer's View* and the *Resident's View*. The *Topographer's View* is, in essence, a geometrising theorem. It shows that the space in question is, up to homotopy equivalence, a countable union of disjoint geometric pieces, joined together by handles. The most typical geometric pieces are varieties of rational maps, and tori. The *Resident's View* is a view of the whole parameter space from the dynamical plane of a map (a resident) in the parameter space. This is necessarily a two-dimensional view, in which the geometric pieces of the parameter space appear as disjoint convex regions in the dynamical plane.

***Résumé* (Points de vue sur l'espace de paramètres: le topographe et le résident)**

Dans ce travail, nous étudions la structure de certains espaces de paramètres. L'objectif est de comprendre les variations de dynamique — en particulier de dynamique hyperbolique — dans certains espaces paramétrant des applications rationnelles. Pour cela, nous examinons la structure topologique et géométrique d'espaces plus grands paramétrant des revêtements ramifiés de la sphère de Riemann $\overline{\mathbf{C}}$, où plusieurs points critiques sont contraints à avoir une orbite positive finie.

Nous obtenons une description topologique complète des espaces considérés, de deux points de vue, que nous appelons la *vue du topographe* et la *vue du résident*. La vue topographique est, en somme, un théorème de géométrisation. Elle montre que l'espace en question est, à une équivalence d'homotopie près, une réunion dénombrable de morceaux géométriques disjoints, reliés ensembles par des anses. Les morceaux géométriques les plus typiques sont des variétés d'applications rationnelles et des tores. La *vue du résident* est une vue de l'espace des paramètres tout entier depuis le plan dynamique d'une application (un résident) situé dans l'espace des paramètres. C'est nécessairement une vue en dimension 2, dans laquelle les morceaux géométriques de l'espace des paramètres apparaissent comme des régions convexes disjointes dans le plan dynamique.

CONTENTS

PART I

TOPOLOGY, COMBINATORICS, VIEWS

INTRODUCTION

The objects of study in this paper are rational maps of the Riemann sphere $\overline{\mathbf{C}}$, considered as dynamical systems. The basic problem is to understand variation of dynamics in a given family of rational maps. The total dynamics of a rational map is greatly influenced by the dynamical behaviour of its critical points. So it is natural to consider families of rational maps in which some critical points are constrained to be periodic, or eventually periodic. Thus, we wish to study a parameter space of dynamical systems, with specified dynamics on some invariant set which varies isotopically throughout the parameter space. A rational map with a finite invariant set is a holomorphic map of a marked Riemann surface. So our object of study is a topological space in which the points are both dynamical systems and geometric structures.

Paths are important in topology. When the points in a topological space M are themselves mathematical objects, then paths in the space reflect this additional structure. For example, let S_0 be a compact topological surface, $M = M(S_0)$ the moduli space of Riemann surfaces homeomorphic to S_0, and let $S \in M$. We get different views of S from the endpoints of a closed homotopically nontrivial path in M based at S. If we wish to understand a space of mathematical objects, then we need an understanding of the different views of each mathematical object. This involves understanding the extra structure inherited by paths in M when the points in M have additional structure. For example, if M is as above, then closed paths in M based at S, which avoid singular points, give rise to homeomorphisms of S, modulo isotopy.

Study of any parameter space of dynamical systems involves looking at relative movement of points in the dynamical plane as a point moves in parameter space. This simple-minded idea manifests itself in virtually every paper written on dynamical systems. Sometimes the study of relative movement is local, as in, for example, basic theory of persistence (or otherwise) of fixed points and corresponding local dynamics. Sometimes it is global, as, for example in study of the Mandelbrot set for quadratic polynomials ([**D-H1**], [**D-H2**], [**T2**]), when, roughly speaking, it is possible to trace the movement of the unit circle (invariant under $z \mapsto z^2$) as the parameter varies throughout the Mandelbrot set (and even outside it). In fact, the idea is especially prevalent in holomorphic dynamical systems, where dynamics of all points are largely influenced by the dynamics of critical points. It is also especially true in any parameter space where the dynamics are constant on some isotopically varying set.

Now we consider movement in the dynamical plane as we move along a path in parameter space. In many parameter spaces (as in many topological spaces), the choice of path is important. It is possible to lose sight of this, because in the examples just cited — of local movement, and of movement in the Mandelbrot set — the choice of path is not important. Suppose for simplicity that we wish to consider the movement of one continuously varying point in the dynamical plane, relative to some other continuously varying point, where both are in the complement of an isotopically varying set on which dynamics are constant. Then a closed path in parameter space gives rise to a path in the dynamical plane minus the set with constant dynamics. (We shall assume for the moment that this path in the dynamical plane is closed: it depends on exactly how we make our definitions.) This gives rise to what I call the *Resident's View* of parameter space. This is a view which is comprehensive, but is in terms of the dynamical plane of some fixed dynamical system *resident* in the parameter space. It is really a view of the fundamental group of the parameter space V, that is, a set-theoretic identification of this group with a subset of the fundamental group of the complement $\overline{\mathbf{C}} \setminus Z$ in the dynamical plane of a set Z on which dynamics are constant. This is a very simple identification, and as such can be made for other of parameter spaces of dynamical systems. I do not know if there are other situations in which it has the far-reaching properties which it possesses for the parameter spaces considered here. It turns out that this map from one group to the other is injective. The map then gives rise to a map from the universal cover of V to a subset of the universal cover of $\overline{\mathbf{C}} \setminus Z$. It is then possible to obtain information about the variation of dynamics on V — from the *resident's view* of the universal cover. This is our aim. Some specific examples are given below (without full details or proof, of course). This is yet another application of the idea that, for many topological spaces, the fundamental group contains all the essential information about the topology, and that, in a topological space with additional structure, all the essential information can be obtained from a fundamental group with additional structure.

Given a topological space, a common aid to understanding structure is to show that a topological space has a stronger structure up to homeomorphism or homotopy equivalence — such as a geometric structure — of a particular type. Let us call the establishing of a such a structure a *Geometrizing Theorem*. Well-known examples of such results are Thurston's theorems [**T1**] producing hyperbolic structures on 3-manifolds (under various conditions). I want to highlight two Geometrizing Theorems which concern spaces of maps. One is the Nielsen-Thurston classification of compact surface homeomorphisms up to isotopy ([**F-L-P**], [**Cas**]). This can be interpreted as: a connected component of homeomorphisms of a compact surface contracts to a set of homeomorphisms preserving a geometric structure. This structure is either a hyperbolic metric on the surface, or a finite disjoint set of simple loops, or a pair of transverse measured foliations. The details are unimportant. The point is simply that the space of homeomorphisms is homotopically equivalent to a smaller space which has a more exact description. The other result is Thurston's Geometrizing Theorem for critically finite branched coverings of the sphere. This result can be regarded as a description of the topology of a connected component of orientation-preserving branched coverings with forward orbits of a fixed finite cardinality, modulo Möbius conjugation. In the presence of a certain combinatorial condition, the component is contractible to the unique rational map within it. In the absence of the combinatorial condition, the component is contractible to a space (usually a torus) of maps preserving a stronger geometric structure. Note that the basic intention of this result is arguably to use the larger space to study the smaller, so that this can be regarded as a Geometrization Theorem in reverse. The smaller set was initially known to be finite, and is shown to be singleton, but, of course, there is more to it than that.

There are other cases in which enlarging a topological space gives crucial information about the original. A large part of the long tradition of studying topological spaces of geometric structures comes from algebraic geometry. In such spaces, singularities are important, and it is often appropriate to *blow up* the singularity in order to understand the structure near it. In these examples, the enlarged space is often homotopically larger than the original.

In summary, both enlarging and reducing have been found useful in studying topological spaces, in a wide variety of situations. Sometimes, a space is found to be endowed with — or to be homotopically equivalent to a space endowed with — a geometric structure whose existence was not originally known. Sometimes a space of geometric structures is found to embed homotopically in a larger space, of which the points might be more easily constructed. This gives rise to what I shall call the *Topographer's View* of a Parameter space. Again, specific examples are given below, without full details.

Some Key Examples. — Both the Topographer's View and the Resident's View are important in understanding topological structure of a parameter space of rational maps, and, more importantly, the influence of the topology on the scope and type of variation in dynamical behaviour. These views are complementary. Neither is more important than the other. Now we view a number of examples, in an attempt to get a feel for the shape of the general results, which will be stated later. For the formal statements of results, see Chapter 5.

Example 1. — Let $V_{3,0}$ be the family of maps

$$g_a : z \longmapsto \frac{(z-a)(z-1)}{z^2}, \quad a \in \mathbf{C},\ a \neq 0,\ \pm 1,$$

where the critical points of g_a are $c_1(g_a) = 0$ and $c_2(g_a) = 2a/(a+1)$. The critical values are $v_1(g_a) = \infty$ and $v_2(g_a) = -(a-1)^2/4a$. Dynamics are constant on the set

$$Z = \{0, \infty, 1\},$$

which is the period 3 orbit of 0 under g_a, for any a. Note that for $a = 0$, g_a degenerates to a degree one map, and $a = \pm 1$ give $v_2(g_a) = 0$, 1 respectively. Let $B_{3,0} \supset V_{3,0}$ be the family of orientation-preserving degree two branched coverings g of $\overline{\mathbf{C}}$ with critical points $0 = c_1(g)$ and $c_2(g) \notin \{0, \infty, 1\}$, where $0 \mapsto \infty \mapsto 1 \mapsto 0$ is a periodic cycle under g. We shall see in Chapter 1 that the inclusion of $V_{3,0}$ in $B_{3,0}$ is a homotopy equivalence. This is the Topographer's View: that $B_{3,0}$ is, up to homotopy equivalence, the same as the subspace $V_{3,0}$ of rational maps within $B_{3,0}$. We shall, by abuse of notation, sometimes identify $V_{3,0}$ with $\mathbf{C} \setminus \{0, \pm 1\}$.

Now we consider the Resident's View. The set Z is forward-invariant under g, for any $g \in B_{3,0}$. By definition, $c_2(g) \notin Z$, where $c_2(g)$ is the second critical point of g. This is equivalent to the statement that $v_2(g) \notin Z$, where $v_2(g) = g(c_2(g))$ is the second critical value. The idea of the Resident's View is to describe the space $B_{3,0}$ in terms of $\overline{\mathbf{C}} \setminus Z$, which is the punctured dynamical plane of g, for any fixed $g \in B_{3,0}$. This is done by describing the universal cover of $B_{3,0}$, together with the action of the fundamental group $\pi_1(B_{3,0})$, in terms of the universal cover of $\overline{\mathbf{C}} \setminus Z$ together with the action of the fundamental group $\pi_1(\overline{\mathbf{C}} \setminus Z)$. Since inclusion of $V_{3,0}$ in $B_{3,0}$ is a homotopy equivalence, inclusion induces an isomorphism between $\pi_1(B_{3,0})$ and $\pi_1(V_{3,0})$. We need to choose basepoints in the fundamental groups, and we choose some $g_{a_0} \in V_{3,0}$, and then consider the fundamental groups $\pi_1(V_{3,0}, g_{a_0})$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g))$. The Resident's View includes an injective map

$$\rho : \pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g)) \longrightarrow \pi_1(V_{3,0}, g_{a_0}).$$

The map ρ is not a group homomorphism, and not surjective. The idea of its definition is as follows. If we move along a path in $B_{3,0}$ from g_{a_0} to g, then this determines a path in $\overline{\mathbf{C}} \setminus Z$ from $v_2(g)$ to $c_2(g)$, which (up to endpoint-preserving homotopy), depends only on the endpoint-preserving homotopy class of the path in $B_{3,0}$. If we trace a closed path in $B_{3,0}$ starting and ending at g_{a_0} then we obtain in this way a closed path in $\overline{\mathbf{C}} \setminus Z$ starting and ending at $v_2(g_{a_0})$. This does indeed give a map between the two fundamental groups. For a more formal definition, see 1.12. The fundamental groups act freely on the universal covers. So if we fix lifts $\widetilde{g}_{a_0}$ and $\widetilde{v}_2(g_{a_0})$, we have an injective map, which we also call ρ, from $\pi_1(V_{3,0}, g_{a_0}) \cdot \widetilde{g}$ to $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0})) \cdot \widetilde{v}_2(g_{a_0})$. Now, since $g_{a_0} \in V_{3,0}$, $\pi_1(V_{3,0}, g_{a_0}) \cdot \widetilde{g}_{a_0}$ can be regarded as a subset of the universal cover of $V_{3,0}$, while $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0})) \cdot \widetilde{v}_2(g_{a_0})$ is, of course, a subset of the universal cover of $\overline{\mathbf{C}} \setminus Z$. Both these universal covers are isomorphic to the open unit disc D: not just topologically, but conformally as well, since the universal covers inherit structures of complex manifolds from the complex manifolds $V_{3,0}$, $\overline{\mathbf{C}} \setminus Z$. So we can identify $\widetilde{g}_{a_0}$ and $\widetilde{v}_2(g_{a_0})$ with points in D. The covering-group-actions of the fundamental group $\pi_1(V_{3,0}, g_{a_0})$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0}))$ on the unit disc are by Möbius transformations. These actions extend to $\overline{D}$. Then it turns out that the map ρ, which is now a map between the subsets $\pi_1(V_{3,0}, g_{a_0}) \cdot \widetilde{g}_{a_0}$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0})) \cdot \widetilde{v}_2(g_{a_0})$ of D, extends continuously to a homeomorphism, also called ρ, of the unit circle ∂D. We can then use ρ to define a new action of $\pi_1(V_{3,0}, g_{a_0}) \cong \pi_1(B_{3,0}, g_{a_0})$ on ∂D by

$$g \cdot \rho(z) = \rho(g \cdot z) \tag{1}$$

for $g \in \pi_1(V_{3,0}, g_{a_0})$ and $z \in \partial D$, where the action on the righthand side is by Möbius transformations. The left-hand action is by homeomorphisms. The map ρ is not defined on the open unit disc, only on a discrete subset of it. It is, of course, possible to extend it homeomorphically to $\overline{D}$ and to then define a new action of $\pi_1(B_{3,0})$ on $\overline{D}$ by (1). Although the extension is not unique, the resulting action of $\pi_1(B_{3,0})$ on D is unique up to isomorphism of topological group actions. This is what we mean by the Resident's View in this case: an action of $\pi_1(B_{3,0}) \cong \pi_1(V_{3,0})$ on the universal cover of $\overline{\mathbf{C}} \setminus Z$ which is isomorphic the action of $\pi_1(V_{3,0})$ on the universal covering space of $V_{3,0}$.

In this very simple first example, the homeomorphism $\rho : \partial D \to \partial D$ has another description. Consider the map

$$\rho_1 : a \longmapsto v_2(g_a) : V_{3,0} \longrightarrow \overline{\mathbf{C}} \setminus Z,$$

which is actually a double covering. This then lifts to a map between the universal coverings, and is simply a Möbius transformation. This Möbius transformation coincides on ∂D — and only there — with ρ. The action of $\pi_1(B_{3,0})$ on ∂D defined using ρ is therefore an action by Möbius transformations — and each of these Möbius transformations is a Möbius transformation in some subgroup of $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0})$ acting on ∂D — but the subgroup is not $(\rho_1)_*(\pi_1(V_{3,0}, g_{a_0}))$.

Both $V_{3,0}$ and $\overline{\mathbf{C}} \setminus Z$ are punctured surfaces. The maps $g_{\pm 1}$ are well-defined degree two maps, with $v_2(g_1) = 0$ and $v_2(g_{-1}) = 1$, while g_0 is degree 1. Each puncture has lifts in ∂D. This works as follows. For any $A \subset \{0, \pm 1\}$, let $\pi_1(V_{3,0}, A, g_{a_0})$ denote homotopy classes of paths $\alpha : [0,1) \to V_{3,0}$ where

$$\lim_{t \to 1} \alpha(t) \in A,$$

and homotopy is through paths with this property. Let $\pi_1(\overline{\mathbf{C}} \setminus Z_1, Z, v_2(g_{a_0}))$ be similarly defined for any $Z_1 \subset Z$. Paths represented by elements of $\pi_1(V_{3,0}, A, g_{a_0})$, $\pi_1(\overline{\mathbf{C}} \setminus Z_1, Z, v_2(g_{a_0}))$ lift to paths in D starting from $\widetilde{g}$, $\widetilde{v}_2(g)$ respectively and limiting on points in ∂D lying in a countable set — which are, by definition, the lifts of the points of A, Z_1 to ∂D. The fundamental groups $\pi_1(V_{3,0}, g_{a_0})$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g))$ act on $\pi_1(V_{3,0}, A, g_{a_0})$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, Z_1, v_2(g_{a_0}))$ on the left, respectively, by composition of paths. If we identify a path with the lift of its second endpoint, then these actions identify with the restriction to the sets of lifts of the natural actions of $\pi_1(V_{3,0}, g_{a_0})$ and $\pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0}))$ on ∂D. Now we restrict to $A = \{\pm 1\}$ and $Z_1 = \{0,1\}$. Similarly to the map ρ on fundamental groups, there is a map

$$\rho_2 : \pi_1(V_{3,0}, \{\pm 1\}) \longrightarrow \pi_1(\overline{\mathbf{C}} \setminus Z, \{0,1\}, v_2(g_{a_0})),$$

which is defined in much the same way as ρ: a path from g_{a_0} to g_1, g_{-1} gives rise to a path from $v_2(g_{a_0})$ to 0, 1. So ρ_2 can be regarded as a map from lifts of ± 1 to lifts of 0, 1 again by identifying paths with the lifts of their second endpoints, that is, as a map from one countable subset of ∂D to another. Then ρ_2 is a restriction of $\rho : \partial D \to \partial D$.

We shall return to the Resident's View of this example in 1.15.

Example 2. — Now we consider $V_{3,1}$ and $B_{3,1}$. These are obtained by removing from $V_{3,0}$ and $B_{3,0}$ those maps g for which $v_2(g) \in g^{-1}(\{0,1,\infty\}) = Z(g)$, where $v_1(g) = \infty$ and $v_2(g)$ are the critical values of g. The space $V_{3,1}$ is simply $V_{3,0}$ with finitely many extra punctures. The inclusion map of $V_{3,1}$ in $B_{3,1}$ is no longer a homotopy equivalence, although (as always, it turns out) it induces an injection of fundamental groups. The group $\pi_1(B_{3,1})$ — which, of course, projects onto $\pi_1(B_{3,0}) \cong F_3$ — is now infinitely generated and infinitely presented. Before describing the structure, we consider the punctures 0, ∞ of $V_{3,1}$.

Note that if $a = 0$, then g_a degenerates to the period 3 Möbius transformation

$$z \longmapsto \frac{z-1}{z}.$$

Let γ_1 be a simple positively oriented loop round 0 and close to 0 in $V_{3,1}$. Now we consider g_a in $V_{3,1}$ for large values of a. Let

$$G_a(z) = \frac{1}{\sqrt{a}} g_a(\sqrt{a} z).$$

Then

$$G_a(z) = -\frac{1}{z} + \frac{1}{\sqrt{a}}\Big(1 + \frac{1}{z^2}\Big) - \frac{1}{az}.$$

So

$$\lim_{a\to\infty} G_a(z) = -\frac{1}{z}.$$

The critical values of G_a are $v_1 = \infty$ — which is of period two under G_a with orbit $\{0, \infty\}$ (as for g_a) and

$$v_2(G_a) = -\frac{(a-1)^2}{4a\sqrt{a}} = \frac{-\sqrt{a}}{4} + O\left(1\sqrt{a}\right).$$

Then

$$G_a(v_2) = \frac{5}{\sqrt{a}} + O\Big(\frac{1}{a\sqrt{a}}\Big), \quad G_a^2(v_2) = \frac{4\sqrt{a}}{5} + O\Big(\frac{1}{\sqrt{a}}\Big).$$

Let γ_2 be a simple positively oriented loop round ∞ and close to ∞ in $V_{3,1}$.

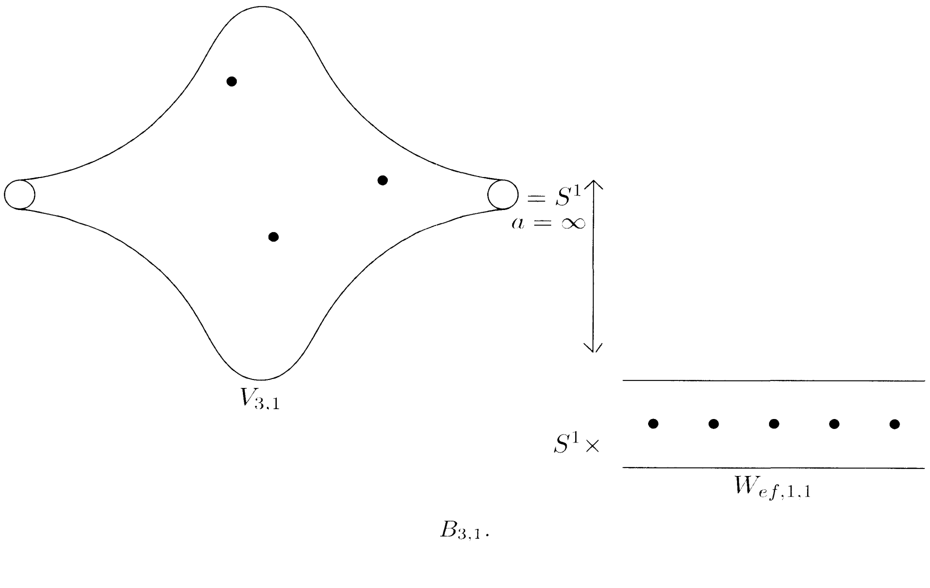
The Topographer's View of $B_{3,1}$ and $V_{3,1}$ is as follows. Up to homotopy equivalence, $B_{3,1}$ is the union of two pieces, joined together along a handle. One of the pieces is $V_{3,1}$, and the other is $S^1 \times W_{ef,1,1}$. Here, $W_{ef,1,1}$ is a cyclic cover of $V_{ef,1,1}$. The space $V_{ef,1,1}$ is the family of maps

$$k_b : z \longmapsto 1 + \frac{b}{z} - \frac{b}{z^2}, \quad b \in \mathbf{C},\ b \neq 0, -4.$$

The critical points of k_b are $c_1(k_b) = 0$ and $c_2(k_b) = 2$. The critical values are $v_1(k_b) = \infty = k_b(0)$, and $v_2(k_b) = 1 + b/4$. Also, $k_b^2(\infty) = k_b(\infty) = 1$. The "ef" is for "eventually fixed". The first index 1 means that the first forward image of $v_1(k_b) = \infty$ is fixed. The second 1 means that $v_2(k_b)$ is not a first preimage of $v_1(k_b) = \infty$, nor of the fixed point $k_b(\infty) = 1$. The restrictions on parameters ensure that for $k_b \in V_{ef,1,1}$, $v_2(k_b) \notin k_b^{-1}(\{\infty, 1\})$, and $\{\infty, 1\} = \{v_1(k_b), k_b(\{v_1(k_b)\})$ is forward invariant. We take $W_{ef,1,1}$ to be the cyclic cover for which the covering map is $\exp : W_{ef,1,1} \to V_{ef,1,1}$. Then we join γ_2 to the loop in $S^1 \times W_{ef,1,1}$ given by

$$t \longmapsto (e^{-it}, x) : [0, 2\pi] \longrightarrow S^1 \times W_{ef,1,1}, \text{ some fixed } x \in W_{ef,1,0},$$

to obtain $B_{3,1}$. We give a sketch of $B_{3,1}$ below.



$B_{3,1}$.

In particular, by the Topographer's View, $\pi_1(B, g_{a_0})$ can be regarded as containing $\pi_1(V_{3,1})$ and $\pi_1(S^1 \times W_{ef,1,1})$.

Now we consider the Resident's View, and we fix a resident $g_{a_0} \in V_{3,1}$. The sets $Z(g) = \{0, \infty, 1, x_0(g), x_1(g)\} = g^{-1}(\{0, 1, \infty\}$ and $Y(g) = Z(g) \cup \{v_2(g)\}$ vary isotopically as g varies over $B_{3,1}$. Here, $x_0(g)$ and $x_1(g)$ are preimages under g of 0 and 1 respectively. (The only preimage of ∞ is 0.) Write $Z = Z(g_{a_0})$, $B = B_{3,1}$ and $V = V_{3,1}$. We again — as in Example 1 — have a map

$$\rho : \pi_1(B, g_{a_0}) \longrightarrow \pi_1(\overline{\mathbf{C}} \setminus Z, v_2(g_{a_0})).$$

This map restricts to a map on $\pi_1(V, g_{a_0})$, which can again be regarded as a subset of the open unit disc, but this time, ρ does not extend continuously to a map on ∂D. What happens is more interesting.

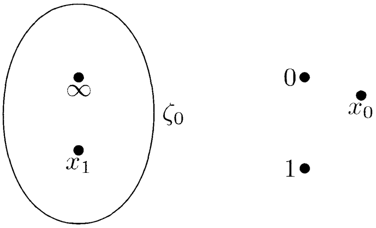
Let C denote the set of ends of B corresponding to degree two branched coverings g for which $v_2(g) \in Z(g)$. We can describe an end by having a family of neighbourhoods U_ε ($\varepsilon > 0$) such that for all g in U_ε, there is a loop in $\overline{\mathbf{C}} \setminus Y(g)$ of length $< \varepsilon$ with respect to the Poincaré metric on $\overline{\mathbf{C}} \setminus Y(g)$, which separates $v_2(g)$, and one particular point of $Z(g) \setminus \{v_1(g)\}$ from all the other points of $Z(g)$. We write $\pi_1(B, C, g_{a_0})$ for the set of paths in $\alpha : [0, 1) \to B$ with $\alpha(0) = g_{a_0}$, and $\lim_{t \to 1} \alpha(t) \in C$, modulo homotopy through paths of this type. Let $\pi_1(\overline{\mathbf{C}} \setminus Z, Zv_2(g_{a_0})$ be similarly defined. Then, again as in Example 1, there is a map

$$\rho_2 : \pi_1(B, C, g_{a_0}) \longrightarrow \pi_1(\overline{\mathbf{C}} \setminus Z, Z, v(g_{a_0})),$$

and by identifying elements of $\pi_1(B, C, g_{a_0})$, $\pi_1(\overline{\mathbf{C}} \setminus Z, Z, v(g_{a_0}))$ with lifts of second endpoints, ρ_2 can be regarded as a map between countable dense subsets of ∂D. But $\pi_1(V, V \cap C, g_{a_0})$, identified with lifts of its second endpoints, is not a dense subset of ∂D. Let $X(V)$ denote its closure. Let $H(V)$ denote the convex hull of $X(V)$, that is, the smallest convex subset of $\overline{D}$ which contains $X(V)$. The homeomorphism of ∂D given by any element of $\pi_1(B, g_{a_0})$ either preserves the set $X(V)$, or maps it to another set whose convex hull has interior disjoint from that of $H(V)$. Such convex hulls are known as *rational convex hulls*. The stabilizer in $\pi_1(B)$ of $H(V)$ is $\pi_1(V, g_{a_0})$. The action of $\pi_1(V)$ on $\partial H(V)$ is semiconjugate to the standard action of $\pi_1(V)$ on the boundary of the universal cover of V, via a map which sends the closure of each component of $\partial H(V) \setminus X(V)$ to a point, but otherwise sends distinct points to distinct points. The components in D of the complement of the rational convex hulls correspond naturally to $S^1 \times W_{ef,1,1}$, or, at least, to its preimages in the universal cover of B. These components are known as *Levy convex hulls*. Thus, the Resident's view of the universal cover of B, with the preimages of V and $S^1 \times W_{ef,1,1}$ in it, is D, decomposed into rational and Levy convex hulls. The correspondence is consistent with the actions of $\pi_1(B)$ on the universal cover of B and on ∂D.

The geodesics in the boundaries of the rational or Levy convex hulls have special properties. They all project down to the same simple closed loop ζ_2 in $\overline{\mathbf{C}} \setminus Z$. The loop ζ_2 is actually well-defined up to isotopy in $\mathbf{C} \setminus Y(g_{a_0})$ (as we shall see in Chapter 3, especially 3.13), and $\zeta_2 \subset g_{a_0}^{-1}(\zeta_2)$ up to isotopy in $\overline{\mathbf{C}} \setminus Z(g_{a_0})$. The loop ζ_2 is related to the cusp at $a = \infty$ in $V_{3,1}$ (which γ_2 is a simple loop round). Assume (without loss of generality) that a_0 is near ∞. Let ζ_2' be a bounded simple loop separating 0 and ∞. The isotopy class with respect to $Y(G_{a_0})$ is uniquely determined, because

$v_2(G_{a_0})$ is near ∞. The image of ζ_2' under the map $z \mapsto (1/\sqrt{a_0})z$ is ζ_2. The loop ζ_2 is a *Levy cycle*. For more details on Levy cycles, see Chapter 2. The isotopy class of ζ_2 is illustrated below.



Period 3 $n = 1$ bounding loop.

Example 3. — Now we consider $B_{3,2}$ and $V_{3,2}$, which are similarly defined, this time using $Z(g) = g^{-2}(\{0, \infty, 1\})$. Then $B_{3,2}$ is a chain of three geometric pieces: the pieces $V_{3,2}$ and $S^1 \times W_{ef,1,1}$ joined as in the case of $B_{3,1}$, and, in addition, a piece $V_{1,1,1}$, which is the family of maps

$$h_b : z \longmapsto bz + z^2, \quad b \in \mathbf{C},\ b \neq 0, 4.$$

The critical points of h_b are $c_1(h_b) = \infty$ (which is fixed) and $c_2(h_b) = -b/2$, so that the second critical value is $v_2(h_b) = -b^2/4$. The set $\{\infty, 0, -b\} = h_b(\{0, \infty\})$ contains two fixed points 0, ∞, and $-b$ is the other preimage of 0 under h_b. The restrictions on parameters in $V_{1,1,1}$ ensure that $v_2(h_b) \notin h_b^{-1}(\{0, \infty\})$, which may explain the notation: there is one fixed critical point, one other fixed point, and $v_2(h_b)$ avoids both these and their first preimages. Let γ_1' be the negatively oriented loop round 0 in $V_{1,1,1}$. Then join $S^1 \times W_{ef,1,1}$ to $V_{1,1,1}$ by joining $S^1 \times \{x\}$ and γ_1', for any fixed $x \in W_{ef,1,1}$ to obtain $B_{3,2}$. This is the Topographer's View.

The Resident's View is similar to Example 2. It comprises an identification of lifts in $\widetilde{B}_{3,2}$ of two different subspaces of $B_{3,2}$ with disjoint convex hulls in the disc D. These subspaces are (up to homotopy equivalence) $V_{3,2}$ and the join of $S^1 \times W_{ef,1,1}$ with $V_{1,1,1}$. It would be possible to make a finer decomposition into convex hulls, giving distinct convex hulls corresponding to the sets $S^1 \times W_{ef,1,1}$ and $V_{1,1,1}$, but we have elected not to do it in the main theorems of Chapter 5. The bounding geodesics are, again, lifts of a simple closed loop called ζ_2. It coincides with the loop of Example 2 up to the appropriate isotopy, but the set $Z(g)$ ha been enlarged, so it is now defined up to a finer isotopy. Once again, the loop generates a Levy cycle.

Example 4. — The space $B_{3,3}$ is a union of four pieces. The spaces $V_{3,3}$ and $S^1 \times W_{ef,1,2}$ replace $V_{3,2}$ and $S^1 \times W_{ef,1,1}$ in $B_{3,2}$. Here, $W_{ef,1,2}$ (like $W_{ef,1,1}$ with $V_{ef,1,1}$) is the cyclic cover of $V_{ef,1,2}$ given by the exponential map and $V_{ef,1,2}$ is the family of maps $k_b \in V_{ef,1,1}$ for which $v_2(k_b) \notin k_{b|}^{-2}(\{1, \infty\})$. In general $V_{ef,1,m}$ will be the family of maps $k_b \in V_{ef,1,1}$ for which $v_2(k_b) \notin k_b^{-m}(\{1, \infty\})$. The spaces $V_{3,3}$,

$S^1 \times W_{ef,1,2}$ and $V_{1,1,1}$ are joined as in $B_{3,2}$. Let $W_{1,1,1}$ be the cyclic cover of $V_{1,1,1}$ such that $\exp : W_{1,1,1} \to V_{1,1,1}$ is the covering. The last geometric piece is a copy of $S^1 \times W_{1,1,1}$ which is joined to $V_{3,3}$ by identifying γ_1 with the loop

$$t \longmapsto (e^{-it}, x) : [0, 2\pi] \longrightarrow S^1 \times W_{1,1,1}, \text{ some fixed } x \in W_{1,1,1}.$$

In the Resident's View the unit disc D is partitioned into convex hulls which are lifts to $\widetilde{B}_{3,3}$ of three different subspaces of $B_{3,3}$. These subspaces are $V_{3,3}$, the join of $S^1 \times W_{ef,1,2}$ and $V_{1,1,1}$, and $W_{1,1,1} \times S^1$. We have $Y(g) = g^{-3}(\{0, \infty, 1\}) \cup \{v_2(g)\}$, and the convex hull boundaries project to simple closed loops ζ_1 and ζ_2, which generate Levy cycles. This loop ζ_2 coincides up to suitable isotopy with the loop of that name in Examples 2 and 3. The loop ζ_1 is associated with the cusp 0 of $V_{3,3}$, in the same way that the loop ζ_2 is associated with the cusp ∞. If we assume (without loss of generality) that a_0 is close to 0, then we simply take ζ_1 to be a simple closed loop bounded from 0, 1, ∞, and separating ∞ from 0 and 1. Such a loop is uniquely determined up to isotopy in $\overline{\mathbf{C}} \setminus Z(g_{a_0})$, because all points of $Z(g_{a_0})$ are close to $\{0, \infty, 1\}$ for a_0 close to 0. We can even choose ζ_1 to be approximately periodic under g_{a_0}, because g_{a_0} is approximately a period three Möbius transformation over most of $\overline{\mathbf{C}}$.

Example 5. — We omit $B_{3,4}$ and consider $B_{3,5}$ briefly. The space $V_{3,5}$ replaces $V_{3,3}$, and $V_{1,1,2}$ replaces $V_{1,1,1}$ — where $V_{1,1,2}$ is simply the family of maps $h_b \in V_{1,1,1}$ such that $v_2(h_b) \notin h_b^{-2}(\{0, \infty\})$. In general $V_{1,1,m}$ will be a family of maps $h_b \in V_{1,1,1}$ such that $v_2(h_b) \notin h_b^{-m}(\{0, \infty\})$. The space $S^1 \times W_{ef,1,2}$ in $B_{3,3}$ is replaced by a union of two geometric pieces: $S^1 \times W_{ef,1,3}$ and $S^1 \times W_{ef,1,1}$. These are joined by identifying $S^1 \times \{x\}$ and $S^1 \times \{y\}$ for some fixed x and y. An extra copy of $V_{ef,1,1}$ occurs, sandwiched between the copy of $S^1 \times W_{ef,1,1}$ and a final copy of $V_{1,1,2}$. Let γ_1'' be the loop round 0 in $V_{ef,1,1}$. Then $S^1 \times \{x\}$ in $S^1 \times W_{ef,1,1}$ is joined to γ_1''. A simple loop round ∞ in $V_{ef,1,1}$ joined to a simple loop round 0 in $V_{1,1,2}$.

For $n > 5$, $B_{3,n}$ is always a union of finitely many geometric pieces. We always have copies of $S^1 \times W_{ef,1,m}$ and $S^1 \times W_{1,1,m'}$ (for varying m and m') joined directly to $V_{3,n}$. The chain of spaces between $S^1 \times W_{ef,1,m}$ and the copy of $V_{1,1,r}$ (for varying r) increases in length. The space $V_{1,1,r}$ is always the end of a chain of spaces.

The Resident's View is basically unchanged from that of Example 4, because we have chosen not to introduce a finer partition of convex hulls corresponding to the lifts up to homotopy convex hull of the complement of $V_{3,n}$: there are more convex hulls, simply because the fundamental group is larger, but they are of the same types as before.

Justification of String of Examples. — One of the reasons for considering this string of examples is to try to indicate that *not only* topological structure is being described. The topological changes in this string of examples are not important. Successively more critically finite maps become associated with cusps in the parameter space.

The Resident's View shows which of these are given by rational maps, and tells us how to determine which are distinct, by using the set-theoretic identification of the fundamental group $\pi_1(B_{3,n})$ into $\pi_1(\overline{\mathbf{C}} \setminus Z)$. The Resident's View is actually similar to a view obtained in [**R3**] and [**R4**], but is a finite version, which makes a better transfer of information possible, including the identification of one fundamental group into another.

Example 6. — We consider $B_{4,0}$. Note that $V_{4,0}$ identifies with the space of maps

$$g_{c,d} : z \longmapsto 1 + \frac{c}{z} + \frac{d}{z^2}$$

where $(c,d) \in \mathbf{C}^2$ and various points are removed. We compute these. The critical points are $0 = c_1(g_{c,d})$ and $-2d/c = c_2(g_{c,d})$, and $\infty = g_{c,d}(0)$, $1 = g_{c,d}(\infty)$, $1+c+d = g_{c,d}(1)$, and the condition $g_{c,d}^2(1) = 0$ yields

$$(1+c+d)(1+2c+d)+d = 0. \tag{1}$$

It follows that $V_{4,0}$ is a 10-times punctured sphere. Degeneracy of $g_{c,d}$ to a Möbius transformation occurs when $(c,d) = (-1,0)$ or $(-1/2,0)$, giving, respectively, period 3 and 4 Möbius transformations, and also as c, $d \to \infty$ with either $c+d = o(c)$ or $2c+d = o(c)$. In both cases, the conjugate of $g_{c,d}$ by $z \mapsto \sqrt{c}z$ is close to a period two Möbius transformation. The condition that v_2 is not in the periodic orbit of v_1 yields another 6 punctures.

Then $\pi_1(B_{4,0})$ is infinitely generated. The space $B_{4,0}$ is still the union of pieces with a clearly-defined geometric structure, but this time there are infinitely many pieces. Fortunately, the pieces are of only finitely many different homeomorphism types. One of these pieces is, of course, $V_{4,0}$. Countably infinitely many are of the form $S^1 \times S^1$. There is just one more, of the form $S^1 \times W_{2,1,0}$, where $W_{2,1,0}$ is an infinite cyclic cover of $V_{2,1,0}$. Here, $V_{2,1,0}$ is the space of maps

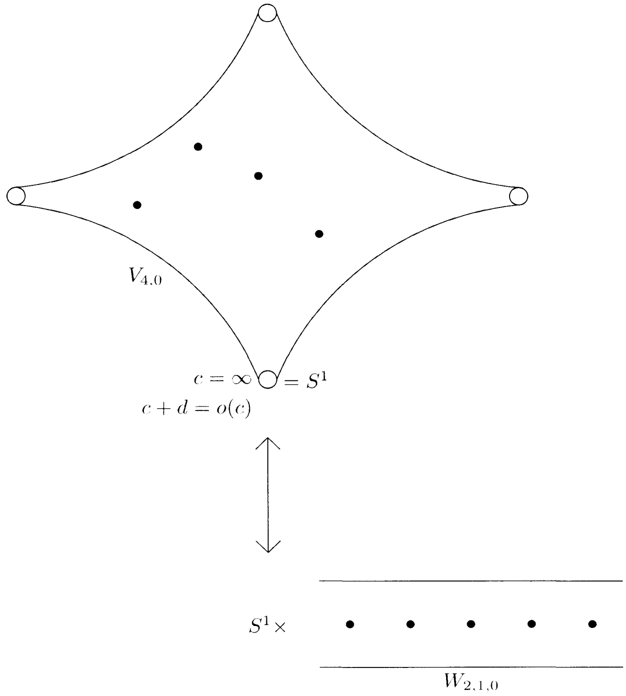
$$m_b : z \longmapsto \frac{(1-b)+bz}{z^2}, \quad b \in \mathbf{C},\ b \neq 0,1,2.$$

The critical points are $c_1(m_b) = 0$ and $c_2(m_b) = -2(1-b)/b$. Then $v_1(m_b) = \infty$, $c_1(m_b)$ is period 2 under m_b, and 1 is fixed by m_b. Then $W_{2,1,0}$ is the infinite cyclic cover for which $z \mapsto e^z + 1$ is the covering map. The join between $V_{4,0}$ and $S^1 \times W_{2,1,0}$ is similar to those in previous examples. Let γ_3 be a simple positively oriented loop round the end of $V_{4,0}$ with c, d large and $c+d = o(c)$. Then we identify γ_3 with the loop in $S^1 \times W_{2,1,0}$ given by

$$t \longmapsto (e^{-it}, x), \quad t \in [0,2\pi] \text{ some fixed } x \in W_{2,1,0}.$$

We write T_0 for this union of $V_{4,0}$ and $S^1 \times W_{2,1,0}$ (in line with notation that we shall adopt later). The space $B_{4,0}$ is an increasing union of spaces T_{κ_n}, where $\{\kappa_n\}_{n \geqslant 0}$ is an unbounded increasing sequence with $\kappa_0 = 0$. For $n \geqslant 1$, the $\kappa_n = \log \lambda_n$, where $\lambda_n > 1$ comprise the eigenvalues > 1 of hyperbolic matrices in $\mathrm{SL}(2,\mathbf{R})$, in increasing order. The infinitely many tori $S^1 \times S^1$ are indexed by κ_n for $n \geqslant 1$. For the n'th torus, the loop $e^{it} \mapsto (1, e^{it})$ ($t \in [0, 2\pi]$) is joined to some closed (probably not simple) loop δ_n in $T_{\kappa_{n-1}}$ which intersects $V_{4,0}$ nontrivially. I know nothing about the loops δ_n, except that they are all homotopically distinct.

We give a somewhat symbolic sketch of the subset T_0 of $B_{4,0}$ below. This is the Topographer's View.



Period 4 $n = 0$: the subset T_0.

Now we consider the Resident's View. Once again, the unit disc D is a union of convex hulls which are associated in a natural way with lifts to $\widetilde{B}_{4,0}$ of homotopically distinct subspaces of $B_{4,0}$. But this time we have infinitely many such subspaces, although all but two of them are tori. The boundary of any convex hull corresponding to a lift of $S^1 \times W_{2,1,0}$ is, similarly to before, a union of lifts of a simple closed loop generating a Levy cycle. The boundary of a convex hull corresponding to a torus is different. Any geodesic in such a boundary projects to a simple geodesic in $\overline{\mathbf{C}} \setminus Z(g)$ (for a suitable basepoint g), which can also be regarded as a simple geodesic $\overline{\ell}$ in $\overline{\mathbf{C}} \setminus Y(g)$, but this simple geodesic is not closed. Instead, its closure is a geodesic lamination, which up to isotopy is contained in its preimage under g. One can then use $g \mid \overline{\ell}$ to define an isotopy class of a pseudo-Anosov homeomorphism on $\overline{\mathbf{C}} \setminus Z(g)$.

Example 7. — We consider briefly $B_{4,n}$ for $n \geqslant 1$. For $n \geqslant 1$ there is an additional geometric piece homeomorphic to $S^1 \times W_{ef,1,m}$ (for varying m) joined to $V_{4,0}$ round the puncture $2c + d = o(c)$ for c, d large. This is within the space T_0 for $B_{4,n}$

For $n \geqslant 2$ there is another copy of $S^1 \times W_{ef,1,m}$ (for varying m) joined round the puncture $(c, d) = (-1, 0)$, and the space $S^1 \times W_{2,1,0}$, which is replaced by a union of finitely many pieces. These are also within the space T_0 for $B_{4,n}$.

For $n \geqslant 4$, there is a copy of $S^1 \times W_{1,1,m}$ (or, for larger n, this is replaced by a chain of finitely many pieces) joined round the puncture $(c, d) = (-1/2, 0)$, and the infinitely many copies of $S^1 \times S^1$ become $S^1 \times S^1 \times W_{1,1,m'}$ (which is replaced by a chain of finitely many pieces for larger n). No chains develop between $V_{4,n}$ and the first adjacent spaces. This is the Topographers View. The Resident's View is similar to that in Example 6.

Example 8. — We consider one feature of $B_{6,0}$. Stimson's calculations [**Sti**] indicate that $V_{6,0}$ is a punctured surface of genus 6. This time, one of the geometric pieces joined to $V_{6,0}$ in $B_{6,0}$ is $S^1 \times W_{3,1,0}$, where $W_{3,1,0}$ is a cyclic cover of $V_{3,1,0}$. In the Topographer's View, there are three handles joining $S^1 \times W_{3,1,0}$ to $V_{6,0}$. The join is from loops round three different punctures in $V_{6,0}$ to the loop $S^1 \times \{x\}$ in $S^1 \times W_{3,1,0}$ for any fixed x in $W_{3,1,0}$. This has consequences for the Resident's View also. This is one of the simplest parameter spaces in which convex hulls corresponding to lifts of one of the "subsidiary spaces" — $S^1 \times W_{3,1,0}$ in this case — have boundary geodesics of more than one type, in fact of three different types in this case. The geodesics project to simple closed geodesics which generate three nonisomorphic Levy cycles.

Example 9. — We consider an example of a different type. For $c \in \mathbf{C}$ let

$$f_c(z) = z^2 + c,$$

and, for $p \geqslant 2$ let V be an irreducible component of

$$V = \{g = f_c \circ f_d : c, d \in \mathbf{C},\ 0 \text{ has period } p,\ c \notin \{g^i(0) : 0 \leqslant i \leqslant p\}\},$$

quotiented by conjugation by Möbius transformations. (This actually means quotienting by the equivalence relation $(c, d) \sim (\omega c, \omega^2 d) \sim (\omega^2 c, \omega d)$, where $\omega^3 = 1$, $\omega \neq 1$.) Let B be the larger space of degree four orientation-preserving branched coverings with fixed critical point ∞ of multiplicity 3 and three other critical points, including 0 and two others which map to the same critical value $v_2(g)$, such that $0 = c_1(g)$ has period p, $v_2(g) \notin \{g^i(0) : 0 \leqslant i \leqslant p\} = Z(g)$. Such a space B is *of polynomial type* because of the presence of a fixed critical point of maximal multiplicity. The Topographer's View is that the inclusion of V in B is a homotopy equivalence. The Resident's View is that the identification of $\pi_1(V) \cong \pi_1(B)$ into $\pi_1(\overline{\mathbf{C}} \setminus Z(g))$ is injective and extends continuously to a homeomorphism of the boundary ∂D of the universal cover D of V to the boundary ∂D of the universal cover D of $\overline{\mathbf{C}} \setminus Z(g)$, so that the covering group $\pi_1(V)$ can be regarded as acting on either circle ∂D. (The

Resident's View is not actually proved for polynomial type in this paper, but it is true.)

We give a very rough statement of the Topographer's View and the Resident's View below. The results concern a space B of branched coverings of either *polynomial* or *degree two* type. It is hoped that the examples above indicate roughly what these words mean. Full statements are deferred until Chapter 5. The results are split between five theorems in Chapter 5. All necessary concepts are defined in Chapters 1–4.

Topographer's View

(1) *Let B be of polynomial periodic type. Then the inclusion of the subspace V of polynomials in B is a homotopy equivalence.*

(2) *Let B be of degree two periodic or eventually fixed type and V the subspace of rational maps. Then the inclusion of V_1 in B is injective on fundamental groups, for any component V_1 of V. The space B is an increasing union of spaces B_{κ_n}. In some cases, we have $\kappa_0 = 0$ only, but in the other cases, $\{\kappa_n\}_{n\geqslant 0}$ is an increasing unbounded sequence with $\kappa_0 = 0$ and $V \subset B_0$. For each $n \geqslant 0$, B_{κ_n} is homotopy equivalent to a finite ordered graph of topological spaces. For each such graph, the edge topological spaces are all tori. For $n = 0$, the node topological spaces include the components of V. For $n \geqslant 1$, the node topological spaces include the components of $B_{\kappa_{n-1}}$. The graphs and their nodes and edges can be computed from the* Resident's View*.*

Resident's View. — *Let B be of degree two type, with one critical point periodic or eventually fixed for all maps in B. Let V denote the space of rational maps in B. Let $G = \pi_1(B)$. Let $\widetilde{B}$ denote the universal cover of B. Let D denote the closed unit disc. There is a partition $\mathcal{P}(B)$ of D into convex regions, and an action of G on ∂D, such that the following hold. The action of G on ∂D extends to an action on the convex regions. There is a G-invariant coarsening $\mathcal{P}_n(B)$ of the partition $\mathcal{P}(B)$ restricted to a subset of D, such that the quotient by the G-action of the dual graph of $\mathcal{P}_n(B)$ is the graph used to describe B_{κ_0} from V if $n = 0$, and B_{κ_n} from $B_{\kappa_{n-1}}$ if $n > 0$. Each lift $\widetilde{V}_1$ to $\widetilde{B}$ of a component V_1 of V corresponds to a single convex region C in the partition $\mathcal{P}(B)$, whose stabiliser in G is $\pi_1(V_1)$. There is a monotone map from $\partial D \cong \partial\widetilde{V}_1$ to $\partial D \cap \partial C$ minus countably many points, which is homomorphism of the two $\pi_1(V_1)$-actions. The inverse of this map extends continuously monotonically to map ∂C to $\partial D \cong \partial\widetilde{V}_1$.*

All structure can be computed from any $f \in B$.

There is, implicit in the Resident's View, a combinatorial condition for the space V of rational maps in B to be connected

The Topographer's and Resident's Views are proved using an iteration on a finite-dimensional Teichmüller space which is analogous to the technique used by Thurston to prove a geometrisation theorem [**D-H3**], [**T2**] for critically finite branched coverings of $\overline{\mathbf{C}}$. Both the method of proof and the results obtained can be regarded as a special case of a generalisation of Thurston's Theorem. We summarize a reformulation of Thurston's Theorem in Chapter 1. It is clear from the present work that the degree two result that we obtain is indeed only a special case of a generalisation of Thurston's Theorem to families (rather than single maps) of branched coverings in which some critical points are constrained to have finite orbits. In order to deal with the higher degree case, one has to understand Thurston obstructions better than I do at present. I believe that work of Shishikura (dating back to at least 1988) in which he derives trees with expanding metrics from Thurston obstructions, essentially carries out the classification of Thurston obstructions, but he has not yet written up this work.

Thurston's theorem gives a way of projecting a critically finite branched covering to a rational map which is actually semiconjugate to it, if a certain combinatorial condition is satisfied. (The method also gives much information even in the absence of the combinatorial condition.) The present method also involves projecting branched coverings to branched coverings which are either rational maps or preserve some geometric structure. But there is not, nor was there ever intended to be, any relation between the dynamics of individual branched coverings and their projections. However, projection along an entire path can, and often does, preserve dynamical information in some sense. The present work, like Thurston's Theorem, involves finite-dimensional Teichmüller spaces.

The organization of this paper is as follows.

Chapters 1–7. — We develop the basic concepts and theory which enable us to state the main results, the *Topographer's View* and the *Resident's View*, in Chapter 5. The theory developed in these chapters is mostly combinatorial or topological in nature. In Chapter 1, we summarize some basic material about mapping class groups of punctured spheres, and adapt it to our purposes. We develop a theory of invariant loop sets satisfying a *Levy Condition* in Chapters 2-4. This theory is suggested by the theory of Thurston obstructions [**T2**] for critically finite branched coverings, which has been extensively studied by many authors ([**TL**], [**L**] for example). In Chapters 6-7 we start to reduce the proofs and identify the main steps.

Chapters 8–16. — We develop the theory of Teichmüller distance which we will need in the proofs. The theory developed is specific to Teichmüller spaces of marked spheres, although some of it extends without any difficulty to the Teichmüller space of any compact marked surface. Some of the material appears to be new. Other material undoubtedly is not, but in any case work is needed to get it into our context. We start (Chapter 8) with a formula for the first derivative of Teichmüller distance, which is clearly related to Earle's original formula [**Ear**] but in a special case where very simple coordinates can be used. The proof also bears a resemblance to that of F. Gardiner [**Gar1**]. We then derive a formula (Chapters 10–13) for the second derivative. I am not aware of any existing analogue in the literature, although I am sure the distance function is known to be generically C^2. We actually show that the distance function is C^2 everywhere at nonzero distance for marked spheres, a result which appears to be new. Other topics covered include: an analysis of the distance function in the thin part of Teichmüller space (Chapters 9, 14–15); a characterization of points on or near a geodesic between two points in Teichmüller space (Chapters 14–15); triangles of geodesics in Teichmüller space (Chapter 15); and a number of results about when quadratic differentials can be expected to have the "same shape" in the thin part of Teichmüller space.

Chapters 17–24: Proof of the Topographer's View. — The steps needing proof were isolated in Chapter 7. Basically there are two of them, covered in Chapters 17–21 and 22–24 respectively. The first involves using a natural analogue of Thurston's pullback function [**T2**], [**D-H3**], together with *pushing* (a technique which forced much of the development of the Teichmüller theory) to homotope certain branched covering spaces to subsets with a geometric structure, joined together by handles. The second step is then an analysis of these handles. This analysis made necessary the interpretation of the second derivative of the Teichmüller distance function in Chapters 10–13.

Chapters 25–31: Proof of the Resident's View of Rational Maps Space. — The proof of the *Resident's View* was reduced to this step in Chapter 7. This is basically a result about extending a map from one disc to another to a map of one boundary to another. Such a result is a key point in the proof of the Mostow rigidity theorem [**Mos**], a result to be borne in mind, although in the present case the map of discs is very far from a quasi-isometry. The proof involves extensive study of how covers of certain moduli spaces -moduli spaces of rational maps — sit inside Teichmüller space. Thurston pullback has to be employed again, although this time it is "foreshortened" in a certain way. The theory of triangles of geodesics, and the analysis of points near geodesics between two given points (Chapters 14-15) is important in this proof.

Acknowledgements. — This article has been "in progress" since the autumn of 1990. I have therefore sought and received help and encouragement in this work from a number of different people at various times. I am very grateful to all of them, but I should particularly like to thank C. Earle, L. Keen, N. Lakic, M. Lyubich, C. McMullen, P. Roesch, D. Schleicher, M. Shishikura, J. Stimson, Tan Lei and C. Taylor. In addition, I am very grateful to P. Roesch for her help with the French version of my abstract. I should like to thank SUNY at Stony Brook for financial and other support in 1991, and MSRI at Berkeley for support in 1995. Finally, I am enormously grateful to the referee, who was presented with an almost impossible task.

CHAPTER 1

THE TOPOLOGY OF SPACES OF HOMEOMORPHISMS AND BRANCHED COVERINGS

1.1. Homeomorphisms of the Sphere. — It may be as well to start by stating the obvious. We shall use $\overline{\mathbf{C}}$ throughout to denote the two-dimensional sphere, even when we are not considering the complex structure. We denote by $\mathrm{Hom}_+(\overline{\mathbf{C}})$ the topological group of orientation-preserving homeomorphisms of the sphere. This space is path connected and locally contractible. The inclusion within it of the group of Möbius transformations is a homotopy equivalence. Thus, the fundamental group of $\mathrm{Hom}_+(\overline{\mathbf{C}})$ is that of $\mathrm{PSL}_2(\mathbf{C})$, which is $\mathbf{Z}_2$. Hence, the quotient of $\mathrm{Hom}_+(\overline{\mathbf{C}})$ by left (or right) composition by Möbius transformations is simply connected. Let $X \subset \overline{\mathbf{C}}$ be a finite set. We can also consider the set $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ of orientation-preserving homeomorphisms fixing X pointwise. This topological group is, of course, not connected, but locally path-connected and locally contractible. If X has one or two points, without loss of generality, $X = \{\infty\}$ or $\{0, \infty\}$, the inclusions of, respectively, affine transformations or scalings $z \mapsto \lambda z$ ($\lambda \in \mathbf{C}^*$) in the identity component of $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ are homotopy equivalences. In these cases, the quotient of the components of $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ by right or left composition by these subgroup of Möbius transformations are again simply-connected. If X has three or more points, the components of $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ are simply connected. The group $\pi_0(\mathrm{Hom}_+(\overline{\mathbf{C}}, X))$ is, of course, the pure modular group $\mathrm{PMG}(\overline{\mathbf{C}}, X)$, more usually described as the group $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ modulo isotopies fixing X. The larger group $\mathrm{MG}(\overline{\mathbf{C}}, X)$ is the quotient, by isotopies that are constant on X, of the group of homeomorphisms of $\overline{\mathbf{C}}$ which map the set X to itself.

1.2. $\boldsymbol{\sigma_\alpha}$**.** — As in [**R3**], [**R4**], we use the notation σ_α for a homeomorphism of $\overline{\mathbf{C}}$ which is defined by a path α, where α is any locally injective path in $\overline{\mathbf{C}}$, parametrised, for convenience, by $[0,1]$. If α is injective, the homeomorphism σ_α is defined to be the identity outside a small disc neighbourhood of $\mathrm{Im}(\alpha)$, and to map $\alpha(0)$ to $\alpha(1)$. In general, we write α as a product $\alpha_1 * \cdots * \alpha_r$ of simple paths, and define $\sigma_\alpha = \sigma_{\alpha_r} \circ \cdots \circ \sigma_{\alpha_1}$. Although this is not a very precise definition, if, for example,

α is a closed loop with $\alpha((0,1)) \subset \overline{\mathbf{C}} \setminus X$ but $\alpha(0) = \alpha(1) \in X$ for some finite subset X, then σ_α is well defined up to isotopy constant on X, and depends only on the homotopy class of α in $\overline{\mathbf{C}} \setminus X$. Such homeomorphisms are commonly used to generate $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ ([**Mag**], [**F-B**], for example). Similarly, if X_t $(0 \leqslant t \leqslant 1)$ is a continuous family of finite subsets of $\overline{\mathbf{C}}$, and $\#(X_t)$ is constant, then there is a path φ_t of homeomorphisms with $\varphi_0 =$ identity and $\varphi_t(X_0) = X_t$, and if $X_0 = X_1$, the isotopy class of φ_1, relative to isotopies which are constant on X_0, is determined by the path X_t. Therefore, $\mathrm{Hom}_+(\overline{\mathbf{C}}, X)$ is locally contractible.

1.3. Homomorphisms from Fundamental Groups and Pure Braid Groups into Pure Modular Groups. — Let $X \subset \overline{\mathbf{C}}$ be a fixed finite set, and first fix $x_0 \in X$. Then $[\alpha] \mapsto [\sigma_\alpha^{-1}]$ defines a homomorphism of $\pi_1(\overline{\mathbf{C}} \setminus (X \setminus \{x_0\}), x_0)$ into $\mathrm{PMG}(\overline{\mathbf{C}}, X)$. Of course, if X has $\leqslant 3$ points, the image of the homomorphism must be trivial, since $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ is trivial in these cases. But it is not hard to show that, if X has > 3 points, $[\alpha] \mapsto [\sigma_\alpha^{-1}]$ is an isomorphism onto its image. (Here, [,] denotes homotopy class and isotopy class respectively. In future, we shall frequently confuse loops and their homotopy classes, homeomorphisms and their isotopy classes.) For example, we can use the following. If β is a simple closed nontrivial, nonperipheral loop in $\overline{\mathbf{C}} \setminus X$ which has essential intersections with α, then $\sigma_\alpha(\beta)$ and β are nonisotopic. If X has 4 points, then $[\alpha] \mapsto [\sigma_\alpha^{-1}]$ is an isomorphism onto $\mathrm{PMG}(\overline{\mathbf{C}}, X)$, which is thus a free group on two generators. For given an arbitrary $[\varphi] \in \mathrm{PMG}(\overline{\mathbf{C}}, X)$, there is an isotopy φ_t of $\overline{\mathbf{C}}$ which is constant on $X \setminus \{x_0\}$ with $\varphi_0 =$ identity, $\varphi_1 = \varphi$. We then take $\alpha(t) = \varphi_t(x_0)$, and find that $[\sigma_\alpha] = [\varphi]$.

The path X_t of 1.2 is a path in the topological space $\mathcal{Y}$ of ordered sets in $\overline{\mathbf{C}}$ of cardinality n (for some n). The fundamental group of this space is the so-called pure spherical braid group. We can find a path σ_t through Möbius transformations such that the first three elements of $\sigma_t X_t$ are constant in t. This induces a homeomorphism between $\mathcal{Y}$ and $\mathcal{Y}/\mathrm{PSL}_2(\mathbf{C}) \times \mathrm{PSL}_2(\mathbf{C})$, where $\mathcal{Y}/\mathrm{PSL}_2(\mathbf{C})$ denotes the quotient by left Möbius composition, which is the pure modular space. We also have an isomorphism between the corresponding fundamental groups. But the fundamental group of $\mathcal{Y}/\mathrm{PSL}_2(\mathbf{C})$ is isomorphic to $\mathrm{PMG}(\overline{\mathbf{C}}, X)$, where $\#(X) = n$, and the isomorphism is given by $\{\sigma_t X_t\} \mapsto [\varphi_1^{-1}]$, using the notation of 1.2. Recalling that the fundamental group of $\mathrm{PSL}_2(\mathbf{C})$ is $\mathbf{Z}_2$, we thus have the well-known isomorphism between the pure spherical braid group and $\mathrm{PMG}(\overline{\mathbf{C}}, X) \times \mathbf{Z}_2$.

1.4. The Presentation of the Pure Modular Group. — Although it is well-known (see [**F-B**], for example), it seems a good idea to introduce here the presentation of $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ for a finite subset X. This entire paper is concerned with certain subgroups of the modular group or pure modular group for varying X, since these are the fundamental groups of the branched covering spaces that we are studying.

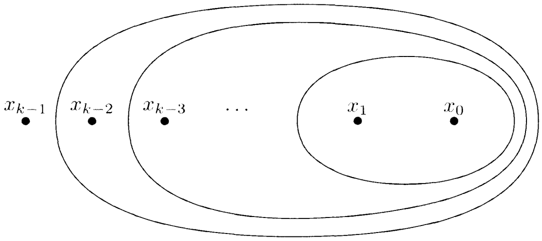
The very standard type of argument of 1.3 is all that is needed to derive the presentation of $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ in general. If $X = \{x_i : 0 \leqslant i \leqslant k-1\}$, let D_i $(1 \leqslant i \leqslant k-2)$ be a decreasing sequence of closed topological discs, such that D_i contains precisely the points x_j $(j \leqslant i)$.

Let H_i $(2 \leqslant i \leqslant k-2)$ be the subgroup of $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ generated by

$$\{[\sigma_\alpha] : \alpha(0) = \alpha(1) = x_i, \ \mathrm{Im}(\alpha) \subset D_i\}.$$

Then, as we have already seen, H_i is a free group on i $(\geqslant 2)$ generators. It is easy to choose a generating set, in terms of loops α, but we leave that to individual choice. It is also clear that, for $j < i$, conjugation by and element of H_j is an automorphism of H_i. For if β and α are closed loops based at possibly different points and $\varphi = \sigma_\beta$,

$$\sigma_\beta \circ \sigma_\alpha \circ \sigma_\beta^{-1} = \sigma_{\varphi\alpha}.$$



Increasing Discs.

Once again, it is easy (but very tedious) to write down the automorphism explicitly in terms of any generating set of H_i, for any generating set of H_j. Finally, given any isotopy class $[\varphi]$, we can compose successively by unique $\psi_i \in H_i$ so that $\psi_j \circ \cdots \circ \psi_{k-2} \circ \varphi$ preserves the disc boundaries ∂D_ℓ $(\ell \geqslant j-1)$ up to isotopy. (Actually, ∂D_{n-2} is automatically preserved.) Then we can compose successively with Dehn twists $\xi_i \in H_i$ round ∂D_{i-1} $(2 \leqslant i \leqslant n-2)$ so that $\xi_j \circ \cdots \circ \xi_2 \circ \psi_2 \circ \cdots \circ \psi_{n-2} \circ \varphi$ is isotopic to the identity restricted to D_j, modulo Dehn twist round ∂D_j. Then using the normalisation of H_i by H_j $(j < i)$, we can rewrite

$$[\varphi] = [\psi_{n-2}^{-1} \circ \cdots \circ \psi_2^{-1} \circ \xi_2^{-1} \circ \cdots \circ \xi_{n-2}^{-1}]$$

as

$$[\varphi] = [\varphi_{n-2} \circ \cdots \circ \varphi_2],$$

with $\varphi_i \in H_i$. Moreover, this representation is unique, because if

$$[\varphi_{n-2} \circ \cdots \circ \varphi_2] = [\text{identity}],$$

then we see by induction (for decreasing i), first, that φ_i preserves D_{i-1}, and then that φ_i is isotopic to the identity. This defines the presentation.

1.5. Some Useful Facts about the Presentation. — I prefer not to give an explicit presentation of the pure modular group, but here are a couple of useful facts concerning it.

(1) If X has $\geqslant 5$ elements, then $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ has nontrivial commuting elements. For example, σ_α and σ_β commute whenever the loops α and β are disjoint. Thus $\mathrm{PMG}(\overline{\mathbf{C}}, X)$ cannot be isomorphic to a Fuchsian group if $\#(X) \geqslant 5$. (This might seem irrelevant.)

(2) If α is a simple closed loop, then σ_α is, up to isotopy, a composition of Dehn twists round the loops bounding an annulus neighbourhood of α. If α encloses a single point of X (but, of course, is based at another point of X), then one of these twists is trivial. If X has 4 points, this enables us to write σ_α, for α based at x_0, in the form σ_β, for some β based at x_1, given $x_1 \in X$. We shall use this in 1.11.

1.6. Branched Coverings of the Sphere. — Let $f : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ be any degree d orientation-preserving branched covering with critical value set Y. Then

$$f : \overline{\mathbf{C}} \smallsetminus f^{-1}(Y) \longrightarrow \overline{\mathbf{C}} \smallsetminus Y$$

is a covering. Let U denote the universal covering space of $\overline{\mathbf{C}} \smallsetminus Y$ with covering map $\pi : U \to \overline{\mathbf{C}} \smallsetminus Y$. Then the fundamental group $\pi_1(\overline{\mathbf{C}} \smallsetminus Y)$ acts freely on U, and for any subgroup K of $\pi_1(\overline{\mathbf{C}} \smallsetminus Y)$, the quotient space U/K is also a covering of $\overline{\mathbf{C}} \smallsetminus Y$ with well-defined covering map π_K given by $\pi_K([u]_K) = \pi(u)$, where $[u]_K$ denotes the orbit of $u \in U$ in U/K. The point of covering space theory is that all coverings of $\overline{\mathbf{C}} \smallsetminus Y$ are of this form. In particular, f determines a conjugacy class of subgroup H of $\pi_1(\overline{\mathbf{C}} \smallsetminus Y)$ of index d, consisting of those based loops whose based lifts to $\overline{\mathbf{C}} \smallsetminus f^{-1}(Y)$ are closed. Then by the standard covering space theory there is a homeomorphism

$$\psi_2 : \overline{\mathbf{C}} \smallsetminus f^{-1}(Y) \longrightarrow U/H$$

such that $f = \pi_H \circ \psi_2$. Now any covering space U/K of $\overline{\mathbf{C}} \smallsetminus Y$ also has the structure of a complex manifold in such a way that π_K is holomorphic. If K is of finite index in $\pi_1(\overline{\mathbf{C}} \smallsetminus Y)$, then U/K must be $\overline{\mathbf{C}} \smallsetminus W_0$, up to holomorphic equivalence, for some finite set W_0. Any finite degree holomorphic map from one punctured Riemann sphere to another extends to a rational map of the Riemann sphere. So there is a finite set $W_0 \subset \overline{\mathbf{C}}$, a rational map $f_0 : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ with $f_0^{-1}(Y) = W_0$ and a homeomorphism

$$\psi_0 : \overline{\mathbf{C}} \smallsetminus Y_1 \longrightarrow U/H$$

such that $\pi_H = f_0 \circ \psi_0$. Moreover (again by standard covering space theory), f_0, W_0, ψ_0 are unique in the sense that if f_1, W_1, ψ_1 have similar properties, then $\psi_1^{-1} \circ \psi_0$ is a Möbius transformation. So then write $\varphi = \psi_0 \circ \psi_2$. Then $f = f_0 \circ \varphi$, and f and φ are unique up to right composition with a Möbius transformation, subject to this identity, f_0 being rational and φ a homeomorphism. Conversely, all maps of this form have critical value set Y and determine the same conjugacy class of subgroup H. If

$d = 2$, so that Y has exactly two points, there is obviously only one choice for H in any case.

If f and g are close branched coverings whose critical value sets have the same multiplicities, then $g = \psi \circ f \circ \varphi^{-1}$, for some homeomorphisms φ and ψ which are close to the identity. Recall that f is *critically finite* if $\#(X(f) < \infty$, where the *postcritical set* $X(f)$ is defined by:

$$X(f) = \{f^n(c) : c \text{ critical}, n > 0\}.$$

Now, if f and g are close critically finite branched coverings, and $\#(X(f)) = \#(X(g))$, and φ, ψ are as above, then φ, ψ must map $X(f)$ to $X(g)$. Conversely, any map $\psi \circ f \circ \varphi^{-1}$ with $\varphi(X(g)) = \psi(X(f))$ is critically finite. Thus, by 1.1, the set of critically finite branched coverings with postcritical set of cardinality k is locally contractible.

1.7. Thurston Equivalence. — Thurston equivalence for critically finite branched coverings f, g was introduced in [**T1**] (probably the first time that the study of purely topological critically finite branched coverings was introduced into the study of complex dynamics). In [**R3**], two equivalent conditions for Thurston equivalence were given. In particular, recall that $f \simeq g$ if and only if $f \simeq_\varphi g$ for some homeomorphism $\varphi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ with $\varphi(X(f)) = X(g)$, and $f \simeq_\varphi g$ if and only if there is a path g_t of branched coverings with $X(g_t) = X(g)$ for all t, and $\varphi \circ f \circ \varphi^{-1} = g_0$, $g = g_1$. Although $\simeq_\varphi$ is not an equivalence relation, $f \simeq_\varphi g$ if and only if $g \simeq_{\varphi^{-1}} f$, and $f \simeq_{\psi\circ\varphi} h$ whenever $f \simeq_\varphi g$ and $g \simeq_\psi h$. By 1.3, we can give yet another definition of the Thurston equivalence class of f. It is the connected component of critically finite branched coverings g with $\#(X(g)) = \#(X(f))$ which contains f. This set is path-connected. There is also an alternative definition of the set $\{g : g \simeq_{\text{identity}} f\}$, which can be described as the connected component of $\{g : X(g) = X(f)\}$ which contains f. Again, by 1.3 it is clear that this set is path connected. By 1.1, if $X(f)$ has at least 3 points, the set is also simply-connected. Even if $X(f)$ has only one or two points, the quotient of $\{g : g \simeq_{\text{identity}} f\}$ by conjugation by the appropriate subgroup of Möbius transformations is simply connected, again by 1.1.

In 1.3 of [**R3**], we refined the concept of equivalence and defined $(f, Y_0) \simeq (g, Y_1)$ for sets Y_i that contain the postcritical sets and are forward invariant under f, g. The definition of $(f, Y_0) \simeq_\varphi (g, Y_1)$ differs from $f \simeq_\varphi g$ only in that $g_t = g$ on Y_1, for all t. If $f \simeq_\varphi g$ then $(f, f^{-n}(X(f)) \simeq_{\varphi_n} (g, g^{-n}(X(g))$, where φ_n is defined inductively by $\varphi_0 = \varphi$, $\varphi_n \circ f = g \circ \varphi_{n+1}$ and $\varphi_n = \varphi$ on $X(f)$.

1.8. Thurston's Theorem for Critically Finite Branched Coverings

Thurston's theorem for critically finite branched coverings of $\overline{\mathbf{C}}$ can then be stated as follows. We do not define the term *associated orbifold*, but we recall that the associated orbifold of f is always hyperbolic if either $\#(X(f)) > 4$ or the forward orbit of every critical point contains a periodic critical point.The term *Thurston obstruction* will be defined — actually in a more general context, but with attention drawn to the critically finite case — in 2.4. For now, we recall that a Thurston obstruction for f is a positive linear combination of disjoint simple loops in $\overline{\mathbf{C}} \setminus X(f)$ which is an eigenvector with eigenvalue $\geqslant 1$ for some linear map defined in terms of f. For more detail, see [**D-H3**], [**T2**] (or 2.4).

Theorem. — *Let $f : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ be a critically finite branched covering such that the associated orbifold is hyperbolic. Then the connected component of*

$$\{g : \#(X(g)) = \#(X(f))\}$$

containing f contains a rational map if and only if f has no Thurston obstruction. In that case, the rational map is unique up to Möbius conjugation, and the quotient of the component by Möbius conjugation is simply connected.

Thurston's theorem is stated in this way to make clear that the Topographer's View (see Chapter 5) is regarded as a special case of a generalisation of it. This is not quite the usual statement but is certainly what is implied by the usual proof. For notes towards the proof see [**T2**], and see [**D-H3**] for a complete proof. We shall also be discussing the proof of Thurston's theorem in Chapter 6. In fact the "if" direction is essentially proved in 6.6-15, when a number of properties of the "pullback map" are proved in more generality. (The "if" direction of the result has more implications and is therefore more powerful, but the "only if" direction is surprisingly hard to prove in full generality.) The proof of Thurston's Theorem reveals a lot about the structure of critically finite branched coverings and their Thurston equivalence classes even when the associated orbifolds are not hyperbolic, or when Thurston obstructions exist.

1.9. Definition of $B(Y, f_0)$. — It is time to specify the parameter spaces with which this paper is concerned. It seems a good idea to make definitions more general than we need at present. So let $f_0 : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ be an orientation-preserving branched covering. Let $Z \subset \overline{\mathbf{C}}$ be finite, with $f_0(Z) \subset Z$. Let $Y = Y(f_0)$ be the union of Z and the critical values of f_0. It may be that some of the critical values are already in Z. In the cases studied in detail in this work, there will always be at least one critical value in Z. We consider pairs $(f, Y(f))$ such that:

(1) $f : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ is an orientation-preserving branched covering of the same degree as f_0;

(2) $Y(f)$ is the union of a finite set $Z(f)$ with $f(Z(f)) \subset Z(f)$ and the critical values of f;

(3) there exists a (not necessarily unique) bijection $\tau : Y \to Y(f)$ which maps Z to $Z(f)$, and maps any critical value v of f_0 to a critical value $v(f)$ of f. Moreover, $\tau \circ f_0 = f \circ \tau$ on Z, and $v(f_0)$ and $v(f)$ have the same numbers of preimages, and of the same multiplicities.

Then we define an equivalence relation $\sim$ by

$$(f, Y(f)) \sim (g, Y(g))$$

if and only if there is a Möbius $\tau : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ such that:

$$g = \tau \circ f \circ \tau^{-1}, \quad \tau(Z(f)) = Z(g).$$

Let $[f, Y(f)]$ denote the equivalence class of $(f, Y(f))$. Then we let $B(Y, f_0)$ denote the connected component of classes of $[f, Y(f)]$ containing $[f_0, Y(f_0)]$. We write $f \in B$ for short, whenever possible. If $f \in B(Y, f_0)$ then $B(Y, f_0) = B(Y(f), f)$. Of course, this notation does not make very clear that the structure of Y includes the named subset Z with dynamics and critical values of fixed topological type, but it is at least reasonably simple. We write $V(Y, f_1)$ for the connected component of $[f, Y(f)]$ for which f is rational and containing $[f_1, Y(f_1)]$.

As in 1.6, $g \in B(Y, f)$ if and only if g can be written in the form

$$\psi \circ f \circ \varphi^{-1}$$

where φ, ψ are orientation-preserving homeomorphisms with $\psi(Y(f)) = Y(g)$, $\varphi(f^{-1}(Y(f))) = g^{-1}Y(g)$. Thus, as in 1.7, B is locally path-connected.

In present applications, $B(Y, f_0)$ will usually be a family of degree two maps, that is, f_0 will usually be a degree two branched covering. In that case, Y will contain two critical values v_1, v_2, with $v_1 \in Z$ and $v_2 \notin Z$. We then say that $B(Y, f_0)$ is of *degree two type*. Actually, our setting is more specific. The critical value in Z will be either periodic or eventually fixed. There may be one other periodic orbit in Z, but if so, it will be fixed. In these cases, we shall say that $B(Y, f_0)$ is of *periodic degree two type* or *eventually fixed degree two type*. We shall also have to consider the case when $B(Y, f_0)$ is a family of degree d maps (for some d) and Z contains one fixed critical point of local degree d, and any other critical value in Z is periodic. In this case, we shall say that $B(Y, f_0)$ is *of polynomial type*. Our real interest is in the degree two periodic type, but studying this case leads to the study of the other cases.

We shall continue to use the notation of the introduction for some periodic degree two type spaces. We write

$$B_{p,n},\ B_{p,1,n},\ B_{ef,m,n}$$

for the spaces of degree two branched coverings $B(Y, f_0)$ where f_0 has critical points v_1, v_2 with $v_1 \in Z$, $v_2 \notin Z$. Furthermore the set Z and its dynamics are described respectively by:

$$Z = f_0^{-n}(\{f_0^i(v_1) : 0 \leqslant i < p\}),\ f_0^p(v_1) = v_1,\ f_0^i(v_1) \neq v_1,\ 0 < i < p,$$

$$Z = f_0^{-n}(\{f_0^i(v_1) : 0 \leqslant i < p\} \cup \{x\}),\ f_0(x) = x,\ v_1 \text{ as above},$$

$$Z = f_0^{-n}(\{f_0^i(v_1) : 0 \leqslant i \leqslant m\}),\ f_0^{m+1}(v_1) = f_0^m(v_1),\ f_0^{i+1}(v_1) \neq f_0^i(v_1),\ 0 \leqslant i \leqslant m.$$

For the spaces of rational maps within these spaces, we write

$$V_{p,n},\ V_{p,1,n},\ V_{ef,m,n}.$$

It would be possible to vary the definition of $B(Y, f_0)$, and add extra structure by numbering some of the points in Z or Y. We could then require the bijection τ in the definition of equivalence relation to preserve numbering. The examples we are chiefly interested in do not necessitate such numbering. However, for degree two type, we shall habitually refer to the critical values as v_1 and v_2 with $v_1 \in Z$ and $v_2 \notin Z$.

1.10. Definition of $(B(Y, f_0), N)$. — It will also be necessary for us to consider pairs of spaces $(B(Y, f_0), N)$, where $B = B(Y, f_0)$ is as in 1.9 and $N \subset B$ is a finite union of neighbourhoods of ends as follows. Fix ε_0 less than the Margulis constant for surfaces. Each component N_1 of N will be an open neighbourhood of a single end, and if N_1 and N_2 are two such, then $\overline{N}_1 \cap \overline{N}_2 = \varnothing$. For $f \in \overline{N}_1$ there is a set $\Gamma(f, N_1) \subset \overline{\mathbf{C}} \setminus Y(f)$ of geodesics of length $\leqslant \varepsilon_0$ with respect to the Poincaré metric on $\overline{\mathbf{C}} \setminus Y(f)$, which varies continuously for $f \in \overline{N}_1$, such that $\Gamma(f, N_1) \subset f^{-1}(\Gamma(f, N_1))$ modulo $Z(f)$-preserving isotopy, $f^{-1}(\Gamma(f, N_1)) \subset \Gamma(f, N_1)$ modulo $Z(f)$-preserving isotopy and trivial and peripheral loops, such that the lengths of all geodesics in $\Gamma(f, N_1)$ have length $< \varepsilon_0$ for $f \in N_1$, and at least one of these geodesics has length exactly ε_0 for $f \in \partial N_1$. Then N_1 and ∂N_1 are invariant under Möbius conjugation — as they need to be because, strictly speaking, $B(Y, f_0)$ is a set of Möbius conjugacy classes. Furthermore the set $\Gamma(f, N_1)$ is to have additional properties. Each loop $\gamma \in \Gamma(f, N_1)$ bounds a disc $D(\gamma) \subset \overline{\mathbf{C}}$ such that $D(\gamma) \cap D(\gamma') = \varnothing$ if $\gamma \neq \gamma'$, γ, $\gamma' \in \Gamma(f, N_1)$. If, for at least one $f \in N_1$, each component of $f^{-1}(D(\gamma))$ is a disc for all $\gamma \in \Gamma(f, N_1)$, then the same is true for all $g \in N_1$. Suppose that this is true for N_1. We can then define a branched covering g_1 and sets $Z(g_1)$, $Y(g_1)$, associated with N_1. Fix any $f \in N_1$. Choose a set $Z(g_1)$ to consist of one point in each disc $D(\gamma)$ ($\gamma \in \Gamma(f, N_1)$) intersected by $Z(f)$, together with all the points of $Z(f)$ which are not in any such disc $D(\gamma)$. Extend this to a set $Y(g_1)$ by adding in a point from each disc $D(\gamma)$ which contains no point of $Z(f)$, and any remaining points of $Y(f) \setminus Z(f)$ which are not in any disc $D(\gamma)$. The critical values of g_1 will then be the points of $Y(g_1)$ which are either contained in discs $D(\gamma)$ containing at least one critical value of f, or are critical values of f. We now choose the critical points of g_1: one in each component D_1 of some $f^{-1}(D(\gamma))$ which contains a critical point of f, in which case the local degree at this critical point is the degree of $f : D_1 \to D(\gamma)$, and the remaining ones are the critical points of f which are not in any $f^{-1}(D(\gamma))$. So the critical points, critical values and local degrees at critical points have now been specified. The map g_1 then maps points of $Y(g_1)$ so that a point in $D(\gamma)$ maps to a point in $f(D(\gamma))$, and $g_1 = f$ near any point of $Y(g_1)$ which is not in any disc $D(\gamma)$. Then $B(Y(g_1), g_1)$ depends only on N_1, not on the choice of $f \in N_1$. If g_1 exists, then N_1 is the intersection with $B(Y, f_0)$ of a neighbourhood of $B(Y(g_1), g_1)$ in a larger space of branched coverings.

We are mostly interested in spaces $B(Y, f_0)$ when f_0 is degree two, one critical point v_1 is in Z and $Y \setminus Z = \{v_2\}$. In that case, a natural example is when $\Gamma(f, N_1)$ consists of a simple loop bounding a disc containing v_2 and exactly one point of $Z \setminus \{v_1\}$, and a simple loop bounding a disc round each of the other points of Z, for $f \in N_1$.

We shall always consider the situation when all but at most one component of N is like N_1 above, that is, g_1 exists. The exception will only occur as follows. The space $B(Y, f_0)$ will be of polynomial type with v_1 a fixed critical value of maximal multiplicity. For the exceptional neighbourhood N_0 and $f \in N_0$, $\Gamma(f, N_0)$ will consist of a single loop γ bounding a disc $D(\gamma)$ containing v_1 and at least one other critical value of f. Since v_1 has maximal multiplicity, $f^{-1}(D(\gamma))$ will not be a disc.

1.11. Modular Subgroup Interpretations of $\pi_1(B)$. — Let $B = B(Y, f_0)$. Assume that $Z = Z(f_0)$ contains at least 3 points. Then each (not necessarily closed) path $t \mapsto g_t$ in B ($t \in [0,1]$) defines paths through sets $Y(g_t)$, $g_t^{-1}Y(g_t)$ starting from Y, $g_0^{-1}Y$, with $Y = Y(g_0)$. Thus, by 1.2, we have paths through homeomorphisms starting with the identity and ending with homeomorphisms fixing Y, $g_0^{-1}Y$. The end homeomorphisms are uniquely determined up to isotopies constant on Y, $g_0^{-1}(Y)$ by the homotopy class of the path in B, where we take homotopies fixing endpoints. Note that the paths through homeomorphisms can be taken of the form

$$\varphi_t,\ g_t^{-1} \circ \varphi_t \circ g_0,$$

and φ_t and $g_t^{-1} \circ \varphi_t \circ g_0$ are isotopic via an isotopy constant on Z. We write $\Phi_1(\{g_t\})$, $\Phi_2(\{g_t\})$ for the isotopy classes of the end homeomorphisms φ_1, $g_1^{-1} \circ \varphi_1 \circ g_0$.

We now restrict to closed paths in B based at some $g_0 \in B$, that is, elements of $\pi_1(B, g_0)$. Again by 1.2, Φ_1 and Φ_2 are antihomomorphisms of $\pi_1(B, g_0)$ into $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ and $\mathrm{MG}(\overline{\mathbf{C}}, g_0^{-1}Y)$. We prefer to keep these as anti-homomorphisms rather than take inverses to get homomorphisms. Write $\varphi_1 = \varphi$. We define

$$G_1 = \{[\varphi] \in \mathrm{MG}(\overline{\mathbf{C}}, Y) : g_0^{-1} \circ \varphi \circ g_0 \text{ is well defined and isotopic to } \varphi \bmod Z(g_0)\}.$$

Then $\Phi_1(\pi_1(B, g_0)) \subset G_1$.

The following lemma will imply that if two paths in B with the same endpoints give homeomorphisms which are isotopic via isotopies constant on Y (or $g_0^{-1}Y$), then the two paths are homotopic in B.

Lemma. — *Φ_1 is an anti-isomorphism onto G_1, and Φ_2 is injective.*

Proof. — First we prove that Φ_1 is injective. Injectivity of Φ_2 then follows immediately. So suppose that $\Phi_1(\{g_t\}) = $ [identity]. Then we have a path φ_t through homeomorphisms with $\varphi_t(Y(g_0)) = Y(g_t)$ and φ_1 is isotopic to the identity modulo $Y(g_0)$. We can assume without loss of generality that φ_1 is the identity. We can also assume that $\{0, 1, \infty\} \in Y(g_t)$ for all t and that these are fixed by all φ_t. Since $\mathrm{Hom}_+(\overline{\mathbf{C}}, \{0, 1, \infty\})$ is simply-connected (1.2), there is a continuous family $\varphi_{s,t}$ ($(s,t) \in [0,1] \times [0,1]$) with $\varphi_{0,t} = \varphi_t$, $\varphi_{1,t} = \varphi_{0,s} = \varphi_{1,s} =$ identity for all s, t. Then define

$$g_{s,t} = \varphi_{s,t}^{-1} \circ g_t \circ \varphi_{s,t}.$$

Then $g_{1,t} = g_t$ but $Y(g_{0,t}) = Y(g_0)$ for all t. So there is a continuous family ψ_t fixing $Z(g_0)$ with

$$g_{0,t} = g_0 \circ \psi_t.$$

Then, since $\mathrm{Hom}_+(\overline{\mathbf{C}}, Z(g_0))$ is simply-connected, $\{\psi_t\}$ is a trivial path in $\mathrm{Hom}_+(\overline{\mathbf{C}}, Z(g_0))$, and hence both $\{g_{0,t}\}$ and $\{g_{1,t}\} = \{g_t\}$ are trivial in B.

Now we show that Φ_1 is surjective. Let $[\varphi] \in G_1$. Let φ_t $(t \in [0,1])$ be a path from the identity to φ. Then since φ and $g_0^{-1} \circ \varphi \circ g_0$ are isotopic modulo $Z(g_0)$, we can choose a path ψ_t from the identity to $g_0^{-1} \circ \varphi \circ g_0$ with $\psi_t(Z(g_0)) = \varphi_t(Z(g_0))$ for all t. Then

$$g_t = \varphi_t \circ g_0 \circ \psi_t^{-1}$$

is the required closed loop with $\Phi_1(\{g_t\}) = [\varphi]$. □

1.12. A Set-theoretic injection of $\pi_1(B)$ into $\pi_1(\overline{\mathbf{C}} \setminus Z)$. — Let $Y \setminus Z = \{v\}$ be a critical value of g_0, with just one critical preimage c, as happens in all the cases (mainly degree two) which most interest us. Then the isotopy between homeomorphisms φ_t and $g_t^{-1} \circ \varphi_t \circ g_0 = \psi_t$ given by the path $\{g_t\}$ in 1.11 can be sharpened: there is a path α_t starting from $v(g_0)$ such that

$$\varphi_t = \psi_t \circ \sigma_{\alpha_t} \text{ in } \mathrm{MG}(\overline{\mathbf{C}}, Y).$$

We are particularly interested in this in two special cases.

Case 1. — Let $g_0(v) = v$ and let $\{g_t : t \in [0,1]\}$ be a closed path. Then $\alpha_1 = \alpha$ is a closed path. So G_1 can be described as

$$(1) \quad G_1 = \{\varphi \in \mathrm{MG}(\overline{\mathbf{C}}, Y) : g_0^{-1} \circ \varphi \circ g_0 \circ \sigma_\alpha = \varphi \text{ in } \mathrm{MG}(\overline{\mathbf{C}}, Y), \quad \text{for some } \alpha \in \pi_1(\overline{\mathbf{C}} \setminus Z, v)\}.$$

Now using the above write $\psi = g_0^{-1} \circ \varphi \circ g_0$. Then the image of Φ_2 is G_2, where

$$(2) \quad G_2 = \{\psi \in \mathrm{MG}(\overline{\mathbf{C}}, g_0^{-1}Y) : g_0^{-1} \circ \psi \circ \sigma_\alpha \circ g_0 = \psi \text{ in } \mathrm{MG}(\overline{\mathbf{C}}, g_0^{-1}Y), \quad \text{for some } \alpha \in \pi_1(\overline{\mathbf{C}} \setminus Z, v)\}.$$

Then we define $\rho : \pi_1(B, g_0) \to \pi_1(\overline{\mathbf{C}} \setminus Z, v(g_0))$ by

$$\rho(\{g_t\}) = \alpha.$$

We can also regard ρ as a map on G_1, since G_1 and $\pi_1(B, g_0)$ are anti-isomorphic

Case 2. — Let $g_1 \in N$, with $v(g_1)$ near a point of $Z(g_1)$ which is not fixed. We can choose the homeomorphisms φ_t $(t \in [0,1])$ so that φ_t^{-1} is bounded except near $v(g_t)$, when $v(g_t)$ is near a point of $Z(g_t)$. This means that ψ_t^{-1} is bounded near $v(g_t)$, and the second endpoint of α_t is near the corresponding point of $Z(g_0)$. Then we define $\rho_2 : \pi_1(B, N, g_0) \to \pi_1(\overline{\mathbf{C}} \setminus Z, Z, v(g_0))$ by

$$\rho_2(\{g_t\}) = \alpha.$$

Lemma. — *The maps ρ and ρ_2 are injective.*

Proof. — Although ρ is not a group homomorphism, if $\Phi_1(\beta_1) = [\varphi_1]$, $\Phi_2(\beta_1) = [\psi_1]$, and $\overline{\beta}$ denotes the reverse of the path β, we have

$$\rho(\beta_1 * \beta_2) = \rho(\beta_1) * \psi_1^{-1}(\rho(\beta_2)) = \rho(\beta_1) * (\sigma_{\rho(\beta_1)} \circ \varphi_1^{-1})(\rho(\beta_2)) = \varphi_1^{-1}(\rho(\beta_2)) * \rho(\beta_1),$$
$$\rho(\overline{\beta}_1) = \psi_1(\overline{\rho(\beta_1)}) = \varphi_1(\overline{\rho(\beta_1)}).$$

Now let

$$\rho(\beta_1) = \rho(\beta_2).$$

Then

$$\rho(\overline{\beta}_1 * \beta_2) = \varphi_1(\rho(\beta_2) * \overline{\rho(\beta_1)})$$

is trivial. So write

$$[\varphi] = \Phi_1(\overline{\beta}_1 * \beta_2)$$

Then

$$[\varphi] \in \{[\chi] \in \mathrm{MG}(\overline{\mathbf{C}}, Y)\colon\ [\chi] = [g_0^{-1} \circ \chi \circ g_0]\} = H.$$

But g_0 is a polynomial and by Thurston's Theorem (1.8) the Thurston equivalence class B_0 of g_0 is simply connected. But by 1.11, $\pi_1(B_0, g_0)$ is anti-isomorphic to H, which is therefore a trivial group. So $[\varphi]$ is the identity and $\beta_1 = \beta_2$

The proof for ρ_2 is similar. Suppose that

$$\rho_2(\beta_1) = \rho_2(\beta_2)$$

for two paths β_1, β_2 in $\pi_1(B, N, g_0)$ with second endpoints in ends N_1, N_2 of B near critically finite maps h_1, h_2. Then $\rho_2(\overline{\beta}_1 * \beta_2)$ is trivial, and the isotopy class $\Phi_1(\overline{\beta}_1 * \beta_2)$ is represented by a homeomorphism φ mapping $Y(h_1)$ to $Y(h_2)$ so that $h_2^{-1} \circ \varphi \circ h_1$ and φ are isotopic via an isotopy constant on $Y(h_1)$, that is, $h_1 \simeq_\varphi h_2$ in the notation of 1.7. Then $N_1 = N_2$, so we can take $h_1 = h_2$. Then by Thurston's Theorem as in the case of ρ, φ is isotopic to the identity, and so $\overline{\beta}_2 * \beta_1$ is a trivial path, as required. □

1.13. An action of $\pi_1(B, g_0)$ on the unit disc boundary. — Once again, let $Y \setminus Z = \{v\}$ and $g_0(v) = v$. Let $[\varphi] \in G_1$ and $\alpha = \rho([\varphi])$ and $\psi = \varphi \circ \sigma_\alpha^{-1}$. Then, in the notation of 1.7,

$$\sigma_\alpha \circ g_0 \simeq_\psi g_0.$$

Similarly, for any $g_0 \in B$, if $\alpha = \rho_2(\beta)$ for some $\beta \in \pi_1(B, N, g_0)$ and h_0 is the critically finite map (up to equivalence) determined by N, then

$$\sigma_\alpha \circ g_0 \simeq h_0$$

if $v(h_0)$ is preperiodic, and if $v(h_0)$ is periodic then

$$\sigma_\zeta^{-1} \circ \sigma_\alpha \circ g_0 \simeq h_0,$$

where ζ is a path in $\overline{\mathbf{C}} \setminus Z(g_0)$ with $g_0 \circ \zeta = \alpha$ up to homotopy preserving endpoints and $Z(g_0)$. To see this, we first note that the result is obviously true if we take g_0

in B close to h_0, because then α (and ζ) can be taken as short paths. Then we note that the equivalence remains true if we vary g_0 continuously.

Let D be the unit disc and let $\pi_2 : D \to \overline{\mathbf{C}} \setminus Z$ be the universal covering map. We assume that $\pi_2(0) = v$. We are now going to show how $\pi_1(B, g_0)$ acts naturally on ∂D. Let $\beta \in \pi_1(B, g_0)$. Let $\varphi = \Phi_1(\beta)$ and $\psi = \Phi_2(\beta)$. Let $\rho_2(\beta) = \alpha$ and let $\widetilde{\alpha} : [0,1] \to D$ be the lift of α with $\widetilde{\alpha}(0) = 0$. Let $\widetilde{\psi}^{-1}$ be the lift of ψ^{-1} such that $\widetilde{\psi}^{-1}(0) = \widetilde{\alpha}(1)$. Then $\widetilde{\psi}^{-1} \mid \partial D$ depends only on the class of ψ^{-1} in $\mathrm{MG}(\overline{\mathbf{C}}, Z)$ (which is the same as the class of φ^{-1}). Then for $z \in \partial D$, we define

$$\beta \cdot z = \widetilde{\psi}^{-1}(z).$$

Note that $\pi_1(\overline{\mathbf{C}} \setminus Z, Z, g_0)$ can be regarded as a subset of ∂D, because if we have a path $\alpha \in \pi_1(\overline{\mathbf{C}} \setminus Z, Z, g_0)$ then we can take the lift $\widetilde{\alpha} : [0,1] \to \overline{D}$ with $\widetilde{\alpha}(0) = 0$, and $\widetilde{\alpha}(1) \in \partial D$. Then if $\beta_1 \in \pi_1(B, g_0)$ and $\beta_2 \in \pi_1(B, N, g_0)$, and we regard $\mathrm{Im}(\rho_2)$ as a subset of ∂D,

$$\beta_1 \cdot \rho_2(\beta_2) = \rho_2(\beta_1 * \beta_2) = \widetilde{\alpha}_1 * \widetilde{\psi}_1^{-1}(\widetilde{\alpha}_2),$$

where $\Phi_2(\beta_1) = \psi_1$, $\alpha_1 = \rho(\beta_1)$, $\alpha_2 = \rho_2(\beta_2)$ and $\widetilde{\alpha}_i$ is the lift of α_i with $\widetilde{\alpha}_i(0) = 0$. This follows from the multiplicative properties of ρ_2, as in the lemma in 1.12. Thus

$$\rho_2 : (\pi_1(B, g_0), \pi_1(B, N, g_0)) \longrightarrow (\pi_1(B, g_0), \partial D)$$

is a homomorphism of left-$\pi_1(B, g_0)$-actions.

1.14. When Z has 1 or 2 points and B is degree two. — Now we restrict to the case when g_0 is a degree two branched covering with critical values v_1, v_2, and critical points c_1, c_2, with $v_1 \in Z$ and $v_2 \notin Z$. Write

$$Y = Z \cup \{v_2\}.$$

If Z has one or two points, then every branched covering represented in B is of the form $g_0 \circ \varphi$ up to equivalence, where φ fixes Z. This representation is unique if Z has two points, but if Z has one point — so that $g_0(z) = z^2$ without loss of generality, and $Z = \{\infty\}$ — then φ can be replaced by $z \mapsto \lambda^{-1}\varphi(\lambda^2 z)$ for any $\lambda \in \mathbf{C}^*$. It follows that $G = \pi_1(B)$ is trivial if Z has one point, and infinite cyclic if Z has two points.

1.15. Return to Example 1. — We return to Example 1 of the Introduction. Thus, any element of $B_{3,0}$ is represented by a branched covering f with critical values $v_1(f) = \infty$ with period 3 orbit $\infty \mapsto 1 \mapsto 0 \mapsto \infty$, and $v_2(f) \notin \{\infty, 1, 0\}$, and any element of $V_{3,0}$ is represented by a rational map with precisely the same properties. Recall that $V_{3,0}$ identifies with $\mathbf{C} \setminus \{0, \pm 1\}$, where $a \in \mathbf{C} \setminus \{0, \pm 1\}$ identifies with the map

$$g_a : z \longmapsto \frac{(z-a)(z-1)}{z^2}.$$

The second critical point and second critical value are

$$c_2 = c_2(a) = \frac{2a}{a+1}, \quad v_2 = v_2(a) = g_a(c_2) = \frac{-(a-1)^2}{4a}.$$

Now we claim that the inclusion of $V_{3,0}$ in $B_{3,0}$ is a homotopy equivalence. In particular, $\pi_1(B_{3,0}) = \pi_1(V_{3,0})$. It suffices to construct a homotopy inverse to this inclusion. Fix $g \in B_{3,0}$ with $c_1(g) = 0$, $v_1(g) = \infty$ and $g(v_1(g)) = 1$. If $v_2 \neq 0, 1, \infty$ (which is true throughout $B_{3,0}$), then there are exactly two values of a, $a = a_j$, $j = 1, 2$ such that g_{a_j} has the same critical values as g. Then, as explained in 1.6, by standard covering space theory, there is $\psi_j \in \mathrm{Hom}_+(\overline{\mathbf{C}})$ fixing the critical point 0 such that

$$g = g_{a_j} \circ \psi_j, \quad j = 1, 2.$$

In fact, since $\infty \in g_{a_j}^{-1}(1)$, the covering space theory says that given either point $x \in g^{-1}(1)$, we can choose ψ_j so that $\psi_j(\infty) = x$. So since we can take $x = \infty$, we can choose ψ_j so that $\psi_j(\infty) = \infty$. Write $g^{-1}(0) = \{1, x'\}$. Then ψ_j must map $\{1, x'\}$ to $\{1, a_j\} = g_{a_j}^{-1}(0)$. We have

$$g_{a_2} = g_{a_1} \circ \psi_1 \circ \psi_2^{-1}.$$

It follows that $\psi_1 \circ \psi_2^{-1}$ is a Möbius transformation, but not the identity, because $g_{a_1} \neq g_{a_2}$. So we cannot have $\psi_1 \circ \psi_2^{-1}(1) = 1$, because if so then $\psi_1 \circ \psi_2^{-1}$ fixes the three points 0, ∞ and 1, and must be the identity. If $\psi_j(1) = a_j$ for both $j = 1$ and 2 then $\psi_j(x') = 1$ for both $j = 1$ and 2 and $\psi_1 \circ \psi_2^{-1}$ does fix 1, a contradiction. So exactly one of ψ_1, ψ_2, say ψ_1, must fix 1 also. It follows that

$$B_{3,0} = \{f \circ \psi_1 : f \in V_{3,0},\ \psi_1 \in \mathrm{Hom}_+(\overline{\mathbf{C}}),\ \psi_1 \text{ fixes } 0,\ 1,\ \infty\},$$

and both f and ψ_1 are uniquely determined by $f \circ \psi_1$.

Then the map

$$f \circ \psi_1 \longmapsto f : B_{3,0} \longrightarrow V_{3,0}$$

is well-defined and is the required homotopy inverse to inclusion.

Now we shall give an explicit description of the map

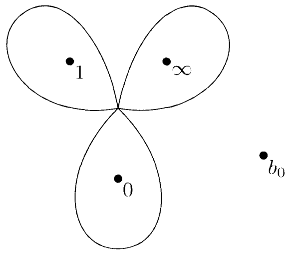
$$\rho : \pi_1(B_{3,0}) = \pi_1(V_{3,0}) \longrightarrow \pi_1(\overline{\mathbf{C}} \setminus \{0, 1, \infty\}).$$

Three values of a, one real and negative, and a pair of complex conjugates, give polynomials, up to Möbius conjugation. For the sake of concreteness, we denote by a_0 the value in the upper half-plane, and by a_1 the real and negative one. Thus,

$$\frac{-(a_0 - 1)^2}{4a_0} = \frac{2a_0}{a_0 + 1} = b_0, \text{ say.}$$

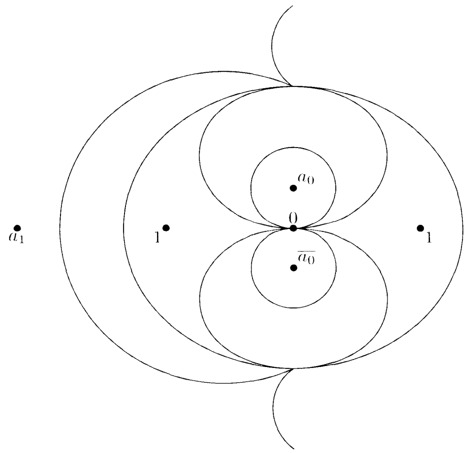
The Julia set of g_{a_0} includes the rabbit-like set sketched very roughly below.

Indeed, this Julia set is sometimes known as *the rabbit*. The important feature of it is that g_{a_0} fixes the intersection point of the three ears and maps homeomorphically arcs connecting the points 0, ∞, 1 to the intersection point, permuting these arcs (rotating in an anticlockwise direction). The Julia set of $g_{\overline{a}_0}$ is the complex conjugate of that for g_{a_0}, and is sometimes known as *the antirabbit*. The polynomial which is Möbius-conjugate to g_{a_0} is in the upper half of the quadratic Mandelbrot set.



Rabbit.

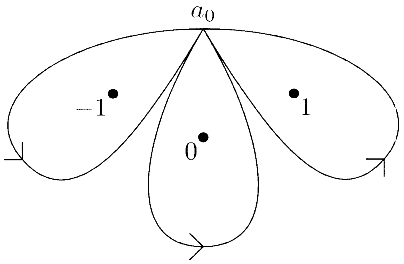
We give below a very rough sketch of the parameter space $\mathbf{C} \smallsetminus \{0, \pm 1\}$ with the hyperbolic components containing a_0, $\overline{a}_0$, ± 1 and a_1. In particular, note that the first four of these have boundaries meeting at zero, while the last three have boundaries meeting at two points, which are complex conjugates. A similar sketch is given in [**R3**], and much better pictures in [**W**].



$V_{3,0}$.

We consider a generating set for $\pi_1(V, a_0)$ consisting of simple loops round 0, ± 1 as drawn below. We shall call these loops γ_0, γ_1, γ_{-1}.

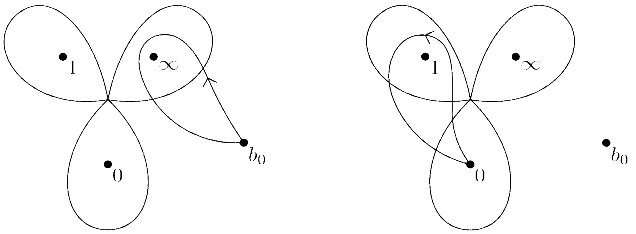
Even though $\rho : \pi_1(V_{3,0}, a_0) \to \pi_1(\overline{\mathbf{C}} \smallsetminus \{0, \infty, 1\}, b_0)$ is not a homomorphism, it is sufficient to describe the map on generators, if we also describe the antihomomorphisms Φ_1 and Φ_2' on generators in $\pi_1(V_{3,0})$. Here, Φ_2' is Φ_2 followed by the natural projection to $\mathrm{MG}(\overline{\mathbf{C}}, Y)$, where $Y = \{0, \infty, 1, b_0\}$. Note that Φ_2' is not injective, although both Φ_1 and Φ_2 are. The images G_1 of Φ_1 and of Φ_2' are both in $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$. By 1.3, $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$ is naturally anti-isomorphic both to $\pi_1(\overline{\mathbf{C}} \smallsetminus \{0, \infty, 1\}, b_0)$ and to $\pi_1(\overline{\mathbf{C}} \smallsetminus \{0, \infty, b_0\}, 1)$ (for example). Here, we are using $\#(Y) = 4$. Thus, we can regard Φ_1 and Φ_2' as homomorphisms from $\pi_1(V_{3,0}, a_0)$ to $\pi_1(\overline{\mathbf{C}} \smallsetminus \{0, \infty, 1\}, b_0)$, or



Generators for $\pi_1(V_{3,0})$.

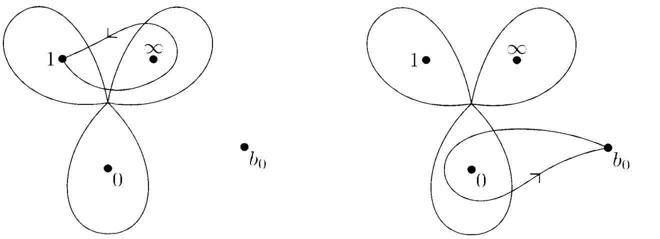
equivalently to $\pi_1(\overline{\mathbf{C}} \setminus Y', y_0)$ for some other decomposition $Y = Y' \cup \{y_0\}$. Write $g = g_{a_0}$. Then clearly the image of Φ_1 in $\pi_1(\overline{\mathbf{C}} \setminus \{0, \infty, b_0\}, 1)$ (for example) is the subgroup of loops whose lifts under g are closed. This gives reassurance that Φ_1 is an isomorphism onto its image, because such a subgroup is the fundamental group of a double cover of $\overline{\mathbf{C}} \setminus \{0, \infty, b_0)$ branched over ∞, b_0, which is a four-times punctured sphere (as is $\mathbf{C} \setminus \{0, \pm 1\}$).

The description of ρ in terms of Φ_1 and Φ_2' is now particularly simple. Let $\gamma_j \in \pi_1(V_{3,0})$ $(j = 0, \pm 1)$ with $\Phi_1(\gamma_j) = \beta_j$ and $\Phi_2'(\gamma_j) = \zeta_j$ for $\beta_j, \zeta_j \in \pi_1(\overline{\mathbf{C}} \setminus \{0, \infty, 1\}, b_0)$. Then $\rho(\gamma_j) = \beta_j * \overline{\zeta}_j$, where (as usual) $\overline{\zeta}$ denotes the reverse of ζ. It is easy to compute β_j by considering the image of γ_j under the map $a \mapsto v_2(a)$. Then one can compute the loop in $\pi_1(\overline{\mathbf{C}} \setminus \{0, \infty, b_0\}, 1)$ or $\pi_1(\overline{\mathbf{C}} \setminus \{1, \infty, b_0\}, 0)$ giving the same element of $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$, take its preimage under g in $\pi_1(\overline{\mathbf{C}} \setminus \{1, 0, b_0\}, \infty)$ or $\pi_1(\overline{\mathbf{C}} \setminus \{\infty, 0, b_0\}, 1)$, and compute the loop in $\pi_1(\overline{\mathbf{C}} \setminus \{0, \infty, 1\}, b_0)$ giving the same element in $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$. This loop is ζ_j. We carry out this process for each of the loops γ_0, $\gamma_{\pm 1}$ below.

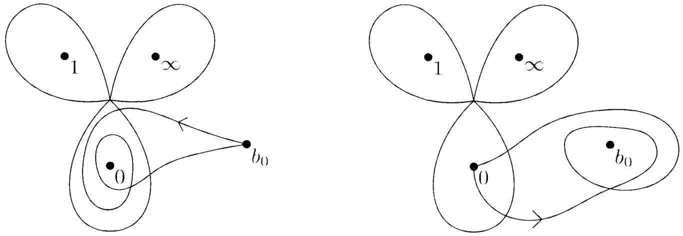


β_0 and equivalent.

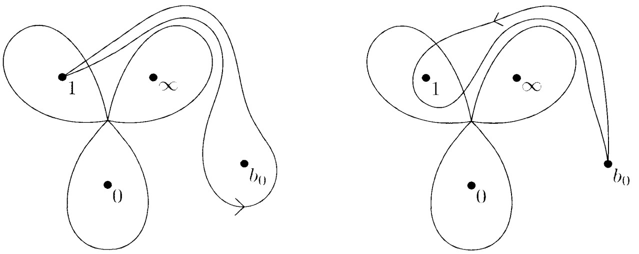
It seems useful to interpret the path $\rho(\gamma_0)$ in terms of the *inadmissable shared mating* which is associated with the point 0. The concept of *mating* of critically finite polynomials (of the same degree) is due to Douady and Hubbard, and subsequently studied by Tan Lei [**TL**], and Wittner [**W**]. One takes two polynomials, each restricted to a topological disc containing the finite critical orbit, and pastes the two together along the bounding circle, to get a critically finite branched covering, which may or



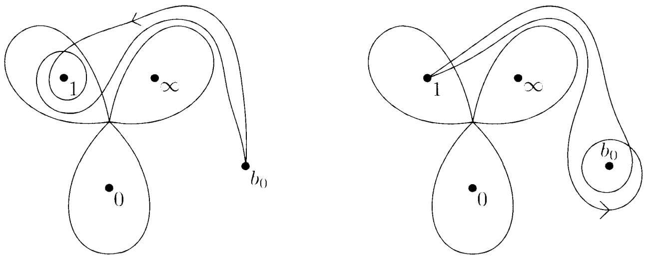
Pullback and ζ_0.



β_1 and equivalent.

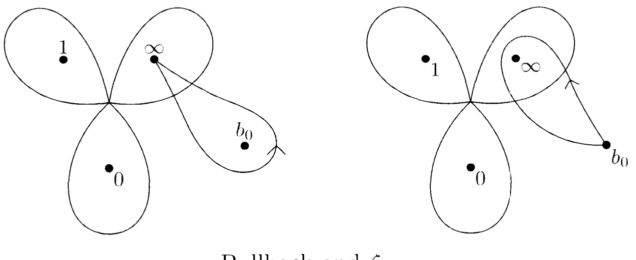


Pullback and ζ_1.



β_{-1} and equivalent.

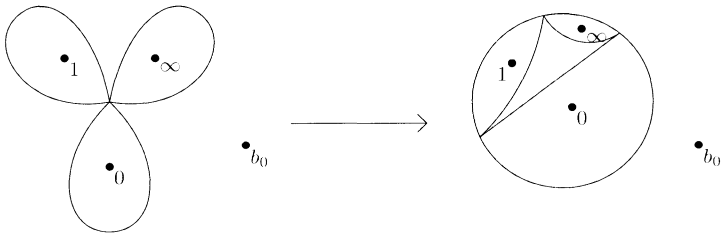
may not be Thurston equivalent to a rational map. Note that g_0 is a period three Möbius transformation (as has already been mentioned in the introduction). We can construct a critically finite branched covering which is C^0 close to g_0, with two distinct period 3 critical orbits. The branched covering thus obtained has two invariant circles



Pullback and ζ_{-1}.

up to isotopy (as we shall see) which makes it a *shared* mating, and is not Thurston equivalent to a rational map, which makes it *inadmissible*. We are going to show that the path $\rho(\gamma_0)$ helps us to identify the two matings involved.

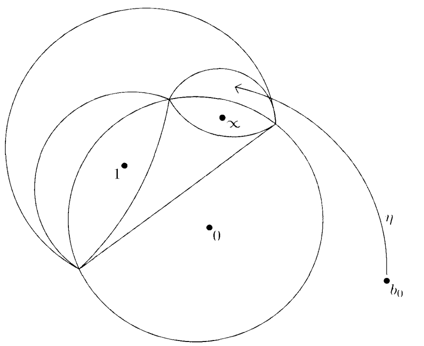
In order to define the associated matings, it is easier to draw pictures if the rabbit is converted to its disc model, as shown.



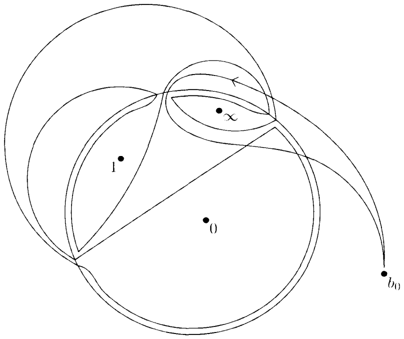
Rabbit and its disc model.

As well as converting the rabbit to its disc model, we can change g_{a_0} up to Thurston equivalence to a critically finite branched covering which preserves the round circle and permutes the set of three arcs bounding the triangle. In future, we shall call this branched covering $s_{1/7}$ (as we have in previous papers) and this is further explained in Chapter 2. A similar branched covering which we call $s_{6/7}$ is equivalent to the antirabbit polynomial. Now we draw the antirabbit disc model on the exterior of the round disc containing the rabbit disc model. Note that if $\eta : [0,1] \to \overline{\mathbf{C}}$ is the path indicated, with $\eta(0) = b_0$, then taking inverse images under $\sigma_\eta \circ s_{1/7}$, the circle is preserved up to isotopy in $\overline{\mathbf{C}} \setminus \{0, 1, \infty, b_0\}$, and so are the two sets of arcs, one set interior to the circle, and one exterior to it. We can take $\eta(1)$ close to the triangle vertex and perturb $s_{1/7}$ so that $\eta(1)$ has period 3 under $s_{1/7}$. Let $\xi : [0,1] \to \overline{\mathbf{C}}$ be the path with $s_{1/7} \circ \xi = \eta$ and $\xi(1) = s_{1/7}^3(\eta(1))$. Then $\sigma_\xi^{-1} \circ \sigma_\eta \circ s_{1/7}$ is Thurston equivalent to the critically finite branched covering we constructed from g_0, and also equivalent to a branched covering which we denote by $s_{1/7} \amalg s_{6/7}$. We summarize this notation (which was used in [**R3**] and [**R4**]) in 2.3.

Now we extend the path η above to the closed path β_0, we see that a new circle is left invariant by $\sigma_{\beta_0} \circ s_{1/7}$, as are the arcs of the old circle meeting the new circle



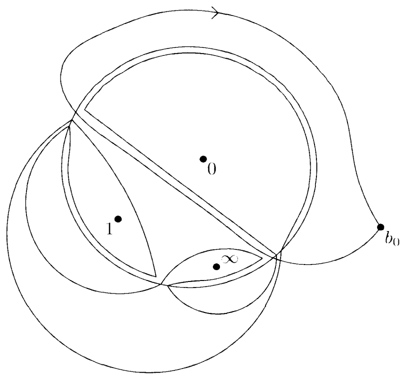
Rabbit and antirabbit.



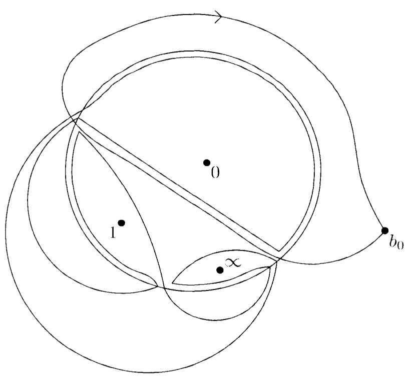
Rabbit and new antirabbit circle.

which cut separate the points 0, 1, ∞ from a central triangular region. The points $0, \infty, 1$ are now arranged in a clockwise direction round this triangle, so that $\sigma_{\beta_0} \circ s_{1/7}$ is equivalent to the antirabbit polynomial, that is, to $g_{\overline{a}_0}$, and so is $\sigma_{\beta_0} \circ g_{a_0}$. Thus we have seen that the combinatorial properties reflect that we have traced a path from g_{a_0} to $g_{\overline{a}_0}$ which went past g_0.

In a similar way we can move round further from the antirabbit to the rabbit, as shown in the diagram of the antirabbit and the new rabbit circle.



Antirabbit and new rabbit circle.



Antirabbit with previous rabbit circle.

The closed loop drawn above goes clockwise round 0. The same is true for $\overline{\zeta}_0$, but this in itself does not prove anything. However, we now draw the antirabbit with the previous rabbit circle, and we see that with reference to this, the closed loop goes clockwise round 0 and does not cut across the central triangular region (for the original circle) except adjacent to 0. This shows that with reference to the original rabbit puncture, the path being traced is $\overline{\zeta}_0$ and $\sigma_{\overline{\zeta}_0} \circ \sigma_{\beta_0} \circ g_{a_0}$ is equivalent to g_{a_0}.

CHAPTER 2

LOOP SETS SATISFYING THE INVARIANCE AND LEVY CONDITIONS

2.1. Let $B = B(Y, f_0)$ be one of our branched covering spaces. More precisely, the elements of $B(Y)$ are $(f, Y(f))$ (up to Möbius conjugation) where f is a branched covering and $Y(f) = Z(f) \cup V(f)$ is a union of the critical values of f and a finite set $Z(f)$ invariant under f. For more precise definitions, see 1.10. We usually write f for $(f, Y(f))$, and Y, Z etc. for $Y(f)$, $Z(f)$ where no confusion can arise. In this chapter we consider pairs (f, Γ), where Γ is a set of simple disjoint nontrivial loops in $\overline{\mathbf{C}} \setminus Y(f)$. The nodes and edges of the graph $\mathcal{B}$ which will ultimately be used to describe the topology of B are described in terms of such pairs satisfying certain conditions. The topological spaces associated to the nodes and edges, and identifications between them, are also described in terms of pairs (f, Γ). Important conditions which can be satisfied by a pair (f, Γ) are the *Invariance* and *Levy* conditions. Important objects associated to pairs satisfying these conditions are *periodic homeomorphic gaps* and *conjugacy of isotopy classes* on these and one or two *reduced map spaces*.

2.2. The Invariance and Levy Conditions. — Let (f, Γ) be such that $f \in B$ and Γ is a finite set of simple disjoint nontrivial loops in $\overline{\mathbf{C}} \setminus Y(f)$. We do not insist that all loops in Γ be isotopically distinct in $\overline{\mathbf{C}} \setminus Y(f)$. We consider the following two conditions which can be satisfied by (f, Γ).

The Invariance Condition. — (i) For any $\gamma \in \Gamma$ there exists $\gamma' \in \Gamma$ such that γ is isotopic in $\overline{\mathbf{C}} \setminus Z$ to a component of $f^{-1}(\gamma')$.

(ii) For any $\gamma \in \Gamma$, any component of $f^{-1}(\gamma)$ is either trivial or peripheral in $\overline{\mathbf{C}} \setminus Z$, or homotopic in $\overline{\mathbf{C}} \setminus Z$ to some loop of Γ.

The Levy Condition. — For $\gamma \in \Gamma$, there are $m \geqslant 1$ and a finite sequence γ_i $(1 \leqslant i \leqslant m)$ in Γ with $\gamma_m = \gamma$, γ_i is isotopic in $\overline{\mathbf{C}} \setminus Z$ to a component γ_i' of $f^{-1}(\gamma_{i+1})$, at least one γ_i is *nonperipheral*, that is, does not bound a disc containing just one point of $Y(f)$, and for some $r > 1$, γ_i $(1 \leqslant i \leqslant r)$ is a *Levy cycle for* $(f, Y(f)$, that is, $f \mid \gamma_i'$ is a homeomorphism and $\gamma_1 = \gamma_r$.

Remark

(1) The definition of *Levy cycle for* $(f, Y(f))$ coincides with the usual definition of Levy cycle if f is critically finite and $Z = X(f)$ $(= Y)$, where $X(f)$ is, as usual, the postcritical set of f. We allow Γ to include peripheral loops (round a single point in Y) because $Y \setminus Z$ is not invariant under f in general.

(2) The following definition will be useful. It extends a definition in the critically finite case. We say that a Levy cycle $\{\gamma_i : 1 \leqslant i \leqslant r\}$ $(\gamma_1 = \gamma_r)$ is *degenerate* if there is a disc D_i bounded by γ_i such that D_i is isotopic via a Z preserving isotopy to a component D'_i of $f^{-1}D_{i+1}$. It follows that each D_i contains the same number $(\geqslant 2)$ of periodic points of Z, and none of these points is critical.

2.3. Examples

(1) Let f be degree 2 such that the forward orbit of v_1 is finite. Let $Z(f)$ contain the forward orbit of v_1, but let $v_2(f) \notin Z(f)$. So $Y(f) \setminus Z(f)$ contains the single point $v_2(f)$. Take any simple loop γ_0 which bounds a disc Δ_0 containing v_1 and v_2 but no other points of $Y(f)$. We shall show how to use Γ_0 to generate a loop set satisfying the Invariance and Levy Conditions.

First, extend γ_0 to a set Γ_0 of disjoint simple loops by adding simple loops bounding disjoint discs round all the other points of Z. Then $f^{-1}(\Delta_0)$ is an annulus which contains, up to Z-preserving isotopy, just the loop of Γ_0 bounding a disc containing c_1. Then Γ_1, which is $\Gamma_0 \cup f^{-1}(\Gamma_0)$ up to Z-preserving isotopy is a set up disjoint simple loops up to Y-preserving isotopy, is a set of simple disjoint loops in $\overline{\mathbf{C}} \setminus Y$. Similarly, by induction, Γ_{n+1}, which is $\Gamma_n \cup f^{-1}(\gamma_n)$ up to Z-preserving isotopy, is a set of disjoint simple loops in $\overline{\mathbf{C}} \setminus Y$. Then for some n we must have $\Gamma_{n+1} = \Gamma_n$, modulo trivial loops and copies. Then $f^{-1}(\Gamma_n) \subset \Gamma_n$, modulo trivial loops and copies and Z-preserving isotopy. We claim that Γ_n contains a nonempty loop set satisfying the Invariance and Levy Conditions. If v_1 is not periodic then this is clear, because if $\Gamma'_0 \subset \Gamma_0$ is the set of loops bounding discs round the periodic orbit in the forward orbit of v_1, then $\Gamma'_0 \subset f^{-1}(\gamma'_0)$ modulo Z-preserving isotopy, and if Γ'_{i+1} is isotopic to $f^{-1}(\Gamma'_i)$ via Z-preserving isotopy with $\Gamma'_i \subset \Gamma'_{i+1}$, $\Gamma'_{m+1} = \Gamma'_m \subset \Gamma_n$ satisfies the Invariance Condition for some m, and is generated by a degenerate Levy cycle in a natural sense. If v_1 is periodic of period p, then we consider the annuli A_i such that A_i is an annulus containing $f^{p-i}(v_1)$, $1 \leqslant i < p$, and A_i is a component of $f^{-i}(\Delta_0)$. Let $Z_1 = \{f^i(v_1) : 1 \leqslant i \leqslant p\}$. Then all the A_i are disjoint up to Z-preserving isotopy. So at least one must have one trivial boundary component modulo Z_1-preserving homotopy. That means that there is a loop γ'_0 bounding a disc D'_0 round a point z_0 of Z_1, and γ'_{i+1} isotopic to a component of $f^{-1}(\gamma'_i)$ via Z-preserving isotopy, such that γ'_p and γ'_0 are isotopic via Z_1-preserving isotopy. Then similarly we can define γ'_i for all $i \geqslant 0$ so that all the loops γ'_i are disjoint modulo Y-preserving isotopy, and γ'_i and γ'_{i+p} are isotopic via Z_1-preserving isotopy, and γ'_{ip} bounds a disc D'_{ip} round

z_0 with $D'_{ip} \subset D'_{(i+1)p}$ and such that the only points of Y in $D'_{(i+1)p} \setminus D'_{ip}$ are in $Z \setminus Z'_1$. Then for some m we have $\gamma'_{(m+1)p} = \gamma'_{mp}$. Then we have a Levy cycle $\Gamma'_0 = \{\gamma'_{mp+i} : 0 \leqslant i < p\}$ which is nondegenerate, and contains Γ_0 up to Y-preserving isotopy if $Z = Z_1$. In general it only contains Γ_0 up to isotopy preserving $Z_1 \cup \{v_1\}$.

(2) Let f be critically finite, and $X(f) \subset Z \subset f^{-i}X(f)$ for some $i \geqslant 0$. Let Γ_0 be any union of Levy cycles for f. Then we can find a loop set containing Γ_1 which satisfies satisfies the Invariance and Levy conditions, and is *generated by* Γ_0 in a natural sense. This is done in exactly the same way as was done in Example 1 in more specific cases. We have $\Gamma_0 \subset f^{-1}(\Gamma_0)$ up to Z-preserving isotopy. Inductively, we define Γ_{n+1} to be $f^{-1}\Gamma_n$ up to Z-preserving isotopy and trivial loops, so that $\Gamma_n \subset \Gamma_{n+1}$. This uniquely determines Γ_n up to Y-preserving isotopy. Then for some n, $\Gamma_n = \Gamma_{n+1}$. For this n, Γ_n satisfies the Invariance and Levy Conditions.

(3) Now we specialise the second example. We use Thurston's theory of quadratic laminations [**T2**] (also described in 1.10 of [**R3**]), which was used to describe critically finite quadratic polynomials (and many others) up to topological conjugacy. We also use the notation we used to describe this theory in [**R3**] 1.10. For each odd denominator rational q in $(0,1)$, we have an *invariant lamination* L_q on $\{z :| z |< 1\}$ with *minor leaf* μ_q, where μ_q has one endpoint at $e^{2\pi iq}$ and other endpoint at a point $e^{2\pi ip}$, where $p \neq q$ is another odd denominator rational in $(0,1)$, and p and q have the same period k under the map $x \mapsto 2x$ mod 1. We shall then have $L_q = L_p$ and $\mu_q = \mu_p$. A *lamination* is a closed set of chords in the unit disc with disjoint interiors. The lamination L_q is *forward invariant* in the sense that if $\ell \in L_q$ is a chord in $\{z : |z| \leqslant 1\}$ with endpoints z_1, z_2 then there is also a leaf $\ell^2 \in L$ where ℓ^2 has endpoints z_1^2 and z_2^2. If ℓ_1 is a chord with endpoints w_1 and w_2, then $-\ell_1$ is the chord with endpoints $-w_1$, $-w_2$. The lamination L_q is also *backward invariant* in the sense that if $\ell \in L_q$ then there is $\ell_1 \in L_q$ with $-\ell_1 \in L_q$ such that $\ell_1^2 = \ell$. In any invariant lamination L the *minor leaf* is the image of the (one or two) longest leaves in L. Then L_q is by definition the smallest invariant lamination which has μ_q as minor leaf. Given q, μ_q is also uniquely determined: p is the only odd denominator rational of the same period as q such that the leaf μ_q with endpoints $e^{2\pi iq}$ and $e^{2\pi iq}$ generates a forward invariant lamination in which it is the shortest leaf — and hence the square of the longest leaf. In general, invariant laminations are allowed to have diameter leaves, and hence also degenerate leaves (a single point on the unit circle) since the square of a diameter is a degenerate leaf. But for q an odd denominator rational in $(0,1)$, L_q has no diameter or degenerate leaves. It is an important result from [**T2**] (and very useful for computation) that the different minor leaves μ_q, for odd denominator rationals q, are disjoint — with endpoints included.

There is a critically finite branched covering s_q of $\overline{\mathbf{C}}$ which is simply $z \mapsto z^2$ outside the unit disc and on the unit circle, maps leaves of L_q to leaves of L_q as dictated by the map on the endpoints, and then maps each *gap* of the lamination — component of the complement of $\cup L_q$ in the unit disc — to another gap, and is a branched covering on each gap of degree one or two, with at most one critical point, with degree as dictated by the map on the boundary. If μ_q has endpoints $e^{2\pi iq}$ and $e^{2\pi ip}$ then $s_q = s_p$. It is again part of Thurston's theory [**T2**] that every gap is eventually periodic, and, for q an odd denominator rational, there is exactly one periodic orbit of gaps for which the return map is degree two. The only gap in the orbit which is mapped with degree two by s_q is the gap containing 0 and s_q is defined so that 0 is the critical point. Then s_q is chosen so that 0 is periodic, of the same period as the gap itself. The minor leaf μ_q is in the boundary of one of the gaps in this orbit, the one containing the critical value, and the period of this gap under s_q, and hence also the period of the critical point, is the same as the period k of q under $x \mapsto x \bmod 1$. We also write $\mu_0 = \mu_1 = 1$ (the point on the unit circle), take $L_0 = L_1$ to be the empty lamination, and $s_0(z) = s_1(z) = z^2$. Each critically periodic quadratic polynomial is Thurston equivalent to s_q, for exactly one minor leaf μ_q and corresponding odd denominator rationals q, p with $\mu_q = \mu_p$. Conversely, for each odd denominator rational q, s_q is Thurston equivalent to exactly one critically periodic polynomial in the family $z \mapsto z^2 + c$.

An even denominator rational q also determines a lamination L_q, and there is an associated map s_q which can be modified slightly to be a branched covering which is Thurston equivalent to a critically finite quadratic polynomial. Any critically finite quadratic polynomial is Thurston equivalent to s_q for some odd or even denominator rational q in $[0, 1]$. The situation is only slightly more complicated than in the odd denominator case. For each even denominator rational q there is a *clean invariant lamination* L_q with a *minor gap* or *minor leaf* μ_q, which might be *degenerate*, with vertex or endpoint at $e^{2\pi iq}$. There is a map s_q preserving L_q, which is not a branched covering if μ_q is a minor leaf rather then a minor gap, but has arbitrarily small perturbations to a critically finite branched covering equivalent to a polynomial. The polynomial corresponding to p and q is the same if and only if L_q and L_p have the same minor leaf or gap. Since minor gaps are always finite-sided, the map from rational to polynomial is finite-to-one.

In [**R3**] 1.10, we also introduced the inverted laminations L^{-1} on $\{z :| z |> 1\}$ — which is simply the image of L under $z \mapsto z^{-1}$. The critically finite branched covering $s_q \amalg s_{q'}$ preserving $L_q \cup L_{q'}^{-1}$ is defined to be equal to s_q on $\{z : |z| \leqslant 1$ and $(s_{q'}(z^{-1}))^{-1}$ for $\{z : |z| \geqslant 1\} \cup \{\infty\}$. We did this in the case when q and q' are both odd denominator rationals. This was our terminology for the *mating* of polynomials introduced by Douady and Hubbard. If q or q' is an even denominator rational, we have a map from $L_q \cup L_{q'}^{-1}$ to itself such that a leaf with endpoint at $a \in S^1$ maps to a leaf with endpoint at a^2. Thus, we have a self-map on the set Γ of closed loops formed from the closures of leaves in $L_q \cup L_{q'}^{-1}$. If some leaf with a periodic endpoint is part of a closed loop, then all the leaves in the loop will have endpoints of the same period, and thus there are finitely many leaves on the loop. Perturbing s_q and

$z \mapsto (s_{q'}(z^{-1}))^{-1}$, we can make a branched covering f with $\Gamma \subset f^{-1}\Gamma$ up to isotopy. So, for varying rationals q and q' the sets $L_q \cup L_{q'}^{-1}$, and maps $s_q \amalg s_{q'}$, or some minor modification of this if q or q' has even denominator, give abundant examples of Levy cycles, and hence of loop sets satisfying the Invariance and Levy Conditions.

For minor leaves or gaps μ_p, μ_q (the gaps arise in some of the critically preperiodic cases), we write $\mu_p < \mu_q$ if μ_p separates μ_q from 0 in $\{z : \mid z \mid < 1\}$, and $\mu_0 < \mu$ for all $\mu \neq \mu_0$. Since minor leaves never intersect transversally, this is a partial order. The minimum $\mu \wedge \mu'$ of two minor leaves always exists, although it might be μ_0. It is not μ_0 if and only if μ and μ' *are in the same limb*. (This can be taken as a definition of being in the same limb.) It is a fact that if $\mu_p \leqslant \mu_q$, all periodic leaves in L_p (and many others) are in L_q. So if p is an odd denominator rational with $\mu_p \leqslant \mu_q$ and $\mu_{1-p} \leqslant \mu_{q'}$ then $\mu_p \cup \mu_{1-p}^{-1}$ generates a Levy cycle for $s_q \amalg s_{q'}$, or for a branched covering f obtained by modification if q or q' is an even denominator rational.

Now let $1 - q'$ be either in the full orbit of q under $z \mapsto z^2$, or a vertex of a finite-sided gap of L_q, with p such that

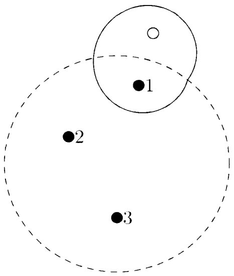
$$\mu_p = \mu_q \wedge \mu_{1-q'}, \ \mu_p \neq \mu_0.$$

Then we can generate a loop set Γ so that $(s_q \amalg s_{q'}, \Gamma)$ satisfies the Invariance and Levy Conditions by using periodic loops in $L_p \cup L_{1-p}^{-1}$, for any choice of Z.

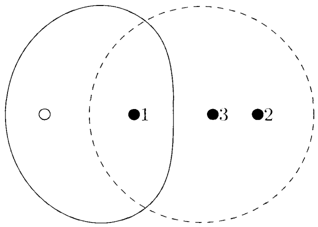
Below are all such possible Γ *up to homeomorphism* in the case $Z = \{f^i(c_1) : i \geqslant 0\}$ and c_1 has period 3 or 4. Here, $(q, q') = (1/7, 6/7)$, $(3/7, 1/3)$, $(1/15, 14/15)$, $(4/15, 11/15)$, $(2/5, 1/3)$, $(2/5, 2/5)$ or $(7/15, 1/3)$. The black dots denote the v_1-orbit, with r denoting $f^{r-1}v_1$, and the white dot denotes v_2. The unit circle is shown (dashed) to indicate how the examples are computed.

(4) *Extreme Pairs* Let f be degree two, let $Z(f)$ contain the critical value v_1 and $Y(f) \setminus Z(f) = \{v_2\}$. Let Γ consist of a simple loop bounding a disc containing v_2, and a single point of $Z \setminus \{v_1\}$, and simple loops bounding discs containing each of the remaining points of Z. Then each component $f^{-1}(\Gamma)$ is a single peripheral or trivial loop and $\Gamma \subset f^{-1}\Gamma$ modulo $Z(f)$-preserving isotopy. So Γ satisfies the Invariance Condition — but not the Levy Condition.

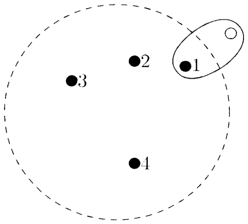
More generally the condition imposed on $\Gamma(f, N_1)$, for N_1, f and $\Gamma(f, N_1)$ in 1.10, is precisely that it should satisfy the Invariance Condition. It will not satisfy the Levy Condition if for every $\gamma \in \Gamma(f, N_1)$ with corresponding disc D bounded by γ and disjoint from $\Gamma(f, N_1)$, every component of $f^{-1}(D)$ is a disc. These were the circumstances (in 1.10) in which we could define a branched covering g_1 associated to N_1, such that N_1 was the intersection of $B(Y, f_0)$ with a neighbourhood of $B(Y(g_1, g_1)$ in $B(Y, f_0) \cup B(Y(g_1, g_1)$. We can then choose a path $t \mapsto g_t :: [0, 1] \to N_1 \cup \{g_1\}$ with $g_0 = f$, $g_t \in N_1$ for $t \in [0, 1)$. Then for $t \in [0, 1)$, $\Gamma(g_t, N_1)$ varies isotopically in $\overline{\mathbf{C}}$ and perturbing $\Gamma(g_t, N_1)$ to Γ_t in the same isotopy class via an isotopy preserving $Y(f_t)$, we can ensure that Γ_t converges to a set Γ_0 of peripheral loops in $\overline{\mathbf{C}} \setminus Y(g_0)$. We call (g_0, Γ_0) *extreme*.



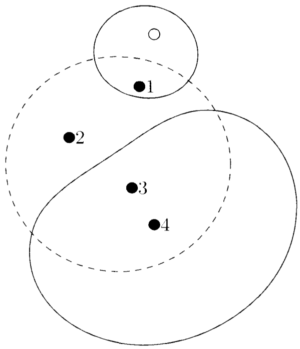
$(1/7, 6/7)$



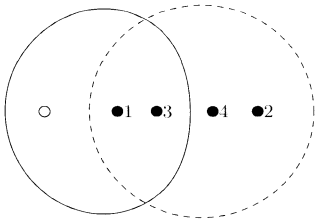
$(3/7, 1/3)$



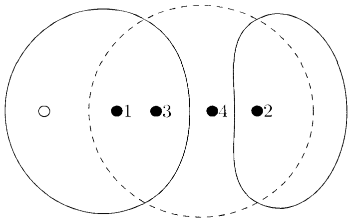
$(1/15, 14/15)$



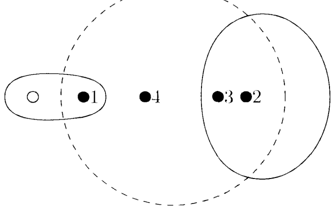
$(4/15, 11/15)$



$(2/5, 1/3)$



$(2/5, 2/5)$



$(7/15, 1/3)$

2.4. The Thurston Obstruction Condition. — The Levy Condition is mainly useful for $(f, Y(f)) \in B$ if B is of degree two type. For critically finite maps f, a loop set can be a *Thurston Obstruction*, and a Levy cycle is a special case of a Thurston obstruction. The generalisation of Thurston obstruction to the partially critically finite case might be necessarily messy. We give a definition below which is suggested by the line of proof in Chapter 6. (See 6.13.)

Let (f, Γ) satisfy the Invariance condition. For $\gamma \in \Gamma$, let $\pi(\gamma)$ denote the same loop up to isotopy in $\overline{\mathbf{C}} \setminus Z$. So π is the identity in the critically finite case, when $Y = Z$. Let $[\Gamma]$ denote the equivalence class of Γ for the equivalence relation $\sim$ such that $\Gamma_1 \sim \Gamma_2$ if every nontrivial nonperipheral loop of Γ_1 is isotopic via Y-preserving isotopy to a loop of Γ_2 and vice-versa, and Γ_1, Γ_2 have the same numbers of loops in each nontrivial nonperipheral Y-preserving isotopy class. We define $[\pi(\Gamma)]$ similarly, using Z-preserving isotopy.

Then following Thurston [**T2**], we can construct a linear map $f^* : \mathbf{R}^{[\Gamma]} \to \mathbf{R}^{[\pi(\Gamma)]}$, regarding the isotopy classes $[\gamma]$ ($[\gamma] \in [\Gamma]$) as basis elements. If $f^{-1}\gamma$ has nontrivial nonperipheral components δ_j, and $f \mid \delta_j$ is degree n_j, then we set

$$f^*[\gamma] = \sum_j \frac{1}{n_j}[\delta_j],$$

and extend linearly. The only difference from the critically finite case is that f^* does not have a square matrix. We say a matrix $A = A([\gamma], [\delta])$ with all entries $\geqslant 0$ is *compatible with* f^* if

$$f^*[\gamma] = \sum_{[\delta] \in [\Gamma]} A([\gamma], [\delta])[\pi(\delta)].$$

If $Y = Z$ is then there is only one matrix A compatible with f^*. In general, there might be more than one compatible matrix, because two loops δ_1, $\delta_2 \in \Gamma$ separated only by points of $Y \setminus Z$ will satisfy $[\pi(\delta_1)] = [\pi(\delta_2)]$. We define $\|A\|_\infty$ for $A = (a_{i,j})$ by

$$\|A\|_\infty = \mathrm{Max}_i \Big(\sum_j |a_{i,j}| \Big)$$

Then $\|\cdot\|_\infty$ is an operator norm, related to the $\|\cdot\|_\infty$-norm on the domain and range.

The Thurston Obstruction Condition. — For any $m \geqslant 1$ there are matrices A_i ($1 \leqslant i \leqslant m$) compatible with f^* such that

$$\Big\| \prod_{i=1}^m A_i \Big\|_\infty \geqslant 1.$$

If $Y = Z$ — that is, f is critically finite — then this coincides with the original condition for a Thurston Obstruction. We then have $A_i = A$ for a fixed matrix A for all i, and therefore $\prod_{i=1}^m A_i = A^m$. Since all entries of A are $\geqslant 0$, the condition $\liminf_{m\to\infty} \|A^m\| < 1$ is equivalent to the largest-modulus eigenvalue of A (which is positive) being < 1. If Γ contains a Levy cycle for $(f, Y(f))$ then it satisfies the

Thurston Obstruction Condition. Our aim is to show that, in a number of cases, the Thurston Obstruction Condition implies the Levy Condition. This generalizes known results for critically finite maps ([**L**], [**TL**]).

2.5. Definition of $\Gamma_1(f,\Gamma)$**,** $\Gamma_2(f,\Gamma)$**.** — Let (f,Γ) satisfy the Invariance condition. If D is some set with boundary in $\cup\Gamma$, then the Z-preserving isotopy class of $f^{-1}D$ is determined by the Y-preserving isotopy class of D. There are finitely many sets D_1 such that $\partial D_1 \subset \Gamma$ and such that D_1 contains the union of nontrivial nonperipheral components of $f^{-1}(D)$ up to Z-preserving isotopy, and is contained in the union of nontrivial components of $f^{-1}(D)$ up to Z-preserving isotopy. The number of such sets D_1 is bounded by $\#(Y \setminus Z)$. Similarly and inductively, for each choice of D_n, $n \geqslant 1$ there are finitely many choices for D_{n+1} such that $\partial D_{n+1} \subset \Gamma$ and such that D_{n+1} contains the union of nontrivial nonperipheral components of $f^{-1}(D_n)$ up to Z-preserving isotopy and is contained in the union of nontrivial components of $f^{-1}(D_n)$ up to Z-preserving isotopy. We define $\Gamma_1 = \Gamma_1(f,\Gamma)$ to be the set of loops γ in Γ with the following property:

There is a disc D bounded by γ such that, for all $n \geqslant 1$, all components of D_n are discs, however defined.

If $(f, Y(f))$ is in a degree two type space, this is equivalent to saying that the components of $f^{-n}D$ never contain both critical values, however defined.

Note that Γ_1 contains no loop from any nondegenerate Levy cycle. We define $\Gamma_2 = \Gamma_2(f,\Gamma)$ by

$$\Gamma_2 = \Gamma \setminus \Gamma_1.$$

2.6. Lemma. — *Let* (f,Γ) *satisfy the Invariance Condition and let* $\Gamma = \Gamma_1(f,\Gamma)$*. Then* (f,Γ) *satisfies the Thurston Obstruction Condition if and only if* Γ *contains a degenerate Levy cycle for* (f,Y)*.*

Proof. — The following argument is essentially the same as in the critically finite case in [**TL**].

Let Z_1 denote the set of periodic points of Z. For $\gamma \in \Gamma$, since $\Gamma = \Gamma_1(f,\Gamma)$, there is at least one component D of $\overline{\mathbf{C}} \setminus \gamma$ such that all components of $f^{-n}D$ are discs, however defined. (See 2.4.) Let $n(\gamma, D)$ be the number points of Z_1 in D and let $n(\gamma)$ be the minimum of the one or two possible numbers $n(\gamma, D)$. Let B_r be the set of isotopy classes of nonperipheral loops γ in Γ with $n(\gamma) = r$. Then for any A compatible with f^*,

$$A([\gamma],[\delta]) \neq 0 \text{ only if } \gamma \in B_r \text{ and } \delta \in B_s \text{ for } s \leqslant r.$$

Let A_i be compatible with f^*. Let $A_{i,(r,s)}$ be the submatrix defined by

$$A_{i,(r,s)} = (A_i([\gamma],[\delta]))\ ([\gamma] \in B_r,\ [\delta] \in B_s).$$

Let $A_{i,r} = A_{i,(r,r)}$. Then A_i has a triangular form with respect to the submatrices $A_{i,(r,s)}$, with the submatrices $A_{i,r}$ down the diagonal. Then any product $\prod_{i=1}^{m} A_i$ is also of triangular form, and the submatrix of $([\gamma],[\delta])$ entries for $[\gamma]\in B_r$ and $[\delta]\in B_s$ is 0 if $s>r$, and if $s\leqslant r$ it is a sum of $\leqslant m^{r-s}$ products of matrices of the form

$$\prod_{i=1}^{m} A_{i,(t_i,u_i)},$$

with $t_{i+1}\leqslant u_i\leqslant t_i$. So we need to show that such products decrease in $\|\cdot\|_\infty$ like λ^m for some $0\leqslant\lambda<1$. So it suffices to show that either Γ contains a degenerate Levy cycle for (f,Y) or, for some m and any choice of A_i, and any r,

$$\Big\|\prod_{i=1}^{m} A_{i,r}\Big\|_\infty < 1. \tag{1}$$

We do this by showing that $A_{i,r}$ splits further into subblocks corresponding to isotopy classes in $\overline{\mathbf{C}}\smallsetminus Z_1$, so that it looks like a permutation matrix. Write $[[\gamma]]$ for the isotopy class of γ in $\overline{\mathbf{C}}\smallsetminus Z_1$. Write

$$A([[\gamma]],[[\delta]]) = (A([\gamma'],[\delta']))\ (\gamma'\in[[\gamma]],\ \delta'\in[[\delta]]).$$

Then if A is compatible with f^* and $[\gamma]\in B_r$, there exists at most one isotopy class $[[\delta]]$, for $[\delta]\in B_r$ such that

$$A([[\gamma]],[[\delta]])\neq 0.$$

For such a $[[\delta]]$, write $[[\delta]]=\tau([[\gamma]])$. Then it suffices to show that either $[[\gamma]]$ contains a loop in a degenerate Levy cycle, or, for sufficiently large m,

$$\Big\|\prod_{i=1}^{m} A_i(\tau^{i-1}([[\gamma]]),\tau^{i}([[\gamma]]))\Big\|_\infty < 1, \tag{2}$$

where this is interpreted as 0 if $\tau^p([[\gamma]])$ is not defined for some $p\leqslant m$. If $\tau^p([[\gamma]])$ is defined for all $p\geqslant 0$ then we can take m to be the period of $[[\gamma]]$ under τ. Now we notice that the row sums of each matrix $A_i(\tau^{i-1}([[\gamma]]),\tau^i([[\gamma]])$ are $\leqslant 1$, and $\leqslant\frac{1}{2}$ if a disc D bounded by $\delta\in\tau^{i-1}([[\gamma]])$, for which all components of $f^{-n}(D)$ are discs, contains a periodic critical value of Z_1. So either (2) holds, or, iterating backwards, some $\gamma'\in f^{-n}(\gamma)$ with $[[\gamma']]=[[\gamma]]$ intersects Z in Z_1, and is in a degenerate Levy cycle. □

2.7. Corollary. — *If B is of polynomial type and (f,Γ) satisfies the Invariance Condition, it satisfies the Thurston Obstruction Condition if and only if Γ contains a degenerate Levy cycle for (f,Y).*

Proof. — Fix a critical value $v_1\in Z$ of multiplicity $d=\text{degree}(f)$. For any $\gamma\in\Gamma$ let $D=D(\gamma)$ be the component of $\overline{\mathbf{C}}\smallsetminus\gamma$ not containing v_1. Then all components of $f^{-n}D$ are discs, however defined, and $\Gamma=\Gamma_1(f,\Gamma)$. So then we apply 2.6. □

2.8. Definition of P. — For the rest of this chapter, let $B = B(Y, f_0)$ be of degree two periodic or eventually fixed type and let $f \in B$. Then Z contains at most two periodic orbits. If it contains two periodic orbits then one of these is a fixed point. If it contains a periodic orbit of period > 1 then this periodic orbit contains the critical point c_1. Also, $Y \setminus Z = \{v_2\} = \{f(c_2)\}$.

Lemma. — *Let $f \in B$ and let (f, Γ) satisfy the Invariance Condition. Then there exists a set P with the following properties.*

(1) *P is either a loop of Γ or a component of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ which is not a disc.*

(2) *There is a component P' of $f^{-1}P$ such that $P' \subset P$ up to isotopy preserving Z.*

(3) *If P is a loop of Γ, then $f : P' \to P$ either reverses orientation or is degree two. If P is a component of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ and $f \mid P'$ is a homeomorphism, then the boundary component of P which separates it from v_1 and v_2 is not fixed up to isotopy.*

If in addition P satisfies

(4) *$f \mid P'$ is a homeomorphism*

then P is the unique set satisfying properties 1 to 3.

Proof. — After composing f on the right with a homeomorphism isotopic via a Z-preserving isotopy to the identity, we can assume that $\Gamma \subset f^{-1}\Gamma$. Let Δ_0' be the component of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ which contains v_1 — this is consistent with notation we shall use later. Then $f^{-1}(\Delta_0')$ is connected. If $\Delta_0' \cap f^{-1}(\Delta_0') \neq \varnothing$ then we can take $P = \Delta_0'$. So now suppose that $\Delta_0' \cap f^{-1}(\Delta_0') = \varnothing$. The following argument is basically a discrete version of the result that a continuous order-reversing map of the interval contains a fixed point and is similar to arguments used previously, by Tan Lei for example [**TL**]. Number the loops of Γ and components of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ which separate Δ_0' and $f^{-1}(\Delta_0')$ in order, Δ_i', $1 \leqslant i \leqslant r$, with Δ_1' nearest to Δ_0'. Then there is a component Δ_1'' of $f^{-1}\Delta_1'$ which separates Δ_0' and $f^{-1}(\Delta_0')$. For if not, the component D of $\overline{\mathbf{C}} \setminus f^{-1}(\Delta_0')$ containing Δ_0' must be mapped strictly inside itself by f — to a component $f(D)$ of $\overline{\mathbf{C}} \setminus \Delta_0'$ disjoint from $f^{-1}(\Delta_0')$. Then ∂D and $\partial f(D)$ are distinct loops in Γ separated by v_1, and the sequence of discs $f^i(D)$ would have to be strictly decreasing, with $f^i(D)$ and $f^{i+1}(D)$ separated by $f^i(v_1)$, which is impossible. Now inductively we define a component Δ_{i+1}'' of $f^{-1}\Delta_{i+1}'$, if Δ_i'' a component Δ_i'' of $f^{-1}\Delta_i'$ has been defined, does not intersect Δ_i' and separates Δ_i' and $f^{-1}(\Delta_0')$. We take Δ_{i+1}'' to be the component of $f^{-1}\Delta_{i+1}'$ which is adjacent to Δ_i''. If $\Delta_{i+1}' \cap \Delta_{i+1}'' = \varnothing$ then Δ_{i+1}'' must separate Δ_{i+1}' and $f^{-1}(\Delta_0')$, because Δ_{i+1}'' is adjacent to Δ_i''. We cannot have Δ_i'' defined and separating $f^{-1}(\Delta_0')$ and Δ_i' for all $1 \leqslant i \leqslant r$, because Δ_r' is adjacent to $f^{-1}(\Delta_0')$ So for some $i \leqslant r$ we must have $\Delta_i' \cap \Delta_i'' \neq \varnothing$. Then since $\Gamma \subset f^{-1}\Gamma$, we must have $\Delta_i'' \subset \Delta_i'$. We then take $\Delta_i' = P$ and $\Delta_i'' = P'$, and property 2 holds. We have that $f \mid P'$ is a homeomorphism if and only if Δ_0' either contains v_2 or separates v_2 from $f^{-1}(\Delta_0')$. In that case, the component of $\overline{\mathbf{C}} \setminus P'$ containing $f^{-1}\Delta_0'$ is mapped to the component containing Δ_0'. So property 3 holds.

If, in addition, P satisfies 4, and U is a component of $\overline{\mathbf{C}} \setminus P$ with $U \cap f^{-1}U \neq \varnothing$, then by 3 for P, U does not contain v_1, v_2, and hence cannot contain any set satisfying 1 to 3. It follows that P is unique satisfying properties 1 to 3. □

2.9. Corollary. — *Let $f \in B$ and (f, Γ) satisfy the Invariance Condition. Let P be defined for (f, Γ) as above.*

(1) *Suppose that $f \mid P'$ is not a homeomorphism. Then $\Gamma = \Gamma_1(\Gamma)$. In particular, (f, Γ) does not satisfy the Thurston Obstruction Condition.*

(2) *Suppose that $f \mid P'$ is a homeomorphism. Then ∂P contains a Levy cycle for f and (f, Γ) does satisfy the Thurston Obstruction Condition.*

Proof

(1) For any $\gamma \in \Gamma$ take $D(\gamma)$ to be the component of $\overline{\mathbf{C}} \setminus \gamma$ disjoint from P. Then by induction on n, any nontrivial component of $f^{-n}D(\gamma)$ is disjoint from P. So $\gamma \in \Gamma_1(\Gamma)$. Since B is of periodic or eventually fixed type, there are no degenerate Levy cycles in Γ for (f, Y).

(2) Each component of ∂P is isotopic in $\overline{\mathbf{C}} \setminus Y$ to a unique component of $\partial P'$. So for each $\gamma \in \Gamma$ in ∂P there is a unique loop that we call $f_*\gamma \in \Gamma$ in ∂P with $\gamma \subset f^{-1}(f_*\gamma)$ up to Y-preserving isotopy. There is at least one periodic cycle under f_* in ∂P, which is a Levy cycle. □

2.10. Definition of Δ_0. — Let (f, Γ) satisfy the Invariance Condition and let $\Gamma_2 = \Gamma_2(f, \Gamma) \neq \varnothing$. Then $\Gamma_2 \subset f^{-1}\Gamma_2$, up to isotopy preserving Z. Let Δ be a disc component of $\overline{\mathbf{C}} \setminus (\cup\Gamma_2)$. Then for some least $n \geqslant 0$, a component of $f^{-n}\Delta$ is a disc component of $\overline{\mathbf{C}} \setminus (\cup\Gamma_2)$ containing v_1 and v_2, up to isotopy preserving Z. We call this component Δ_0, and $\Delta_0 \cap P = \varnothing$, $\partial\Delta_0 \in \Gamma_2$. Then we have the following lemma, very similar to both 1.4 of [**R4**] and an argument in [**TL**]

2.11. Lemma. — *Let $\gamma \neq \partial\Delta_0$, $\gamma \subset \Delta_0$, and let $D \subset \Delta_0$ be the disc bounded by γ. Then either $\gamma \in \Gamma_1$, or γ is periodic, and $\gamma = \partial\Delta_0$ and $D = \Delta_0$ up to isotopy preserving Z (but not Y).*

Proof. — Let $D = D(\gamma)$ be the disc in Δ_0 bounded by γ. We define the *nesting depth* of γ to be the maximum number n such that there exist discs D_i $(0 \leqslant i \leqslant n)$ with boundaries isotopically distinct in $\overline{\mathbf{C}} \setminus Y$ with $\partial D_i \subset \Gamma$ and $D_{i+1} \subset D_i$. If $\gamma \neq \gamma_0$ up to isotopy preserving Z then the nesting depth for γ is less than that for any loop which is γ_0 up to Z-preserving isotopy, any component D' of $f^{-1}D$ is a disc, we can define the nesting depth of $\partial D'$ using D', and it is $\leqslant$ the nesting depth for γ. Then by induction we prove that all components of $f^{-n}D$ are discs whose boundaries have nesting depth less than that of any loop which equals γ_0 up to Z-preserving isotopy. We can also use this induction for γ equal to γ_0 up to isotopy preserving Z but not Y, unless for some least $n > 0$ a component of $f^{-n}\gamma$ is isotopic to γ up to Z-preserving isotopy. □

2.12. Corollary. — *The loop set Γ_2 satisfies the Levy Condition, and Γ satisfies the Levy Condition if and only if $\Gamma = \cup_{n\geqslant 0} f^{-n}\Gamma_2$, up to isotopy preserving Z, copies, trivial and peripheral loops. Consequently, if (f,Γ) satisfies the Invariance Condition, then (f,Γ) satisfies the Thurston Obstruction Condition if and only if (f,Γ') satisfies the Invariance and Levy Conditions for some $\Gamma' \subset \Gamma$.*

Proof. — Let $\gamma \in \Gamma_2$. Write $\gamma = \gamma_0$. Then since $\Gamma_2 \subset f^{-1}\Gamma_2$, we can inductively choose $\gamma_i \in \Gamma_2$ $(i \geqslant 0)$ so that γ_i is isotopic in $\overline{\mathbf{C}} \setminus Z$ to a component of $f^{-1}\gamma_{i+1}$ Then $\gamma_r = \gamma_s$ for some $0 \leqslant r < s$, and $\{\gamma_i : r \leqslant i < s\}$ is a Levy cycle. Thus, $\cup_{n\geqslant 0} f^{-n}\Gamma_2$ satisfies the Invariance and Levy Conditions. As we have already noted, any Levy cycle in Γ is contained in Γ_2, so the proof is finished. □

2.13. Gaps. — Now let $\Gamma \subset \overline{\mathbf{C}} \setminus Y$ be any set of simple disjoint loops.

Definition. — Δ is a *gap for* Γ if Δ is a component of $\overline{\mathbf{C}} \setminus \cup\Gamma$ such that $\Delta \setminus Z$ is not an annulus.

Now suppose that (f,Γ) satisfies the Invariance Condition. If Δ is a gap for Γ, then choose Δ' such that $f^{-1}(\Delta')$ contains a component Δ'' intersecting Δ essentially modulo isotopy preserving Z. By the Invariance Condition, every component of $\partial\Delta''$ is either trivial or peripheral isotopic to a loop of Γ via a Z-preserving isotopy. Again by Invariance, no component of $\partial\Delta$ can be in the interior of Δ'' modulo Z-preserving isotopy. So the nontrivial nonperipheral components of Δ and Δ'' coincide modulo Z-preserving isotopy. So Δ' is unique with this property and we write $\Delta' = f_*\Delta$. If $f \mid \Delta''$ is a homeomorphism, then $f_*\Delta$ has at least as many boundary components as Δ, and we say Δ *is mapped homeomorphically*. Since there are only finitely many gaps, each is eventually periodic under f_*. If Δ has period k and each $f_*^i\Delta$ is mapped homeomorphically, we call Δ a *periodic homeomorphic gap*. Then there is φ isotopic to the identity relative to Y such that $\psi_\Delta = f^k \circ \varphi$ maps Δ homeomorphically to itself, and $[\psi_\Delta] \in \mathrm{MG}(\Delta)$ is uniquely determined, where $\mathrm{MG}(\Delta)$ is the group of homeomorphisms of Δ modulo isotopies that preserve $\partial\Delta$, but do not necessarily fix it pointwise. We call this the *conjugacy of isotopy class* of Δ.

We say that a connected union Δ of loops of Γ, gaps of Γ and annulus components of $\overline{\mathbf{C}} \setminus (\cup\Gamma \cup Y)$ is a *maximal periodic homeomorphic union* if all gaps in Δ are periodic and homeomorphic and Δ is a maximal union satisfying these properties. Then we can define $f_*\Delta$ in the same way as for gaps, Δ is periodic under f_* of some period k, we can again define a *conjugacy of isotopy class* $[\psi_\Delta]$.

2.14. Classification of Gaps. — *Let (f,Γ) be invariant and let $\Gamma_2(\Gamma, f) \neq \varnothing$. Let Δ be any periodic gap for Γ. Then there are three possibilities:*

(1) *Δ is homeomorphic,*

(2) *the gap Δ_0' with $\Delta_0' \subset \Delta_0$, $\partial\Delta_0 \subset \Delta_0'$ up to isotopy preserving Z, is periodic, and Δ is in the forward orbit of Δ_0',*

(3) Δ *is contained in a disc* D *bounded by a loop of* Γ, *where* D *intersects the forward orbit of* c_1, *which is periodic. If* Γ *satisfies the Levy Condition, then* $\Delta = D$.

Proof. — We shall consider separately periodic gaps with a boundary component in $\Gamma_2 = \Gamma_2(\Gamma, f)$, and those with no such boundary component.

First, suppose that Δ has a boundary component in Γ_2. Then the same is true for $f_*^i(\Delta)$, $i \geqslant 0$. So $f_*^i\Delta$ is never properly contained in Δ_0 up to Z-preserving isotopy, by 2.11. So either $f_*^i(\Delta) = \Delta_0'$ for some least $i \geqslant 0$, or $f_*^i(\Delta) \cap \Delta_0 = \varnothing$ for all $i \geqslant 0$. In the latter case, Δ is a homeomorphic gap.

Now suppose that no boundary component of Δ is in Γ_2, that is, every boundary component is in Γ_1. Then there is a disc D with $\Delta \subset D$, $\partial D \subset \partial\Delta$ such that all components of $f^{-n}(D)$ are discs. Then if $\Delta = f_*^n(\Delta)$ for a least integer $n > 0$, we have that D is a component of $f^{-n}(D)$ up to Z-preserving isotopy. Suppose that Δ is not homeomorphic. Then D is not either. If $n = 1$ then $D = f^{-1}(D)$ up to Z-preserving isotopy. But then $\overline{\mathbf{C}} \setminus D = f^{-1}(\overline{\mathbf{C}} \setminus D)$ modulo Z-preserving isotopy, and $\Gamma_2(\Gamma, f) = \varnothing$. So $n > 1$. Then D must contain a point in Z of period > 1, which must be in the forward orbit of v_1, and since v_1 is either periodic or eventually fixed, v_1 must be periodic.

Finally, if Γ satisfies the Levy Condition, $\Delta = D$, because any loops in the interior of D are not in $f^{-n}(\Gamma_2)$ for any $n > 0$. □

2.15. Important Gaps. — As above, we denote by Δ_0' the component of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ with $\Delta_0' \subset \Delta_0$, $\partial\Delta_0 \subset \Delta_0'$ up to isotopy preserving Z, and such that Δ_0' is either a gap or intersects Z. Then we might have $v_2 \in \Delta_0 \setminus \Delta_0'$. It can also happen that v_1 and v_2 are both in Δ_0 but in different components of $\Delta_0 \setminus \Delta_0'$. If Δ_0' has period n_0 under f_*, let Δ_0'' be the component of $f^{-n_0}\Delta_0'$ with $\Delta_0'' \subset \Delta_0'$, $\partial\Delta_0' \subset \partial\Delta_0''$. Let E_i be the component of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ containing v_i. Then E_i might not be a gap. But if it is, we write E_i' for the component of $f^{-p}E_i$ containing v_i. We also write $\Delta_0(f, \Gamma)$, $E_i(f, \Gamma)$ etc. if confusion can arise.

2.16. Extra Conditions on Loop Sets. — Let (f, Γ) satisfy the Invariance and Levy Conditions throughout this section. We now define some extra conditions on (f, Γ).

Node Condition. — The set $E_2(f, \Gamma)$ is a gap, that is, the component of $\overline{\mathbf{C}} \setminus (\cup\Gamma \cup Z)$ containing v_2 is not an annulus. If this holds, we also say that (f, Γ) *is a node.*

Edge Condition. — The set $E_2(f, \Gamma)$ is not a gap, that is, $E_2(f, \Gamma) \setminus Z$ is an annulus. in this case, we also say that (f, Γ) *is an edge.* We say that (f, Γ) is an *extreme edge* if $E_2(f, \Gamma)$ is a disc.

Maximal Reduced Condition. — Any maximal periodic homeomorphic union is a single gap. If $\Delta_0' \neq E_2$, and (f, Γ') is obtained from (f, Γ) by adding additional loops to the full orbit of Δ_0', then (f,Γ') does not satisfy the Levy Condition. If two loops are isotopic modulo Z but not modulo Y, then neither is peripheral.

We remark that if (f,Γ) does not satisfy the Maximal Reduced Condition, then we can remove loops from the full orbits of interiors of maximal periodic homeomorphic unions, remove at most one peripheral loop, and possibly add loops to the full orbit of Δ_0', so that it does. We shall see in 3.3 that there is exactly one such loop set, which we call the *Maximal Reduced Version* of (f,Γ).

For the remaining conditions, we take (f,Γ) satisfying the Maximal Reduced Condition, and we shall define (f,Γ) to have one of the following properties if its Maximal Reduced Version does.

If E_1 is periodic and $E_2 \neq E_1$ but $f_*^i E_2 = E_1$ for some $i > 0$, then (f,Γ) is a *capture*. We say this capture is *periodic* or *preperiodic*, depending on whether E_2 is periodic or preperiodic. These alternatives can only occur if B (with $(f, Y(f)) \in B$) is of periodic type rather than eventually fixed type. If (f,Γ) is a capture then $E_2 \neq \Delta_0'$. By 2.14, if (f,Γ) is not a capture then either $E_2 = \Delta_0'$ is periodic or E_2 is preperiodic. In the latter case, $f_*^i E_2 = f^i E_2$ up to isotopy for all $i \geqslant 0$, and $f \mid f^i E_2$ is a homeomorphism for all $i \geqslant 0$, unless $v_1 \in f^{i+1} E_2$, which can only happen in the fixed type case. In any case, $f \mid f^i E_2$ is a homeomorphism for all sufficiently large i. We then say that (f,Γ) is *preperiodic homeomorphic*. We give no special name to the case when $E_2 = \Delta_0'$ is periodic, although this is arguably the most important case.

If $\partial\Delta_0(f,\Gamma) \subset \partial P$ then (f,Γ) is *minimal nonempty*. (The reason for this terminology will become clear in Chapter 3.) If this is not true, but

$$\Gamma = \Gamma', \quad \text{where } \Gamma' = \cup_{n\geqslant 0} f^{-n}\partial P \tag{1}$$

up to Y-preserving isotopy, modulo copies of loops and trivial and peripheral loops, then we say that (f,Γ) is *primitive*.

Let (f,Γ') also satisfy the Invariant, Levy and Maximal Reduced Conditions. We say that (f,Γ') *is a tuning of* (f,Γ) if $\Gamma \subset \Gamma'$ and $\Gamma' \neq \Gamma$. If this happens then there must be loops of Γ' in the interior of $\Delta_0'(f,\Gamma)$, which must then be periodic and equal to $E_2(f_0,\Gamma)$, since (f,Γ) is Maximal Reduced. If (f,Γ') satisfies just the Invariance and Levy Conditions then we say that (f,Γ') is a tuning of (f,Γ) if its Maximal Reduced Version is.

If $\Gamma \neq \varnothing$, and (f,Γ) satisfies the Invariance and Levy Conditions, then (f,Γ) is exactly one of: nonempty minimal, primitive, or tuning. This follows from the definitions.

2.17. Nodes and Edges. — The following lemma gives a simple passage from an edge to a node, and vice versa.

Lemma. — *Let (f,Γ) satisfy the Invariance, Levy and Edge Conditions, but such that (f,Γ) is not an extreme edge. Let γ_1 be the loop which is isotopic to $\gamma_0 = \partial\Delta_0(f,\Gamma)$*

up to Z-preserving isotopy, but such that v_2 is between γ_0 and γ_1. Let Γ' be obtained from Γ by removing those loops in the backward orbit of γ_1 which are not also in the backward orbit of $\Gamma \setminus \{\gamma_1\}$. Then (f, Γ') satisfies the Invariance, Levy and Node Conditions.

Conversely, let (f, Γ) satisfy the Invariance, Levy and Node Conditions. If $E_2 = \Delta_0'$, let γ_1 be isotopic to γ_0 up to Z-preserving isotopy, but let v_2 be between γ_0 and γ_1. If $E_2 \neq \Delta_0'$, so that E_2 is a disc, let γ_1 be isotopic to ∂E_2 up to Z-preserving isotopy, but let v_2 be between γ_1 and ∂E_2. Let Γ' be the union of Γ and the backward orbit of γ_1. Then (f, Γ') satisfies the Invariance, Levy and Edge Conditions.

Proof. — Immediate. □

2.18. Reduced Branched Coverings and Map Spaces. — Let (f, Γ) satisfy the Invariance, Levy and Maximal Reduced Conditions. Suppose that Δ_0' exists as in 2.15. If Δ_0' is periodic of period ℓ under f_*, we define a *reduced branched covering g of Δ_0'* by

$$g = f^{\ell} \text{ on } \Delta_0'',$$

and then g is extended by mapping components of $\overline{\mathbf{C}} \setminus \Delta_0''$ to components of $\overline{\mathbf{C}} \setminus \Delta_0'$ as dictated by the boundary maps, and each component of $\overline{\mathbf{C}} \setminus \Delta_0'$ contains exactly one point of a set $Y(g)$, where the remaining points of $Y(g)$ are $\Delta_0' \cap Y$. We also choose g so that f, g have the same critical values v_1, v_2, and so that $g(Z(g)) \subset Z(g)$, where $Z(g) = Y(g) \setminus \{g(c_2)\}$. Thus we continue to have a labelling of critical points and their orbits. A *reduced branched covering for E_i* is similarly defined whenever E_i is a periodic gap.

This construction for Δ_0' is the same as that of critical branched covering in 1.9 of [**R4**] for critically finite branched coverings — except that there we used $f^{-1}\Delta_0'$ rather than Δ_0', but it makes no difference.

The *reduced map space $B(f, \Gamma, \Delta_0')$ for Δ_0'* (or $B(f, \Gamma, E_j)$ for E_j) is $B(Y(g), g)$, where g is a reduced branched covering. Note that, if the Edge Condition is satisfied, then $B(f, \Gamma, \Delta_0')$ is an equivalence class of critically finite branched coverings if Δ_0' is a periodic gap, and similarly for $B(f, \Gamma, E_1)$. If E_2 is a periodic gap, then either $E_2 = \Delta_0'$ or E_2 is in the periodic orbit of E_1. In the latter case, $B(f, \Gamma, E_1)$ and $B(f, \Gamma, E_2)$ are naturally identified. Note also that if Δ is periodic and $g \in B(f, \Gamma, \Delta)$, $\Delta = \Delta_0'$ or E_j, then

$$3 \leqslant \#(Z(g)) < \#(Z(f)).$$

Now we consider $B(f, \Gamma, E_1)$, in the cases where E_1 is a periodic gap. If (f, Γ) is preperiodic homeomorphic or a preperiodic capture, then $B(f, \Gamma, E_1)$ is critically finite, v_1 is periodic and v_2 is eventually fixed under $g \in B(f, \Gamma, E_1)$. If (f, Γ) is a periodic capture, then $B(f, \Gamma, E_1) = B(Y', f')$, where f' is degree 4, $Z(f') = \{v_0\} \cup \text{orbit}(v_1)$, $Y' = Y(f') = Z(f') \cup \{v_2\}$, v_0 is fixed and of local degree four, v_1 is periodic and of local degree two, v_2 has two critical points in its preimage, both mapping with local degree two. If (f,Γ) is minimal nonempty — or, more generally, whenever $\Delta_0' = E_1 = E_2$ — then $B(f,\Gamma,E_1)$ is of degree two periodic type. (This includes some cases when (f,Γ) is primitive.) If (f,Γ) is not a periodic capture, and $E_1 \neq \Delta_0'$ then $B(f,\Gamma,E_1)$ is critically finite and equivalent to a polynomial.

Now we consider $B(f,\Gamma,\Delta_0')$ when Δ_0' is a periodic gap, but $E_1 \neq \Delta_0'$. If (f,Γ) is a capture, or preperiodic homeomorphic with $E_2 \neq \Delta_0'$, then $B(f,\Gamma,\Delta_0')$ is critically finite with both critical points eventually fixed. The Maximal Reduced Condition ensures that the critically finite maps in this space are *irreducible*, that is, either equivalent to a rational map or to

$$[x] \longmapsto [Ax+b] : \mathbf{R}^2/\sim \;\longrightarrow \mathbf{R}^2/\sim$$

where $[x]$ is the equivalence class for $\sim$, $x \sim y$ if $x = \pm y + m$ $(m \in \mathbf{Z}^2)$, and A is an integer matrix with determinant 2 and eigenvalues λ, μ, $|\lambda| < 1 < 2 < |\mu|$, $2b = 0 \bmod \mathbf{Z}^2$. If (f,Γ) is primitive and $E_1 \neq \Delta_0'$ — or more generally whenever $E_2 = \Delta_0' \neq E_1$ — then $B(f,\Gamma,\Delta_0')$ is of degree two eventually fixed type.

2.19. Examples of Reduced Branched Map Spaces

(1) We consider the cases listed in 2.3, so that $f = s_q \amalg s_{q'}$ and $\Gamma \subset L_q \cup L_{q'}^{-1}$. Then $B(f,\Gamma,\Delta_0')$ contains the subset

$$\{g_\lambda : z \mapsto \lambda z + z^2,\ \lambda \neq 0\}$$

whenever $\mu_q = \mu_{1-q'}$, and is homotopically equivalent to this subset. These maps have a distinguished noncritical fixed point at 0. Now let $(q,q') = (3/7, 1/3)$ or $(4/15, 6/7)$ or $(7/15, 1/3)$ or $(7/15, 3/7)$. Then $B(f,\Gamma,\Delta_0')$ contains the subset

$$\left\{h_\lambda : z \mapsto \frac{z^2+\lambda z-\lambda}{z^2}\right\},$$

for which the second iterate of the critical point 0 is fixed. In all these cases, $B(f,\Gamma,E_1)$ is the Thurston equivalence class of $z \mapsto z^2$. Now let $(q,q') = (2/5, 1/3)$. The reduced map space for Δ_0' contains the subset

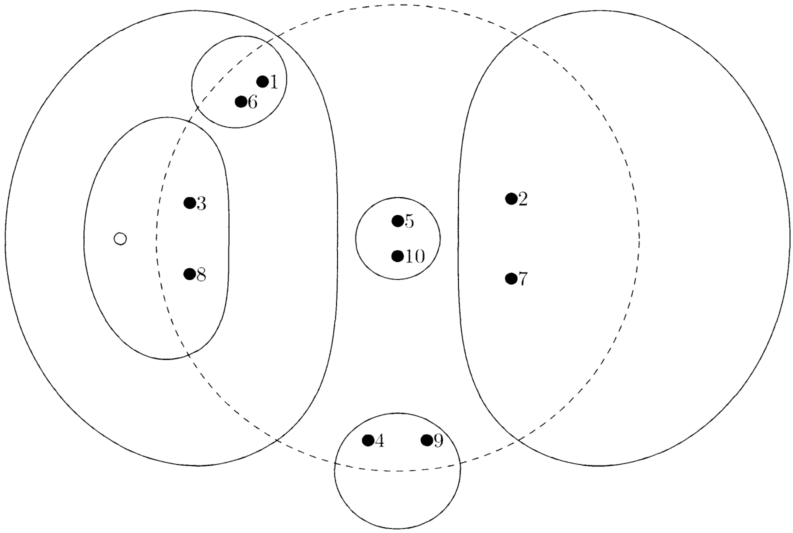
$$\left\{k_\lambda : z \mapsto \frac{1+\lambda z}{z^2}\right\}$$

— for which $c_1 = 0$ has period 2.

In all these cases, it can be shown that the reduced branched map space is actually homotopically equivalent to the given subset.

(2) We give an example of a reduced map space of polynomial degree four type for $B_{10,0}$. There is no such example for $B_{p,n}$ with $p < 10$. We consider $s_{11/31}$, which is equivalent to a polynomial with critical point of period 5. Let s_q denote the tuning of this, with critical point of period 10. Write $[[a,b]]$ for the leaf with endpoints $e^{2\pi ia}$, $e^{2\pi ib}$. Then $\mu_{11/31} = [[11/31, 12/31]]$. This leaf is separated from the minor gap of $L_{1/3}$ by the leaf $[[17/48, 19/48]]$ of $L_{1/3}$, while the forward image $[[13/31, 17/31]]$ is separated from the minor gap by $[[5/12, 7/12]]$. Let $f = s_q \amalg s_{q'}$ for

any odd denominator rational $q' \in [5/12, 7/12]$. Let Γ be obtained by taking inverse images under f of the loop $\mu_{1/3} \cup \mu_{1/3}^{-1}$. See the diagram. As in 2.3, the unit circle is shown (dashed) to indicate how the example is computed. Then (f, Γ) satisfies the Invariance and Levy Conditions and $B(f, \Gamma, E_2)$ is of degree four polynomial type.



Degree four polynomial type.

CHAPTER 3

NODES, EDGES AND ENHANCED LEVY SETS

3.1. Throughout this chapter, $f \in B(Y) = B$, where $B(Y)$ is of degree two periodic or eventually fixed type. We shall use the group $G = \pi_1(B, f)$ in this chapter, which, as we recall from 1.11, has interpretations as subgroups of $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ and $\mathrm{MG}(\overline{\mathbf{C}}, f^{-1}Y)$. The purpose of this Chapter is to examine the structure of the set of pairs (f, Γ). We define an equivalence relation on the set of pairs and an action of G on the set of pairs. The equivalence classes are called *enhanced Levy sets.* We also associate pairs (f, Γ) to convex regions in the unit disc, which are disjoint (after restricting the set of pairs slightly), and show that boundary convex regions in components of the union correspond to minimal nonempty pairs. This is the groundwork for the *Resident's View* of B.

3.2. We use the gap notation of 2.10, 2.15.

Lemma. — *Let (f, Γ), (f, Γ') satisfy the Invariance and Levy Conditions, with $E_2(f, \Gamma) = E_2(f, \Gamma')$. Let (f, Γ) satisfy the Maximal Reduced Condition. Then a loop of Γ' is contained either in Γ or in the full orbit of a periodic homeomorphic gap of Γ, up to isotopy preserving Y.*

Proof. — Write $\gamma_0 = \partial\Delta_0(f, \Gamma)$. If Γ satisfies the Edge Condition, let γ_1 be the unique loop of Γ which is isotopic to γ_0 in $\overline{\mathbf{C}} \setminus Z$ but not in $\overline{\mathbf{C}} \setminus Y$. Let $\delta = \delta_1 \in \Gamma'$, and let $\delta_i \in \Gamma'$ $(1 \leqslant i \leqslant n)$ satisfy: $\delta_i \subset f^{-1}(\delta_{i+1})$, $\delta_p = \delta_n$ up to Z-preserving isotopy, with $1 \leqslant p < n$, and such that $\{\delta_i : p \leqslant i < n\}$ is a Levy cycle. By the Levy Condition, any $\delta \in \Gamma'$ satisfies these properties. Suppose that any intersections of δ_i with $\cup\Gamma$ are essential in $\overline{\mathbf{C}} \setminus Y$ — and hence also in $\overline{\mathbf{C}} \setminus Z$, since $E_2(f, \Gamma) = E_2(f, \Gamma')$. The number of transverse intersections with $\cup\Gamma$ can only decrease under taking inverse images. So the number of intersections of δ_i with $\cup\Gamma$ must be constant for $p \leqslant i \leqslant n$. Similarly, the numbers of intersections with gaps, periodic homeomorphic gaps, and periodic gaps, must be constant. So for $p \leqslant i$, δ_i can intersect only periodic unions or gaps, and only periodic sides of periodic gaps which are not separated from the fixed set by $\Delta_0(f,\Gamma)$. Suppose $\Delta_0' = \Delta_0'(f,\Gamma)$ is periodic (and hence not homeomorphic). Suppose δ_i has an essential arc of intersection with Δ_0'. Then the arc must have both endpoints in γ_0 (or possibly γ_1 if Γ satisfies the Edge Condition), since this is the only periodic boundary component. Then $\Delta_0' \neq E_2$ and $v_2 \notin \Delta_0'$. So then the reduced branched covering for Δ_0' has no Levy cycle. But we can form a Levy cycle by joining essential arcs of δ_i along γ_0 (or γ_1), and taking the full orbit. This is a contradiction. So δ_i has no essential intersections with Δ_0', if Δ_0' is periodic. So the only possibilities are that $\delta_i \in \Gamma$, or δ_i lies in a periodic homeomorphic gap of Γ. Then we obtain the required conclusion for $\delta = \delta_1$ □

3.3. Corollary. — *If (f,Γ'') satisfies the Invariance and Levy Conditions, then there is a unique (f,Γ) satisfying the Invariance, Levy and Maximal Reduced Conditions with $E_2(f,\Gamma'') = E_2(f,\Gamma)$ such that:*

(1) *every loop of $\Gamma \setminus \Gamma''$ is in the full orbit of a periodic homeomorphic gap of Γ'',*

(2) *$\Gamma'' \setminus \Gamma = \varnothing$ unless $\Delta_0' \neq E_2$ and Δ_0' is periodic, in which case every loop of $\Gamma'' \setminus \Gamma$ is in the full orbit of Δ_0'.*

Proof. — If Δ_0' is periodic and $\Delta_0' \neq E_2$, then the reduced branched covering f_1 for Δ_0' is critically finite degree two (2.18). If there is a loop set satisfying the Invariance and Levy Conditions for f_1, then we can add the full orbit of this loop set to Γ'' to form a larger loop set satisfying the Invariance and Levy Conditions. There is only space to do this finitely many times. So now we assume that Γ'' cannot be augmented in this way. Any connected union of homeomorphic gaps is again homeomorphic. So first form Γ_1 by removing from Γ'' any loops in the interior of a maximal homeomorphic union. It is possible that the preimage Δ of a maximal homeomorphic union might include E_1 or E_2 and not be homeomorphic. If that is the case, we also need to remove loops in the interior of Δ and the backward orbits of these in Γ, to obtain Γ_2. If $\Delta_0'(\Gamma_2)$ is periodic then $\Delta_0'(\Gamma_2) = \Delta_0'(\Gamma'')$ and if $\Delta_0' \neq E_2$. So Γ_2 satisfies the Maximal Reduced Condition as well as the Invariance and Levy Conditions.

So suppose that both (f,Γ) and (f,Γ') satisfy the Invariance, Levy and Maximal Reduced Conditions with $E_2(f,\Gamma'') = E_2(f,\Gamma) = E_2(f,\Gamma')$. Apply 3.2. □

3.4. Definition. — We call (f,Γ') as in the Lemma the *Maximal Reduced Version* of (f,Γ).

3.5. The next two lemmas are preliminary to defining the group associated to a pair (f,Γ).

Lemma. — *Let (f,Γ) satisfy the Invariance and Levy Conditions. Let Δ be a periodic gap of Γ of period k under f, such that $f^k \mid \Delta$ is a homeomorphism. Then either f^k cyclically permutes all the components of $\partial\Delta$, or it fixes one and cyclically permutes the rest. If Δ contains a point of Z, then this point is fixed, and the components of $\partial\Delta$ are cyclically permuted.*

Proof. — Replacing Δ by some $f^j\Delta$ if necessary, we can assume that all $f^i\Delta \neq \Delta$ are in the same component of $\overline{\mathbf{C}} \setminus \Delta$. If some component of $\overline{\mathbf{C}} \setminus \Delta$ intersects Z in only a fixed noncritical point of Z, then Δ is fixed by f. Take any component D of $\overline{\mathbf{C}} \setminus \Delta$ bounded by $\gamma \subset \partial\Delta$ which contains no $f^i\Delta$, and which intersects the forward orbit of v_1. Thus, at most one component of $\overline{\mathbf{C}} \setminus \Delta$ is excluded. Inductively, let $D_0 = D$, and let D_{j+1} be defined as the component of $f^{-1}D_j$ whose boundary is common to some $f^\ell\Delta$, if $f(c_1) \notin D_i$. Then for a minimal j, D_j contains $f(c_1)$ and no $f^i\Delta$. Thus, the boundary is uniquely determined, and has γ in its forward orbit. Hence all but at most one component of $\partial\Delta$ are in the same cyclic orbit under f. By the nature of Z, Δ can contain at most one point z of Z, which then has to be fixed, and by the specified properties of Z, is the only fixed point of f in Z. Suppose that Δ does contain such a point. So Δ is also fixed. Then D, as above, can be any component of $\overline{\mathbf{C}} \setminus \Delta$ and all components of $\partial\Delta$ are in the same orbit. □

3.6. Classifying the Gap Homeomorphisms. — The next Lemma uses Thurston's isotopy classification of surface homeomorphisms [**F-L-P**]. Let S be any compact surface with boundary and $f : S \to S$ an orientation-preserving homeomorphism. Then f is *irreducible* if there is no set Γ of simple closed disjoint non-boundary-homotopic loops such that $f(\Gamma)$ is isotopic to Γ. For every (S, f), either f is irreducible, or there is a set of Γ of simple closed disjoint non-boundary-homotopic loops such that $f(\Gamma)$ is isotopic to Γ and such that if S_i is any component of $S \setminus \cup\Gamma$ and m is the least integer > 0 with $f^m(S_i)$ isotopic to S_i, then $f^m \mid S_i$ is irreducible. An irreducible homeomorphism is either isotopic to an isometry for some complete hyperbolic metric on S or is isotopic to a *pseudo-Anosov*. This means that f^m is isotopic to a homeomorphism which leaves invariant two transverse *measured foliations* on S_i which have finitely many singularities of specified types [**F-L-P**]. One of the foliations is called *stable* (or *contracting*) and the other is called *unstable* (or *expanding*). The fact that the foliations are measured and transverse means that there is a measurement of length on all leaves of both foliations. Then there is $\lambda > 1$ such that f^m multiples length of leaves of the unstable foliation by λ and length of leaves of the stable foliation by λ^{-1}.

3.6. Lemma. — *Let $F : \Delta \to \Delta$ be a homeomorphism of a holed sphere with boundary, which either cyclically permutes boundary components or fixes one and cyclically permutes the rest.*

(1) *Write*

$$\Delta = \bigcup_{i=1}^{t} S_i,$$

where the S_i are closed subsurfaces with disjoint interiors that are permuted by F, the loops ∂S_i are all isotopically nontrivial and distinct, and such that the first return *map F^m to $S = S_i$ is, up to isotopy, either an isometry for some complete hyperbolic metric on S or a pseudo-Anosov. Then this decomposition is unique.*

(2) *Let* PMG(Δ) *denote the group of isotopy classes preserving each component of $\partial\Delta$, where isotopies preserve $\partial\Delta$ but do not fix it pointwise. The centraliser of F in* PMG(Δ) *is free abelian, of rank $n_1 + n_2$, where n_1 is the number of cycles of S_i in the above decomposition for which the return map $F^m \mid S_i$ is pseudo-Anosov, and n_2 is the number of cycles of loops ∂S_i in the interior of Δ.*

Proof

(1) This uses a technique that we have applied previously in both Chapter 8 of [**R3**] and 3.3 of [**R4**]. Note that at most one subsurface $S_i = S$ is fixed by F and that $F^m \mid S$ has the same property as $F : \Delta \to \Delta$, namely it either cyclically permutes the boundary components or fixes one and cyclically permutes the rest. If $h = F^m \mid S$ is an isometry, we can find an arc α joining distinct components of ∂S such that all arcs $h^i\alpha$ are (up to isotopies preserving ∂S but not fixing it pointwise) either disjoint or equal, and $\mathrm{Int}(S) \setminus (\cup_{i\geqslant 0} h^i\alpha)$ consists of two discs. We simply take α to be a geodesic arc of minimal length joining any two distinct components of ∂S — if h cyclically permutes them all — or between the fixed component and any of the others — if h fixes one component. Then the set $\cup_{i\geqslant 0} h^i\alpha$ consists (up to isotopy) of simple disjoint cyclically permuted arcs, and h^p = identity, where p is the least integer > 0 such that h^p fixes all components of ∂S. If h is pseudo-Anosov, there is no such arc set in S. So any F-invariant set of simple disjoint loops Γ in Δ can only intersect ∂S essentially, for S in the decomposition, if $F^m \mid S = h$ is an isometry, and then all components of ∂S must be intersected. If ∂S is intersected transversally by Γ for some S, then $\partial S'$ is intersected transversally for some S' with $\partial S' \cap \partial\Delta \neq \varnothing$, which is impossible. So there is S in the decomposition such that $\Gamma \cap S$ is a set of disjoint simple loops invariant under h. This is impossible if h is pseudo-Anosov, unless $\Gamma \cap S \subset \partial S$ up to isotopy. We get the same conclusion if h is an isometry by considering intersections of Γ with the h-orbit of α. Hence no simple loops, apart from the ∂S_i components, are mapped periodically by F and the decomposition of Δ into the S_i is unique.

(2) All components of all ∂S_i are fixed by any $\varphi \in \mathrm{PMod}(\Delta)$ in the centraliser of F, as are any arcs $h^i\alpha$. Since the centraliser of any pseudo-Anosov is cyclic, we obtain the result. □

3.7. The Group $G(f,\Gamma)$. — Let f be our chosen basepoint in B and $G = \pi_1(B,f)$. Then as we have seen in 1.11, $\pi_1(B,f)$ identifies with the set of $[\varphi] \in \mathrm{MG}(\overline{\mathbf{C}},Y)$ arising from closed loops in B based at f. Let (f,Γ) satisfy the Invariance and Levy Conditions. We define $G(f,\Gamma)$ to be the subgroup of those φ in G for which $\varphi(\Gamma) = \Gamma$ up to isotopy. Our Lemmas imply that $G(f,\Gamma) = G(f,\Gamma')$ if (f,Γ') is the Maximal Reduced Version of (f,Γ).

3.8. The Equivalence Relation on Pairs and Enhanced Levy Sets

Let f, $g \in B$ and let (f, Γ) satisfy the Invariance and Levy Conditions. Then any path from f to g induces a homeomorphism $\varphi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ with $\varphi(Y(f)) = Y(g)$. As we have seen in Chapter 1, if φ_0 and φ_1 are two choices for φ, then there is an isotopy φ_t between φ_0 and φ_1 with $\varphi_t(Y(f)) = Y(g)$ for all t. Then $(g, \varphi(\Gamma))$ also satisfies the Invariance and Levy Conditions.

We define $\sim$ to be the equivalence relation generated by: (f, Γ) is equivalent to its Maximal Reduced Version, and, if (f, Γ) satisfies the Invariance, Levy and Maximal Reduced Conditions, then $(f, \Gamma) \sim (g, \Gamma')$ if, for some path, from f to g and associated φ defined up to isotopy constant on $Y(f)$, $\varphi(\Gamma) = \Gamma'$. We let $[f, \Gamma]$ denote the equivalence class of (f, Γ). Thus, if we fix f, $[f, \Gamma]$ is simply the orbit of (f, Γ) under the action of G. We can transfer all the definitions of 2.16 (Edge, Node Conditions and so on) from (f, Γ) to $[f, \Gamma]$. For the later definitions, except for the definition of tuning, we need to take (f, Γ) satisfying the Maximal Reduced Condition. If $[f, \Gamma]$ is a tuning of $[f, \Gamma']$ and both (f, Γ) and and (f, Γ') satisfy the Maximal Reduced Condition, then loops of Γ' are either in Γ or in orbits of periodic homeomorphic gaps of Γ. If we do this, then Reduced Map Spaces and Conjugacy of Isotopy Classes also depend only on $[f, \Gamma]$ (and not on (f, Γ)). The group $G(f, \Gamma)$ depends only on $[f, \Gamma]$ up to isomorphism. Therefore, we shall often write $G[f, \Gamma]$.

We call the equivalence classes $[f, \Gamma]$ *enhanced Levy sets.*

3.9. Theorem. — *Let (f, Γ) and (g, Γ') satisfy the Invariance, Levy and Maximal Reduced Conditions. Then, for some Dehn twist composition τ round loops of Γ, $[f \circ \tau, \Gamma] = [g, \Gamma']$ if and only if the following hold for some orientation-preserving homeomorphism $\varphi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$.*

(1) $\varphi(Y(f)) = Y(g)$, $\varphi(v_i(f)) = v_i(g)$ *($i = 1$, 2), $\varphi(\Gamma) = \Gamma'$ and $\varphi \circ f = g \circ \varphi$ on $Z(f)$.*

(2) *The elements $[\psi_\Delta]$ and $[\psi_{\varphi\Delta}]$ of $\mathrm{MG}(\Delta)$, $\mathrm{MG}(\varphi\Delta)$ are conjugated by some homeomorphism which maps each component b of $\partial\Delta$ to φb, for at least one gap Δ in each periodic homeomorphic cycle.*

(3) *The reduced map spaces for $\Delta_0'(f)$, $\Delta_0'(g)$ are the same, as are the reduced map spaces for $E_i(f)$, $E_i(g)$, $i = 1$, 2, whenever these gaps are periodic.*

Proof. — Suppose that 1, 2 and 3 hold for (f, Γ) and (g, Γ'). Then we can change the definition of φ on each periodic homeomorphic gap Δ so that φ conjugates $[\psi_\Delta]$ and $[\psi_{\varphi\Delta}]$. Then we can change the definition of φ on all but one gap in each cycle so that $\varphi \circ f = g \circ \varphi$ up to isotopy constant on $\partial\Delta$. Then we can change the definition of φ on preimage gaps so that $\varphi \circ f = g \circ \varphi$ on any gap in the backward orbit of a periodic homeomorphic cycle. Then we can change the definition of g on whichever of $\Delta_0'(g)$ $E_1(g)$, $E_2(g)$ are periodic, and change the definition of φ on gaps in the full orbits of these, so that $\varphi \circ f = g \circ \psi$ on each gap Δ, for φ and ψ isotopic via an isotopy constant on $\partial\Delta$ and Z. Then

$$g \circ \psi = \varphi \circ f \circ \tau,$$

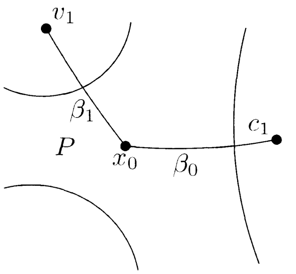
where τ is a composition of Dehn twists round loops of Γ. □

Remark. — In general, if τ is a Dehn twist round loops of Γ, $[f,\Gamma]$ and $[f\circ\tau,\Gamma]$ can be distinct. However, we do have the following.

3.10. Minimal Isometric Nodes and Edges: characterization

Lemma. — *Let (f,Γ) be minimal nonempty isometric. Then $[f,\Gamma] = [s_p \amalg s_q, \Gamma']$, where $s_p \amalg s_q$ preserves Γ', $\mu_p \cup \mu_{1-p}^{-1} \in \Gamma$, μ_p is minimal and $\mu_{1-p} \leqslant \mu_q$, and $c_1(s_p \amalg s_q) = \infty$, $v_1(s_p \amalg s_q) = (s_p \amalg s_q)(\infty)$.* (See 2.3 for the notation.)

Proof. — Assume without loss of generality that f preserves its fixed set P and that $f \mid P$ is an isometry with respect to some metric. Let ∂P have k components. Let α be an arc of minimal length joining components of ∂P. Then $\{f^i\alpha : 0 \leqslant i < k\}$ is a set of disjoint simple cyclically permuted arcs with endpoints on ∂P (as in 3.6). Choose a fixed point x_0 of f in P. Let β_1 be an arc from the critical value v_2 of f to x_0, which crosses ∂P exactly once, and is disjoint from the arc set $\cup_i f^i\alpha$. Let $\beta_k = \beta_0 \subset f^{-1}(\beta_1)$ be the arc joining x_0 to the critical preimage c_2 of v_2, which is in a different component of $\overline{\mathbf{C}}\setminus P$. Then similarly define an arc $\beta_i \subset f^{-1}(\beta_{i+1})$ for $k > i \geqslant 2$, with one endpoint at x_0. Now $[f,\Gamma] = [\sigma_\gamma \circ f, \Gamma]$, for any path $\gamma : [0,1] \to \overline{\mathbf{C}}\setminus(Z \cup (\cup\Gamma))$ with $\gamma(0) = v_2$. To see this, note that $[\sigma_{\gamma_t} \circ f, \Gamma] = [\sigma_\gamma \circ f, \Gamma]$, where $\gamma_t = \gamma \mid [0,t]$. Replacing f by $\sigma_\gamma \circ f$ if necessary for suitable γ, we can assume that $\beta_1 \subset f^{-1}\beta_2$. Thus we have made f critically finite. There is an invariant circle separating the critical orbits, formed as follows.Take the boundary of a small disc neighbourhood of $\cup_i\beta_i$ where $\beta_0 = \beta_k$ and $\beta_i \subset f^{-1}\beta_{i+1}$, $0 \leqslant i \leqslant k$. The loop set ∂P gives $L_{1-p} \cup L_p^{-1}$ up to isometry, for some p with μ_p minimal. See the diagram in the case $k = 3$. The arcs β_0 and β_1 need not be adjacent in general.



Minimal Isometric.

Hence $[f,\Gamma] = [s_p \amalg s_q, \Gamma']$, for some q and Γ'. □

3.11. Equivalence and the Edge Group. — Again, let μ_p be minimal and $\mu_{1-p} \leqslant \mu_q$. For the moment, write $[s_p \amalg s_q]$ for the (unique) equivalence class $[s_p \amalg s_q, \Gamma]$ where Γ contains $\mu_p \cup \mu_{1-p}^{-1}$ and $(s_p \amalg s_q, \Gamma)$ satisfies the Invariance and Levy Conditions, and the Edge Condition. Let Γ' be the loop set defined similarly for $s_{p'} \amalg s_{q'}$. The question which naturally arises is: when is $[s_p \amalg s_q] = [s_{p'} \amalg s_{q'}]$? We must have $p' = p$ or $1-p$, by considering the isotopy class $[\psi_P]$ (2.13) where $P = P(s_p \amalg s_q)$ is the fixed gap. This corresponds to looking at the rotation order of the arcs β_i of 3.10. Equally, given $[s_p \amalg s_q]$, there is at least one q' such that $[s_p \amalg s_q] = [s_{1-p} \amalg s_{q'}]$. The invariant circle which realises $[s_p \amalg s_q]$ as $[s_{1-p} \amalg s_{q'}]$ can be drawn in exactly the same way as the rabbit and antirabbit circles in 1.15, that is, by a half-twisting the circle for $s_p \amalg s_q$ round the fixed set. See the pictures in 1.15.

So now suppose that $p = p'$. It turns out, in the lemma below, that $[s_p \amalg s_q] = [s_p \amalg s_{q'}]$ if and only if $s_p \amalg s_q$ and $s_p \amalg s_{q'}$ are Thurston equivalent as critically finite branched coverings. This equivalence is considered in Stimson's thesis [**Sti**]. It is claimed incorrectly in [**Sti**] that if $s_p \amalg s_q \simeq s_p \amalg s_{q'}$ then $\mu_q = \mu_{q'}$. The true result is a little different, and is shown by the same method (after correction) that is employed by Stimson.

Lemma. — *The group $G[s_p \amalg s_q]$ is cyclic. The number of different leaves $\mu_{q'}$ such that $[s_p \amalg s_q] = [s_p \amalg s_{q'}]$ is the same as the number of $\mu_{q'}$ such that $s_p \amalg s_q$ is Thurston equivalent to $s_p \amalg s_{q'}$ as a critically finite branched covering, and is N, where $N \geqslant 1$ is the integer such that the generator of $G[s_p \amalg s_q]$ is $\sigma_2 \circ \sigma_1$, where σ_1 is an N-fold Dehn twist round the forward orbit Γ_1 of $\mu_p \cup \mu_{1-p}^{-1}$ and σ_2 is a composition of Dehn twists round loops isotopically distinct from and disjoint from the loops of Γ_1.*

Remark. — Write $f = s_p \amalg s_q$. Let $\Gamma \subset \overline{\mathbf{C}} \setminus X(f)$ be the smallest loop set containing Γ_1 such that $f^{-1}(\Gamma) \subset \Gamma$ modulo trivial loops and homotopy in $\overline{\mathbf{C}} \setminus X(f)$. Let $f^* : \mathbf{R}^{\Gamma} \to \mathbf{R}^{\Gamma}$ be the usual linear map for critically finite maps, (with the same formula as in 2.4). Then N is also the least integer $\geqslant 1$ such that there is an eigenvector of f^* with entry N in the entries corresponding to Γ_1 and integers in all entries.

Proof. — The proof, a slight modification of Stimson's method, is combinatorial. We first note that if $s_p \amalg s_q \simeq_\varphi s_p \amalg s_{q'}$, where this denotes Thurston equivalence as branched coverings, then $[s_p \amalg s_q] = [s_p \amalg s_{q'}]$. The converse takes a little more work. So now suppose that $[s_p \amalg s_q] = [s_p \amalg s_q, \Gamma] = [s_p \amalg s_{q'}, \Gamma'] = [s_p \amalg s_{q'}]$. Let φ be the homeomorphism with $\varphi(\Gamma) = \Gamma'$, $\varphi \circ (s_p \amalg s_q) \circ \psi^{-1} = s_p \amalg s_{q'}$ for ψ isotopic to φ via an isotopy constant on $Z(s_p \amalg s_q) \supset \{(s_p \amalg s_q)^i(\infty) : i \geqslant 0\}$.

We have

$$P(s_p \amalg s_q) = P(s_p \amalg s_{q'}) = P, \quad E_2(s_p \amalg s_q) = E_2(s_p \amalg s_{q'}) = E_2,$$
$$\varphi(\Delta_0'(s_p \amalg s_q)) = \Delta_0'(s_p \amalg s_{q'}), \quad \varphi(E_1(s_p \amalg s_q)) = E_1(s_p \amalg s_{q'}).$$

The reduced branched coverings for Δ_0' — which are critically finite since we are considering edges — are Thurston equivalent, and similarly for E_1. (See 2.15, 2.18). If $\Delta_0' = E_1$, then the reduced branched covering is a critically periodic polynomial, and we deduce immediately that $\mu_q = \mu_{q'}$. If $\Delta_0' \neq E_1$, then the reduced branched covering for Δ_0' is a polynomial with the finite critical value eventually fixed. The Thurston equivalence classes of these are in one-to-one correspondence with rationals in $(0,1)$ of the form $r/2^n$ for $n > 0$ and integers r, $0 < r < 2^n$. It follows that $\Delta_0'(s_p \amalg s_q) = \Delta_0'(s_p \amalg s_{q'}) = \Delta_0'$. Then since $\pi_1(B(\Delta_0')$ is trivial, the homeomorphism φ is the identity on Δ_0' up to isotopy, and fixes all boundary components. Consider

$$\Omega = P \cup E_2 \cup \big(\cup_{i \geqslant 0} (s_p \amalg s_q)^i (\Delta_0')\big).$$

Then $\varphi(\Omega) = \Omega$, and $\varphi \mid \Omega$ must be isotopic to a composition of Dehn twists t_i, t_0' times round the loops γ_i $(0 \leqslant i \leqslant k-1)$, γ_0', where $\gamma_i \subset \partial P$ with $\gamma_{i+1} \subset (s_p \amalg s_q)^{-1}(\gamma_i)$ and γ_0, γ_0' are separated by $v_2 = 0$, but are isotopic in $\overline{\mathbf{C}} \smallsetminus Z(s_p \amalg s_q)$. Then we obtain $t_i = t_{i+1}$ for $0 \leqslant i \leqslant k-2$, $t_{k-1} = t_0 + \frac{1}{2}t_0'$, giving $t_0' = 0$, $t_i = t$ for $0 \leqslant i \leqslant k-1$. So then we can adjust φ to induce a Thurston equivalence $s_p \amalg s_q \simeq_\varphi s_p \amalg s_{q'}$. So now, given p and μ_q we need to determine how many different $\mu_{q'}$ there are with $s_p \amalg s_q$ Thurston equivalent to $s_p \amalg s_{q'}$.

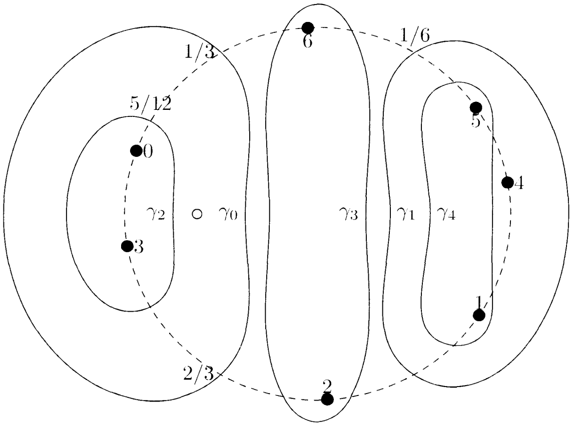
The claim in Stimson's thesis was that, by taking inverse images under $s_p \amalg s_q$ and $s_p \amalg s_{q'}$, φ is isotopic to the identity on all gaps in the full orbit of Δ_0'. But this is only true so long as gaps are pulled back homeomorphically, or so long as a gap is in the *forward* orbit of Δ_0'. (This is how the mistake arose.) The map φ is the identity on the full backward orbit of $\Delta_0'(s_p \amalg s_q)$ if and only if $N \mid t$. Distinct homeomorphisms are given by $0 \leqslant t < N$. We write φ_t for the homeomorphism corresponding to t, so that φ_0 is the identity. Each such t does determine a $\mu_{q'} = \mu(t)$, because if γ is the loop of Γ nearest to μ_q, then γ is homotopic to $\partial E_1(s_p \amalg s_q)$ and $\varphi_t(\gamma)$ must be homotopic to $\partial E_1(s_p \amalg s_{q'})$. If s_q is not a tuning then these boundaries are homotopic to $\mu_{1-q} \cup \mu_q^{-1}$, $\mu_{1-q'} \cup \mu_{q'}^{-1}$. If s_q is a tuning, then the fact that the reduced branched coverings for $E_1(s_p \amalg s_q)$ and $E_1(s_p \amalg s_q)$ are equivalent under φ_t shows that $\mu_{q'}$ is uniquely determined by φ_t. □

3.12. Examples. — This integer N must always be 2^m for some $m \geqslant 0$. We can see this by analysing the types of matrix A which can occur for $f^* : \mathbf{R}^\Gamma \to \mathbf{R}^\Gamma$ with respect to the standard basis. Take the first few rows and columns to be indexed by the loops of Γ_1. There may or may not be a second periodic cycle of loops Γ_2 in Γ. If Γ_2 exists, it is not a Levy cycle, and we take the last few rows and columns to be indexed by the loops of Γ_2. Then

$$A = \begin{pmatrix} P_1 & 0 \\ A_{21} & A_{22} \end{pmatrix} \text{ or } \begin{pmatrix} P_1 & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & P_2 \end{pmatrix}$$

where A_{22} is lower triangular with 0's on the diagonal, P_1 is a permutation matrix and P_2 almost is, but one of the entries is multiplied by 1/2. All entries of the matrix A are integers or half-integers. Then $\det(A_{22} - I) = \pm 1$, and (if P_2 exists) $\det(P_2 - I)$ has two terms, one of which is $(-1)^r$ and the other is $\frac{1}{2}(-1)^{r+1}$, if P_2 is an $r \times r$ matrix. So $\det(P_2 - I) = \pm\frac{1}{2}$ and $(P_2 - I)^{-1}$ has integer entries. So the inverse of the matrix $A_{22} - I$ has all entries in $\mathbf{Z}/2^{m'}$ for some $m' \geqslant 0$. It follows that $N = 2^m$ for some $m \geqslant 0$

The picture illustrates L_q for $q = 55/127$ or equivalently $q = 56/127$, since these are ends of the same minor leaf. In this case $m = 1$. The point $2^n q$ mod 1 on the circle is labelled n for the least possible $n \geqslant 0$ The period of q under $x \mapsto 2x$ mod 1 is 7, while the period of $p = 1/3$ is 2. This gives the lowest possible period of q with $m \geqslant 1$.



$q = 55/127$.

Writing $f = s_p \amalg s_q$, and letting $\gamma_0 = \mu_p \cup \mu_{1-p}^{-1}$ as before, there is one nonperiodic component γ_1 of $f^{-1}(\gamma_0)$ and for $i \leqslant 3$ we can inductively find one nontrivial nonperipheral component of $f^{-1}(\gamma_i)$. Taking these five loops γ_i $(0 \leqslant i \leqslant 4)$ to index the rows and columns, the matrix of f^* is

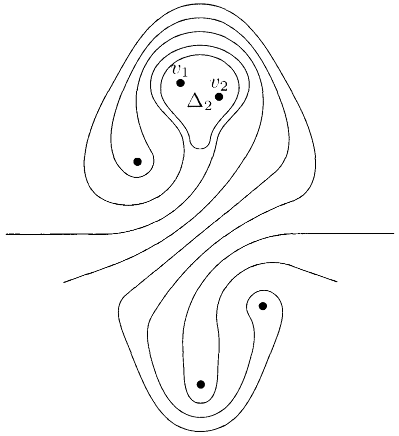
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The smallest eigenvector for eigenvalue 1 with integer entries, written as a row vector, is $(2, 2, 2, 1, 1)$, that is, $N = 2 = 2^m$ and $m = 1$. The q' with $\mu_{q'} \neq \mu_q$ and $s_{1/3} \amalg s_q \simeq s_{1/3} \amalg s_{q'}$ is $q' = 72/127$ (or equivalently $q' = 71/127$).

3.13. The sets $C(f,\Gamma)$. — Fix a basepoint f. We are going to construct a *convex hull* $C(f,\Gamma)$ in the unit disc D, regarded as the universal cover of $\overline{\mathbf{C}}\setminus Z$, for each (f,Γ) satisfying the Invariance and Levy Conditions, or Invariance and Extreme Conditions, such that $C(f,\Gamma')=C(f,\Gamma)$ if (f,Γ') is the Maximal Reduced Version of (f,Γ), and such that $C(f,\Gamma_2)\subset C(f,\Gamma_1)$ if $(f,\Gamma_1)\leqslant(f,\Gamma_2)$. We shall call $C(f,\Gamma)$ a *Levy convex hull* if (f,Γ) satisfies the Levy Condition.

Let $\pi_2 : D\to\overline{\mathbf{C}}\setminus Z$ be the covering map, with $\pi_2(0)=v_2(f)$. Let $\widetilde{E}_2(f,\Gamma)$ be the lift of $E_2(f,\Gamma)$ with $0\in\widetilde{E}_2(f,\Gamma)$. Let $\Delta_2(f,\Gamma)$ be the disc with $E_2\subset\Delta_2$, $\partial\Delta_2\subset\partial E_2$ and Δ_2 disjoint from the fixed set of (f,Γ). Let $\widetilde{\Delta}_2$ be the lift to D with $\widetilde{E}_2\subset\widetilde{\Delta}_2$. If (f,Γ) is a node, we simply take $C(f,\Gamma)$ to be the geodesic version of $\widetilde{E}_2$, that is, $C(f,\Gamma)$ is the geodesic convex hull of $\overline{\widetilde{E}_2}\cap S^1$. If (f,Γ) is a edge, which is not minimal or a minimal tuning we take $C(f,\Gamma)$ to be the geodesic version of $\partial\widetilde{\Delta}_2$. The geodesic version of an extremal edge (f,Γ) is a point on S^1. In fact, this point coincides with $\rho_2([\beta])$, where ρ_2 is the map defined in 1.12 and $[\beta]\in\pi_1(B,N,f)$ is defined as follows: $\beta(t)=\sigma_{\alpha_t}\circ f$ where α is the (unique) path in $\overline{\mathbf{C}}\setminus(\cup\Gamma)$ from v_2 to a point of Z, and $\alpha_t=\alpha\mid[0,t]$, where α is parametrised by $[0,1]$.

Now suppose that (f,Γ) is an edge which is minimal or a minimal tuning. Then E_2 is adjacent to a periodic homeomorphic gap Q such that $[\psi_Q]$ is irreducible. If $[\psi_Q]$ is isometric, we define $C(f,\Gamma)$ as for the other edges. Now let $[\psi_Q]$ be pseudo-Anosov. Then there are two transverse measured geodesic laminations on $\overline{\mathbf{C}}\setminus Y$ which are preserved by ψ_Q^{-1} up to isotopy with transverse measures multiplied by a scalars λ, λ^{-1}, $1<\lambda$. We take the lamination whose transverse measure is multiplied by λ and let $\Omega_2=\Omega_2(f,\Gamma)$ be the complementary component which contains both critical values. Up to Y-preserving isotopy, we have $\Delta_2\subset\Omega_2$. See the diagram, which is a sketch in the case when v_1 has period four.



Ω_2 and Δ_2.

Then we take $C(f,\Gamma)$ to be the geodesic version of the lift $\widetilde{\Omega_2 \setminus \Delta_2}$ of $\Omega_2 \setminus \Delta_2$ which is adjacent to $\widetilde{E}_2$. Then $\partial C(f,\Gamma)$ contains countably many geodesics projecting to $\partial\Omega_2$, and one projecting to $\partial\Delta_2$.

3.14. Lemma. — *If $C(f,\Gamma) \cap C(f,\Gamma') \neq \varnothing$, then one is contained in the other, in which case one is a tuning of the other, or $C(f,\Gamma)$, $C(f,\Gamma')$ intersect at most in a boundary component.*

Proof. — This is similar to the Parameter Laminations Theorem 1.16 of [**R3**], but in the present finite situation the proof is simpler. It is also similar to 3.2. The basic idea, in any case, is that Levy cycles for a critically finite degree two branched covering do not intersect.

We can assume, replacing f by $\sigma_\beta \circ f$ for $\beta : [0,1] \to \overline{\mathbf{C}} \setminus Z$ with $\beta(0) = v_2$ if necessary, that $\widetilde{E}_2(f,\Gamma)$ and $\widetilde{E}_2(f,\Gamma')$ have geodesic boundaries. So suppose that $\widetilde{E}_2(f,\Gamma) \cap \widetilde{E}_2(f,\Gamma') \neq \varnothing$. We can assume that Γ and Γ' have only transverse intersections, and that both satisfy the Maximal Reduced Condition. Then replacing f by $f \circ \psi$ for a homeomorphism $f \circ \psi$, we can assume that f preserves both Γ and Γ'. We can also assume that f is critically finite. Then a periodic loop of Γ can only intersect periodic loops and periodic homeomorphic gaps of Γ', and vice versa. Then, by the Maximal Reduced Condition, there are no transversal intersections. So $(f,\Gamma \cup \Gamma')$ satisfies the Invariance and Levy Conditions. So

$$\Delta_0'(f,\Gamma \cup \Gamma') = \Delta_0'(f,\Gamma') \subset \Delta_0'(f,\Gamma),$$

without loss of generality. So either $(f,\Gamma \cup \Gamma')$ satisfies the Edge Condition, in which case $C(f,\Gamma) \cap C(f,\Gamma')$ is a common boundary component, or $\Delta_0'(f,\Gamma') = \Delta_0'(f,\Gamma)$ and $\Gamma = \Gamma'$ up to isotopy, or

$$\Delta_0'(f,\Gamma') = E_2(f,\Gamma') \subset \Delta_0'(f,\Gamma) = E_2(f,\Gamma),$$

and $\Gamma' \subset \Gamma$ up to isotopy, in which case, (f,Γ') is a tuning of (f,Γ), without loss of generality. □

3.15. An action of G on Convex Hull Boundaries. — Fix a basepoint f_0 of B. Remember (1.11) that $G = \pi_1(B)$ is antiisomorphic via Φ_1 to the subgroup G_1 of $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$ of φ such that φ and $f_0^{-1} \circ \varphi \circ f_0 \circ \sigma_\alpha$ are isotopic via an isotopy fixing Y for a closed path α in $\overline{\mathbf{C}} \setminus Z$ based at v_2. Then $[\alpha] = \rho([\varphi])$. Then if (f_0,Γ) satisfies the Invariance Condition and $\varphi \in G_1$, so does $(f_0,\varphi^{-1}(\Gamma))$, and similarly for the Levy Condition or the Extreme Condition. Let $\psi^{-1} = \sigma_\alpha \circ \varphi^{-1}$ as in 1.13, and let $\widetilde{\psi}$ be the lift to D as defined there. Then $\widetilde{\psi}^{-1}(\widetilde{E}_2(f_0,\Gamma))$ is homotopic to $\widetilde{E}_2(f_0,\varphi^{-1}(\Gamma))$. It follows that, with the notation of 1.13, if $[\varphi] = \Phi_1([\beta])$,

$$C(f_0,\varphi^{-1}(\Gamma)) \cap S^1 = [\beta] \cdot C(f_0,\Gamma) \cap S^1.$$

In particular, the action of G on S^1 defined in 1.13 preserves the set of boundaries of convex hulls.

3.16. Adjacent Convex Hulls. — We now know that if each of (f,Γ) and (f,Γ') is primitive or minimal nonempty, then the sets $C(f,\Gamma)$ and $C(f,\Gamma')$ have disjoint interiors. The following theorem gives us important information about the union of the sets $C(f,\Gamma)$. It is an analogue of the Parameter Gaps Theorem of 1.16 of [**R3**], and the idea of the proof is the same.

Theorem. — *The boundary of a connected component of Levy convex hulls is a union of convex hulls $C(f,\Gamma)$ where (f,Γ) is a minimal nonempty isometric edge, and of components of $\partial C(f,\Gamma)$ lifting to geodesic lamination leaves, where (f,Γ) is a minimal nonempty pseudo-Anosov edge.*

Proof. — Suppose that (f,Γ) is minimal nonempty or primitive and satisfies the Node Condition. Fix a boundary component $\widetilde{\gamma}'$ of $\widetilde{E}_2(f,\Gamma)$ projecting to a boundary component γ' of $E_2(f,\Gamma)$. Let Ω be the gap on the other side of γ'. Let $\widetilde{\Omega}$ be the lift of Ω adjacent to $\widetilde{E}_2$. If (f,Γ) is minimal nonempty and Ω is the fixed set, then γ' is also in the boundary of (f,Γ') for (f,Γ') minimal nonempty satisfying the Edge Condition. The corresponding convex hull is either $\widetilde{\gamma}'$ or $\widetilde{\Omega}$, depending on whether the fixed set is isometric or pseudo-Anosov.

So now suppose that Ω is not the fixed set. Then we shall construct (f,Γ') such that

$$\widetilde{\gamma}' \subset C(f,\Gamma') \quad \text{and} \quad \widetilde{\Omega} \cap C(f,\Gamma') \smallsetminus \widetilde{\gamma}' \neq \varnothing.$$

Let Γ_1 be the set of loops $\gamma \in \Gamma$ with the following property. There exist n and $\gamma_i \in \Gamma$ $(1 \leqslant i \leqslant n)$ with $\gamma = \gamma_1$, $\gamma_i \subset f^{-1}(\gamma_{i+1})$ up to Z-preserving isotopy, $\gamma_i \neq \gamma'$ up to Z-preserving isotopy for $1 < i \leqslant n$, and γ_n is not in the interior of $\Delta_2(f,\Gamma)$. Then $\gamma' \in \Gamma_1$, because we are assuming that (f,Γ) is either primitive or minimal nonempty. In the latter case, we are assuming that γ' is not in the boundary of the fixed set. Let α be a path with first endpoint at v_2 and the second in the gap of Γ_1 adjacent to γ' and intersecting Ω, such that α intersects γ' just once. Then inductively define

$$\Gamma_{i+1} = (\sigma_\alpha \circ f)^{-1}(\Gamma_i)$$

We assume inductively that α has second endpoint in the gap of Γ_i adjacent to γ' and intersecting Ω. We have $\Gamma_1 \subset \Gamma_2$ and then $\Gamma_i \subset \Gamma_{i+1}$ by induction. Then for some i, $\Gamma_i = \Gamma_{i+1}$. So (f,Γ_i) satisfies the Invariance and Levy Conditions, and we take $(f,\Gamma') = (f,\Gamma_i)$. □

3.17. Tunings of (f,Γ) and convex hulls of pairs for $B(f,\Gamma,\Delta_0')$. — Let (f,Γ) satisfy the Invariance, Levy and Maximal Reduced Conditions. Recall that tunings of (f,Γ) exist only if $E_1(f,\Gamma) = E_2(f,\Gamma) = \Delta_0'(f,\Gamma)$, in which case this set is periodic. Recall that $B(f,\Gamma,\Delta_0')$ is a component of $B(Y(f,\Gamma,\Delta_0'))$, where $Y(f,\Gamma,\Delta_0') = Y'$ is obtained from $\Phi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ which is a homeomorphism on Δ_0' but collapses each component of $\overline{\mathbf{C}} \smallsetminus \Delta_0'$ to a point in Y'. We can assume that f^p maps Δ_0'' into Δ_0', where $\Delta_0'' \subset \Delta_0'$ and all loops of $\partial\Delta_0'' \smallsetminus \partial\Delta_0'$ are trivial or peripheral. Let g be any branched covering which equals f^p on Δ_0'' and preserves Y' and maps $\overline{\mathbf{C}} \setminus \Delta_0''$ to $\overline{\mathbf{C}} \setminus \Delta_0'$. So we have a natural map

$$(f, \Gamma') \longmapsto (g, \Gamma' \cap \Delta_0'')$$

which maps tunings of (f, Γ) to pairs for $B(f, \Gamma, \Delta_0')$.

We can lift Φ to $\widetilde{\Phi} : \widetilde{E}_2(f, \Gamma) \to D$, since D is the universal cover of $\overline{\mathbf{C}} \setminus Z(f, \Gamma, \Delta_0')$. Assume without loss of generality that $\widetilde{E}_2(f, \Gamma) = C(f, \Gamma)$ and suppose that $\widetilde{\Phi}$ is chosen with $\widetilde{\Phi}(0) = 0$. Then from the definitions, it follows that

$$\widetilde{\Phi}(C(f, \Gamma')) = C(g, \Gamma' \cap \Delta_0'').$$

3.18. A Partial Ordering on Enhanced Levy Sets. — We now define a partial ordering on enhanced Levy sets. We take

$$(f, \Gamma) < (f, \Gamma')$$

if

a) $C(f, \Gamma')$ is properly contained in $C(f, \Gamma)$ and (f, Γ') is not an edge; or

b) if $C(f, \Gamma')$ is adjacent to $C(f, \Gamma)$ or in $\partial C(f, \Gamma)$ and is properly contained in $\Delta_2(f, \Gamma)$; or

c) $C(f, \Gamma) = \partial \Delta_2(f, \Gamma')$.

We then take the partial order generated by this, and then define $[f, \Gamma] \leqslant [f, \Gamma']$ if this is true for some representatives (f, Γ) and (f, Γ'). This is a well-defined partial order, because the original ordering is clearly G-invariant.

3.19. Examples in period three. — Now we describe the nodes and extreme edges and the ordering on enhanced Levy sets, for $B(Y)$ when $Y(f) = Z(f) \cup \{v_2(f)\}$, $v_1(f) \in Z(f)$ has period 3, and

$$Z(f) = f^{-n}\{f^i(v_1) : i \geqslant 0\}$$

for $0 \leqslant n \leqslant 5$. Recall that we described the structure of $B(Y)$ in these cases in the Introduction. All our enhanced Levy sets will be of the form $[f, \Gamma]$ where f is critically finite. This means that $Y(f)$ can be regarded as a subset of the full orbit of the postcritical set of f, but $Y(f)$ may not be forward invariant under f. Let $Y'(f)$ denote the forward orbit of $Y(f)$. Then (f, Γ) will satisfy the Invariance, Levy and Maximal Reduced Conditions for $(f, Y'(f))$. By 3.2 (and 3.3-3.5 in general but 3.2 suffices in the present case), Γ is uniquely determined up to $Y'(f)$-preserving isotopy by f. So we shall simply write $[f]$ rather than $[f, \Gamma]$, remembering that for the moment, these are all either nodes or extreme edges. For this current definition, we can have $[f] = [g]$ for two critically finite branched coverings which are not Thurston equivalent and may even have postcritical sets of different cardinalities. In fact (since we are only dealing with primitive Enhanced Levy Sets) it is easy to generate Γ by considering the backward orbit of ∂P where P is the fixed set.

We use the notation for matings $s_q \amalg s_p$ as in 2.3 (recalled from 1.10 of [**R3**]). We shall also need *captures* originally introduced by Wittner [**W**]. I shall use a notation slightly adapted from [**R3**]. Given s_q, for an odd denominator rational q, suppose first that the endpoints of μ_q are not in the same orbit under $z \mapsto z^2$. Take q' in the full orbit of q under $x \mapsto 2x$ mod 1. Then $e^{2\pi iq'}$ is in the boundary of a unique gap G of the invariant lamination L_q in the full orbit of the central gap. Let $\alpha_{q'}$ be a simple path starting at ∞, crossing S^1 at $e^{2\pi iq'}$ into G and ending at the point x_0 in G which is in the full orbit of 0 under q. Write

$$\sigma_{\alpha_{q'}} = \sigma_{q'}.$$

Then $\sigma_{q'} \circ s_q$ is a *capture*. If the endpoints of μ_q lie in a single orbit, then $e^{2\pi iq'}$ will be in the boundary of two gaps in the full orbit of the central gap. We then simply use more points to uniquely determine the gap, and write, for example $\sigma_{q,q'}$ for a homeomorphism given by a path which crosses into a gap with $e^{2\pi iq'}$ and $e^{2\pi iq''}$ in its boundary (if this uniquely determines the gap).

It will also be necessary to consider matings in which the critical points are not both periodic. For example, if $e^{2\pi iq'}$, $e^{2\pi iq''}$ and $e^{2\pi iq'''}$ are the vertices of the preperiodic triangular minor gap T of the lamination $L_{q'}$ then we write $s_{q',q'',q'''}$ for a branched covering preserving $L_{q'}$ and with preperiodic critical point whose eventual period is the eventual period of T.

In all our examples, the fixed set P is either the closed loop $\mu_{1/3} \cup \mu_{1/3}^{-1}$, or (up to homotopy) $\Delta \cup \Delta^{-1}$, where Δ is the triangle with vertices at $e^{2\pi i(1/7)}$, $e^{2\pi i(2/7)}$, $e^{2\pi i(4/7)}$.

Case $n = 0$. — The 3 enhanced Levy sets, and the ordering, are:

$$[s_{3/7} \amalg s_{1/3}] \leqslant [s_{3/7} \amalg s_{1/3}] \quad \text{and} \quad [s_{1/7} \amalg s_{6/7}] = [s_{6/7} \amalg s_{1/7}].$$

(The last equality arises because $s_{1/7} \amalg s_{6/7}$ is a *shared mating*. There is a homeomorphism φ which fixes both critical orbits of $s_{1/7} \amalg s_{6/7}$, but for which $s_{1/7} \amalg s_{6/7} \simeq_\varphi s_{6/7} \amalg s_{1/7}$.) These are all extreme edges.

Case $n = 1$. — The ELS $[s_{1/3} \amalg s_{3/7}]$ is a node for for $n \geqslant 1$, and $E_2[s_{1/3} \amalg s_{3/7}]$ is periodic with degree two eventually fixed type reduced map space. The following enhanced Levy sets and orderings are added:

$$[s_{3/7} \amalg s_{1/3}] \leqslant [\tau \circ \sigma_{5/14} \circ s_{3/7}],$$

where τ is any integral Dehn twist round the loop $\mu_{1/3} \cup \mu_{1/3}^{-1}$. All integral twists do give distinct enhanced Levy sets. The ELS $[\tau \circ \sigma_{5/14} \circ s_{3/7}]$ is an extreme edge, and remains so for all $n \geqslant 1$.

Case $n = 2$. — The ELS $[s_{3/7} \amalg s_{3/7}]$ is a node for $n \geqslant 2$, $E_2[s_{3/7} \amalg s_{3/7}]$ is periodic and the corresponding reduced map space if of periodic degree two type. We add in one more extreme edge and ordering:

$$[s_{3/7} \amalg s_{3/7}] \leqslant [\sigma_{13/28} \circ s_{3/7}].$$

It will become clear in Chapter 4 why we do not need a Dehn twist in this example.

Case $n=3$. — The ELS $[s_{1/7} \amalg s_{6/7}]$ is a node with $E_2[s_{1/7} \amalg s_{6/7}]$ periodic, and the corresponding reduced map space is of periodic degree two type.The new ELS's and orderings are

$$[s_{3/7} \amalg s_{1/3}] \leqslant [\tau \circ \sigma_{19/56} \circ s_{3/7}], \quad [s_{3/7} \amalg s_{1/3}] \leqslant [\tau \circ \sigma_{23/56} \circ s_{3/7}],$$
$$[\sigma_{13/28} \circ s_{3/7}] \leqslant [\sigma_{27/56} \circ s_{3/7}],$$
$$[s_{1/7} \amalg s_{6/7}] \leqslant [\tau' \circ s_{1/7} \amalg s_{9/56,11/56,15/56}] \leqslant [\tau' \circ \sigma_{9/56,11/56} \circ s_{1/7}],$$
$$[\tau' \circ s_{1/7} \amalg s_{9/56,11/56,15/56}] \leqslant [\tau' \circ \sigma_{11/56,15/56} \circ s_{1/7}],$$

where τ' is any integral Dehn twist round $\mu_{1/7} \cup \mu_{6/7}^{-1}$, and τ is as in case $n=1$. All of the new ELS's are extreme edges, except for $[\tau' \circ s_{1/7} \amalg s_{9/56,11/56,15/56}]$, which are preperiodic homeomorphic. The ELS $[\sigma_{13/28} \circ s_{3/7}]$ is now a preperiodic capture, and $E_2[\sigma_{13/28} \circ s_{3/7}] \smallsetminus Z$ is a once-punctured annulus although it was a disc for $n=2$. It develops more topology as n increases. For $n=6$, for example, it will be a once-punctured four-holed sphere (since it maps under the degree two map $(s_{3/7} \amalg s_{3/7})^4$ to the annulus $E_2[s_{3/7} \amalg s_{3/7}]$ for $n=2$).

Case $n=4$. — It is time to stop listing all the new extreme edges. There are 5 new ones.

Case $n=5$. — We have 8 new extreme edges, and some of the previous new ones become preperiodic captures. We do not give a systematic list, but

$$[s_{3/7} \amalg s_{1/3}] \leqslant [s_{3/7} \amalg s_{13/31}] \leqslant [s_{3/7} \amalg s_{3/7}],$$
$$[s_{3/7} \amalg s_{13/31}] \leqslant [\sigma_{93/224} \circ s_{3/7}].$$

Here, $[s_{3/7} \amalg s_{13/31}]$ is of degree two eventually fixed type, while $[\sigma_{93/224} \circ s_{3/7}]$ is one of the extreme edges. As indicated by the ordering, $\Delta_2[s_{3/7} \amalg s_{3/7}]$ and $\Delta_2[\sigma_{93/224} \circ s_{3/7}]$ are in different components of $\overline{\mathbf{C}} \smallsetminus E_2[s_{3/7} \amalg s_{13/31}]$. This confirms that $E_2[s_{3/7} \amalg s_{13/31}]$ is not an annulus: it is, in fact, a 3-holed sphere. There are also tunings of $[s_{3/7} \amalg s_{1/3}]$ between $[s_{3/7} \amalg s_{1/3}]$ and $[s_{3/7} \amalg s_{13/31}]$, which we have not listed.

3.20. Examples in Period 4. — We consider only the case when $B = B(Y)$, and for $[f, Y(f)] \in B$, $Y(f) = Z(f) \cup \{v_2(f)\}$, where $Z(f)$ is the period four orbit of $v_1(f)$. Then the primitive node Enhanced Levy Sets, and the extreme edges, are

$$[s_{4/15} \amalg s_{6/7}] \leqslant [s_{4/15} \amalg s_{11/15}], \quad [s_{2/5} \amalg s_{1/3}], [s_{7/15} \amalg s_{1/3}] \leqslant [s_{7/15} \amalg s_{7/15}],$$
$$[\varphi \circ s_{1/15} \amalg s_{14/15}],$$

where φ is the identity off the four holed sphere P which is the fixed set for $s_{1/15} \amalg s_{14/15}$, and is either the identity or such that $\varphi \circ s_{1/15} \amalg s_{14/15} \mid P$ is pseudo-Anosov. All the pseudo-Anosov ones contribute to the topology of $B(Y)$, (without increasing Y to $f^{-n}Y$ for some $n > 0$), as we shall see.

CHAPTER 4

THE GROUP OF AN ENHANCED LEVY SET

4.1. Let B be of degree two periodic or eventually fixed type. In this chapter, we examine the subgroup structure of $G = \pi_1(B)$. We introduced the subgroup $G[f, \Gamma]$ associated to an Enhanced Levy Set in 3.6. We describe this subgroup in 4.5 below in terms of a topological space $B[f, \Gamma]$. It turns out that $G[f, \Gamma]$ is the fundamental group of a fibration over a covering space of $B[f, \Gamma]$, with torus fibres. We also examine, in 4.16, intersections between subgroups, arising from Edges (i.e. Enhanced Levy Sets satisfying the Edge Condition) joining Nodes. For such an intersection we obtain a handle between the Node topological spaces. Finally, we make some remarks about computing examples, and give some examples.

4.2. Characterization of $G[f, \Gamma]$. — In this chapter, it will be convenient to consider only (f, Γ) satisfying the Invariance, Levy and Maximal Reduced Conditions. It will also be convenient to choose f to preserve the point set $\cup\Gamma$ — not just up to isotopy. We may need to have parallel copies of some loops in Γ to achieve this. Throughout this chapter, we use the characterisation of G as a subgroup of $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ (1.11). Then we can characterize $G[f, \Gamma]$ as the set of $[\varphi] \in \mathrm{MG}(\overline{\mathbf{C}}, Y)$ such that:

φ preserves $\cup\Gamma$ (pointwise) and φ is isotopic to $f^{-1} \circ \varphi \circ f$ via an isotopy preserving $Z \cup (\cup\Gamma)$.

We can take the isotopy to preserve $\cup\Gamma$, because f does.

4.3. The spaces $B(f, \Gamma)$ and $(B, N)(f, \Gamma)$. — We continue with the standing hypotheses on (f, Γ). We describe the topological space $B[f, \Gamma] = B(f, \Gamma)$.

Periodic Case. — Suppose that $E_2(f, \Gamma)$ is periodic. Then $B(f, \Gamma)$ is simply the reduced map space $B(f, \Gamma, E_2)$. See 2.17.

Preperiodic Capture. — $B(f, \Gamma) = E_2(f, \Gamma) \setminus Z$.

Preperiodic Homeomorphic. — Let Δ be any gap, and let φ be a homeomorphism of Δ. Then we use $[\varphi \mid \Delta]$ to denote the isotopy class of $\varphi \mid \Delta$ modulo isotopies which leave $\partial\Delta$ and $\Delta \cap Y$ invariant. Let $\mathrm{MG}(\Delta, \Delta \cap Y)$ denote the group of such isotopy classes, and $\mathrm{PMG}(\Delta, \Delta \cap Y)$ denote the group of isotopy classes of homeomorphisms which fix $\Delta \cap Y$ and each component of $\partial\Delta$ (not necessarily pointwise).

First, suppose that B is degree two periodic type. Then $f \mid f^i E_2(f, \Gamma)$ is a homeomorphism for all $i \geqslant 0$. Let p be the eventual period of E_2 under f. Then we can regard the centraliser of $[f^p \mid f^i(E_2)]$ on $f^i E_2$ as a subgroup of $\mathrm{MG}(E_2, E_2 \cap Y)$. Then we let H be the intersection of this centraliser with $\mathrm{PMG}(E_2, E_2 \cap Y)$. We recall from 3.6 that the centraliser is free abelian. We can choose representatives in $\mathrm{Hom}_+(E_2, E_2 \cap Y)$ so that this group — which we continue to call H — is also abelian, using the classification of surface homeomorphisms [**F-L-P**]: we can take free generators φ_j of H $(1 \leqslant j \leqslant r)$ so that for $j \neq k$, the sets where φ_j and $\varphi_k \neq$ identity are disjoint. Then we take $B[f, \Gamma]$ to be the mapping torus of H on E_2: that is

$$B[f, \Gamma] = (E_2 \times \mathbf{R}^r)/\sim$$

where $\sim$ is the group orbit equivalence relation generated by:

$$(\varphi_j(z), \underline{t}) \sim (z, \underline{t} + \underline{e}_j)$$

where $\underline{e}_j \in \mathbf{R}^r$ is the vector with 1 in the j'th position and zeros elsewhere.

If B is degree two eventually fixed then the only difference is that $f^i : E_2 \to f^i E_2$ may be degree two, and $f^i E_2$ may be fixed and contain a fixed point z_0 of Z, in which case we have to consider $\mathrm{PMG}(f^i E_2, \{z_0\})$ and lift to the double cover $\mathrm{PMG}(E_2, \{z_1\})$, where z_1 is the preimage of z_0 in E_2.

Edge Condition. — If (f, Γ) satisfies the Edge Condition, $B(f, \Gamma)$ is a point

Note that, in all cases, $\pi_1(B[f, \Gamma])$ identifies with a subgroup of $\mathrm{PMG}(E_2, Y \cap E_2)$. In the preperiodic capture case, we identify $[\alpha] \in \pi_1(B[f, \Gamma]) = \pi_1(E_2 \setminus Z, v_2)$ with $[\sigma_\alpha]$. In the preperiodic homeomorphic case, $Y \cap E_2 = \{v_2\}$ or $\{z_1, v_2\}$, and the subgroup of $\mathrm{PMG}(E_2, Y \cap E_2)$ is a semidirect product of H with $\pi_1(E_2, v_2)$ (or $\pi_1(E_2 \setminus \{z_1\}, v_2)$). This makes sense even though H is initially described up to isotopies which can move v_2.

4.4. The groups H_i. — Excluding the forward orbit of E_2 if E_2 is preperiodic homeomorphic, let Δ_i $(1 \leqslant i \leqslant r)$ be a set of connected periodic homeomorphic gaps for (f, Γ) which includes precisely one gap from each periodic orbit. (Forget about the use of Δ_2 to denote a disc containing E_2 in the last chapter — we do not need it in this chapter.) Let n_i be the period of Δ_i under f, so that $f^{n_i} : \Delta_i \to \Delta_i$ is a homeomorphism. Then H_i is the intersection of the centraliser of $[f^{n_i}]$ with $\mathrm{PMG}(\Delta_i, \Delta_i \cap Y)$. Of course, $\Delta_i \cap Y$ is at most a single fixed noncritical point of Z. Then, as we have seen in 3.6, H_i is free abelian, and the rank of H_i can be computed in terms of the number of irreducible cycles of components of $f^{n_i} \mid \Delta_i$, and the number of pseudo-Anosov cycles.

4.5. Elementary Structure of $G[f,\Gamma]$. — *Let* $\Gamma_2=\Gamma_2(f,\Gamma)$ *as in 2.5. There is an injective homomorphism*

$$\Theta=\Theta_1\times\Theta_2\times\Theta_3: G[f,\Gamma]\longrightarrow\pi_1(B[f,\Gamma])\times\prod_{i=1}^{r}H_i\times\mathbf{Z}^{\Gamma_2}$$

with the following properties.

(1) a) $\operatorname{Ker}\Theta_1$ *is central in* G_Γ.

b) $\operatorname{Ker}(\Theta_1\times\Theta_2)$ *is a group of Dehn twists round loops of* Γ, *and is isomorphic under* Θ_3 *to* $\operatorname{Ker}(f^*-I)\cap\mathbf{Z}^\Gamma$, *where* f^* *is as in 2.4. The kernel is infinite cyclic if* (f,Γ) *is minimal nonempty, and the projection* $[\varphi]\mapsto\Theta_3([\varphi],\gamma)\in\mathbf{Z}$ *is an isomorphism on the kernel, for any loop* γ *in the boundary of the fixed set.*

(2) *The image of* Θ_1 *has at most finite-by-abelian index. More precisely we have the following.*

a) *It is of at most infinite-cyclic-by-finite index if* $E_2(f,\Gamma)$ *is periodic degree two.*

b) *It is of at least infinite cyclic index if* B *is periodic degree two and* $[f,\Gamma]$ *is minimal nonempty satisfying the Node Condition.*

c) *It is of finite index* ($\leqslant 3$) *if* $[f,\Gamma]$ *is a periodic capture.*

d) *If* $[f,\Gamma]$ *is preperiodic homeomorphic, then the image contains an index two subgroup of the normal subgroup* $\pi_1(E_2[f,\Gamma])$.

(3) *The image of* Θ_2 *is cyclic if* $[f,\Gamma]$ *is minimal nonempty pseudo-Anosov.*

Remark. — It does not seem to be possible to say anything more precise about the image of Θ_2, nor about $\operatorname{Ker}(\Theta_1)$ in general. See the discussion of an example in 4.17.

4.6. First step in the proof of Elementary Structure: Construction of Θ

Let $[\varphi]\in G$. We can assume without loss of generality that φ preserves $\cup\Gamma$ pointwise. Also, φ preserves Z, and must fix all periodic points of Z. So $\varphi(\Delta)=\Delta$ if Δ is any periodic gap, or if $\Delta=E_1$, E_2 or Δ_0'. Let Δ be any periodic homeomorphic gap. Then each component of $\overline{\mathbf{C}}\smallsetminus\Delta$ contains periodic points of Z. So $[\varphi\mid\Delta]\in\operatorname{PMG}(\Delta,\Delta\cap Y)$.

Let Δ be any connected periodic gap union, of period p under f, and let $\Delta'\subset\Delta$ be the set with $\partial\Delta\subset\partial\Delta'$ such that $f^p(\Delta')=\Delta$. Then our characterization of $G[f,\Gamma]$ in 4.2 implies that

$$\text{(1)}\qquad \varphi\mid\Delta' \text{ is isotopic to } f^{-p}\circ\varphi\circ f^p$$

modulo isotopies of Δ which are constant on $\Delta\cap Z$. Thus, we can regard $[\varphi\mid E_2]$ as an element of $\pi_1(B[f,\Gamma])$ (since this identifies with a subgroup of $\operatorname{MG}(E_2,Y\cap E_2)$).

Then we define

$$\Theta_1([\varphi])=[\varphi\mid E_2],$$
$$\Theta_2([\varphi])=([\varphi\mid\Delta_1],\ldots,[\varphi\mid\Delta_r]).$$

Let Y_1 denote the union of the periodic points of Z and v_2. These points are fixed by all elements of G. Let $\gamma \in \Gamma_2$. Fix two points of Y_1 on each side, giving 4 in all, and call this set $Y(\gamma)$: it does exist. Take any nonperipheral loop γ' in $\overline{\mathbf{C}} \setminus Y(\gamma)$ which cuts γ exactly twice. Orient the loops γ, γ', and fix the crossing where the angle between the directions of γ and γ' is in $(0,\pi)$. Take any $[\varphi] \in G(f,\Gamma)$. Then regarded as an element of $\mathrm{PMG}(\overline{\mathbf{C}}, Y(\gamma))$, φ fixes γ and twists γ' an integer number $\Theta_3([\varphi],\gamma)$ times round γ. This integer does depend on the choice of $Y(\gamma)$, but then not on the choice of γ' and on the choice of orientations of γ, γ' (although it does depend on the orientation on $\overline{\mathbf{C}}$). Define

$$\Theta_3([\varphi]) = (\Theta_3([\varphi],\gamma)).$$

Then $\Theta = \Theta_1 \times \Theta_2 \times \Theta_3$ has been defined.

4.7. Lemma. — *Let $\Delta = \Delta_0'$ or E_1 be periodic of period p, with $v_2(f) \notin f_*^i\Delta$, for any $i \geqslant 0$. Then if $\varphi \mid \Delta$ satisfies (1) of 4.6,*

$$[\varphi \mid \Delta] = [\text{identity} \mid \Delta].$$

Proof. — Note that, under these hypotheses, the reduced map space $B(f,\Gamma,\Delta)$ consists of critically finite maps of degree two. If $\Delta = E_1$, this reduced map space contains a polynomial, and branched coverings in this space admit no Thurston obstruction. If $\Delta = \Delta_0'$, the Maximal Reduced Condition implies that branched coverings in this space admit no Thurston Obstruction. In both cases, it follows that $\pi_1(B(f,\Gamma,\Delta))$ is trivial. If $\Delta = \Delta_0'$, $B(f,\Gamma,\Delta)$ might be a Thurston equivalence class with a Euclidean, rather than hyperbolic, orbifold. But that happens only in one special case previously referred to in 2.17. In that case we can identify $\overline{\mathbf{C}}$ with $\mathbf{R}^2/\sim$ (where $\sim$ denotes the relation $\underline{x} \sim \pm\underline{x} + \underline{m}$, and $\underline{m} \in \mathbf{Z}^2$) and maps in $B(f,\Gamma,\Delta)$ are equivalent to $g : [\underline{x}] \mapsto [A\underline{x} + \underline{b}]$ for a fixed $A \in GL(2,\mathbf{Z})$ with irrational eigenvalues and determinant 2 and $\underline{b}$ with $2\underline{b} = \underline{0} \bmod \mathbf{Z}^2$. There are no homeomorphisms of $\mathbf{R}^2/\mathbf{Z}^2$ which commute with the endomorphism given by A, even up to homotopy. But any element of the subgroup of $\mathrm{MG}(\overline{\mathbf{C}}, Y(g))$ which identifies with $\pi_1(B(f,\Gamma,\Delta))$ would have to lift to such a homeomorphism of the torus. So $\pi_1(B(f,\Gamma,\Delta))$ is trivial in this case also. Thus, in all cases, using the identification of $\pi_1(B(f,\Gamma,\Delta))$ with a subgroup of $\mathrm{MG}(\Delta, \Delta \cap Y)$, we have the result. □

4.8. Elementary Structure: identifying the kernel of Θ_1

Proof of 1a) of 4.5. — First, we need to show that Θ is injective. Let $[\varphi] \in \mathrm{Ker}(\Theta)$. Then φ fixes all boundaries of all periodic gaps, and in addition preserves E_1, E_2 and Δ_0'. Since $\varphi \in \mathrm{Ker}(\Theta)$, $[\varphi \mid \Delta_i]$ and $[\varphi \mid E_2]$ are trivial in $\mathrm{PMG}(\Delta_i)$ and $\mathrm{PMG}(E_2, E_2 \cap Y)$. By 4.7, if E_1 or Δ_0' is periodic and not in the forward orbit of E_2, $[\varphi \mid E_1]$ and $[\varphi \mid \Delta_0']$ are trivial in $\mathrm{MG}(E_1, E_1 \cap Z)$, $\mathrm{MG}(\Delta_0', \Delta_0' \cap Z)$. Using the characterization 4.2, we obtain that $\varphi \mid \Delta$ is isotopic to the identity for all periodic gaps Δ. By continuity it follows that φ fixes all gaps and loops of Γ. In particular,

φ fixes all boundary components of all gaps. Using 4.2 again, we deduce that $\varphi \mid \Delta$ is isotopic to the identity for all gaps Δ. Hence φ must be a Dehn twist composition round loops of Γ. Since $[\varphi] \in \mathrm{Ker}(\Theta_3)$, the twists round all loops of Γ_2 are zero. Hence, using the characterization of 4.2, the twists round all loops of Γ are trivial. If $[f, \Gamma]$ satisfies the Edge Condition and α is the simple loop (in either direction) based at v_2 in E_2, then we use the fact that σ_α is isotopic to a composition of Dehn twists round the components of ∂E_2. Hence $[\varphi] = [\text{identity}]$, as required.

Let $[\varphi] \in \mathrm{Ker}(\Theta_1)$ and $[\psi] \in G[f, \Gamma]$. Then $[\varphi]$ is the identity on the full orbit of E_2, Δ_0', E_1, and $[\varphi]$ and $[\psi]$ commute on the full orbit of any periodic homeomorphic gap — where we take isotopies which are constant on gap boundaries. Hence $[\varphi]$ and $[\psi]$ commute.

Proof of 1b) of 4.5. — The first part is obvious. The claims in the case of (f, Γ) minimal nonempty are proved in 3.11. □

4.9. The image of Θ. — Let $E_2[f, \Gamma] = \Delta_0'$ be periodic of period p. We need to exhibit a subgroup of $\pi_1(B[f, \Gamma])$ of at most finite-by-cyclic index which will be in the image of Θ_1. Recall that $\pi_1(B[f, \Gamma])$ is regarded as a subgroup of $\mathrm{MG}(\Delta_0', \Delta_0' \cap Y)$. Therefore, the elements are regarded as isotopy classes $[\varphi \mid \Delta_0']$. We define L to be the subgroup of elements of $\pi_1(B[f, \Gamma])$ satisfying (1) to (3) below. We shall then show that L is in the image of Θ_1.

The first property of an element of L is the simplest:

$$[\varphi \mid \Delta_0'] \in \mathrm{PMG}(\Delta_0', \Delta_0' \cap Y). \tag{1}$$

This property is even a little stronger than is needed. Since $\Delta_0' \cap Y$ is finite and $\partial \Delta_0'$ has only finitely many components, this is a finite index condition.

Now let $\gamma_0 = \partial \Delta_0 \subset \partial \Delta_0'$, and let $A_0 = A(f^{-1}E_2)$ be the annulus bounded by the two components of $f^{-1}\gamma_0$. The second defining property for an element of L is

(2) Composing φ with a Dehn twist round γ_0 if necessary, in A_0, $f^{-1} \circ \varphi \circ f$ is isotopic to the identity.

Suppose simply that $f^{-1} \circ \varphi \circ f$ is isotopic to τ^{2n}, some $n \in \mathbf{Z}$, where τ denotes a single Dehn twist round A_0. This is clearly an index two condition. Note, however, that if this holds, and τ_0 denotes a single Dehn twist round γ_0, then $f^{-1} \circ (\tau_0^{-n} \circ \varphi) \circ f$ is isotopic to the identity. So the representative $\tau_0^{-n} \circ \varphi$ satisfies (2). So (2) is an index two condition.

To give the third property, recall that there is a continuous map $\Phi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ which is injective except on $\overline{\mathbf{C}} \setminus \Delta_0'$, and that each component of $\overline{\mathbf{C}} \setminus \Delta_0'$ maps to a point in $Y(f, \Gamma, \Delta_0')$, while the points of $Y \cap \Delta_0'$ map to the remaining points of $Y(f, \Gamma, E_2)$. Let φ be a homeomorphism of $\partial \Delta_0'$ which fixes boundary components and is semiconjugate under Φ to an element of $\pi_1(B[f, \Gamma])$ (under the natural correspondence). Then we have:

$$\varphi \text{ is isotopic to } \tau \circ f^{-p} \circ \varphi \circ f^p \tag{3_τ}$$

via an isotopy which is constant on $Z \cap \Delta_0'$ *and* on $\partial\Delta_0'$ where τ is a composition of integral twists round all but one loop of $\partial\Delta_0'$, and a half integral twist round at most one loop γ_2 of $\partial\Delta_0'$ (which is at least a second preimage of γ_0). The loop γ_2 (if it exists) is in the preimage under f^p of the loop in $\partial\Delta_0'$ separating v_1 from Δ_0', if this exists. (Note that $f^{-p} \circ \varphi \circ f^p$ is initially only defined on Δ_0'', but can be extended to Δ_0' by mapping disc components of $\Delta_0' \setminus \Delta_0''$ as dictated by $f^{-p} \circ \varphi \circ f^p$ on their boundaries.) Then our third defining condition of $[\varphi \mid \Delta_0'] \in L$ is

(3) composing φ with integral twists round nonperiodic loops of $\partial\Delta_0'$ if necessary,

$$\varphi \text{ is isotopic to } f^{-p} \circ \varphi \circ f^p$$

via an isotopy which is constant on $Z \cap \Delta_0'$ *and* on $\partial\Delta_0'$.

Now γ_0 is fixed by f^p and all other loops in $\partial\Delta_0'$ are in the full orbit of γ_0. For any homeomorphism φ of Δ_0' fixing boundary components, (3_τ) must hold for some τ. The condition of τ having zero twist round γ_0 is clearly of at most cyclic index. Now suppose that τ has 0 twist round γ_0, and let (3_τ) hold for φ. Then τ corresponds naturally to a vector $\underline{m} = (m_\gamma) \in \mathbf{R}^{\Gamma_1}$, where Γ_1 is the set of components of $\partial\Delta_0'$, excluding γ_0, $m_\gamma \in \mathbf{Z}$ for $\gamma \neq \gamma_2$ and $m_{\gamma_2} \in \mathbf{Z}/2$. Then (3) holds for $\tau' \circ \varphi$, if τ' corresponds to the vector $\underline{n} \in \mathbf{Z}^r$ and $\underline{n}$ is a solution to

$$(A - I)\underline{n} = \underline{m},$$

where A is f^* restricted to $\mathbf{R}^{\Gamma_1}$. The matrix A is strictly triangular, with zero diagonal entries, and all entries are integers, except for one in the row indexed by γ_2, if γ_2 exists. Then the set of $\underline{m} \in \mathbf{Z}^{\Gamma_1}$ for which there is an integer solution is of finite index. So (3) defines a subgroup of at most cyclic-by-finite index.

4.10. Proof of the image property of Θ_1 when (f, Γ) is periodic degree two

The following lemma completes the proof of 2a) of 4.5.

Lemma. — *L is in the image of Θ_1.*

Proof. — Let $[\varphi \mid \Delta_0'] \in L$, that is $[\varphi \mid \Delta_0']$ satisfies the defining properties (1) to (3) of 4.9. Then we can choose the homeomorphism φ so that (2) and (3) of 4.9 are satisfied. This is possible, because for (2) we choose an appropriate Dehn twist round γ_0, and then given that, for (3) we choose an appropriate Dehn twist round the remaining loops. We can assume that φ is the identity on $\partial\Delta_0'$ by property (1). We are going to extend φ to fix all loops and gaps of Γ. We define φ inductively on the full orbit of Δ_0'. Suppose that φ has already been defined on the gap Δ, and Δ' is a gap which contains a component of $f^{-1}\Delta$ with $\partial\Delta' \subset \partial f^{-1}\Delta$. Then we define φ on $\Delta' \cap f^{-1}\Delta$ by

$$f \circ \varphi = \varphi \circ f$$

and extend φ to discs of $\Delta' \setminus f^{-1}\Delta$ if necessary. This uniquely determines φ on Δ' unless $\Delta = \Delta_0'$ when there are two choices: we make the choice which fixes all boundary components (rather than inducing a permutation of order two). Then φ fixes all boundary components of all these gaps, and we extend φ to be the identity elsewhere. By property (2), if Δ is a gap in the full orbit of Δ_0' and Δ' is a component of $f^{-1}\Delta$ such that all components of $\partial\Delta'$ are either trivial or homotopic to a single loop of Γ, then $f^{-1}\circ\varphi\circ f$ is the identity. Using this and property (3), we see that characterization 4.2 does hold for φ. So $[\varphi]\in G[f,\Gamma]$, and $[\varphi\mid\Delta_0']\in \mathrm{Im}(\Theta_1)$, as required. □

4.11. Proof of the Image Property of Θ_1 in the degree two periodic minimal nonempty case (2b) of 4.5. — Now let B be period two and let $[f,\Gamma]$ be minimal nonempty satisfying the Node Condition. We want to show that (3) of 4.9 defines a subgroup of $\pi_1(B[f,\Gamma])$ which is of infinite index. The gap E_2 is periodic, of some period p. Let α_1 be the simple nontrivial loop in $E_2\smallsetminus Z$ based at v_2 which is freely isotopic in $\overline{\mathbf{C}}\smallsetminus Z$ to $\gamma_0=\partial\Delta_0$. Let γ_1 be a simple loop in $E_2\smallsetminus Y$ such that α_1 lies in the annulus bounded by γ_0 and γ_1. Let τ_j denote Dehn twist round γ_j, appropriately orientated. Let $\alpha=\alpha_1*\alpha_1$. Then σ_α is isotopic via a Y preserving isotopy to $\tau_0^2\circ\tau_1^{-2}$. Of course this is isotopic to the identity via an isotopy preserving $\Delta_0'\cap Z$ and $\partial\Delta_0'$. But $f^{-p}\circ\sigma_\alpha\circ f^p$ is isotopic to $\tau_0^2\circ\tau_1^{-1}$, and hence to τ_0 via an isotopy preserving $\Delta_0'\cap Z$ and $\partial\Delta_0'$. Thus, σ_α^n satisfies (3_τ) of 4.9, where τ has a nonzero twist round γ_0, for any $n\neq 0$. If τ' is any twist round loops of Γ such that $f^{-p}\circ(\tau'\circ\sigma_\alpha^n)\circ f^p\mid\Delta_0'$ is isotopic to $\tau'\circ\sigma_\alpha^n\mid\Delta_0'$, then $\tau'\circ\sigma_\alpha^n$ also satisfies (3_τ). This is because (f,Γ) is minimal, and hence $f^{-p}(\Gamma\smallsetminus\partial P)\cap\partial P=\varnothing$. Thus, there is no choice of τ' such that $\tau'\circ\sigma_\alpha^n$ satisfies (3) of 4.9. It follows that $[\sigma_\alpha^n]\notin\mathrm{Im}(\Theta_1)$ for $n\neq 0$. Hence $\mathrm{Im}(\Theta_1)$ has infinite index in $\pi_1(B[f,\Gamma])$, as required.

4.12. Proof of the Image Property of Θ_1 in the Periodic Capture case (2c) of 4.5. — We can prove a stronger property this time, partly because E_2 is a disc. Take any element of $\pi_1(B(f,\Gamma,E_2))$ represented by φ on E_2. Let E_2 be of period p. Initially, we only have

$$f^{-p}\circ\varphi\circ f^p \text{ is isotopic to } \tau^n\circ\varphi \text{ on } E_2$$

via an isotopy fixing $E_2\cap Z$, where τ is a single Dehn twist round ∂E_2. However, f^p is degree four on ∂E_2. So now assume $n=3m$, and we can replace φ by $\tau^{-4m}\circ\varphi$ and hence eliminate τ^n. We can take φ equal to the identity on ∂E_2 and extend the definition of φ to the full orbit of E_2 as in the previous case. Since all components of the full orbit are discs, there are no awkward annuli. We then take φ to be the identity outside the full orbit of E_2. Then the characterization of 4.2 holds for φ, so $[\varphi]\in G$. So the image of Θ_1 is of index at most three. □

4.13. Proof of the Image Properties of Θ in the preperiodic capture cases (2d) of 4.5. — Take any closed loop in $E_2\smallsetminus Z$ based at v_2 which lifts to a closed loop in $f^{-1}E_2$. Then σ_α defines a homeomorphism of E_2 which is isotopic to the identity via Z-preserving isotopy, and lifts under f to another such homeomorphism of $f^{-1}E_2$. We can take σ_α to be the identity outside E_2, and then $\Theta_1([\sigma_\alpha]) = [\alpha]$, as required. □

4.14. Proof of the Image Property of Θ_2 in the Minimal Nonempty Pseudo-Anosov case (3 of 4.5). — Let P be the fixed set for (f, Γ) which is minimal nonempty pseudo-Anosov. Then there is n such that for any component of $f^{-(n+1)}P$ which is not in $f^{-n}P$, all boundary loops are trivial. (This is not true for a general (f, Γ), when components can degenerate to loops of Γ.) Take any element $[\varphi]$ of $\mathrm{PMG}(P)$ in the centraliser of $f \mid P$, and define φ on the full orbit of P so that $f \circ \varphi = \varphi \circ f$, fixing all components. Then extend φ to be the identity elsewhere, and $\Theta_2([\varphi]) = [\varphi]$ as required. □

4.15. Edge Subgroups and Adjacent Nodes. — Let $[f, \Gamma]$ satisfy the Edge Condition and let $[f, \Gamma']$ be an immediate successor or immediate predecessor node, using the ordering of 3.16. Then, replacing (f, Γ) by an equivalent loop set if necessary, we can assume that $\Gamma' \subset \Gamma$, and therefore $G[f, \Gamma] \subset G[f, \Gamma']$. We can also assume that (f, Γ) and (f, Γ') satisfy the Maximal Reduced Condition. Now we define a set $C[f, \Gamma, \Gamma'] \subset B[f, \Gamma']$.

Periodic Case. — Suppose that $E_2(f, \Gamma')$ is periodic of period p. Then we can use Γ' to define a critically finite branched covering g_0 as follows. Let

$$\Omega = E_2(f, \Gamma') \smallsetminus E_2(f, \Gamma),$$

and let Ω' be the component of $f^{-p}\Omega$ in $E_2(f, \Gamma')$. Then let $g_0 = f^p$ on Ω' and map the disc components of $\overline{\mathbf{C}} \smallsetminus \Omega'$ to components of $\overline{\mathbf{C}} \smallsetminus \Omega$ as indicated by the map on their boundaries, so that the following holds for a finite set $Z(g_0)$ which is the union of $Z(f) \cap E_2(f, \Gamma')$ and one point in each component of $\overline{\mathbf{C}} \smallsetminus \Omega$: $Z(g_0)$ is invariant under g_0 and contains all critical values of g_0. If $[f, \Gamma']$ is a successor of $[f, \Gamma]$ then g_0 is simply the reduced branched covering of (f, Γ) for $\Delta_0'(f, \Gamma)$ and is a polynomial of degree two or four. If $[f, \Gamma']$ is a predecessor of $[f, \Gamma]$ then, naming as $v_2(g_0)$ the critical value of g_0 in the same component of $\overline{\mathbf{C}} \smallsetminus \Omega$ as $v_2(f)$, $v_2(g_0)$ is eventually fixed.

Then $B(g_0, Y(g_0))$ is disjoint from $B(f, \Gamma', \Delta_0')$, but $B(g_0, Y(g_0)) \cup B(f, \Gamma', \Delta_0')$ is connected and locally contractible. Let $C[f, \Gamma, \Gamma']$ be the intersection with $B(f, \Gamma', \Delta_0')$ of a tubular neighbourhood of $B(g_0, Y(g_0))$ in $B(g_0, Y(g_0)) \cup B(f, \Gamma', \Delta_0')$. The construction of such a tubular neighbourhood is given in 1.10.

Preperiodic Capture. — $C[f, \Gamma, \Gamma'] = E_2(f, \Gamma) \subset E_2(f, \Gamma')$.

Preperiodic homeomorphic. — Recall that $B(f, \Gamma')$ is a fibration over a torus with fibre $E_2(f, \Gamma')$. Then $C[f, \Gamma, \Gamma']$ is the subspace fibration with fibre $E_2(f, \Gamma)$.

4.16. Proposition. — *The image $\Theta_1(G[f, \Gamma])$ is in $\pi_1(C[f, \Gamma, \Gamma'])$.*

Remark. — The Topographer's View implies that the inclusion of $C[f, \Gamma, \Gamma']$ in $B[f, \Gamma']$ is injective on π_1.

Proof of Proposition. — This is basically definition-chasing. Take a path (f_t, Γ_t) realising an element of $G[f, \Gamma]$. Let $\Gamma'_t \subset \Gamma_t$ be the corresponding path through loop sets starting and ending at Γ'. We can assume that f_t preserves Γ_t and Γ'_t (not just up to isotopy). The periodic case is obviously the most substantial. We consider the case when $E_2(f, \Gamma')$ is periodic of period p degree two. We think about our explicit construction of the tubular neighbourhood in 1.10. Let Φ be the map which collapses components of $\overline{\mathbf{C}} \setminus E_2(f, \Gamma')$ to form the reduced map space for $E_2(f, \Gamma')$. We can embed Φ in a closed path of maps Φ_t where Φ_t collapses components of $\overline{\mathbf{C}} \setminus E_2(f_t, \Gamma'_t)$. We can also assume Φ_t maps $E_2(f_t, \Gamma_t)$ close to the corresponding collapsed point. If $[f, \Gamma']$ is a successor of $[f, \Gamma]$, then the annulus $E_2(f, \Gamma)$ is periodic, as is $E_2(f_t, \Gamma_t)$ and $E_2(f_t, \Gamma_t)$ is equal to a component of $f_t^{-p}(E_2(f_t, \Gamma_t))$. This gives us a path of the reduced branched coverings in the tubular neighbourhood that we constructed. □

4.17. Computation of Examples. — The statement of the Elementary Structure Theorem is a little vague. In many examples, the following holds:

(1) $\Theta_1 \times \Theta_2$ is injective, and the image is of finite index.

We have seen that this fails to be true — but in an orderly way — if $[f, \Gamma]$ is minimal non-empty. I initially hoped that it would be possible to characterize those $[f, \Gamma]$ for which (1) fails to be true, but this hope now seems ill-founded. To see why, it is necessary to discuss the computation of examples.

First, we consider the computation of $\mathrm{Ker}(\Theta_1 \times \Theta_2)$ for a node $[f, \Gamma]$. A Dehn twist composition σ is in the kernel if and only if

$$f^{-1} \circ \sigma \circ f = \sigma. \qquad (2)$$

This is a straightforward linear equation in variables indexed by Γ. In fact, we obtain the kernel up to finite index by considering Dehn twists round loops of Γ_2, where $\Gamma_2 = \Gamma_2(f, \Gamma)$ is as in 2.5. In fact, one can reduce further, considering loops to be equivalent if they are not separated by a periodic gap. Hence we obtain annuli R_i $(1 \leqslant i \leqslant m)$. Then let $A = (a_{i,j})$ where $a_{i,j}$ is the number of components of $f^{-1}R_j$ in R_i (assuming f leaves Γ pointwise invariant). Then a solution to (2) corresponds to an integral solution $\underline{x}$ to

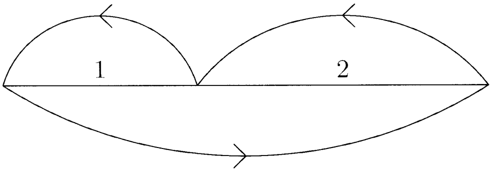
$$(A - I)\underline{x} = \underline{0}. \qquad (3)$$

This is precisely the equation we considered (with $\Gamma_2 = \partial P$) in 3.8. Of course, kernel and image are related in linear algebra and $(A - I)\mathbf{Z}^m$ is of finite index in $\mathbf{Z}^m$ if and only if the $\underline{x} = \underline{0}$ is the only solution to (3). Further, $(A - I)\mathbf{Z}^m = \mathbf{Z}^m$ if and only if $\det(A - I) = \pm 1$. These statements imply, correspondingly, that $\mathrm{Im}\,\Theta_1 \times \Theta_2$ is of finite index and that $\Theta_1 \times \Theta_2$ is surjective. The converse implication is not clear, but in examples computation of $\det(A-I)$ plus some auxiliary calculation leads to a description of $G[f,\Gamma]$

There is a neat and quick way to compute $\det(A-I)$. First, we need a quick computation of A. We have such a computation for the examples $[s_p \amalg s_q]$ that we considered at the end of Chapter 3 for $B_{3,n}$ for various $n \geqslant 0$. To compute A, take r with $\mu_r = \mu_p \wedge \mu_{1-q}$ (as in 2.3), and consider the finite set $s_r^k(\mu_r)$ $(k \geqslant 0)$. If μ_r is minimal, equivalence classes of loops of Γ_2 correspond to the leaves $s_r^k\mu_r$ Otherwise, equivalence classes of loops of Γ_2 correspond to rectangular regions of $D \setminus (\cup_k s_r^k \mu_r)$, that is, with two bounding leaves. If these are numbered R_i then $a_{i,j}$ is the number of components of $s_r^{-1}(R_j) \cap R_i$. In fact, this is always either 0 or 1. For the explicit examples we consider, $\mu_{1/3} \leqslant \mu_r \leqslant \mu_{1/2}$, μ_r joins $e^{2\pi i r}$ and $e^{2\pi i(1-r)}$, and $s_r^k\mu_r$ joins $e^{2\pi i x}$ and $e^{2\pi i(1-x)}$ where $x = 2^k r \bmod 1$. We can project these leaves to points on an interval, and the rectangles between them to intervals between the points. In the examples below, we draw $\{y : 0 \leqslant y \leqslant 1/2\}$ with $1/2$ to the *left* of 0, because we are considering this interval to be the projection of the circle $\{e^{2\pi i y} : 0 \leqslant y \leqslant 1/2\}$. The orbit of μ_r then becomes the orbit of the critical value of a unimodal interval map. The critical value is the leftmost point, and the image of the critical value is the rightmost point. We number the intervals I_i $(1 \leqslant i \leqslant n)$ from right to left. We write $i \to j \cup \cdots \cup j+k$, or $i \to \ell$ for $j \leqslant \ell \leqslant j+k$, if the image of I_i is $I_j \cup \cdots \cup I_{j+k}$ This is equivalent to $a_{i,j} = \cdots = a_{i,j+k} = 1$.

Example 1. — Consider $[s_{3/7} \amalg s_{3/7}]$ or any capture $[\sigma_p \circ s_{3/7}]$ with $3/7 < p < 4/7$. Then $r = 3/7$. This gives orbits

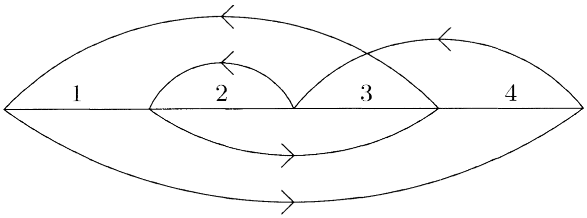
$$3/7 \to 6/7 \to 5/7 \to 3/7$$
$$4/7 \to 1/7 \to 2/7 \to 4/7$$



$r = 3/7$.
$1 \to 1 \cup 2,\ 2 \to 1.$

Example 2. — Consider $[s_{3/7} \amalg s_{13/31}]$ (considered in 3.19 in the period 3, $n = 5$ case). Then $r = 13/31$. This gives orbits

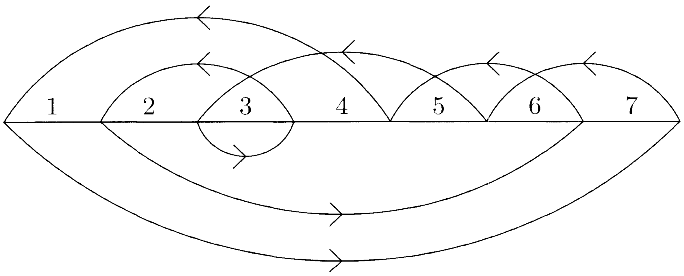
$$13/31 \to 26/31 \to 21/31 \to 11/31 \to 22/31 \to 13/31$$
$$18/31 \to 5/31 \to 10/31 \to 20/31 \to 9/31 \to 18/31$$



$r = 13/31$.
$1 \to 4,\ 2 \to 2 \cup 3,\ 3 \to 1,\ 4 \to 1 \cup 2.$

Example 3. — Consider $[s_{7/15} \amalg s_{116/255}]$. Then $r = 116/255$ (which has period 8). This is an extension of the examples of 3.20 — and would occur in the period 4 case for n sufficiently large. It is *primitive* in the sense of 2.16. We omit the denominator 255 below.

$$116 \to 232 \to 209 \to 163 \to 71 \to 142 \to 29 \to 58 \to 116$$
$$139 \to 23 \to 46 \to 92 \to 184 \to 113 \to 226 \to 193 \to 139$$



$r = 116/255$.
$1 \to 7,\ 2 \to 4 \cup 5 \cup 6,\ 3 \to 2 \cup 3,\ 4 \to 1,\ 5 \to 1 \cup 2,\ 6 \to 3 \cup 4,\ 7 \to 5.$

Now we give a formula for $\det(A - I)$ which is simply derived from the usual definition of determinant in terms of permutations of indices. Our formula actually gives $\det(B)$, where B is any matrix which has only 0's and 1's off the diagonal and only 0's and -1's on the diagonal. For this we need to define a set of permutations $\Sigma(B)$ of $\{i : 1 \leqslant i \leqslant m\}$. The condition for σ to be in $\Sigma(B)$ is

$$b_{i,\sigma(i)} \neq 0.$$

In general, it is possible for the set $\Sigma(B)$ to be empty. If $B = A - I$ for A as above, then $b_{i,j} \neq 0$ equivalent to: exactly one of $i = j$, $i \to j$ holds. We remark that for the above examples there is precisely one t with $t \to t$, and this always happens if $\mu_r \geqslant \mu_{1/3}$ — but this is a peculiarity of the leaf $\mu_{1/3}$. If μ_r is $\geqslant$ any other minimal leaf, there is no such t. Note that $\Sigma(A - I)$ is always nonempty. In general, let n_s be the number of elements of $\Sigma(B)$ with s disjoint nontrivial cycles. Then our formula is

$$\det(B) = (-1)^m \sum_s (-1)^s n_s.$$

The sum is interpreted as 0 if $\Sigma(B) = \varnothing$.

Example 1. — We have $t = 1$ and the only element of $\Sigma(A-I)$ is the cycle $2 \to 1 \to 2$, and $\det(A-I) = -1$ (as we can easily check, of course). As in all three examples, there are no groups H_i, so Θ_2 is trivial and $G[s_{3/7} \amalg s_{3/7}]$ is isomorphic to $\pi_1(B[s_{3/7} \amalg s_{3/7}])$ (for $n \geqslant 2$).

If we take one of the captures $[\sigma_p \circ s_{3/7}]$ with $3/7 < p < 4/7$, then we can check that $G[\sigma_p \circ s_{3/7}]$ is isomorphic to $\pi_1(E_2[\sigma_p \circ s_{3/7}])$. This actually has nothing to do with $\det(A-I)$. It has more to do with the fact that, if $f = \sigma_p \circ s_{3/7}$ then only two components of $\overline{\mathbf{C}} \setminus f^{-1}E_2$ intersect Z.

Example 2. — We have $t = 2$. The only nontrivial cycle containing 2 is $2 \to 3 \to 1 \to 4 \to 2$. So this is the only permutation in $\Sigma(A-I)$, and $\det(A-I) = 1$. So, again $G[s_{3/7} \amalg s_{13/31}]$ is isomorphic to $\pi_1(B[s_{3/7} \amalg s_{13/31}])$.

Example 3. — We have $t = 3$. The only nontrivial cycle containing 3 is $2 \to 6 \to 3 \to 2$. The only cycle disjoint from this is $1 \to 7 \to 5 \to 1$, which is included in one of the two elements of $\Sigma(A-I)$ and not in the other. So $n_1 = n_2 = 1$ and $\det(A-I) = 1 - 1 = 0$. We can easily check that the kernel has dimension 1 by showing that the determinant of the matrix obtained by replacing some column of $A-I$ by another column is non-zero. If we can insert the column with ± 1 in the last entry and 0's elsewhere then this shows that the image of Θ_1 is of infinite-cyclic by finite index — and the image has infinite cyclic index if the determinant is ± 1. If we replace the seventh column of $A-I$, and call this matrix B, then elements of $\Sigma(B)$ must fix 7, and otherwise the rules are the same as for $\Sigma(A-I)$: there is then one element of $\Sigma(B)$, with nontrivial cycle $2 \to 6 \to 3 \to 2$, and so the determinant of B is 1. So $\mathrm{Ker}(\Theta_1)$ is one-dimensional and $\mathrm{Im}(\Theta_1)$ is of infinite cyclic index.

There are also examples with similar features to this one of the form $[s_{3/7} \amalg s_p]$, which occur in the period 3 case for n sufficiently large, but higher period of p is required. I found one of period 12. The nature of the example is similar to Example 3 in the following sense (although Example 3 was originally found by trial and error). The orbit determined by 116/255 can be constructed by taking the standard period 3 orbit — generated by 3/7 — taking a period two tuning of this, and then giving an extra "twist" to destroy the tuning but keep the property of two disjoint cycles of intervals. Similarly, the period 12 orbit can be constructed by taking the period 5 orbit determined by 13/31, taking a period two tuning of this and then giving an extra "twist" to destroy the tuning.

CHAPTER 5

GRAPHS OF TOPOLOGICAL SPACES AND THE TOPOGRAPHER'S AND RESIDENT'S VIEWS

5.1. In this chapter, we state the Topographer's and Resident's Views concerning the topological structure of the space $B(Y, f_0)$ of branched coverings, where $B = B(Y, f_0)$ is of polynomial of degree two periodic or eventually fixed type. More generally, in order to carry out an inductive process, we describe the topological structure of the pair (B, N) (see 1.10), where N is a deleted tubular neighbourhood of finitely many spaces of critically finite branched coverings with one critical value eventually fixed. From now on in this chapter, except in the statement of the Polynomial Type Theorem in 5.9, we take $B(Y, f_0)$ to be of degree two type.

5.2. Some graphs and their duals. — A topological space S is a *graph of topological spaces* if there is a finite graph G such that the following hold.

(1) To each edge node ν, and edge ϵ of G, there are associated closed subspaces S_ν and S_ϵ of S.

(2) There is a closed equivalence relation $\sim$ on

$$\cup_\nu\big((S_\nu \times \{\nu\}) \cup (\cup_\epsilon S_\epsilon \times \{\epsilon\})\big)$$

generated by relations of the form $(x, \nu) \sim (f_{\nu\cdot\epsilon}(x), \epsilon)$, where ν and ϵ are an adjacent node and edge and $f_{\nu,\epsilon}$ is a homeomorphism from a closed subset of S_ν to a closed subset of S_ϵ.

(3) The space S is homeomorphic to the quotient space

$$(\cup_\nu(S_\nu \times \{\nu\}) \cup (\cup_\epsilon S_\epsilon \times \{\epsilon\}))/\sim .$$

Up to homotopy equivalence, S is also the space

$$S_1 = (\cup_\nu(S_\nu \times \{\nu\}) \cup (\cup_\epsilon S_\epsilon \times [0,1] \times \{\epsilon\}))/\sim_1,$$

where, if we call the nodes attached to ϵ ν_0, ν_1 (in whichever order we like) then $\sim_1$ is such that $(x, \nu_j) \sim_1 (f_{\nu,\epsilon}(x), j, \epsilon)$ for $j = 0$, 1. A possible advantage of the space S_1 is that there is a continuous map from S_1 onto G which maps each space $S_\nu \times \{\nu\}$ to ν and each space $S_\epsilon \times (0,1) \times \{\epsilon\}$ to ϵ.

A graph of topological spaces is a useful way of describing T if the sets S_ν and S_ϵ have a particular structure. There is no distinction, in the definition above, between the forms that can be taken by the topological spaces S_ν and S_ϵ for nodes and edges ν and ϵ.

There are two complementary views of the topological space $B = B(Y, f_0)$. The topographer's view is that $B(Y, f_0)$ is homotopic to an increasing union of graphs $\mathcal{G}_n(B)$ of topological spaces, each with a clearly recognizable geometric structure. The resident's view is that the universal cover $\widetilde{B}$ of $B(Y, f_0)$ is a union of pieces projecting down to sets in a partition $\mathcal{P}(B)$ of the resident's space — which is the unit disc D, regarded as the universal cover of $\overline{\mathbf{C}} \setminus Z$. Each set in the partition is the geodesic-convex hull of a subset of ∂D (using the Poincaré metric). We need to describe $\mathcal{P}(B)$, and to obtain the graphs from it. In order to do this, we shall use the pairs (f_0, Γ) examined in Chapters 2-4.

In Chapter 1, we described an action of $G = \pi_1(B, f_0)$ on ∂D. This clearly gives an action on the set of geodesic-convex hulls of subsets of ∂D. We shall see that $\mathcal{P}(B)$ is invariant under the action of G. We shall then take coarsenings $\mathcal{P}_n(B)$ $(n \geqslant 0)$ on various subsets, and dual graphs $\widetilde{\mathcal{G}}_n(B)$. We then define

$$\mathcal{G}_n(B) = \widetilde{\mathcal{G}}_n(B)/G,$$

with some nodes removed, as described below in 5.6: after quotienting by the G-action, only two edges meet at these nodes, so it makes sense to remove them. The tree $\widetilde{\mathcal{G}}_n(B)$ is not the universal cover of $\mathcal{G}_n(B)$, because the stabilizer in G of a node of $\widetilde{\mathcal{G}}_n(B)$ is always nontrivial, and the stabilizer of an edge is often nontrivial.

5.3. The Partition $\mathcal{P}(B)$. — By partition, we mean that the interior of a set in the partition is disjoint from any other set: some of the sets in the partition are closures of single geodesics, or points on ∂D. We also mean that the union includes the whole open unit disc. The partition may not be locally finite: it is likely that sets in the partition accumulate on others. The sets in $\mathcal{P}(B)$ include all closed sets $C(f_0, \Gamma)$ (3.13) where (f_0, Γ) satisfies the Invariance and Levy Conditions and is minimal nonempty or primitive (see 2.16), or (f_0, Γ) satisfies the Invariance Condition and is extreme (2.3.4). We have seen in 3.14 that the interior of any such set $C(f_0, \Gamma)$ is disjoint from any other such set. The remaining sets in $\mathcal{P}(B)$ are the closures of complementary components. The collection $\mathcal{P}(B)$ of sets is then clearly invariant under G. We shall see in the Resident's View Theorem that these complementary components are in natural one-to-one correspondence with components $\widetilde{V}_1$ in $\widetilde{B}$ of the preimage of the set $V \subset B$ of rational maps. We shall also see in 5.5 that there is a natural way to define the component $C(\widetilde{V}_1)$ from $\widetilde{V}_1$, although it is not at all obvious that is then a complementary component of the union of sets $C(f_0, \Gamma)$.

5.4. Putative Nodes and Edges of the covering graphs: Singletons and Pairs. — At least some of the putative nodes and edges of our graphs $\widetilde{G}_n(B)$ can be defined independently of the partition $\mathcal{P}(B)$, as follows.

We shall refer to (f_0, Γ) satisfying the Invariance, Levy and Maximal Reduced Conditions as a *pair*. Let Γ' be obtained from Γ by adding in loops to the full orbit of the fixed gap P of Γ, so that the fixed gap P' of Γ' is irreducible. Let $\psi_{P'}$ be the homeomorphism of P' as in 2.13. If $\psi_{P'}$ is isotopic to an isometry, define $\kappa(f, \Gamma) = 0$. Otherwise, $\psi_{P'}$ is isotopic to a pseudo-Anosov with invariant transverse measured foliations, which are expanded and contracted by $\psi_{P'}$ by $\lambda^{\pm 1}$ for some $\lambda > 1$. Then let $\kappa(f_0, \Gamma) = \log \lambda$. The possible numbers $\kappa(f_0, \Gamma)$ (for $f \in B$) form an increasing discrete sequence $\{\kappa_n\}_{n \geqslant 0}$ with $\kappa_0 = 0$. If $\#(Z) \leqslant 3$, then $\kappa = 0$ is the only possibility. Otherwise, the sequence $\{\kappa_n\}_{n \geqslant 0}$ is infinite and unbounded. If (f_0, Γ) is minimal nonempty or primitive, and $\kappa(f_0, \Gamma) = \kappa_n$, then (f_0, Γ) is a *putative node or edge* of the graph $\widetilde{\mathcal{G}}_n(B)$. The alternatives node or edge depend on whether or not it satisfies the Node or Edge Condition, except in the case of minimal nonempty pseudo-Anosov edge pairs (f_0, Γ), which are putative nodes of $\widetilde{\mathcal{G}}_n(B)$. This is because the associated convex hull $C(f_0, \Gamma)$ has nonempty interior. These give all the edges of $\widetilde{\mathcal{G}}_n(B)$, and some of the nodes.

We now introduce *singletons*, which give the remaining nodes of $\widetilde{\mathcal{G}}_0(B)$. Let $\widetilde{B}$ denote the universal cover of B, and V the set of rational maps in B. *Singletons* are simply the components $\widetilde{V}_1$ of the preimage of V in $\widetilde{B}$. In future, we shall tend to use $\widetilde{V}$ to denote the full preimage of V in $\widetilde{B}$.

5.5. Convex hulls $C(\widetilde{V}_1)$. — We are now going to make a connection between the singletons and the remaining sets in the partition $\mathcal{P}(B)$. For each singleton $\widetilde{V}_1$, we are now going to define a subset $C(\widetilde{V}_1)$ of the closed unit disc D, which will be the geodesic-convex-hull of its intersection $\partial C(\widetilde{V}_1) \cap \partial D$ with ∂D. So we only need to define $\partial C(\widetilde{V}_1) \cap \partial D$.

Recall from 1.10 that there is an associated set of ends $N \subset B = B(Y, f)$, such that $\overline{N} \smallsetminus N$ consists of critically finite maps. We need to know that the space of rational maps V_1 covered by $\widetilde{V}_1$ intersects at least one of the components of N. Using the notation of 1.9, this is equivalent to: $\overline{V}$ contains at least one map g with critical point $c_2(g)$ in the forward orbit of $c_1(g)$, where V is a component of $V_{p,0}$, $V_{p,1,0}$ or $V_{ef,m,0}$. Something similar was proved in 2.4 of [**R3**], where we showed that $\overline{V}$ contained a polynomial. We can use exactly the same idea here. Our standard normalisation is to regard V as a space of rational maps

$$g_{c,d} : z \longmapsto 1 + \frac{c}{z} + \frac{d}{z^2}$$

satisfying a polynomial equation in c and d. This normalisation gives $c_1(g_{c,d}) = 0$ and $v_1(g_{c,d}) = \infty$. If $v_1(g_{c,d})$ is never a critical point for $(c, d) \in V$, then the function

$$\lim_{z \to \infty} z^2 g'_{c,d}(z) = -c$$

never vanishes on V. Then $(c, d) \mapsto 1/c$ is bounded and holomorphic on V, and hence constant. So there is c_0 such that $c = c_0$ on V, and d varies freely on V. Taking $d = -1 - c_0$ gives an element of V with 0 of period 3, which means that V must consist of maps with 0 of period 3. This is clearly not true with $c = c_0$ and d varying freely. So we have obtained the required contradiction.

By the definition of universal covers, $\widetilde{V}_1$ corresponds to a homotopy class of some path γ from f $(B = B(Y, f_0))$ to V_1, up to homotopies keeping first endpoint at f_0 and second endpoint in V_1. Fix such a path γ, with second endpoint f_1. Then $\pi_1(V_1, V_1 \cap N, f_1) \neq \varnothing$. Then define

$$\partial C(\widetilde{V}_1) \cap \partial D = \overline{\{\rho_2([\gamma * \alpha]) : [\alpha] \in \pi_1(V_1, V_1 \cap N, f_1)\}}.$$

This definition is clearly independent of the precise definition of γ

We shall refer to the sets $C(\widetilde{V}_1)$ as *rational convex hulls*. It is not clear that the set $C(\widetilde{V}_1)$ is a single set in the partition $\mathcal{P}(B)$, nor that the union of all sets $C(f_0, \Gamma)$ and $C(\widetilde{V}_1)$ covers the interior of D, but this is part of our statement of results, and will be proved. Note that, in analogy, lifts of cusps are dense in the limit set of a Fuchsian group. Note also that $\partial C(\widetilde{V}_1) \cap \partial D$ is invariant under the action of the image of $\pi_1(V_1, f_0)$ in $G = \pi_1(B, f_0)$. (This action was defined in 1.13.)

5.6. The partitions and the graphs. — We define partitions $\mathcal{P}_n(B)$ of subsets of the disc inductively, then take $\widetilde{\mathcal{G}}_n(B)$ to be the dual graph (actually a union of trees) and $\mathcal{G}_n(B)$ to be the quotient of $\widetilde{\mathcal{G}}_n(B)$ by the G-action (with one minor modification). We take

$$\mathcal{P}_0(B) = \{C(\widetilde{V}_1) : \widetilde{V}_1 \text{ singleton}\} \\ \cup \{C(f_0, \Gamma) : (f_0, \Gamma) \text{ putative node or edge of } \widetilde{\mathcal{G}}_0(B)\}.$$

This is clearly G-invariant. Then we define $\mathcal{P}_n(B)$ inductively by

$$\mathcal{P}_n(B) = \{C(f_0, \Gamma) : (f_0, \Gamma) \text{ putative node or edge of } \widetilde{\mathcal{G}}_n(B)\} \\ \cup \{U : U \text{ is a component of } \cup_{0 \leqslant i < n} \mathcal{P}_i(B)\}.$$

We may as well call such components U *singleton*. The graph $\mathcal{G}_n(B)$ is $\widetilde{\mathcal{G}}_n(B)/G$ with the minor modification that as follows. If (f_0, Γ) is pseudo-Anosov satisfying the Edge Condition, with $\kappa_n = \kappa(f_0, \Gamma)$, then it is a *node* of $\widetilde{\mathcal{G}}_n(B)$, because the corresponding set $C(f_0, \Gamma)$ has nonempty interior. But after quotienting by the G-action, only two edges meet at this node. So in $\mathcal{P}G(B)$, $\mathcal{G}_n(B)$, we delete this node, and denote the resulting edge by $[f_0, \Gamma]$.

Part of our statement of the Resident's View (in 5.10) is that $\mathcal{P}_n(B)$ is a locally finite partition of any component of $\cup \mathcal{P}_n(B)$, and hence $\widetilde{\mathcal{G}}_n(B)$ is a locally finite graph.

5.7. The Topological Spaces associated to Nodes and Edges. — We describe topological spaces associated to the singletons and pairs and which are the nodes and edges of the graphs $\mathcal{G}_n(B)$. Let $\nu = [\widetilde{V}_1]$ be a singleton node of $\widetilde{\mathcal{G}}_0$, where $\widetilde{V}_1$ projects to a component V_1 of rational maps in B. Then $S_\nu = V_1$.

Let $\nu = [f, \Gamma]$ be a node or edge of $\mathcal{G}_n$ for some $n \geqslant 0$. The definition of the topological space S_ν is suggested by the Elementary Structure Theorem 4.5, which gives an injective homomorphism Θ from $G[f_0, \Gamma]$ to $\pi_1(B(f_0, \Gamma)) \times \mathbf{Z}^q$ $(q \geqslant 0)$, where we can obviously choose q so that the projection of $\mathrm{Im}(\Theta)$ to $\mathbf{Z}^q$ has finite index. We choose S_ν to be the corresponding covering space of $B(f_0, \Gamma) \times T^q$, where T^q denotes the q-dimensional torus. The Elementary Structure Theorem shows that the appropriate space is an abelian-by finite covering space. Thus, S_ν is a torus of some dimension (actually $\geqslant 1$) precisely when ν is an edge. Note that this includes extreme edges. If ν is a minimal isometric or pseudo-Anosov edge, then S_ν is a circle or two-dimensional torus respectively. If $\nu = [f_0, \Gamma]$ is a node then S_ν is a fibre bundle over a covering space of $B[f_0, \Gamma]$ with a (possibly 0-dimensional) torus as fibre.

The topological spaces corresponding to the singleton nodes of $\mathcal{G}_n(B)$ for $n > 0$ have to be defined inductively, using the Topographer's View of 5.11. The components of $\mathcal{G}_{n-1}(B)$ are in 1-1 correspondence with the singleton nodes of $\mathcal{G}_n(B)$. The Topographer's View at level $n-1$ associates a topological space S_ν to each component ν of the graph of $\mathcal{G}_{n-1}(B)$. These are also the topological spaces associated to the singleton nodes of $\mathcal{G}_n(B)$.

5.8. Natural identifications between Node and Edge Spaces. — Let $\nu = [f_0, \Gamma]$ and $\varepsilon = [f_0, \Gamma']$ be a node and edge of $\mathcal{PG}(B)$ or $\mathcal{G}(B)$, with ν and ε adjacent, which means that $\Gamma \subset \Gamma'$. Write $B_1 = B[f_0, \Gamma]$ and $G_1 = G[f_0, \Gamma]$, $G_1' = G[f_0, \Gamma']$. Then the abelian group G_1' identifies with a conjugacy class of subgroups of G_1 containing the centre of $G[f, \Gamma]$, that is, the preimage under $\Theta : G_1 \to \pi_1(B_1) \times \mathbf{Z}^q$ of $\{1\} \times \mathbf{Z}^q$ (see 4.5, 4.6), and this preimage is of at most cyclic index in G_1'. Further, by 4.16, the image under projection to $\pi_1(B_1)$ — if B_1 is a map space — is $\pi_1(N)$, for some deleted neighbourhood of critically finite maps in $\overline{B}_1 \setminus B_1$. So the torus S_ε identifies with a subspace of S_ν. Recall that S_ν is a fibre bundle with torus fibres over a covering space of B_1. Then S_ε identifies with the subbundle over either a point or homotopically nontrivial loop in the covering space of B_1. We thus have an equivalence relation $\sim$, which is a union of equivalence relations $\sim_n$, where $\sim_n$ is an equivalence relation on on

$$\coprod\{S_\mu : \mu \text{ is a pair node or edge of } \widetilde{\mathcal{G}}_n\}.$$

The Topographer's View will involve an extension of the equivalence relation $\sim_n$ to

$$\coprod\{S_\mu : \mu \text{ is a node or edge of } \widetilde{\mathcal{G}}_n\}.$$

5.9. Polynomial Type Theorem. — This is relatively easy to prove, and its proof is completed in Chapter 6.

Polynomial Type Theorem. — *Let $B = B(Y, f_0)$ be of polynomial type. Then under inclusion, B is homotopy equivalent to $V = V(Y, f)$, where V is the space of rational maps in B.*

Let B be of the polynomial degree four type with fixed critical value v_0 of maximal multiplicity, v_1 periodic, and v_2 the image of two critical points. Let $N \subset B$ be the subset in which the Poincaré distance between v_0 and v_2 is less than ε, for some fixed ε less than the Margulis constant. Then under inclusion, (B, N) is homotopy equivalent to (V, N'), where N' is a tubular neighbourhood of ∞ in V.

This theorem implies that the set of polynomials in B is connected. The second part of the theorem for polynomial type is included simply because it is needed for the results about degree two type.

5.10. Theorems for Degree Two Type

Theorem: Injective on π_1. — *Let $B = B(Y, f_0)$ be of degree two type. Let V_1 be any component of the rational maps in B. Then the inclusion $V_1 \hookrightarrow B$ induces an injection $\pi_1(V_1) \to \pi_1(B)$.*

Let $\widetilde{B}$ denote the universal cover of B. The theorem implies that any component $\widetilde{V}_1$ of the preimage of V_1 in $\widetilde{B}$ is simply connected, and is (as the notation suggests) the universal cover of $\widetilde{V}_1$. In 5.5, we chose a point $f_1 \in V_1$ and a path γ in B from f_0 to f_1 so that the lift of the path identifies a component in $\widetilde{B}$ of the preimage of V_1. By Injective on π_1, this preimage component is simply connected and identifies with $\widetilde{V}_1$. Then using γ, $\rho_2 : \pi_1(B, N, f) \to \partial D$ restricts to a map $\rho_2 : \pi_1(V_1, V_1 \cap N, f_1) \to \partial D$. The map ρ_2 on $\pi_1(B, N, f_0)$ was shown to be injective in 1.12. By Injective on π_1, it is also injective on $\pi(V_1, V_1 \cap N, f_1)$, which we can regard as a countable subset of ∂D, identifying $\widetilde{V}_1$ with the disc D.

Theorem: Resident's View of Rational Maps Space. — *Consider*

$$\rho_2 : \pi_1(V_1, V_1 \cap N, f_1) \longrightarrow \partial D.$$

(1) *ρ_2 extends continuously to ∂D except at countably many points, where right and left limits exist.*

(2) *A point of discontinuity of the extension on $\partial D = \partial \widetilde{V}_1$ is either an endpoint of a path lifting a geodesic to a puncture of V_1, or is an endpoint of a lift of a closed geodesic in V_1.*

(3) *The right and left limits at such a point are either the endpoints of $C(f, \Gamma)$ for some minimal nonempty isometric (f, Γ) satisfying the Edge Condition, or are the endpoints of a geodesic in $\partial C(f, \Gamma)$ projecting to a geodesic lamination leaf, for a minimal nonempty pseudo-Anosov (f, Γ) satisfying the Edge Condition.*

(4) *The inverse map* ρ_2^{-1} *of the extension extends across such geodesics to map* $\partial C(\widetilde{V}_1)$ *continuously and monotonically to* ∂D*, and injectively except for mapping the closure of each geodesic in* $\partial C(f,\Gamma)$ *to a point. Moreover*

$$\rho_2^{-1}(g\cdot z) = g\cdot\rho_2^{-1}(z) \quad \text{for all } z\in\partial C(\widetilde{V}_1)\cap\partial D,\ g\in\pi_1(V_1,f).$$

Remark. — We can also regard $\pi_1(V_1, f_1)$ as a subset of $\widetilde{V}_1 = D$, $\pi_1(\overline{\mathbf{C}}\setminus Z, v_2)$ as a subset of D. Then 1.12 gives a map $\rho:\pi_1(V_1,f_1)\to D$. The theorem holds with ρ_2 replaced by ρ, that is, we can show that ρ also extends continuously to ∂D. The proof for ρ is not greatly more difficult than that for ρ_2, but one has to prove the version for ρ_2 first, and then the proof for ρ has an extra step at each stage. I have therefore omitted the proof for ρ. The proof for ρ_2 will be given in Chapters 25-31.

We call this theorem the Resident's View, because it gives an identification of $\widetilde{V}_1$ with a subset $C(\widetilde{V}_1)$ of the universal cover of the dynamical plane $\overline{\mathbf{C}}\setminus Z$. In particular, the topology of V_1 is recorded in the action of $\pi_1(V_1)$ on the boundary of $C(\widetilde{V}_1)$.

We call the following theorem the Topographer's View, because it identifies B, up to homotopy equivalence, with a union of spaces with a strong geometric structure.

Theorem: Topographer's View of B. — *Let* B *be of degree two type and* $N\subset B$ *a union of tubular neighbourhoods of critically finite spaces* $B'\subset\overline{B}\setminus B$*. Then* (B,N) *is homotopically equivalent to*

$$(\cup_{n=0}^{\infty}S_n(B), \cup_{n=0}^{\infty}S_n(N))$$

with either $S_n(B)=S_n(N)=\varnothing$ *for* $n>0$*, or* $\{S_n(B)\}_{n\geqslant 0}$*,* $\{S_n(N)\}_{n\geqslant 0}$ *are increasing families of spaces, and the following hold.*

$$(S_n(B), S_n(N)) = \Big(\Big(\coprod_{\nu} S_\nu\Big)/\sim_n, \coprod_{\mu} S_\mu\Big),$$

where ν *runs over the nodes and edges of a graph* $\mathcal{G}_n(B)$*,* μ *runs over the extreme edges* $[f,\Gamma]$ *corresponding to components of* N*, and the equivalence relation* $\sim_n$ *relates only points from spaces indexed by adjacent nodes and edges of* $\mathcal{G}_n(B)$ *and is an extension of the equivalence relation of 5.8, as follows.*

(1) *Let* $\nu=[\widetilde{V}_1]$ *and let* $\varepsilon=[f,\Gamma]$*, where* (f,Γ) *is minimal nonempty satisfying the Edge Condition. Let* ε *be isometric. Then* S_ε *is a circle and identifies with a simple loop round a puncture in* $V_1=S_\nu$*.*

(2) *Let* $\varepsilon=[f,\Gamma]$ *be pseudo-Anosov, with* $\kappa(f,\Gamma)=\kappa_n$*. Then* S_ε *is a two-dimensional torus, and a simple homotopically nontrivial loop in* S_ε *(corresponding to an element of* $G(\varepsilon)\cong\pi_1(S_\varepsilon)$ *with a pseudo-Anosov component) identifies under* $\sim_n$ *with some homotopically nontrivial loop in* S_ν *for some singleton node* ν *in* $\mathcal{G}_n(B)$*. [Remember that* S_ν *is a component of* $S_{n-1}(B)$*.]*

The structure of $\mathcal{G}_n(B)$ is described completely by the following.

Theorem: Resident's View of $\widetilde{B}$. — *Consider the sets* $C(\widetilde{V}_1)$, $C(f_0, \Gamma)$, *for all singletons* $\widetilde{V}_1$, *and pairs* (f_0, Γ) *satisfying the Invariance, Levy and Maximal Reduced Conditions, and which are either primitive or minimal nonempty. Thus,* $\mathcal{P}(B)$ *is a* G-*invariant partition of* D. *Furthermore,* $\mathcal{P}_n(B)$ *is a locally finite partition, restricted to any component of* $\cup\mathcal{P}_n$. *Then the quotient* $\mathcal{G}_n(B) = \widetilde{\mathcal{G}}_n(B)/G$ *of the dual* $\widetilde{\mathcal{G}}_n(B)$ *of* $\mathcal{P}_n(B)$ *is the same graph as in the Topographer's View.*

CHAPTER 6

AN ITERATION ON A TEICHMÜLLER SPACE

6.1. Definition of standard Teichmüller space. — Let $Y \subset \overline{\mathbf{C}}$ be a finite set, $\#(Y) \geqslant 3$. Then we write $\mathcal{T}(Y)$ for the *Teichmüller space of* $\overline{\mathbf{C}}$ *with marked set* Y. This space (as usual) is described as follows. Let $\varphi_i : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ be homeomorphisms $(i = 1, 2)$. We say that $\varphi_1 \sim \varphi_2$ if, for some Möbius transformation σ, φ_1 and $\sigma \circ \varphi_2$ are isotopic via an isotopy which is constant on Y. Let $[\varphi]$ denote the equivalence class of φ with respect to $\sim$. Then $\mathcal{T}(Y)$ is the set of all $[\varphi]$, and has a natural topology. If $\#(Y) = n$, then $\mathcal{T}(Y)$ is homeomorphic to $\mathbf{R}^{2n-6}$. We have a natural projection from the quotient of $\mathrm{Hom}_+(\overline{\mathbf{C}})$ by Möbius left composition onto $\mathcal{T}(Y)$. Note that the inverse images of points are homeomorphic to quotients of $\mathrm{Hom}_+(\overline{\mathbf{C}}, Y)$, and hence contractible, by Chapter 1.

If $W \subset Y$ with $\#(W) \geqslant 3$, then there is a natural projection from $\pi : \mathcal{T}(Y) \to \mathcal{T}(W)$, given by $\pi([\varphi]) = [\varphi]_W$, where $[\varphi]_W$ denotes φ modulo isotopies constant on W, and left composition by Möbius transformations.

6.2. The Teichmüller space of B, and the Projection of the universal cover. — Now let $B = B(Y, f_0)$ be as in 1.9. Suppose that $\#(Y) = \#(Y(f_0)) \geqslant 4$. Then $\mathcal{T}(Y) = \mathcal{T}(Y(f_0))$ is the *Teichmüller space of* B, also denoted $\mathcal{T}(B)$. If we fix basepoints, then there is a natural projection from $\widetilde{B}$ to $\mathcal{T}(B)$, as follows. We identify the elements of $\widetilde{B}$ as homotopy classes of paths in B starting at f_0. Then, as we have seen in 1.11, a path f_t $(t \in [0,1])$ gives rise to paths $[\varphi_t] \in \mathcal{T}(Y)$, $[\varphi_t'] \in \mathcal{T}(f_0^{-1}Y)$, with $[\varphi_0] = [\text{identity}]$ and $\varphi_t(Y(f_0)) = \varphi_0(Y(f_t))$, $\varphi_t'(f_0^{-1}(Y(f_0))) = f_t^{-1}(Y(f_t))$, $[\varphi_t]_Z = [\varphi_t']_Z$, $f_t = \varphi_t \circ f_0 \circ \varphi_t'^{-1}$. In fact the map

$$\pi : \widetilde{B} \longrightarrow \mathcal{T}(B) : [f_t] \longmapsto [\varphi_1]$$

is well defined, and is the *natural projection.*

Let V denote the space of rational maps in B, with preimage $\widetilde{V}$ in $\widetilde{B}$. Then $\pi \mid \widetilde{V}$ is injective, so we can, and shall, regard $\widetilde{V}$ as a subspace of $\mathcal{T}(Y)$, which, in general will be disconnected. According to Injective on π_1 of 5.11, the components of $\widetilde{V}$ are simply connected in the cases that most concern us. But this is very far from being obvious.

6.3. Modular Group Actions. — Classically, $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ can be regarded as a discrete subset of $\mathcal{T}(Y)$, and acts on $\mathcal{T}(Y)$, on the right, by

$$[\varphi] \cdot [\psi] = [\varphi \circ \psi]$$

Here, we are using [] to denote both equivalence classes in $\mathcal{T}(Y)$ and in $\mathrm{MG}(\overline{\mathbf{C}}, Y)$. Recall (1.11) that the group $G = \pi_1(B)$ also identifies with

$$G_1 = \{[\psi] \in \mathrm{MG}(\overline{\mathbf{C}}, Y) : [\psi]_Z = [f_0^{-1} \circ \psi \circ f_0]_Z\},$$

and thus also acts on $\mathcal{T}$ on the right. But $G = \pi_1(B)$ also acts on $\widetilde{B}$ on the left. (See 1.1.) We easily verify that

$$\pi(g \cdot [f_t]) = \pi([f_t]) \cdot g^{-1}.$$

In fact, this is an extension of the fact that the map Φ_1 of 1.11 is an anti-isomorphism. Thus we have a quotient map

$$[\pi] : B \longrightarrow \mathcal{T}/G$$

which is, in fact, also a homotopy equivalence. (To see this, we need to choose the section of π to be G-invariant. This can be done.)

6.4. Teichmüller metrics. — Once again, we recall the standard theory. (See [**Gar2**], for example.) For a quasi-conformal homeomorphism $\chi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$, let $D\chi_z$ denote the derivative of χ at z, and $(D\chi_z)^T$ is its adjoint. Let $K(\chi)(z)$ be the quasi-conformal distortion λ/μ of χ at z, where $\lambda^2 > \mu^2 > 0$ are the eigenvalues of $(D\chi_z)^T D\chi_z$. Let

$$\|\chi\|_{\mathrm{qc}} = \operatorname{ess\,sup}\{K(\chi)(z) : z \in \overline{\mathbf{C}}\}.$$

The Teichmüller metric d_Y on $\mathcal{T}(Y)$ is defined by

$$d_Y([\varphi_1], [\varphi_2]) = \inf\{(1/2) \log \|\chi\|_{\mathrm{qc}} : [\varphi_2] = [\chi \circ \varphi_1]\}.$$

Note that the action of $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ on $\mathcal{T}(Y)$ is isometric with respect to d_Y. In the case when $\#(Y) = 4$, and $\mathcal{T}(Y)$ identifies with the upper half plane H, then the Teichmüller distance is actually half the Poincaré metric.

The following standard information was important in the proof of Thurston's theorem, and will also be important here. The infimum in the definition of $d_Y([\varphi_1], \varphi_2])$ is attained by a unique χ, with quasi-conformal distortion which is constant (say, K) almost everywhere. Moreover, there are coordinate systems on $\overline{\mathbf{C}} \setminus \varphi_1(Y)$, $\overline{\mathbf{C}} \setminus \varphi_2(Y)$ such that, in these coordinates, χ has the form

$$\chi(x + iy) = (K^{1/2}x + iK^{-1/2}y).$$

These coordinates are given by quadratic differentials on $\overline{\mathbf{C}} \setminus \varphi_1(Y)$, $\overline{\mathbf{C}} \setminus \varphi_2(Y)$, and thus are defined except at finitely many singularities. If these quadratic differentials are $q_1 dz^2$, $q_2 dz^2$ respectively, then the coordinates are given locally by

$$\int \sqrt{q_j}\, dz,$$

whenever q_j is finite and $\neq 0$. Here, $q_j(z)dz^2$ is holomorphic on $\overline{\mathbf{C}} \setminus \varphi_j(Y)$, and $q_j(z)dz^2$ has at most simple poles. This ensures that the path integral is bounded. If $z = 1/w$, then $dz = -dw/w^2$. So q_j is a rational function with at most simple poles, occurring only at finite points of $\varphi_j(Y)$, and at least 3 more poles than zeros, up to multiplicity. If ∞ is a simple pole of $q_j(z)dz^2$, then q_j has exactly three more poles than zeros, up to multiplicity. If ∞ is a zero of $q_j(z)dz^2$ of multiplicity m, then q_j has exactly $m+4$ more poles than zeros, up to multiplicity. So $q_j(z)dz^2$ has exactly 4 more poles than zeros, up to multiplicity. It is usual to normalise q_j so that

$$\int |q_j| = 1.$$

In this case, $q_1(z)dz^2$ and $q_2(z)dz^2$ are uniquely determined by χ. We shall say that $q_1(z)dz^2$ *is the quadratic differential at* $[\varphi_1]$ *for* $d([\varphi_1],[\varphi_2])$. Using this definition, $-q_2(z)dz^2$ is the quadratic differential at $[\varphi_2]$ for $d([\varphi_2],[\varphi_1])$. We say that $q_2(z)dz^2$ is the *stretch of* $q_1(z)dz^2$ *at* $[\varphi_2]$.

6.5. Thick and Thin parts of Teichmüller space. — For each $[\varphi] \in \mathcal{T}(Y)$, there is a unique Poincaré metric ρ_φ on $\overline{\mathbf{C}} \setminus \varphi(Y)$. We write $\mathcal{T}_{\geqslant\varepsilon}(Y)$ for the subset of $\mathcal{T}(Y)$ for which all closed geodesics have ρ_φ-length $\geqslant \varepsilon$. It is well-known [**T1**] that the quotient of $\mathcal{T}_{\geqslant\varepsilon}(Y)$ by the action of $\mathrm{MG}(\overline{\mathbf{C}}, Y)$ or $\mathrm{PMG}(\overline{\mathbf{C}}, Y)$ is compact, for any $\varepsilon > 0$.

The Margulis decomposition [**T1**] tells us that for some $\varepsilon = \varepsilon_0$, depending only on $\#(Y)$, for any $[\varphi] \in \mathcal{T}(Y)$, geodesics of length $\leqslant \varepsilon_0$ in $\overline{\mathbf{C}} \setminus \varphi(Y)$ cannot intersect, nor even self-intersect. We then have a thick-and-thin decomposition of the surface $\overline{\mathbf{C}} \setminus \varphi(Y)$ into sets $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$ and $(\overline{\mathbf{C}} \setminus \varphi(Y))_{\geqslant\varepsilon_0}$, where $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$ is the set of points x such that there is a closed nonperipheral geodesic segment in $\overline{\mathbf{C}} \setminus \varphi(Y)$ of length $< \varepsilon_0$ with both endpoints at x, and $(\overline{\mathbf{C}} \setminus \varphi(Y))_{\geqslant\varepsilon_0}$ is the complement of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$. Then $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$ is a union of disjoint homotopically nontrivial and nonperipheral cylinders in $\overline{\mathbf{C}} \setminus \varphi(Y)$. A cylinder homotopic to $\varphi(\gamma)$ has modulus within $O(1/\varepsilon_0)$ of $2\pi^2/\varepsilon$, if ε is the length of the geodesic homotopic to $\varphi(\gamma)$.

Given a set Γ of disjoint simple nonperipheral loops in $\overline{\mathbf{C}} \setminus Y$, we write $\mathcal{T}(\Gamma, \varepsilon)$ for the set of $[\varphi]$ in $\mathcal{T}$ such that $\varphi(\gamma)$ has length $< \varepsilon$ for all $\gamma \in \Gamma$.

6.6. Properties of the Pullback Map. — We now come to the key object of this paper, namely, our pullback map

$$\tau : \mathcal{T} \longrightarrow \mathcal{T}$$

which is a generalisation of Thurston's pullback on a Teichmüller space [**T2**]. Indeed, our pullback reduces to Thurston's (essentially) in the case when $B(Y)$ is a space of critically finite maps (that is, $Y = Z$). The basic properties of τ are more important than its definition, so we give them first. For the third property, we say that $B = B(Y, f_0)$ is *exceptional* if $Y = Z$, $\#(Y) = 4$, Z contains no critical points, but every point of $f_0^{-1}(Z) \setminus Z$ is critical and all critical points of f_0 are in $f_0^{-1}(Z) \setminus Z$. As usual, we let V denote the space of rational maps in B. From now on, given Y we fix a constant $\varepsilon_0 > 0$ such that the Margulis decomposition (6.5) holds.

1. The Commuting Property. — $\tau : \mathcal{T} \to \mathcal{T}$ commutes with the G-action.

2. The Fixed Set. — The set $\{x : \tau x = x\}$ is $\pi(\widetilde{V})$, where $\widetilde{V}$ is the preimage of V in the covering space $\widetilde{B}$, and $\pi : \widetilde{B} \to \mathcal{T}(B)$ is the natural projection defined in 6.2. The fixed set is all of $\mathcal{T}$ if $\#(Z) \leqslant 3$.

3. Distance-Decreasing along Iterates. — For any x, the sequence

$$\{d(\tau^n(x), \tau^{n+1}(x))\}$$

is decreasing. Moreover, if $\#(Z) \geqslant 4$ and B is not exceptional, then for an integer k depending only on Y and some $1 \leqslant m \leqslant k$,

$$d(\tau^m(x), \tau^{m+1}(x)) \leqslant \lambda(x) d(x, \tau(x)),$$

where $\lambda(y) \leqslant \lambda(\varepsilon, M) < 1$ if $d(y, \tau(y)) \leqslant M$, and $y \in \mathcal{T}_{\geqslant \varepsilon}$.

4. The Thin Part Invariance Property. — For L, $\varepsilon_1 < \varepsilon_0$ depending only on $\#(Y)$ and a given constant $M > 0$, the following holds. Suppose $\varepsilon \leqslant \varepsilon_1$, $d(x, \tau(x)) \leqslant M$ and $x \in \mathcal{T}_{<\varepsilon}$. Then there is (f_0, Γ) satisfying the Invariance Condition such that $x \in \mathcal{T}(\Gamma, L\varepsilon)$.

5. The Thin Part Levy Property

Polynomial type. — Let $B(Y, f_0)$ be of polynomial type. If $\varepsilon_1 > 0$ is sufficiently small given ε_0 and M, and $x \in \mathcal{T}_{\geqslant \varepsilon_0}$, $d(x, \tau(x)) \leqslant M$, then $\tau^n(x) \in \mathcal{T}_{\geqslant \varepsilon_1}$ for all $n \geqslant 0$.

Degree two type. — Let $B(Y, f_0)$ be of degree two type. There is an integer $m > 0$ depending only on (f_0, Y) such that, given $M > 0$, there are $\varepsilon_1 \leqslant \varepsilon_0$ and L' such that the following hold. Let $\varepsilon \leqslant \varepsilon_1$, let $d(x, \tau(x)) \leqslant M$, let $\tau^i(x) \in \mathcal{T}_{\geqslant \varepsilon}$ for $0 \leqslant i < m$, and let there be a least integer $n \geqslant m$ with $\tau^n x \in \mathcal{T}_{<\varepsilon}$. Then there is Γ satisfying the Invariance and Levy Conditions such that $\tau^n x \in \mathcal{T}(\Gamma, L'\varepsilon)$.

6. Close to zero distance. — Let $B(Y, f_0)$ be of degree two type. Given $\varepsilon > 0$, there is $\delta > 0$ such that if $d(x, \tau(x)) < \delta$, then either $d(x, y) < \varepsilon$ for some $y \in \widetilde{V}$, or $x \in \mathcal{T}(\Gamma, \varepsilon)$ for some (f_0, Γ) satisfying the Invariance and Levy Conditions.

6.7. Definition of the Pullback Map. — Now we define τ. We fix a space $B(Y, f_0)$ as in 1.9. Let $Z \subset Y$ be as in the definition of $B(Y, f_0)$ in 1.9, and let $\#(Z) \geqslant 3$. Then $\varphi \circ f_0 : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ is a branched covering. Let $[\varphi] \in \mathcal{T} = \mathcal{T}(B(Y, f_0)) = \mathcal{T}(Y)$. So there is a holomorphic branched covering $s : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ and an orientation-preserving homeomorphism $\varphi_1 : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ such that $\varphi \circ f_0 = s \circ \varphi_1$, s is uniquely determined up to left composition with a Möbius transformation, and $[\varphi_1] \in \mathcal{T}(f_0^{-1} Y)$ is uniquely determined. By abuse of notation, we shall often write

$$[\varphi_1] = [s^{-1} \circ \varphi \circ f_0].$$

Now $Z(f_0) = Z \subset Y \cap f_0^{-1}(Y)$ (see 1.9). So $[\varphi_1]_Z \in \mathcal{T}(Z)$ is uniquely determined, where this is the natural projection to $\mathcal{T}(Z)$ (see 6.1). Then $\tau([\varphi])$ will be a lift of $[\varphi_1]_Z$ to $\mathcal{T}(Y)$, as follows.

We consider $[\varphi_1]_Z$ and the projection $[\varphi]_Z$ of $[\varphi]$ to $\mathcal{T}(Z)$. So by 6.4 we can find a unique $\chi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ such that $\varphi_1(Z) = \chi \circ \varphi(Z)$, $[\varphi_1]_Z = [\chi \circ \varphi]_Z$ and

$$d_Z([\varphi], [\varphi_1]) = (1/2)\log \text{ ess sup}\|\chi\|_{\text{qc}}.$$

Then $[\chi \circ \varphi]_Y$ is uniquely determined, and and we define

$$\tau([\varphi]) = [\chi \circ \varphi]_Y.$$

6.8. Remark. — If $Z = Y$ — that is, if $B(Y, f_0)$ is a space of critically finite maps — then the definition of τ is precisely the same as Thurston's original definition [**T2**]. If $Z \neq Y$ we may still have chosen f_0 so that $Y \subset f_0^{-1}(Y)$, and thus $[\varphi_1] \in \mathcal{T}(f_0^{-1}(Y))$ projects to $[\varphi_1]_Y$, but even so, it is definitely *not* the case that in general $[\varphi_1]_Y = [\chi \circ \varphi]_Y$.

6.9. Proof of Property 1. — Let $[\varphi] \in \mathcal{T}(Y) = \mathcal{T}(B(Y, f_0))$ and $[\psi] \in G \leqslant \text{MG}(\overline{\mathbf{C}}, Y)$. Then if s is as in 6.7,

$$[s^{-1} \circ \varphi \circ \psi \circ f_0]_Z = [(s^{-1} \circ \varphi \circ f_0) \circ (f_0^{-1} \circ \psi \circ f_0)]_Z = [(s^{-1} \circ \varphi \circ f_0) \circ \psi]_Z.$$

So if χ is as in 6.7, so that $[\chi \circ \varphi]_Z = [s^{-1} \circ \varphi \circ f_0]_Z$,

$$[s^{-1} \circ \varphi \circ \psi \circ f_0]_Z = [\chi \circ \varphi \circ \psi]_Z,$$

and $\tau([\varphi] \cdot \psi]) = \tau([\varphi]) \cdot [\psi]$, as required.

6.10. Proof of Property 2. — Again, let s be as in 6.7, given $[\varphi] \in \mathcal{T}(Y)$, $Y = Y(f_0)$. It is clear that τ fixes $[\varphi]$ if and only if $[s^{-1} \circ \varphi \circ f_0]_Z = [\varphi]_Z$, that is, if and only if $\varphi \circ f_0 = s \circ \varphi_1$ with φ_1 and φ isotopic via isotopy constant on Z (composing s on the left with a Möbius transformation if necessary). Then

$$s(\varphi(Z)) = s(\varphi_1(Z)) = \varphi(Z), \quad s = \varphi \circ f_0 \circ \varphi_1^{-1}.$$

So $s \in V$ and $\pi(\widetilde{s}) = [\varphi]$ where $\widetilde{s}$ is a preimage of s in $\widetilde{V}$ and π is as in 6.2. The converse is also clear, that is, if $\widetilde{s} \in \widetilde{V}$ then $\tau(\pi(\widetilde{s})) = \pi(\widetilde{s})$. It is also immediate that τ fixes all points of $\mathcal{T}$ if $\#(Z) \leqslant 3$. □

6.11. Proof of Property 3. — Let $[\varphi] \in \mathcal{T}(Y)$. Let s be a holomorphic branched covering whose critical values are in $\varphi(Y)$. Let $q(z)dz^2$ be a quadratic differential on $\overline{\mathbf{C}} \setminus \varphi(Y)$. Then

$$s^* q(z)dz^2 = q(s(z))(s'(z))^2 dz^2$$

is a quadratic differential on $\overline{\mathbf{C}} \setminus s^{-1}\varphi(Y)$ (the *pullback*). If s has degree d then

$$\int |s^* q| = d \int |q|.$$

Now assume that $\#(Z) \geqslant 4$, where, as usual, Z is as in the definition of $B(Y, f_0)$. Write $\tau([\varphi]) = [\varphi_1] \in \mathcal{T}(Y)$. Let $\chi : \overline{\mathbf{C}} \setminus \varphi(Y) \to \overline{\mathbf{C}} \setminus \varphi_1(Y)$ be the homeomorphism, as in 6.7 minimising quasi-conformal distortion, in both $\mathcal{T}(Y)$ and $\mathcal{T}(Z)$, so that $\tau([\varphi]) = [\varphi_1] = [\chi \circ \varphi]$. Let $q_0(z)dz^2 = q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], [\varphi_1])$ and let $p(z)dz^2$ be the stretch of $q(z)dz^2$ at $[\varphi_1]$. (See 6.4.) Let $\frac{1}{2}\log K = d([\varphi], \tau([\varphi]))$ Then χ takes the form

(1) $$(x, y) \longmapsto (K^{1/2}x, K^{-1/2}y)$$

with respect to the singular local coordinates given by $q(z)dz^2$, $p(z)dz^2$ on $(\overline{\mathbf{C}}, \varphi(Z))$, $(\overline{\mathbf{C}}, \varphi_1(Z))$ respectively. Let the holomorphic branched covering $s = s_0$ be as in the definition of $\tau([\varphi])$ in 6.7. Let s_1 be the holomorphic branched covering at $[\tau([\varphi])$ used to define $\tau(\tau([\varphi]))$. This means that the critical values of s_1 are the images under χ of the critical values of s. Then the homeomorphism that we denote by $s_1^{-1} \circ \chi \circ s$ has the form (1) with respect to local coordinates given by $s^*q(z)dz^2$ and $s_1^*p(z)dz^2$, and

$$\pi_Z(\tau^2([\varphi])) = [s_1^{-1} \circ \chi \circ s \circ \psi]_Z,$$

using $\pi_Z : \mathcal{T}(Y) \to \mathcal{T}(Z)$ to denote the natural projection. So

(2) $$d(\tau([\varphi]), \tau^2([\varphi])) \leqslant \log K = d([\varphi], \tau([\varphi])).$$

Let d be the degree of f_0. Let $q_1(z)dz^2$ be the quadratic differential at $\tau([\varphi])$ for $d(\tau([\varphi]), \tau^2([\varphi])$. We have strict inequality in (2) unless

$$\frac{1}{d}s^*q(z)dz^2 = q_1(z)dz^2.$$

Define $[\varphi_i] = \tau^i([\varphi])$ for $i \geqslant 1$. Let s_i be the holomorphic branched covering with $[s_i^{-1} \circ \varphi_i \circ f_0]_Z = [\varphi_{i+1}]_Z$. Let $q_i(z)dz^2$ be the quadratic differential at $\tau^i([\varphi])$ for $d(\tau^i([\varphi]), \tau^{i+1}([\varphi]))$. Proceeding inductively, we have either

(3) $$d(\tau^k([\varphi]), \tau^{k+1}([\varphi])) < d([\varphi], \tau([\varphi]))$$

or

(4) $$\frac{1}{d}s_i^*q_i(z)dz^2 = q_{i+1}(z)dz^2 \quad \text{for } 0 \leqslant i < k.$$

If (4) holds, and if $\varphi_i(W_i)$ is the pole set of $q_i(z)dz^2$, then for all $0 \leqslant i < k$,

(5) $$f_0^{-1}(W_i) \subset W_{i+1} \cup \{c : c \text{ critical }\}.$$

We have $\#(W_i) \geqslant 4$ for all i, as explained in 6.4. Then $f_0^{-1}(W_i)$ contains $\#(W_i)d$ points up to multiplicity. So if $f_0^{-1}(W_i)$ contains n_j points of multiplicity j, for $1 \leqslant j \leqslant r$, then

$$\sum_{j=1}^{r} jn_j = \#(W_i)d, \quad \sum_{j=1}^{r}(j-1)n_j \leqslant 2d-2,$$

because the number of critical points is $2d-2$ up to multiplicity. It follows that

$$\#(W_i)d \leqslant n_1 + \sum_{j=1}^{r} 2(j-1)n_j \leqslant n_1 + 4d - 4,$$

with strict inequality unless $n_j = 0$ for $r > 2$, and unless $f_0^{-1}(W_i)$ contains all critical points of f_0. Then

$$\#(W_{i+1}) = n_1 \geqslant (\#(W_i) - 4)d + 4 \geqslant 2\#(W_i) - 4 \geqslant \#(W_i),$$

with strict inequality if $\#(W_i) > 4$. So if (5) holds for all $0 \leqslant i < k$, and k is large enough given $\#(Z)$, $\#(W_i) = 4$ for all $0 \leqslant i < k$. It also follows that $W_i = W$ for all $0 \leqslant i < k$, $f_0(W) \subset W$, all points in W are noncritical, the points in $f_0^{-1}(W) \setminus W$ are precisely the critical points of f_0, all of local degree 2, and $B(Y, f_0)$ is a space of critically finite maps. In fact, the set W contains either 2 or 4 periodic points, depending on whether d is even or odd. So we are in the exceptional case.

So (3) holds in the nonexceptional case, for k sufficiently large, depending only on $\#(Y)$. To obtain Property 3, we have to refer to an estimate which we shall prove in 8.3. There is a constant $C = C(M)$ such that the following holds. Let

$$\theta_i(z) = \arg s_i^*(q_i)(z) - \arg q_{i+1}(z).$$

Then, by 8.3, using the homeomorphism $s_{i+1}^{-1} \circ \chi_i \circ s_i$, where $[\chi_i \circ \varphi_i] = [\varphi_{i+1}]$ and χ_i minimizes distortion, if $d([\varphi], \tau([\varphi])) \leqslant M$,

$$d(\tau^{i+1}([\varphi]), \tau^{i+2}([\varphi])) \leqslant \Big(1 - C\int |\theta_i|^2 |q_{i+1}|\Big) d(\tau^i([\varphi]), \tau^{i+1}([\varphi])).$$

By compactness of the space of the corresponding space of quadratic differentials, we see that, except in the exceptional case, for some $1 \leqslant i \leqslant k$, if $d([\varphi], \tau([\varphi])) \leqslant M$ and $[\varphi] \in \mathcal{T}_{\geqslant \varepsilon}$,

$$d(\tau^i([\varphi]), \tau^{i+1}([\varphi]) \leqslant \lambda(M, \varepsilon) d(\tau^{i-1}([\varphi]), \tau^i([\varphi]))$$

for $\lambda(M, \varepsilon) < 1$.

This proof is basically the same as Thurston's original proof in the critically finite case [**T2**]. The exceptional case is the case of *Euclidean orbifold*.

6.12. Proof of Property 4. — By 6.5, there is a constant C such that the following holds, for all sufficiently small ε. Given $\varphi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$, any geodesic on $\overline{\mathbf{C}} \setminus \varphi(Z)$ of length $< \varepsilon$ is isotopic in $\overline{\mathbf{C}} \setminus \varphi(Z)$ to at least one geodesic in $\overline{\mathbf{C}} \setminus \varphi(Y)$ of length $< C\varepsilon$. Conversely, a geodesic in $\overline{\mathbf{C}} \setminus \varphi(Y)$ of length $< \varepsilon$ is isotopic in $\overline{\mathbf{C}} \setminus \varphi(Z)$ to a geodesic of length $< C\varepsilon$ in $\overline{\mathbf{C}} \setminus \varphi(Z)$.

So now consider $x = [\varphi] \in \mathcal{T}$ with $d(x, \tau(x)) \leqslant M$ and $\tau(x) = [\varphi_1]$. Let s be the holomorphic branched covering used to define $\tau([\varphi]) = [\psi]$, so that $[\varphi_1]_Z = [s^{-1} \circ \varphi \circ f_0]_Z$. So $\varphi_1(Z) \subset s^{-1}(\varphi(Y))$. So for $\varepsilon > 0$ sufficiently small, a geodesic on $\overline{\mathbf{C}} \setminus s^{-1}(\varphi(Y))$ of length $< \varepsilon$ is isotopic in $\overline{\mathbf{C}} \setminus \varphi_1(Z)$ to at least one geodesic on $\overline{\mathbf{C}} \setminus \varphi_1(Y)$ of length $< C^2\varepsilon$, and vice versa. Now

$$s : \overline{\mathbf{C}} \setminus s^{-1}(\varphi(Y)) \longrightarrow \overline{\mathbf{C}} \setminus \varphi(Y)$$

is a local isometry. So if γ is a loop such that $\varphi(\gamma)$ is isotopic to a geodesic of length $< \varepsilon$, and $\gamma_1 \subset f_0^{-1}(\gamma)$ is nontrivial and nonperipheral in $\overline{\mathbf{C}} \setminus Z$, then $\varphi_1(\gamma_1)$ is homotopic to a geodesic of length $\leqslant C^2\varepsilon$ in $\overline{\mathbf{C}} \setminus \varphi_1(Z)$, and $\varphi(\gamma_1)$ is isotopic to a geodesic of length $\leqslant C^3\varepsilon$ on $\overline{\mathbf{C}} \setminus \varphi(Y)$. Also, $\varphi_1(\gamma)$ is isotopic in $\overline{\mathbf{C}} \setminus \varphi_1(Z)$ to a geodesic of length $\leqslant C^3\varepsilon$ in $\overline{\mathbf{C}} \setminus s^{-1}\varphi(Y)$, and so there must be $\gamma_2 \subset \overline{\mathbf{C}} \setminus Y$ such that $\gamma \subset f_0^{-1}(\gamma_2)$ in $\overline{\mathbf{C}} \setminus Z$ and $\varphi(\gamma_2)$ is isotopic to a geodesic of length $\leqslant C^3\varepsilon$ in $\overline{\mathbf{C}} \setminus \varphi(Y)$. Repeatedly applying this procedure, we obtain an invariant (f_0, Γ) with $\gamma \in \Gamma$ and $[\varphi] \in \mathcal{T}(\Gamma, L\varepsilon)$ for a suitable L depending only on $\#(Y)$. The process has to terminate after finitely many steps, because short loops have to be isotopically disjoint.

6.13. Proof of Property 5. — Once again, this actually mimics Thurston's original proof in the critically finite case [**T2**], [**D-H3**].

Given (f_0, Y), there is an integer m with the following property. Let (f_0, Γ') be any invariant set such that $\Gamma_2(f_0, \Gamma') = \varnothing$ (see 2.5). Let $A_i = (A_i(\gamma, \delta))$, $(\gamma, \delta \in \Gamma, i \geqslant 1)$ be positive entry square matrix with

$$f_0^*(\gamma) = \sum_{\delta \in \Gamma} A_i(\gamma, \delta)\pi(\delta),$$

where $\pi(\delta)$ denotes the isotopy class of δ in $\overline{\mathbf{C}} \setminus Z$. Then for any vector $\underline{v} = (v_j)$, if $\|\underline{v}\|_\infty = \mathrm{Max}_j |v_j|$,

$$\Big\| \prod_{i=1}^{m} A_i \underline{v} \Big\|_\infty \leqslant \frac{\|\underline{v}\|_\infty}{2}.$$

Now choose L_1 so that

$$L_1 \geqslant 4(\#(Y)\deg(f_0))^m.$$

Now let $\tau^n(x) \in \mathcal{T}_{<\varepsilon}$ for a least $n \geqslant m$. Write $[\varphi_1] = \tau^{n-m}(x)$ and $[\varphi_2] = \tau^n(x)$. Let ε be sufficiently small. There is L_2, depending only on L_1 and $\#(Y)$ such that the following holds. We can find ε' with $L_2\varepsilon \geqslant \varepsilon' \geqslant \varepsilon$ and (f_0, Γ') satisfying the Invariance Condition such that $\tau^{n-m}(x) \in \mathcal{T}(\Gamma', \varepsilon')$ and $\tau^{n-m}(x) \notin \mathcal{T}(\gamma, L_1\varepsilon')$ for $\gamma \notin \Gamma'$. This is achieved by using property 4. If $\tau^{n-m}(x) \notin \mathcal{T}(\gamma, L_1\varepsilon)$ for any loop γ then we simply take $\Gamma' = \varnothing$. If at least one loop γ exists with $\tau^{n-m}(x) \in \mathcal{T}(\gamma, L_1\varepsilon)$ then we enlarge to an invariant set Γ_1', and take ε_2 to be the length of the longest of these loops ($\varepsilon_2 \leqslant LL_1\varepsilon$). We then repeat the process if there is a loop $\gamma' \notin \Gamma_1'$ with $\tau^{n-m}(x) \in \mathcal{T}(\gamma', L_1\varepsilon_2)$. Since loops of length $\leqslant$ the Margulis constant are disjoint, the process must terminate. So Γ' exists (even if empty). We then have $\tau^n(x) \in \mathcal{T}(\Gamma', L'\varepsilon)$ for a constant L' depending only on M, f_0 and M. So it suffices to show that $\Gamma_2(f_0, \Gamma') \neq \varnothing$. This will ensure that $\Gamma' \neq \varnothing$. In the Polynomial case, this is impossible, and we will deduce that $\tau^n(x) \in \mathcal{T}_{\geqslant\varepsilon}$.

Suppose on the contrary that the set $\Gamma_2(f_0, \Gamma') = \varnothing$. Let s_m be the holomorphic branched covering of degree $(\deg(f_0))^m$ such that $[\varphi_2]_Z = [s_m^{-1} \circ \varphi_1 \circ f_0^m]_Z$. Now take γ_0 with $\tau^n(x) \in \mathcal{T}(\gamma_0, \varepsilon)$. Consider the component $A(\gamma_0)$ of $(\overline{\mathbf{C}} \setminus \varphi_2(Y))_{<\varepsilon_0}$ homotopic in $\overline{\mathbf{C}} \setminus \varphi_2(Y)$ to $\varphi_2(\gamma_0)$. Then by 6.5,

$$\text{modulus}\,(A(\gamma_0)) \geqslant 2\pi^2/\varepsilon - O(1/\varepsilon_0). \tag{1}$$

But to within $\leqslant (\deg(f_0))^m \,\#(Y)$ annuli of moduli $\leqslant 2\pi^2/L_1\varepsilon + O(1/\varepsilon_0)$ with boundaries of bounded distortion, $A(\gamma_0)$ is a union of preimages under s_m of components of $(\overline{\mathbf{C}} \setminus \varphi_1(Y))_{<\varepsilon_0}$ homotopic to loops of $\varphi_1(\Gamma')$. (A similar argument was used in [**R2**] 7.5. See section 2 of that paper for an explanation, if needed, of why the boundaries have bounded distortion.) So for $\underline{v}$ with $\|\underline{v}\|_\infty \leqslant 2\pi^2/\varepsilon + O(1/\varepsilon_0)$,

$$\begin{aligned}\text{modulus}\,(A(\gamma_0)) &\leqslant \Big\| \textstyle\prod_{i=1}^m A_i \underline{v} \Big\|_\infty + (\deg(f_0))^m \,\#(Y)(2\pi^2/L_1\varepsilon + O(1/\varepsilon_0)) \\ &\leqslant \pi^2/\varepsilon + \pi^2/2\varepsilon + O(1/\varepsilon_0).\end{aligned} \tag{2}$$

But (1) and (2) are incompatible if $\varepsilon \leqslant \varepsilon_1$ for ε_1 small enough. So $\Gamma_2(f_0, \Gamma') \neq \varnothing$, and Property 5 is proved.

6.14. Remark. — The precise statement of Property 5 is a source of some considerable technical difficulty in the case of degree two type. Although we have shown that a sequence $\tau^n x$ which enters the thin part must first enter $\mathcal{T}(\Gamma, L\varepsilon)$ for (f_0, Γ) satisfying the Invariance and Levy Conditions, it is not clear that this property will persist as long as $\tau^n x$ remains in the thin part. This necessitates the construction of a modification τ' of τ in the thin part in the proof of the Topographer's View (and, less directly, in the proof of the Resident's View). This, in turn, is one of the reasons for the development of the theory of Teichmüller distance, in the second part of this paper.

6.15. Proof of Property 6. — Let L_1 and L_2 be as in the proof of Property 5. Let $\varepsilon > 0$ be given. Let $\delta > 0$ be sufficiently small. If $d(x, \tau(x)) < \delta$ and $x \in \mathcal{T}_{\geqslant \varepsilon/L_2}$, then

$$y = \lim_{n\to\infty} \tau^n(x) \in \widetilde{V}$$

exists, and

$$d(y, x) \leqslant \delta/(1-\lambda),$$

where $\lambda = \lambda(\varepsilon, 1)$. So now suppose that $x \in \mathcal{T}_{<\varepsilon/L_2}$ and $d(x, \tau(x)) < \delta$. We proceed much as in the proof of Property 5. Construct Γ', as in the proof of Property 5, such that $x \in \mathcal{T}(\Gamma', \varepsilon')$ and $x \notin \mathcal{T}(\gamma, L_1\varepsilon')$ for $\gamma \notin \Gamma'$. This time, we can ensure that $\varepsilon/L_2 \leqslant \varepsilon' \leqslant \varepsilon$, and, for sure, $\Gamma' \neq \varnothing$. Then we consider $\tau^m(x) = [\varphi_2]$. Now suppose that $\Gamma_2(f_0, \Gamma') = \varnothing$. Arguing as in the proof of Property 5, we see that if $A(\gamma)$ is the component of $(\overline{\mathbf{C}} \setminus \varphi_2(Y))_{<\varepsilon_0}$ homotopic to $\varphi_2(\gamma)$, then the modulus of $A(\gamma)$ is $< 7/8$ of the modulus of any component of $(\overline{\mathbf{C}} \setminus \varphi_1(Y))_{<\varepsilon_0}$. This contradicts $d(x, \tau^m(x)) < m\delta$, if δ is sufficiently small. So $\Gamma_2(f_0, \Gamma') \neq \varnothing$, as required.

6.16. Proof of the Polynomial Type Theorem. — Now we use Properties 1-5 to prove the Polynomial Type Theorem 5.9. We shall use the same basic idea to prove the Topographer's View for degree two type, although the additional technical difficulties in that case are considerable.

We are given a polynomial type space B of branched coverings in which the space of rational maps is V. We are also given an end N with $N' = V \cap N$, of the type described in 5.9. The spaces B, V and the pairs (B, N), (V, N') are homotopy equivalent to locally finite simplicial complexes, or pairs of such, and hence are homotopy equivalent to CW-complexes [**Mi**], or pairs of such. The spaces B, V are also $K(\pi, 1)$'s, and B is trivially homotopy equivalent to $\mathcal{T}/G$ Hence, by 7.8 of [**Spa**], it suffices to show the following. Let $\alpha : (\Delta, \partial\Delta) \to (\mathcal{T}/G, V_0)$ be continuous, where Δ is an interval of disc and $V_0 \subset V$ (possibly $V_0 \subset N'$). Then α is homotopic to map into (V, V_0), via a homotopy which is constant on $\partial\Delta$.

To show this, we have

$$\alpha(\Delta) \subset (\mathcal{T}_{\geqslant \varepsilon_0} \cap \{x : d(x, \tau x)) \leqslant M\})/G$$

for suitable M and $\varepsilon_0 > 0$, since Δ is compact. We can regard τ as a map on $\mathcal{T}/G$ by Property 1. Then by Properties 4 and 5, since B is Polynomial type, there is $\varepsilon_1 > 0$ (depending only on ε_0 and M) such that $\tau^n(\alpha(\Delta)) \subset \mathcal{T}_{\geqslant \varepsilon_1}$ for all $n \geqslant 0$. Then for all $x \in \alpha(\Delta)$, $\lim_{n\to\infty} \tau^n(x)$ exists, by Property 3, and is in $\widetilde{V}$, by Property 2. By Property 2, $\tau(x) = x$ for $x \in \alpha(\partial\Delta)$, and in fact this is true if we lift α to $\mathcal{T}$. So $\tau^n\alpha$ is homotopic to α for all $n \geqslant 0$, and $\lim_{n\to\infty} \tau^n\alpha$ is homotopic to α. This is the required homotopy. □

CHAPTER 7

HOW TO APPROACH THE TOPOGRAPHER'S AND RESIDENT'S VIEWS

7.1. In this chapter we give the first outline of the proofs of the main results for degree two type spaces. In 5.11, we split the results up into four theorems — Injective on π_1, Resident's View of Rational Maps space, Resident's View (of the whole space), and Topographer's View. In this chapter, we shall state other results, and show that the four original theorems are equivalent to: the Resident's View of Rational Maps space, the Level κ tool, and Descending Points.

7.2. Subspaces of $\mathcal{T}/G$ associated to pairs and singletons. — Fix $B = B(Y, f_0)$. Let $\mathcal{T} = \mathcal{T}(B)$. Let (B, N) be of degree two type. As usual, let $\widetilde{B}$ denote the universal cover of B, and $G = \pi_1(B)$. Let V denote the space of rational maps in B, with preimage $\widetilde{V}$ in $\widetilde{B}$. We can also (6.2) regard $\widetilde{V}$ as a subset of $\mathcal{T}$, and V as a subset of $\mathcal{T}/G$. Let $\mathcal{G}_n = \mathcal{G}_n(B)$ be the graphs defined in 5.6. The putative nodes and edges include pairs $[f, \Gamma]$ which are primitive or minimal nonempty, and singletons V_1 in the case of $\mathcal{G}_0$. In fact, these are all the putative nodes and edges in the case of $\mathcal{G}_0$.

We are going to define a space $T_\mu \subset \mathcal{T}/G$ for each pair $[f, \Gamma]$, and for each singleton V_1. Let $\mu = V_1$ be singleton. This means that V_1 is a connected space of rational maps, and $\widetilde{V}_1$ is a connected component of the preimage of V_1 in $\widetilde{B}$. Define

$$T_\mu = V_1.$$

Now let $\mu = [f_0, \Gamma]$ be a pair. Fix any $\delta > 0$ less than the Margulis constant, so that any nontrivial nonperipheral loop of length $\leqslant \delta$ is simple and any two such isotopically simple loops are isotopically disjoint. We define

$$T_\mu = \mathcal{T}(\Gamma, \delta)/G = \mathcal{T}(\Gamma, \varepsilon)/G(f_0, \Gamma),$$

since $G(f_0, \Gamma)$ is the subgroup of $g \in G$ for which $g \cdot \mathcal{T}(\Gamma, \delta) \cap \mathcal{T}(\Gamma, \delta) \neq \varnothing$. Sometimes we may wish to vary δ, in which case we shall write $T_\mu(\delta)$. But the spaces $T_\mu(\delta)$ for δ less than the Margulis constant are all homeomorphic. Also, the definition of T_μ is independent of the choice of (f_0, Γ) in the equivalence class, up to homeomorphism.

Let $\leqslant$ be the ordering on nodes and edges defined in 3.18. For the same μ, we define

$$T_{\geqslant\mu} = \bigcup\{T_\nu : \nu \geqslant \mu\}, \quad S_{\geqslant\mu} = \bigcup\{S_\nu : \nu \geqslant \mu\},$$

where the unions are over those pairs with are primitive or minimal nonempty. Thus if μ is a node or edge of $\mathcal{G}_n$, the union is over some nodes and edges of $\mathcal{G}_n$ Let N be a union of deleted neighbourhoods of codimension one critically finite spaces in $\overline{B} \setminus B$, with associated edges $\mathcal{E}(N)$, and define

$$T_N = \bigcup\{T_\varepsilon : \varepsilon \in \mathcal{E}(N)\}.$$

In 5.4 we defined a number $\kappa(f,\Gamma) \geqslant 0$ for pairs (f,Γ) satisfying the Invariance and Levy Conditions. We define $K[f,\Gamma] = \kappa(f,\Gamma)$. This definition makes sense. Note that if $\mu \leqslant \nu$ for pairs μ, ν then $\kappa(\mu) = \kappa(\nu)$, because $\kappa(\mu)$ is determined by the fixed gap. We also define $\kappa(V_1) = 0$ if V_1 is a singleton node of $\mathcal{G}_0$

We define

$$T_0 = \bigcup\{T_\mu : \kappa(\mu) = 0\}, \quad T_0' = T_0 \setminus V, \quad S_0 = \bigcup\{S_\mu : \kappa(\mu) = 0\}/\sim .$$

Here, $\sim$ denotes the equivalence relation given by identifying edge spaces S_ν into node spaces S_μ, for adjacent μ and ν, as in 5.7. Then we have the following. 1, 2, 3 and 4 are essentially obvious. 5 and 6 will be proved shortly.

***7.3. Theorem**.* *The following hold, if, in the case of 5 and 6, identifications between loops in sets S_μ, for pairs μ with $\kappa(\mu) = 0$, and singleton spaces V, are suitably defined.*

(1) *T_ν and S_ν are homotopy equivalent, for any putative node or edge of $\mathcal{G}_0$ and any putative node or edge pair of $\mathcal{G}_n$, $n > 0$.*

(2) *Let $\nu = [f_0, \Gamma]$ be a pair which is a node of $\mathcal{G}_n$ for some n, and let Σ be any set of adjacent edges, so that we have an inclusion $S_\varepsilon \subset S_\nu$ for $\varepsilon \in \Sigma$. Then the following are homotopy equivalent:*

$$(T_\nu, \bigcup\{T_\varepsilon : \varepsilon \in \Sigma\}) \quad \text{and} \quad (S_\nu, \bigcup\{S_\varepsilon : \varepsilon \in \Sigma\}).$$

(3) *$T_{\geqslant\mu}$ and $S_{\geqslant\mu}$ are homotopy equivalent for any pair node μ*

(4) *Given (B,N), $(T_{\geqslant\mu}, T_{\geqslant\mu} \cap T_N)$ and $(S_{\geqslant\mu}, S_{\geqslant\mu} \cap S_N)$ are homotopically equivalent for any pair node μ.*

(5) *T_0 and S_0 are homotopy equivalent.*

(6) *Given (B,N), $(T_0, T_N \cap T_0)$ and $(S_0, S_0 \cap PS_N)$ are homotopy equivalent.*

Most of the Proof. — We consider 1-4. Since the spaces involved are homotopy equivalent to locally finite simplicial complexes, and hence homotopy equivalent to CW-complexes [**Mi**], and are $K(\pi,1)$'s, it suffices ([**Spa**] 7.8) that the fundamental groups be isomorphic, under an isomorphism that preserves subgroups in the case of 2 and 4. This is true, simply by the construction of the spaces PS_μ, $PS_{\geqslant\mu}$. □

7.4. Proofs of 5 and 6: Intersections between V and T'_0. — The following is basically in Stimson's thesis [**Sti**]. We outline the proof given there.

Proposition

(1) *If δ is sufficiently small, and $V \cap T_\nu \neq \varnothing$ for some pair node ν, then $\kappa(\nu) = 0$, so that $T_\nu \subset T_{\geqslant \mu}$ for some minimal nonempty isometric edge μ.*

(2) *Let μ be a minimal nonempty isometric node. Then the components $T_{\geqslant \mu} \cap V$ are homotopic in $T_{\geqslant \mu}$ to the circles T_ε for the minimal nonempty edges $\varepsilon \leqslant \mu$, and in V to simple loops round punctures of V. This gives a one-to-one correspondence between minimal nonempty isometric edges and punctures in V.*

Proof (outline)

(1) First, we identify punctures in V, as in Stimson's thesis. Consider the maps

$$g_{c,d} : z \longmapsto 1 + \frac{c}{z} + \frac{d}{z^2}.$$

The critical points of $g_{c,d}$ are 0 and $-2d/c$. Fix any integer $k \geqslant 3$. There is a polynomial F such that $F(c,d) = 0$ if and only if 0 has period k under $g_{c,d}$. Similarly, there is a polynomial G such that $G(c,d) = 0$ if and only if $g^k_{c,d}(0)$ is fixed by $g_{c,d}$, for any $k \geqslant 2$. Our space V is either the zero set of such a polynomial, minus finitely many points, or a finite cover of such a zero set, minus finitely many points. The latter occurs if V is one of the spaces $V_{p,1}$ of 1.9. The excluded points include points above (c,d) for which the critical value $g^n_{c,d}(-2d/c)$ lands in $Z(g_{c,d})$, and, more importantly, those values for which $g_{c,d}$ degenerates to a Möbius transformation. Whenever this happens, the Möbius transformation is periodic, with period bounded by the size of the orbit of the critical point 0 of $g_{c,d}$ (which is constant over V). A periodic Möbius transformation is an isometry, and hence we are in $\mathcal{T}(\Gamma, \delta)$ for some isometric $[f_0, \Gamma]$. It only happens when either $d = 0$ or c, d are both large. In the latter case, by conjugating by $z \mapsto \sqrt{c}z$ we get the period two Möbius transformation $z \mapsto 1/z$. So at these singularities, the limiting Möbius transformation has two fixed points whose multipliers are roots of unity ($\neq 1$). Fix one of these fixed points. For some nearby maps in V, the multiplier will be inside the unit circle. Thus, any such puncture of V is in the closure of a hyperbolic component of some polynomial. The main result of Stimson's thesis [**Sti**] is that the puncture is in the closure of precisely two such hyperbolic components if the corresponding Möbius transformation has period $\geqslant 3$, and precisely one if the period is two. This is actually wrong. However, the method of Stimson's thesis (which is outlined below) gives a combinatorial rule for computing the number of polynomial hyperbolic components limiting on any puncture in terms. The number is always a power of two, as we showed in 3.12. This rule and Stimson's method (together with a correction of the mistake in the combinatorics) are outlined below. It is also shown in [**Sti**] that the only singular points of V are these Möbius transformation punctures (and this is correct).

(2) Now suppose that $\mu = [f_0, \Gamma]$ is a minimal isometric enhanced Levy set satisfying the Edge Condition. See 2.3 for a summary of the following notation, or 1.10 of [**R3**]. By 3.10, there are odd denominator rationals p and q such that the corresponding minor leaves μ_p, μ_q satisfy: q has the same period under $x \mapsto 2x$ mod 1 as 0 under maps in V, μ_p is minimal and $\mu_p \leqslant \mu_{1-q}$, and there is Γ' such that

$$[f_0, \Gamma] = [s_p \amalg s_q, \Gamma'].$$

Let f be the polynomial (up to Möbius conjugation) in V which is Thurston-equivalent to s_q. Let g be the polynomial $z \mapsto z^2 + c$ equivalent to s_p. Let h be the polynomial with a parabolic fixed point in the boundary of the hyperbolic component of g, and let ζ be the multiplier at that fixed point. Let H be the intersection with V of the hyperbolic component of f. Let R be the ray in H of maps which have multipliers $r\zeta$ ($r \in (0, 1)$) at the attractive fixed point. Then R limits on a (unique) puncture in V, for which the corresponding Möbius transformation has a fixed point with multiplier ζ. (If the limit of the ray was a rational map, it would have a parabolic fixed point with multiplier ζ, and we could draw a "Levy cycle" round the forward orbit of the critical points which was contracted under inverse images.)

So now we have a correspondence between the "degenerating" punctures of V and minimal isometric edges. It is not yet clear (as is claimed) that the correspondence is one-to-one. But now we analyse maps in V near a degenerating puncture. Let ζ_0 be the corresponding primitive k'th root of unity ($k \geqslant 2$) — not uniquely determined if $k \geqslant 3$, since we can replace ζ_0 by $\overline{\zeta}_0$. Let $p = p(\zeta_0)$ be such that s_p is Thurston equivalent to the unique critically finite polynomial $z \mapsto z^2 + c$ such that the boundary of the hyperbolic component of this polynomial contains a polynomial with parabolic fixed point and multiplier μ_0 at this fixed point. Then up to Möbius conjugation, the maps in V near this puncture are of the form

$$h_{\zeta,\rho} : z \longmapsto \zeta z\Big(1 - \frac{2+\rho}{2(1+\rho)}z\Big)\Big(1 - \frac{2}{2+\rho}z\Big)^{-1} = \zeta z P_\rho(z),$$

where ζ is close to ζ_0 and ρ is close to 0. This map has fixed points at $0, \infty$, multiplier ζ at 0 and critical points at $1, 1+\rho$. If $k = 2$, this representation is not unique, because conjugation by $z \mapsto 1/z$ allows us to replace ζ by $1/\zeta$ and ρ by $(1/1+\rho) - 1$.

For z near 1, the k'th iterate of this map is of the form

$$z \longmapsto \zeta^k z(P_\rho(z) + O(\rho^2)).$$

So for bounded b (and small ρ) we have

$$1 + b\rho \longmapsto \zeta^k\Big(1 + b\rho + \frac{\rho}{2(2b-1)} + O(\rho^2)\Big).$$

Since 1 is supposed to be periodic for maps in V, we see that $\zeta^k = 1 + O(\rho)$. By standard theory of singularities of algebraic curves, ζ has an expansion in powers of $\rho^{1/m}$ for some integer $m \geqslant 1$. So for some $a \in \mathbf{C}$, we can write

(1) $$\zeta^k = 1 + a\rho + o(\rho).$$

So then we have

$$1 + b\rho \longmapsto 1 + \rho\Big(a + b + \frac{1}{2(2b-1)}\Big) + o(\rho).$$

The map

$$b \longmapsto a + b + \frac{1}{2(2b-1)}$$

has fixed parabolic point at ∞ and critical points at 0, 1 (corresponding to 1 and $1+\rho$). If $a = 0$ then both critical points are attracted to 0. But 0 is supposed to be periodic of bounded period (since the same is true for 1 under $h_{\zeta,\rho}$), so $a \neq 0$, and only the critical point 1 is attracted to ∞. It follows that given $M > 0$ there is N such that $h_{\zeta,\rho}^{Nk}(1+\rho)$ is distance $\geqslant M\rho$ from 1. (For more detail see [**Sti**].)

Again, we outline from [**Sti**]. As (ζ, ρ) traces a circle in V round the puncture of V, the straight line segment $\tau_n = \tau_n(\zeta, \rho)$ joining $h_{\zeta,\rho}^{(n+1)k}(1+\rho)$ and $h_{\zeta,\rho}^{nk}(1+\rho)$ does not intersect the periodic orbit of 1 under $h_{\zeta,\rho}$ for $n \geqslant N$. The multiplier of the fixed point 0 of $h_{\zeta,\rho}$ is ζ. So if $|\zeta| < 1$, 0 is an attractive fixed point. So in the case $|\zeta| < 1$, $h_{\zeta,\rho}$ lies in the hyperbolic component of a rational map which is Möbius conjugate to a polynomial. Choose an odd denominator rational $q = q(\zeta, \rho)$ so that this polynomial is Thurston equivalent to s_q. There are actually two choices for q given by the two endpoints of μ_q.) Let $\varphi = \varphi_{\zeta,\rho}$ be the uniformising map so that

$$\varphi(0) = 0, \quad \varphi(\zeta z) = h_{\zeta,\rho}(\varphi(z)), \quad 1 + \rho = \varphi(\zeta).$$

Now suppose, further, that $\zeta = \lambda\zeta_0$ for real λ, $0 < \lambda < 1$. Then for sufficiently large n, the image under φ of the straight line segment joining $\lambda^{(n+1)k}$ and λ^{nk} is isotopic to τ_n relative to the forward orbits of 1, $1+\rho$. It follows by induction that the same is true for all $n \geqslant N$, and that there is a preimage $\tau = \tau(\zeta, \rho)$ of τ_N under $h_{\zeta,\rho}^{Nk-1}$ which joins $h_{\zeta,\rho}(1+\rho)$ and $h_{\zeta,\rho}^{k+1}(1+\rho)$, and is disjoint from the periodic orbit of 1. Moreover, $\tau = \tau_{\zeta,\rho}$ varies continuously with (ζ, ρ). All the critically finite maps

$$\sigma_{\tau_{\zeta,\rho}}^{-1} \circ h_{\zeta,\rho}$$

are Thurston equivalent and are also equivalent to $s_p \amalg s_q$ for $p = p(\zeta_0)$ and $q = q(\zeta, \rho)$ for any (ζ, ρ) in a polynomial hyperbolic component near the puncture. Let Γ be the loop set containing $L_p \cup L_{1-p}^{-1}$ such that $(s_p \amalg s_q, \Gamma)$ satisfies the Invariance, Levy and Edge conditions. By abuse of notation, use Γ to use the same loop set in the complement of the critical orbits of $\sigma_{\tau_{\zeta,\rho}}^{-1} \circ h_{\zeta,\rho}$. It follows that the homeomorphism φ corresponding (under Φ_1, as in 1.11) to a simple closed loop round a puncture preserves Γ, and of course, $1+\rho$ and the forward orbit of 1. It remains to show that the homeomorphism represents a generator of $\pi_1(T_\varepsilon) = G[s_p \amalg s_q, \Gamma]$, where $\varepsilon = [s_p \amalg s_q]$. We have seen in 3.11, 3.12 that $G[s_p \amalg s_q, \Gamma]$ is a cyclic group generated by Dehn twists round the loops of Γ, and that the Dehn twist round each of the loops of ∂P is the same, say L, and L has to be a multiple of 2^m for some $m \geqslant 1$. If $k \geqslant 3$ then L is also the degree of $\rho \mid U'$, where $U' \subset V$ is a neighbourhood of the puncture, and the set U

of maps $h_{\zeta,\rho}$ which are Möbius conjugate to maps in U' under conjugacies mapping 1 to 0 has two components, each mapped homeomorphically to U' under this map of conjugacy identification. So the degrees of ρ and $\zeta - \zeta_0$ on U and U' are all L. If $k = 2$ then the set of maps $h_{\zeta,\rho}$ mapping to U' in this way may have one or two components. We claim that there is one component if $m = 0$, and that there are two if $m \geqslant 1$. Let U be one of these components, and let $\pi : U \to U'$ be the projection. Since twists of the critical point $1 + \rho$ and the critical value $-1(1 + (1 + \frac{1}{2}a)\rho + O(\rho^2)$ round the fixed point $-1(1 + \frac{1}{2}a\rho + O(\rho^2)$ both contribute to the Dehn twist round γ_0, the degree of $\rho \mid U$ and of $(\zeta + 1) \mid U$ is $\frac{1}{2}L\deg(\pi)$. If $m = 0$, then ζ must trace a complete path round -1 on a circuit in U round the puncture. A hyperbolic component is passed through each time either $|\zeta| < 1$ or $|\zeta^{-1}| < 1 - O(\rho^2)$. This is only possible if the two hyperbolic components are conjugate, that is, $\deg(\zeta + 1) \mid U) = 1$, $\deg(\pi) = 2$, and hence $L = 1$. Now let $m \geqslant 1$. Then the homeomorphism φ is a composition $\sigma_1 \circ \sigma_2$, where σ_1 is an L-fold twist round the loop $\gamma_0 = \mu_p \cup \mu_{1-p}^{-1}$, and σ_2 is a composition of Dehn twists round strictly preperiodic loops (3.11). The Dehn twist σ_1 can be written as $\sigma_{1.1} \circ \sigma_{1.2}$, where the $\sigma_{1.1}$ and $\sigma_{1.2}$ are $L/2$-fold twists round loops parallel to, and either side of, γ_0. Then φ and $(s_p \amalg s_q)^{-1} \circ \varphi \circ (s_p \amalg s_q)$ are isotopic via an isotopy fixing $\gamma_0 \cup X(s_p \amalg s_q)$. It follows that the lift of a closed path in U' going once round the puncture and starting in a hyperbolic component is indeed a closed path: because fixed points are preserved, it must end in the same hyperbolic component, not just in a Möbius conjugate of it. Then the map from U to U' is a homeomorphism. So if $k \geqslant 3$, on a simple path round the puncture in U or U', ζ passes through L different regions with modulus < 1 that is, L different hyperbolic components of maps which are Thurston equivalent to maps $s_{q'}$ with $s_p \amalg s_{q'}$ Thurston equivalent to $s_p \amalg s_q$. In the case of $k = 2$, that is, $p = 1/3$, so that $s_p = s_{1/3} = s_{2/3} = s_{1-p}$, and if $m \geqslant 1$, the number of hyperbolic components is actually $2\deg(\zeta + 1) \mid U = L$, because such a hyperbolic component is passed through each time either $|\zeta| < 1$ or $|\zeta^{-1}| < 1 - O(\rho^2)$. But the number of distinct q' is L in each case. So, for all k, the hyperbolic components passed through always give L different q' such that $s_p \amalg s_{q'} \simeq s_p \amalg s_q$. By 3.11, there are exactly 2^m such q', giving $L \leqslant 2^m$. We already have $2^m \mid L$. So $2^m = L$. So the path round the puncture is indeed a generator of $\pi_1(T_\varepsilon)$. □

7.5. Rational Convex Hulls. — We have now completed the proof of the Topographer's View "to level 0", given the Injective on π_1 result of 5.10. The next step is to prove the Resident's View "to level 0" given Injective on π_1 and the Resident's View of Rational Maps Space (RVRMS). So we need the following. This, with RVRMS completes the proof of the Resident's View to level 0.

Lemma. — *Assuming injective on π_1 and RVRMS, the convex set $C(\widetilde{V}_1)$ defined in 5.5 is a single component of the complement of the Levy convex hulls $C(f, \Gamma)$.*

Proof. — Identify $\widetilde{V}_1$ with the open unit disc with boundary ∂D. The RVRMS says that a map ρ_2 maps ∂D to $\partial C(\widetilde{V}_1)$, with right and left limits existing at every point, and these coinciding except when the limits are $\partial C(f_0,\Gamma)$ for some minimal nonempty edge pair (f_0,Γ). (In fact, $\partial C(f_0,\Gamma) = C(f_0,\Gamma)$ if (f_0,Γ) is isometric.)Then every geodesic in $\partial C(\widetilde{V}_1)$ is of the form $\partial C(f_0,\Gamma)$ for some minimal nonempty edge (f_0,Γ). We shall call a connected component of the union of Levy convex hulls a *Levy component*, and shall call a *complementary component* just that. Thus, $C(\widetilde{V}_1)$ is a union of Levy components and complementary components. Their boundaries are permuted by $\pi_1(V_1)$.

Now we need to show that $C(\widetilde{V}_1)$ is a single complementary component. First, we claim that it cannot be contained in a single Levy component or its boundary. This is simply because $C(\widetilde{V}_1)$ is defined as the convex hull of points $C(\mu)$ for extreme pairs μ which do not satisfy the Levy Condition. In fact, these sets $C(\mu)$ must be dense in $\partial C(\widetilde{V}_1) \cap \partial D$. So now suppose that $C(\widetilde{V}_1)$ contains a Levy convex hull C_1. Then $\partial C(\widetilde{V}_1)$ intersects infinitely many components of $\partial D \setminus C_1$, because it cannot have open intersection with ∂C_1. Moreover, for any extreme pair μ, with $C(\mu) \in C(\widetilde{V}_1)$, $C(\mu) \notin \partial C_1$. Now let g be the parabolic element of $\pi_1(V_1)$ fixing the parabolic point x_0 with $\rho_2(x_0) = C(\mu)$ (using the realisation of $\pi_1(V_1)$ as a group of Möbius transformations). Let I be the component of $\partial C(\widetilde{V}_1) \setminus \partial C_1$ containing $C(\mu)$. Then $g \cdot I \cap I \neq \varnothing$ because $C(\mu) \in I$. Then $g \cdot I \subset I$ or $g^{-1}.I \subset I$, because $g \cdot C_1 = C_1$ or $g \cdot C_1 \cap C_1 = \varnothing$. Then $g \cdot \rho_2^{-1}(I) \subset \rho_2^{-1}(I)$ or $g^{-1} \cdot \rho_2^{-1}(I) \subset \rho_2^{-1}(I)$. This is impossible, because $\rho_2^{-1}(I)$ is an open neighbourhood of x_0, and x_0 is a parabolic point. So we have a contradiction, and $C(\widetilde{V}_1)$ is a single complementary component. □

7.6. Corollary. — *If $\rho_2 : \partial D = \partial\widetilde{V}_1 \to \partial C(\widetilde{V}_1)$ is continuous except at the specified exclusions, then ρ_2 is injective.*

Remark. — This shows that 4 of the RVRMS follows from 1-3.

Proof. — We now know that $\partial C(\widetilde{V}_1)$ is a single complementary component. If ρ_2 is not injective, then we can find an open interval $I \subset \partial D$, and, for $x \in I$, a point $y(x) \notin I$ such that $\rho_2(x) = \rho_2(y(x))$. Let $x \in I$ be fixed by a hyperbolic element g of $\pi_1(V_1)$, that is, g has precisely two fixed points on ∂D. We can assume without loss of generality that x is an attractive fixed point of g.Then $\lim_{n\to+\infty} g^{-n}(y(x)) = x'$ is the repelling fixed point of g. We can approximate any point of $I \times \partial D$ by such pairs (x,x'), since pairs of fixed points of hyperbolic elements are dense in $\partial D \times \partial D$ [**G-H**]. So $\rho_2(x) = \rho_2(x')$ for all $x' \in \partial D$, $x \in I$. So $\rho_2(\partial D)$ is a single point. This is impossible, because ρ_2 is injective on parabolic points by 1.12. □

7.7. The Level κ Tool and Descending Points. — We now state two results: the Level κ Tool, and Descending Points. We shall show that these, together with the Resident's View of Rational Maps Space, imply Injective on π_1, and the Topographer's and Resident's Views, that is, all the results of 5.10. Write

$$F(x) = d(x, \tau(x)).$$

This can be regarded as a function on $\mathcal{T}$ or on $\mathcal{T}/G$. Define

$$T'_\kappa(\varepsilon) = \bigcup\{T_\mu(\varepsilon) : \kappa(\mu) \leqslant \kappa\}.$$

The results depend on constants $D_i > 0$ and sets K_i, for $0 \leqslant i \leqslant 2$ which will be defined explicitly in 18.11. The number $\varepsilon_0 > 0$ is smaller than the Margulis constant. The constants E_i satisfy $E_i < E_{i+1}$. We do not define the sets K_i yet, but we write K_i indiscrinately for a subset of $\mathcal{T}/G$ and for the preimage in $\mathcal{T}$. We have $K_i \subset K_{i+1}$ for $0 \leqslant i \leqslant 2$, and $K_i \subset \mathcal{T}_{<\varepsilon_0}$ (or $\subset \mathcal{T}_{<\varepsilon_0}/G$). The K is supposed to indicate compactness in $\mathcal{T}/G$. (We shall use "C" for something else shortly.) If $E_0 > 0$ is large enough, and $\kappa(\mu) > 0$, then κ is bounded from 0, and, as we shall see in 17.4, for all sufficiently small ε, depending only on $\#(Y)$,

$$T_\mu(\varepsilon) \cap \{x : F(x) \leqslant \kappa(\mu) - E_0 e^{-2\pi^2/\varepsilon}\} = \varnothing.$$

The Level κ Tool then says that a certain subset of K_1 acts as a "plug" between the disjoint sets $T'_\kappa(\varepsilon)$ and $\{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon}\}$. Write

$$K_i(\mu, \varepsilon) = K_i \cap T_\mu(\varepsilon/(1 - E_i\varepsilon)) \cap \{x : F(x) \leqslant \kappa(\mu) - E_i^{-1} e^{-2\pi^2/\varepsilon}\}$$

if $\kappa(\mu) > 0$, and, if $\kappa(\mu) = 0$, simply take $K_i(\mu, \varepsilon) = \varnothing$. We shall assume in future that E_1 is chosen given E_0 so that, if $\varepsilon' \leqslant \varepsilon$, then

$$\begin{aligned}K_0(\mu, \varepsilon') &\subset K_1(\mu, \varepsilon) \cup T_\mu(\varepsilon),\\ K_0(\mu, \varepsilon) &\subset \{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon'}\} \cup K_1(\mu, \varepsilon').\end{aligned}$$

We then define

$$\begin{aligned}T''_\kappa(\varepsilon) = \bigcup\{K_1(\mu, \varepsilon) &: \mu \text{ minimal nonempty}, \kappa(\mu) \leqslant \kappa\}\\ &\cup T'_\kappa(\varepsilon) \cup \{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon}\},\\ T''_\kappa(\varepsilon, \varepsilon') = \bigcup\{K_0(\mu, \varepsilon'') &: \mu \text{ minimal nonempty}, \kappa(\mu) \leqslant \kappa, \varepsilon' \leqslant \varepsilon'' \leqslant \varepsilon\}\\ &\cup T'_\kappa(\varepsilon') \cup \{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon}\}.\end{aligned}$$

Then we have

$$T''_\kappa(\varepsilon, \varepsilon') \subset T''_\kappa(\varepsilon) \cap T''_\kappa(\varepsilon').$$

Level κ Tool. — *Let κ_0 be given. Let ε be sufficiently small, and ε/ε' sufficiently large, given κ_0. Let $2E_0 e^{-2\pi^2/\varepsilon} \leqslant \kappa \leqslant \kappa_0$. Let $\alpha : \Delta \to \mathcal{T}/G$ be continuous, where Δ is the unit interval, disc or circle, and $\partial \subset \partial\Delta$ with*

$$\alpha(\partial) \subset T''_\kappa(\varepsilon, \varepsilon'), \qquad d(\alpha(x), \tau \circ \alpha(x)) \leqslant \kappa \text{ for all } x \in \Delta.$$

Then α can be homotoped, via a homotopy constant on $\alpha \mid \partial$, to $\alpha' : \Delta \to \mathcal{T}/G$, with

$$\alpha'(\Delta) \subset T''_\kappa(\varepsilon).$$

Moreover, for κ_1 depending only on κ, ε'', if $\alpha(x) \in \mathcal{T}_{\geqslant \varepsilon''}$,

$$d(\alpha(x), \alpha'(x)) \leqslant \kappa_1.$$

Descending Points. — *The following holds for any minimal nonempty node $\mu = [f_0, \Gamma]$ with $\kappa_0 \geqslant \kappa(\mu) > 0$, and for ε sufficiently small given κ_0. Let P be the fixed set for (f_0, Γ) and let $p > 1$ be the least positive integer such that f_0^p fixes the components of ∂P.*

Then each component C of $K_i(\mu, \varepsilon)$ is contractible within $K_{i+1}(\mu, \varepsilon)$ to a nontrivial closed loop $\gamma_\nu = \gamma_\nu(\varepsilon)$ for a minimal edge $\nu \leqslant \mu$ such that the following holds. The loop $\gamma_\nu(\varepsilon)$ varies continuously with ε. The group $\pi_1(\gamma_\nu)$ (which is determined up to conjugacy in G) is the same central subgroup of $G(f_0, \Gamma)$ for all such ν. Using the usual representation of elements of G as homeomorphisms, the generator g_ν of $\pi_1(\gamma_\nu)$ satisfies $g_\nu = f_0^p$ on P, up to isotopy. This gives a one-to-one correspondence between components of $K_i(\mu, \varepsilon)$ and minimal edges $\nu < \mu$.

7.8. How to obtain Injective on π_1 from the Level κ Tool and Descending Points. — Let $\alpha : \Delta \to \mathcal{T}/G$ be continuous with $\alpha(\partial\Delta) \subset V$. Then there is $\kappa > 0$ such that $d(\alpha(x), \tau \circ \alpha(x)) \leqslant \kappa$ for all $x \in \Delta$. If κ is sufficiently small given $\delta_0 > 0$ and $\varepsilon > 0$, then by Property 6 of 6.6,

$$\alpha(\Delta) \subset T_0'(\varepsilon) \cup \{x : d(x, \widetilde{V}) < \delta_0\}.$$

Then we can homotope α to α', by a homotopy constant on $\partial\Delta$, so that $\alpha'(\Delta) \subset V \cup T_0' = T_0$. Then since the inclusion $V \hookrightarrow T_0$ is injective on π_1, by 7.4, we can ensure that $\alpha'(\Delta) \subset V$.

Now let κ be bounded from 0. Let ε_1 be sufficiently small given κ. It suffices to show that α is homotopic to α'' with $\alpha = \alpha''$ on $\partial\Delta$ and

(1) $$F(x) \leqslant \kappa - \eta_1 \text{ for all } x \in \alpha''(\Delta),$$

where

$$\eta_1 = E_2^{-1} e^{-2\pi^2/\varepsilon_1}.$$

For we can then repeat the argument until we obtain a path homotopic into V. First, we obtain α' as in the Level κ tool. Then we consider the boundary of the set

$$W = \{x : \alpha'(x) \in \bigcup\{T_{\geqslant\mu}(\varepsilon_1) : \mu \text{ minimal nonempty, } \kappa(\mu) > 0\}\}.$$

Perturbing α' slightly if necessary, we can assume that ∂W is a finite union of disjoint topological circles disjoint from $\partial\Delta$, since $\alpha(\partial\Delta) \subset V$. Then there are finitely many disjoint topological discs Δ_i, $1 \leqslant i \leqslant r$ such that

$$W \subset \cup_i \Delta_i, \quad \cup_i \partial\Delta_i \subset \partial W.$$

By the Level κ Tool,

$$\partial\Delta_i \subset \bigcup\{K_1(\mu, \varepsilon_1) : \mu \text{ minimal nonempty}\}.$$

Some components of ∂W may be interior to some of these discs. Then by Descending Points we can redefine α' in each Δ_i to obtain α'', where

$$\alpha''(\Delta_i) \subset \bigcup\{K_2(\mu,\varepsilon_1) : \mu \text{ minimal nonempty}\} \subset \{x : F(x) \leqslant \kappa - \eta_1\},$$

as required. □

7.9. How to obtain the Topographer's View. — We show how to obtain the Topographer's View from the Level κ Tool, Injective on π_1 and Descending Points.

The space $\mathcal{T}/G$ is obviously the increasing union of sets $\mathcal{T}_\kappa/G$ where

$$\mathcal{T}_\kappa = \{x : F(x) \leqslant \kappa\},$$

and F is as in 7.7. Let T_N be the subspace of $\mathcal{T}/G$ defined in 7.2. Recalling that F can equally well be regarded as a function on $\mathcal{T}/G$, we define

$$T_{N,\kappa} = T_N \cap \{x : F(x) \leqslant \kappa\}.$$

Let T_0 be as in 7.2. For $\kappa > 0$, we shall choose T_κ inductively to have the homotopy type of $T''_\kappa(\varepsilon)$ for any ε sufficiently small given κ. By Descending Points, the components of $K(\mu,\varepsilon)$ are in 1-1 correspondence with minimal nonempty edges $\nu \leqslant \mu$. Then by Descending Points of 7.7, we see that $T''_\kappa(\varepsilon)$ is homotopy equivalent to

$$(T'_\kappa(\varepsilon) \cup \mathcal{T}_{\kappa-\eta}/G)/\sim_\kappa,$$

where $\eta = E_0 e^{-2\pi^2/\varepsilon}$ and $\sim_k$ identifies a closed loop $\gamma_\nu(\varepsilon_1)$ from each component of $T'_\kappa(\varepsilon)$ with some closed loop in $\mathcal{T}_{\kappa-\eta}/G$. Then, by induction, $T''_\kappa(\varepsilon)$ is also homotopy equivalent to a space

$$T_\kappa = (T'_\kappa(\varepsilon) \cup T_{\kappa-\eta}/G)/\sim_\kappa .$$

Here, by abuse of notation, we use $\sim_\kappa$ to denote the two naturally corresponding equivalence relations. In fact, this makes sense, because we can regard all the spaces T_κ as being embedded in $\mathcal{T}/G \times (-1,1)$. We also see by Descending Points that $T''_\kappa(\varepsilon)$ is homotopy equivalent to $T''_\kappa(\varepsilon,\varepsilon')$ for all $\varepsilon' \leqslant \varepsilon$, via a homotopy with image in $\mathcal{T}_\kappa/G$.

Let $\alpha : S^1 \to \mathcal{T}/G$ or $\alpha : ([0,1],\{0,1\}) \to (\mathcal{T}/G, T_N)$ be continuous. Then we have $\alpha : S^1 \to \mathcal{T}_\kappa/G$ or $\alpha : ([0,1],\{0,1\}) \to (\mathcal{T}_\kappa/G, T_{N,\kappa})$ for some $\kappa > 0$. Then, by the Level κ tool, for all sufficiently small ε given κ, α can be homotoped to α', via a homotopy constant on $\{0,1\}$ in the second case, with

$$\alpha'(\Delta) \subset T''_\kappa(\varepsilon).$$

Now let $\alpha : \Delta \to \mathcal{T}/G$ be any continuous map with $\alpha(\partial) \subset T''_\kappa(\varepsilon_1)$, where Δ is the unit disc, and $\partial \subset \partial\Delta$. Then by compactness of Δ, there is $\kappa_1 \geqslant \kappa$ such that $\alpha(\Delta) \subset \mathcal{T}_{\kappa_1}/G$. Let ε and ε' be as in the Level κ tool with κ_1 replacing κ_0. Applying a homotopy if necessary (not constant on ∂), we can assume that $\alpha(\partial) \subset T''_\kappa(\varepsilon,\varepsilon')$, $\alpha(\Delta) \subset \mathcal{T}_{\kappa_1}/G$. Then by the Level κ tool, we can homotope α to α', via a homotopy constant on ∂, with $\alpha'(\Delta) \subset T''_{\kappa_1}(\varepsilon)$. Then by Descending Points, and the same technique as in 7.8, we can homotope α', via a homotopy constant on ∂, to α'' with $\alpha''(\Delta) \subset T''_{\kappa_2}(\varepsilon)$, where

$$\kappa_2 = \mathrm{Max}(\kappa, \kappa_1 - E_2^{-1} e^{-2\pi^2/\varepsilon}).$$

Then by applying the Level κ tool and the technique of 7.8 finitely many times, we can homotope α to α''', via a homotopy constant on ∂, so that

$$\alpha'''(\Delta) \subset T''_{\kappa}(\varepsilon).$$

Now $\mathcal{T}/G$ is homotopy equivalent to a locally finite simplicial complex, and hence is homotopy equivalent to a CW complex [**Mi**]. Since $\mathcal{T}/G$ is a $\kappa(\pi, 1)$ and $T_N \hookrightarrow \mathcal{T}/G$ is injective on π_1, by [**Spa**] 7.8, the above homotopies on maps α show that $(\mathcal{T}/G, T_N)$ is homotopy to the increasing union of sets $(T''_{\kappa}, T_{N,\kappa})$ and hence to the increasing union of sets $(T_{\kappa}, T_{N,\kappa})$. This completes the proof of the Topographer's View.

7.10. How to obtain the Resident's View. — We have now realised T_{κ_n} as a graph of topological spaces over a graph $\mathcal{G}_n$, and the universal cover as a graph of topological spaces over a graph $\widetilde{\mathcal{G}}_n$. To prove the Resident's View, it remains to show that $\widetilde{\mathcal{G}}_n$ is the dual of $\mathcal{P}_n$ and that $\mathcal{P}_n$ is a locally finite partition of a subset of the disc, and that $\mathcal{P}$ is a partition of the disc.

By 7.4, and the Resident's View of Rational Maps Space, $\widetilde{\mathcal{G}}_0$ is the dual of $\mathcal{P}_0$, which is a locally finite partition of any connected component of $\cup\mathcal{P}_0$. The stabiliser in G of each component of $\cup\mathcal{P}_0$ is $\pi_1(T_0)$, up to conjugacy. Now suppose inductively that we have proved this for $\widetilde{\mathcal{G}}_n$, $\mathcal{P}_n$ and $\pi_1(T_{\kappa_n})$. Then consider the case $n+1$. So we have to consider identifications between components of $\widetilde{S}_n$ and $\widetilde{S}_{\geqslant\mu}$ in $\widetilde{S}_{n+1}$. Here μ is a minimal nonempty pseudo-Anosov node pair of $\widetilde{\mathcal{G}}_{n+1}$, with $\kappa(\mu) = \kappa_{n+1}$, and $\widetilde{S}_n$, $\widetilde{S}_{\geqslant\mu}$ are the preimages of S_n, $S_{\geqslant\mu}$ in the universal cover $\widetilde{S}_{n+1}$ of S_{n+1}. Let P be the fixed set of μ, with p boundary components. Let $[\psi_P]$ be as in 2.13: this is a pseudo-Anosov isotopy class on P. Let $[\chi] = [\psi_P^{-p}]$. We can choose χ up to isotopy to preserve two transverse geodesic laminations on P, corresponding to the stable and unstable foliations of the pseudo-Anosov. There is a unique $g \in G$ (regarding g as an element of $\mathrm{MG}(\overline{\mathbf{C}}, Z)$, using the anti-isomorphism Φ_1 of 1.11) which equals χ on P, and is the identity off P, with zero Dehn twists round components of ∂P. (The map χ is specified on the two geodesic laminations, which include ∂P up to isotopy. Hence g is also specified on ∂P.) Let C be the component of

$$\bigcup\{C(f_0, \Gamma) : [f_0, \Gamma] \geqslant [\mu]\}$$

which contains $C(\mu)$. Here, $[\mu] = [f_0, \Gamma_1]$, if $\mu = (f_0, \Gamma_1)$. All components of ∂C project down to the same geodesic ℓ^- in the unstable geodesic lamination, adjacent to the component of the complement of the lamination which contains $v_1 \in Z$. This geodesic corresponds to a singular leaf of the unstable foliation of the pseudo-Anosov, and is fixed by χ. Now, considering the action of $g \in G$ on the unit disc D, g acts as a lift of the homeomorphism on $\overline{\mathbf{C}} \smallsetminus Z$, fixes all components of ∂C. Considering g as an element of $\pi_1(S_{\geqslant\mu}) \cong \pi_1(T_{\geqslant\mu}) \leqslant \pi_1(\mathcal{T}/G) = G$, the closed loop in $S_{\geqslant\mu}$ determined by g is the one which is freely homotopic to at least one closed loop in S_n, and in fact to finitely many, in 1-1 correspondence with minimal nonempty edges ν of $\mathcal{G}_{n+1}$ with $\nu \leqslant [\mu]$.

It remains to show that each component of ∂C is adjacent to a component of $\cup\mathcal{P}_n$, which is fixed by g. There is a unique half geodesic ℓ^+ in the stable lamination, starting from ℓ^-, corresponding to a singular leaf in the stable foliation. Let U be any component of $D \setminus C$. Then $\partial U \cap \partial C$ is a lift of ℓ^-, and by abuse of notation we write $\ell^- = \partial U \cap \partial C$. Take the lift of ℓ^+ which starts on $\partial U \cap \partial C$ and lies in U. By abuse of notation we call the lift ℓ^+ also. Let x_U be the (unique) endpoint in $\partial U \cap \partial D$. Then $g \cdot x_U = x_U$, and for all $x \in \partial D \cap \overline{U} \setminus \{x_U\}$,

$$\lim_{n\to\infty} g^n \cdot x \in \partial C. \tag{1}$$

We see this as follows. We can assume that g fixes the start point of ℓ^+ on ℓ^-. We identify x with a half-geodesic ℓ which starts on ℓ^-, at the same point as ℓ^+, and with a lift in U which starts on ∂C and ends at x. We call lift ℓ also. The lifts and projections are shown in the diagram.

U and C.

Regarding g as a homeomorphism of $\overline{\mathbf{C}} \setminus Z$, $g^n\ell$, the initial segment of $g^n\ell$ approaches ℓ_- as n increases. Then regarding g as a homeomorphism of the closed disc, $g^n\widetilde{\ell}$ approaches ∂C in the Euclidean metric. This gives (1). Hence, if U contains a component of $\cup\mathcal{P}_n$ which is fixed by g, then that component must be adjacent to C. There must be such a component, because the orbits (under G) of such components are in 1-1 correspondence with minimal nonempty edges ν of $\mathcal{G}_n$ with $\nu \leqslant [\mu]$. This completes the proof that the graph $\widetilde{\mathcal{G}}_{n+1}$ is the dual of $\mathcal{P}_{n+1}$.

It follows that sets of $\mathcal{P}_{n+1}$ do not accumulate on $C(\mu)$, if $\kappa(\mu) = \kappa_{n+1}$. Hence the partition $\mathcal{P}_{n+1}$ is locally finite restricted to any component of $\cup\mathcal{P}_{n+1}$. Finally, we need to see that $\mathcal{P}$ is indeed a partition of the disc. We already know that any two sets of $\mathcal{P}$ have disjoint interiors. First we show that the open disc contained in a single component of $\cup\mathcal{P}$. We use the map

$$\rho_2 : \pi_1(B, N, f_0) \longrightarrow \pi_1(\overline{\mathbf{C}} \setminus Z, Z, v_1)$$

of 1.12. This is map is trivially surjective, because

$$\rho_2([t \longmapsto \sigma_{\alpha_t} \circ f_0]) = [\alpha].$$

Here, we take $\alpha_t = \alpha \mid [0,t]$, $\alpha : [0,1] \to \overline{\mathbf{C}}$. But regarding $\pi_1(\overline{\mathbf{C}} \setminus Z, Z, v_1)$ as a subset of $\partial D = \widetilde{\overline{\mathbf{C}} \setminus Z}$, it is clearly dense. Up to homotopy, the path $t \mapsto \sigma_{\alpha_t} \circ f_0$ lies in $\mathcal{T}_\kappa/G$ for some $\kappa > 0$. It follows that $\rho_2([\alpha]) \in \overline{U}$, where U is a component of $\cup\mathcal{P}_n$, for some n. Hence a single component of $\cup_n(\cup\mathcal{P}_n)$ contains the whole open disc. Any set in $\mathcal{P}$ is in $\mathcal{P}_n$ for a least n. So $\mathcal{P}$ is indeed a partition.

PART II

TEICHMÜLLER DISTANCE

CHAPTER 8

L^1 ESTIMATES ON THE DISTORTION AND THE FIRST DERIVATIVE OF TEICHMÜLLER DISTANCE

8.1. All of the following nine chapters are concerned with Teichmüller distance on Teichmüller spaces of marked spheres. Some of the theory generalizes easily to other Teichmüller spaces of finite type surfaces. We use the notation introduced in Chapter 6. Thus, if $Y \subset \overline{\mathbf{C}}$ is finite, then

$$\mathcal{T}(Y) = \{[\varphi] : \varphi \in \mathrm{Hom}_+(\overline{\mathbf{C}})\}.$$

Here $[\varphi_1] = [\varphi_2]$ if and only if there is a Möbius transformation σ such that $\sigma \circ \varphi_1$ and φ_2 are isotopic via an isotopy constant on Y. We shall occasionally write $[\varphi]_Y$ if more than one Teichmüller space is being considered. We use the Teichmüller distance, introduced in 6.4, defined by

$$d([\varphi], [\psi]) = \inf\{\tfrac{1}{2} \log \|\chi\|_{\mathrm{qc}} : [\chi \circ \varphi] = [\psi]\},$$

where

$$\|\chi\|_{\mathrm{qc}} = \|K(\chi)\|_\infty, \quad K(\chi)(z) = \lambda(z)/\mu(z),$$

where $\lambda(z)^2 \geqslant \mu(z)^2 \geqslant 0$ are the eigenvalues of $D\chi_z^T D\chi_z$, and $D\chi_z$ is the derivative of χ at z (considered as a 2×2 matrix). We shall occasionally write d_Y if more than one Teichmüller space is being considered.

The theory says, of course, that the infimum in the definition of distance d above is attained uniquely by a quasi-conformal homeomorphism χ with the following form, if $d([\varphi], [\psi]) > 0$. There are quadratic differentials $q(z)dz^2$, $p(z)dz^2$ of unit area with at most simple poles, at most at the points $\varphi(Y)$, $\psi(Y)$ respectively, such that if a coordinate $\zeta = \xi + i\eta$ is given by

$$\zeta(z) = \int_{z_0}^{z} \sqrt{q(t)}dt$$

and similarly for a coordinate ζ' and p, then χ can be expressed in these coordinates as

$$\xi + i\eta \longmapsto \sqrt{K}\xi + i\eta/\sqrt{K}.$$

Thus, $K(\chi)$ is the constant K. In this situation,we shall say that $q(z)dz^2$ is the *quadratic differential at* $[\varphi]$ *for* $d(([\varphi],[\psi])$, and $p(z)dz^2$ *is the stretch of* $q(z)dz^2$ *at* $[\psi]$. We may also say that $p(z)dz^2$ is the stretch of $q(z)dz^2$ *by factor* $\sqrt{K}$. It follows that $-p(z)dz^2$ is the quadratic differential at $[\psi]$ for $d([\psi],[\varphi])$, and $-q(z)dz^2$ is the stretch of $-p(z)dz^2$ at $[\varphi]$, by factor $\sqrt{K}$.

The standard proof that the infimum of quasi-conformal distortion is achieved uniquely gives a bound on the distortion of any quasiconformal homeomorphism χ, in terms of how close it comes to achieving this minimum distortion. The bound is, in fact, an L^1 bound. Some of the results in this chapter elaborate this principle. We also use the principle to give a formula for the first derivative of Teichmüller distance (for marked spheres). While a general formula is known [**Ear**], both the formula given in our special case of marked spheres, and the simple-minded proof, appear to be new. At the end of the chapter, we specialise to consider the first derivative of the function $d(x,\tau x)$, where τ is one of the pullback functions on $\mathcal{T}(Y)$ defined in 6.7.

We want to conserve y for elements of Y. For this reason, we shall write the standard area element in the plane in the form

$$\frac{dz\wedge d\overline{z}}{2i},$$

simply to avoid writing $z=x+iy$ and using $dxdy$. If it seems reasonable, we shall avoid writing the area element altogether.

8.2. An Estimate of Quasi-Conformal Distortion. — The following lemma is gleaned from [**Abi**], and will prove very useful.

Lemma. — *Let* $Y\subset\overline{\mathbf{C}}$ *be finite, and let* $[\varphi]$, $[\psi]\in\mathcal{T}(Y)$. *Let* χ, χ_1 *be two quasi-conformal homeomorphisms with* $[\chi_1\circ\varphi]=[\chi\circ\varphi]=[\psi]$. *Let* χ *be the homeomorphism minimizing distortion* $K(\chi)$, *so that* $K(\chi)=K$ *is constant. Write* $K(\chi_1)(z)=K_1(z)$. *Let* $q(z)dz^2$ *be the quadratic differential at* $[\varphi]$ *for* $d([\varphi],[\psi])$. *Then*

$$(1),\qquad K\leqslant\int K_1(z)|q(z)|\frac{dz\wedge d\overline{z}}{2i}$$

with equality if and only if $\chi_1=\chi$.

Remark. — The proof is much as it is given in [**Abi**], where it is used to show simply that $K\leqslant\text{ess sup}K(\chi_1)$, which clearly follows.

Proof. — We use the coordinate $\xi+i\eta$ as in 8.1, and similarly for the range. Write $f=\chi^{-1}\circ\chi_1$, so that, if $f=f_1+if_2$, $\chi_1=\chi_{1,1}+i\chi_{1,2}$ for real f_1, f_2, $\chi_{1,1}$, $\chi_{1,2}$, then $f_1=\dfrac{1}{\sqrt{K}}\chi_{1,1}$. Then, as in [**Abi**],

$$(2)\qquad 1\leqslant\int\left|\frac{\partial f_1}{\partial\xi}\right|d\xi d\eta,$$

and hence

$$(3)\qquad \sqrt{K} \leqslant \int \frac{\partial \chi_{1,1}}{\partial \xi} d\xi d\eta \leqslant \int \sqrt{\sup\{\langle D(\chi_1)_\zeta^T D(\chi_1)_\zeta v, v\rangle : \|v\| = 1\}} d\xi d\eta$$
$$\leqslant \int \lambda(\zeta) d\xi d\eta,$$

where $0 < \mu(\zeta)^2 \leqslant \lambda(\zeta)^2$are the eigenvalues of $D(\chi_1)_\zeta^T D(\chi_1)_\zeta$. But, for $\lambda = \lambda(\zeta)$, $\mu = \mu(\zeta)$,

$$(4)\qquad \lambda = \sqrt{\frac{\lambda}{\mu}} \cdot \sqrt{\lambda\mu}.$$

The squares of the terms on the right are, respectively, $K(\chi_1)(\zeta)$ and $\det D(\chi_1)_\zeta$. So, by Cauchy-Schwartz,

$$(5)\qquad \Big(\int \lambda(\zeta) d\xi d\eta\Big)^2 \leqslant \int K(\chi_1)(\zeta) d\xi d\eta \int \det D(\chi_1)_\zeta d\xi d\eta.$$

By the choice of coordinates (with q and p both having mass 1),

$$(6)\qquad \int \det D(\chi_1)_\zeta d\xi d\eta = 1.$$

Then (2) to (6) give (1), because $d\xi d\eta = (1/2i)|q(z)| dz \wedge d\overline{z}$. □

8.3. Lemma. — *Continue with the same hypotheses and notation as in 8.2. The following holds for a constant $L > 0$. Let $\theta(z)$ be the angle in $(-\pi/2, \pi/2]$ between the directions of maximum dilatation of χ and χ_1 at z. Then*

$$K \leqslant \int K_1|q| - L \int |\theta|^2 (K_1 - 1)|q|.$$

Remark. — Suppose that K_1 is constant. Let $M > 1$ be given. Then, by taking logs, it follows that there are constants $C_1(M) > 0$, $C_2(M) > 0$ such that, if $K_1 \leqslant M$ then

$$d([\varphi], [\psi]) \leqslant \tfrac{1}{2} \log K_1 \Big(1 - C_1(M) \int |\theta|^2 |q|\Big),$$

and if $K_1 \geqslant M$ then

$$d([\varphi], [\psi]) \leqslant \tfrac{1}{2} \log K_1 - C_2(M) \int |\theta|^2 |q|.$$

The first of these has already been used, in 6.11.

Proof. — The proof is simply a more painstaking version of 8.2. Use the same coordinates ζ, ζ' as before on domain and range. Let $\theta'(\zeta)$ be the angle between the directions of maximum dilatation of χ^{-1} and χ_1^{-1} at $\chi_1(\zeta)$. Then

$$D\chi_\zeta^{-1} = \begin{pmatrix} 1/\sqrt{K} & 0 \\ 0 & \sqrt{K} \end{pmatrix},$$

$$D(\chi_1)_\zeta = \sqrt{\det D\chi_1} \begin{pmatrix} \cos\theta' & -\sin\theta' \\ \sin\theta' & \cos\theta' \end{pmatrix} \begin{pmatrix} \sqrt{K_1} & 0 \\ 0 & 1/\sqrt{K_1} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

So we find that, for f_1 as in the previous lemma,

$$\left|\frac{\partial f_1}{\partial \xi}\right| \leqslant \sqrt{\frac{\det D(\chi_1)_\zeta}{\sqrt{K}}}\left(\sqrt{K_1}\cos(|\theta|-|\theta'|)+\left(\frac{1}{\sqrt{K_1}}-\sqrt{K_1}\right)|\sin\theta\sin\theta'|\right).$$

Write

$$a(\zeta)=\sqrt{\det D(\chi_1)_\zeta}, \quad b(\zeta)=|\cos(|\theta(\zeta)|-|\theta'(\zeta)|)|, \quad c(\zeta)=|\sin\theta(\zeta)\sin\theta'(\zeta)|.$$

Then

$$0\leqslant a,\, b,\, c, \quad b,\, c\leqslant 1,\ \|a\|=1.$$

Write

$$\langle k,m\rangle=\int k(\zeta)m(\zeta)d\xi d\eta,$$

(all functions involved are real-valued), and

$$\|k\|^2=\langle k,k\rangle.$$

So

$$\sqrt{K}+\sqrt{K}\int\left(\left|\frac{\partial f_1}{\partial\xi}\right|-1\right)d\xi d\eta\leqslant\left\langle\sqrt{K_1}a, b+\left(\frac{1}{K_1}-1\right)c\right\rangle.$$

So by (2) of the lemma,

$$\sqrt{K}\leqslant\left\langle\sqrt{K_1}a, b+\left(\frac{1}{K_1}-1\right)c\right\rangle.$$

So we obtain

$$\langle\sqrt{K_1},a\rangle-\sqrt{K}\geqslant\left\langle\sqrt{K_1}a, 1-b+\left(1-\frac{1}{K_1}\right)c\right\rangle.$$

Now for any k, m,

$$\|k\|\,\|m\|-\langle k,m\rangle=\frac{\|\,\|m\|k-\|k\|m\,\|^2}{2\|k\|\,\|m\|}.$$

Apply this with $k=a$ and $m=\sqrt{K_1}$. Remember that $\|a\|=1$. We write

$$\lambda=\|m\|=\sqrt{\int K_1}.$$

So

$$\begin{aligned}\lambda-\sqrt{K}&=\langle\sqrt{K_1},a\rangle-\sqrt{K}+\lambda-\langle a,\sqrt{K_1}\rangle\\&\geqslant\left\langle\sqrt{K_1}a, 1-b+\left(1-\frac{1}{K_1}\right)c\right\rangle+\frac{1}{2\lambda}\left\|\lambda a-\sqrt{K_1}\right\|^2.\end{aligned}$$

The righthand side is a sum of positive terms. We now assume that λ/K is bounded, because otherwise the proof is finished. So if we write $\sqrt{K}\delta=\lambda-\sqrt{K}$, we obtain, for a constant M_1,

$$\left\|\lambda a-\sqrt{K_1}\right\|\leqslant M_1\sqrt{K}\,\delta.$$

So

$$\begin{aligned}\sqrt{K}\delta &\geqslant \langle\sqrt{K_1}a, 1-b\rangle = \left\langle \lambda a, \frac{\sqrt{K_1}(1-b)}{\lambda}\right\rangle \\ &\geqslant \left\langle \sqrt{K_1}, \frac{\sqrt{K_1}(1-b)}{\lambda}\right\rangle - M_1\sqrt{K\delta}\frac{\|\sqrt{K_1}(1-b)\|}{\lambda} \\ &\geqslant \frac{C_1}{\sqrt{K}}\int K_1(1-b) - C_2\sqrt{\delta}\|\sqrt{K_1}(1-b)\|\end{aligned}$$

for constants $C_1 > 0$, $C_2 > 0$. Since $(1-b) \leqslant (1-b)^{1/2} \leqslant 1$, we deduce that

$$\int K_1(1-b) \leqslant M_2\delta K.$$

Similarly, we deduce that

$$\int (K_1 - 1)c \leqslant M_3\delta K.$$

So we have

$$\int K_1||\theta| - |\theta'||^2 \leqslant M_4\delta K, \qquad \int |\theta\theta'|(K_1-1) \leqslant M_4\delta K.$$

Now write

$$|\theta|^2 = |\theta||\theta'| + |\theta|(|\theta| - |\theta'|), \quad I = \int |\theta|^2(K_1 - 1).$$

This gives, using Cauchy-Schwarz,

$$I \leqslant M_4\delta K + \sqrt{M_4\delta K}\sqrt{I}.$$

So $I \leqslant M_5\delta K$, that is,

$$\int |\theta|^2(K_1-1) \leqslant M_5 K\delta \leqslant M_5\sqrt{K}\left(\sqrt{\int K_1} - \sqrt{K}\right) \leqslant L^{-1}\left(\int K_1 - K\right),$$

as required. □

8.4. The Derivative of the Teichmüller Distance Function. — Let $[\varphi] \in \mathcal{T}(Y)$. Let $\underline{h} = (h(y)) \in \mathbf{C}^{Y\smallsetminus\{\infty\}}$ with all $\|\underline{h}\|$ small. We define $[\varphi + \underline{h}]$ to be $[\varphi'] \in \mathcal{T}(Y)$ close to $[\varphi]$ in the uniform topology, with $\varphi'(y) = \varphi(y) + h(y)$. To simplify subsequent notation, if $\infty \in Y$, we define $h(\infty) = 0$. Similarly, we write $[\psi + \underline{h}']$. Thus we have charts round each of $[\varphi]$, $[\psi]$ in $\mathcal{T}(Y)$. Fix any $Y_1 \subset Y$ such that $Y \smallsetminus Y_1$ contains 3 points, and $\infty \notin Y_1$. Then

$$\{[\varphi + \underline{h}] : h(y) = 0 \text{ for } y \notin Y_1\}$$

is a very natural chart for the complex manifold structure of $\mathcal{T}(Y)$ at $[\varphi]$. Of course, such simple charts are not available for Teichmüller spaces of other finite type surfaces.

Using these natural coordinates, we now give a formula for the first derivative of Teichmüller distance. The formula obviously recalls the result of [**Ear**], that the first derivative is essentially given by the quadratic differential. Since Earle's result was for the Teichmüller space of a general finite type surface, it naturally used the formalism of Beltrami differentials, and the duality between quadratic differentials and Beltrami differentials. But there is no reference in the following statement, or in the proof, to Beltrami differentials.

The Derivative Formula. — *Let $[\varphi]$, $[\psi] \in \mathcal{T}(Y)$ with $d([\varphi],[\psi]) > 0$. Let $q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi],[\psi])$, and let $p(z)dz^2$ be the stretch of $q(z)dz^2$ at $[\psi]$. Write*

$$I_1(\underline{h}, q) = 2\pi \operatorname{Re}\Big(\sum_{y\in Y} \operatorname{Res}(q, \varphi(y))h(y)\Big).$$

Then

$$d([\varphi+\underline{h}],[\psi+\underline{h}']) = d([\varphi],[\psi]) + I_1(\underline{h}, q) - I_1(\underline{h}', p) + o(\underline{h}) + o(\underline{h}').$$

In particular, $(\varphi,\psi) \mapsto d(\varphi,\psi)$ is differentiable.

Proof. — It suffices to prove the formula with $\underline{h}' = 0$, provided the $o(\underline{h})$ and $o(\underline{h}')$ terms are uniform on compact subsets of $([\varphi],[\psi])$, because

$$d([\varphi+\underline{h}],[\psi+\underline{h}']) = d([\psi+\underline{h}'],[\varphi+\underline{h}]).$$

This is the familiar fact that a function with continuous partial derivatives is continuously differentiable, together with the fact that $-p(z)dz^2$ is the quadratic differential at $[\psi]$ for $d([\varphi],[\psi])$, with stretch $-q(z)dz^2$ at $[\varphi]$. So we are using the continuity of the map

$$([\varphi],[\psi]) \longmapsto (q(z)dz^2, p(z)dz^2),$$

which is proved, for example, in [**Abi**].

Furthermore, we need only prove that

$$d([\varphi+\underline{h}],[\psi]) \leqslant d([\varphi],[\psi]) + I_1(\underline{h}, q) + o(\underline{h}),$$

again, provided the $o(h)$ term is uniform on compact subsets of φ. For let $q_1(z)dz^2$, $p_1(z)dz^2$ denote the quadratic differentials at $[\varphi+\underline{h}]$, $[\psi]$ for $d([\varphi+\underline{h}],[\psi])$. Since q_1 is close to q for small $\underline{h}$, we obtain

$$d([\varphi],[\psi]) \leqslant d([\varphi+\underline{h}],[\psi]) - I_1(\underline{h}, q) + o(\underline{h}),$$

which gives equality, as required. We shall also assume that all residues of $q(z)dz^2$ at points $\varphi(Y)$ $(y \in Y)$ are $\neq 0$, since the result will then follow at points with zero residues by continuity. (This is just for convenience: it is not hard to do the exact calculation in that case also.)

We use 8.2. Let χ and $K(\chi)$ be as in 8.2. Note that $K(\chi)$ is constant, and, as in 8.2, we write $K = K(\chi)$. We shall construct a family $\chi_{\underline{h}}$ of quasi-conformal homeomorphisms with the following properties.

1. $$\chi_0 = \chi,\ \chi_{\underline{h}}((\varphi+\underline{h})(w)) = \psi(w)$$

2. $$\int (K(\chi_{\underline{h}})(z) - K)|q| \frac{dz\wedge d\overline{z}}{2i} = I(\underline{h})$$

is real-analytic in $\underline{h}$.

3. $$|K(\chi_{\underline{h}})(z) - K| = O(\underline{h})$$

uniformly in z, $\underline{h}$ and $= 0$ near the poles of q.

Let $q_{\underline{h}}$ denote the quadratic differential for the minimising quasi-conformal homeomorphism from $[\varphi + \underline{h}]$ to $[\psi]$. Then 3 will imply

$$\int (K(\chi_{\underline{h}})(z) - K)|q_{\underline{h}}| \frac{dz \wedge d\overline{z}}{2i} = I(\underline{h}) + o(\underline{h}).$$

This will then imply, by 8.2, that

$$d([\varphi + \underline{h}], [\psi]) \leqslant d([\varphi], [\psi]) + \frac{I(\underline{h})}{2K} + o(\underline{h}).$$

It is probably worth noting that this alone implies that $[\varphi_1] \mapsto d([\varphi_1], [\psi])$ is differentiable at $[\varphi_1] = [\varphi]$ with derivative $DI(0)$. So any mistakes in our subsequent calculation do not greatly matter. Anyway, the result will be proved if we can construct $\chi_{\underline{h}}$ so that, in addition,

4. $$I(\underline{h}) \leqslant 2KI_1(\underline{h}, q).$$

We shall, in fact, work locally. Fix $y \in Y$ and assume without loss of generality that $\varphi(y) = \psi(y) = 0$. Also assume without loss of generality that $\mathrm{Res}(q, 0) = \mathrm{Res}(p, 0) = 1$. We write h for $h(y)$, and for $\delta > 0$ sufficiently small we consider $|h| < \delta$ only. Then it suffices to construct

$$\chi_h : \overline{\mathbf{C}} \longrightarrow \overline{\mathbf{C}}$$

with the following properties.

5. $$\chi_h = \chi \text{ except in } \{z : |z| \leqslant r_1\}.$$

6. $$\chi_h(z) = \chi(z - h) \text{ in } \{z : |z| \leqslant r_0\}$$

7. $$\int_A (K(\chi_h)(z) - K) \frac{1}{|z|} \frac{dz \wedge d\overline{z}}{2i} = 4\pi K \,\mathrm{Re}\, h + o(h), \quad |K(\chi_h(z) - K| = O(h).$$

Here, $\delta < r_0$ and r_1 is small, and $A = \{z : r_0 \leqslant |z| \leqslant r_1\}$. Then χ_h is defined as follows.

$$\chi_h = (1 - f(z))\chi(z) + f(z)\chi(z - h)$$

where

$$f = 1 \text{ on } |z| \leqslant r_0, \quad f = 0 \text{ on } |z| \geqslant r_1, \quad f(z) = \frac{r_1 - |z|}{r_1 - r_0} \text{ for } r_0 \leqslant |z| \leqslant r_1.$$

It remains to prove 7. It is convenient to consider the functions ζ_h and $\zeta = \zeta_0$, where

$$\zeta_h(z^{1/2}) = \sqrt{\chi_h(z)}.$$

Thus, $K(\zeta_h)(z^{1/2}) = K(\chi_h)(z)$ for $z \neq 0$. The advantage is that ζ is given in $|z| \leqslant r_1^{1/2}$ by a (real)-linear map with matrix

$$\begin{pmatrix} \sqrt{K} & 0 \\ 0 & 1/\sqrt{K} \end{pmatrix} = D.$$

Let A' be the double cover of A. If $t(z) = z^{1/2}$ then the determinant of t is $1/(4|z|)$. So it remains to prove

$$8. \qquad \int_{A'} (K(\zeta_h)(z) - K) \frac{dz \wedge d\overline{z}}{2i} = \pi K \operatorname{Re}(h) + o(h), \quad |K(\zeta_h)(z) - K| = O(h).$$

Now we see that

$$\zeta_h(z) = \zeta(z) + \frac{f(z^2)}{2\zeta(z)} \left(\left(\zeta\left(z - \frac{h}{2z} \right) \right)^2 - (\zeta(z))^2 \right) + O(h^2)$$

We can rewrite this as

$$\zeta_h(z) = \zeta(z) - F(z) D\zeta_z \left(\frac{h}{z} \right) + O(h^2),$$

where

$$F(z) = \frac{f(z^2)}{2},$$

and $D\zeta_z$ denotes the derivative of the (real)-differentiable function ζ at z. Write

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = E(z) = D(\zeta_h - \zeta)_z.$$

Thus, for fixed h, a, b, c, d are functions of z.

8.5. Lemma. — $K(\zeta_h)(z) = K + \sqrt{K}(a - Kd) + O(h^2)$.

Proof

$$\operatorname{trace}(D+E)^T(D+E) = K + \frac{1}{K} + 2\left(a\sqrt{K} + \frac{d}{\sqrt{K}} \right) + O(h^2),$$
$$\det(D+E) = 1 + \frac{a}{\sqrt{K}} + d\sqrt{K} + O(h^2).$$

The ratio of these is $s + s^{-1}$, where $s = K(\zeta_h)(z)$ is the distortion. Then

$$s + \frac{1}{s} = K + \frac{1}{K} + 2a\sqrt{K} - a\sqrt{K} - \frac{a}{K\sqrt{K}} + \frac{2d}{\sqrt{K}} - dK\sqrt{K} - \frac{d}{\sqrt{K}} + O(h^2).$$

So

$$s + \frac{1}{s} = K + \frac{1}{K} + \frac{1}{\sqrt{K}} \left(K - \frac{1}{K} \right) (a - Kd) + O(h^2).$$

Writing $s = K(1 + \mu)$, we have

$$\mu \left(K - \frac{1}{K} \right) = \frac{1}{\sqrt{K}} \left(K - \frac{1}{K} \right) (a - Kd) + O(h^2),$$

which gives the required result. □

8.6. Now we return to the proof of the derivative formula, that is, of 8 of 8.4. Then it remains to prove

$$\int_{A'} \sqrt{K}(a - Kd) = K\pi \operatorname{Re}(h) + o(h), \quad |a - Kd| = O(h).$$

Write

$$G(z) = \zeta_h(z) - \zeta(z).$$

Of course, G depends on h, but we simplify the writing. We have

$$\frac{\partial G}{\partial z} + \frac{\partial G}{\partial \overline{z}} = a + ic, \quad \frac{\partial G}{\partial \overline{z}} - \frac{\partial G}{\partial z} = ib - d.$$

So

$$\sqrt{K}(a - Kd) = \sqrt{K} \operatorname{Re}\Big((1 - K)\frac{\partial G}{\partial z} + (1 + K)\frac{\partial G}{\partial \overline{z}}\Big).$$

Now we have

$$\zeta(z) = \Big(\sqrt{K} + \frac{1}{\sqrt{K}}\Big)\frac{z}{2} + \Big(\sqrt{K} - \frac{1}{\sqrt{K}}\Big)\frac{\overline{z}}{2},$$

and for any complex number w,

$$D\zeta_z w = \Big(\frac{\partial \zeta}{\partial z} w + \frac{\partial \zeta}{\partial \overline{z}} \overline{w}\Big).$$

So

$$D\zeta_z\Big(\frac{h}{z}\Big) = \frac{1}{2\sqrt{K}}\Big((K + 1)\frac{h}{z} + (K - 1)\frac{\overline{h}}{\overline{z}}\Big).$$

It follows that

$$G(z) = L(z) + M(z) + O(h^2),$$

where

$$L(z) = \frac{1}{4(r_1 - r_0)\sqrt{K}}\big((K + 1)h\overline{z} + (K - 1)\overline{h}z\big),$$

$$M(z) = \frac{-r_1}{4(r_1 - r_0)\sqrt{K}}\Big((K + 1)\frac{h}{z} + (K - 1)\frac{\overline{h}}{\overline{z}}\Big).$$

Then

$$\begin{aligned}\sqrt{K} \operatorname{Re}\Big((1 - K)\frac{\partial L}{\partial z} + (1 + K)\frac{\partial L}{\partial \overline{z}}\Big) &= \frac{1}{4(r_1 - r_0)} \operatorname{Re}((K + 1)^2 h - (K - 1)^2 \overline{h}) \\ &= \frac{K}{r_1 - r_0} \operatorname{Re}(h),\end{aligned}$$

independently of z, and

$$\sqrt{K} \operatorname{Re}\Big((1 - K)\frac{\partial M}{\partial z} + (1 + K)\frac{\partial M}{\partial \overline{z}}\Big) = \frac{(K^2 - 1)r_1}{4(r_1 - r_0)} \operatorname{Re}\Big(\frac{h}{z^2} - \frac{\overline{h}}{\overline{z}^2}\Big) = 0.$$

So $a - Kd$ is actually constant on A', up to $o(h)$ (and $\leqslant O(h)$). So

$$\int_{A'} \sqrt{K}(a - Kd) = \pi K \operatorname{Re}(h) + o(h),$$

as required. □

Remark. — An argument in the same spirit can be used to obtain the semi-directional derivatives of $(\varphi, \psi) \mapsto d(\varphi, \psi)$ at a point $\varphi = \psi$.

8.7. Corollary. — *There is $C > 0$ such that the following holds. Let $Y \subset \overline{\mathbf{C}}$ be finite with $\{0, 1, \infty\} \subset Y$ and $\#(Y) \geqslant 4$. Let $[\psi_1]$, $[\psi_2] \in \mathcal{T}(Y)$ and let $d([\psi_1], [\psi_2]) \leqslant \varepsilon$. Given $y \in Y \setminus \{0, 1, \infty\}$, normalize so that 0, 1, ∞ are fixed by ψ_1, ψ_2, and assume without loss of generality that $\psi_1(y)$ is bounded and bounded from 1. Then*

$$\frac{|\psi_1(y) - \psi_2(y)|}{|\psi_1(y)|} \leqslant C\varepsilon.$$

Remark. — Although this is obtained as a corollary of the Derivative Formula, it is, of course, well-known.

Proof. — We can assume without loss of generality that $Y = \{0, 1, \infty, y\}$. Then the quadratic differentials of integral 1 at $[\psi_1]$ are all of the form $\zeta q(z)dz^2$ for $|\zeta| = 1$ and a fixed q. Let $\operatorname{Res}(q, \psi_1(y)) = \lambda$. Then $\lambda\psi_1(y)$ is bounded and bounded from 0. Let $[\varphi_\zeta]$ be such that $d([\psi_1], [\varphi_\zeta]) = 1$ and $\zeta q(z)dz^2$ is the quadratic differential for $d([\psi_1], [\varphi_\zeta])$ at $[\psi_1]$. Then consider the function

$$F_\zeta([\psi]) = d([\psi], [\varphi_\zeta])$$

near $[\psi] = [\psi_1]$. We see that

$$F_\zeta([\psi_2]) - F_\zeta([\psi_1]) \leqslant \varepsilon.$$

By the first derivative formula, we deduce that

$$|\operatorname{Re}(\lambda\zeta(\psi_2(y) - \psi_1(y)))| \leqslant \varepsilon(1 + o(1)).$$

Since this is true for all $|\zeta| = 1$, the result follows. □

8.8. We now give two results which we shall use later. See, in particular, Chapter 17 and 25.2. These are both results about quadratic differentials for nearby geodesics or quasi-geodesics. The results are basically consequences of 8.3, but 8.7 is used in the first result.

Close Points Lemma. — *Let $[\varphi]$, $[\psi_1]$, $[\psi_2] \in \mathcal{T}(Y)$ with $1/M \leqslant d([\varphi], [\psi_1])$ and $d([\psi_1], [\psi_2]) \leqslant \varepsilon$. Let η be the homeomorphism minimizing distortion such that $[\eta \circ \psi_1] = [\psi_2]$. Let $q_j(z)dz^2$ denote the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi_j])$, and let $p_j(z)dz^2$ be the stretch of $q_j(z)dz^2$ at $[\psi_j]$. Let θ, θ' be defined a.e. on $\overline{\mathbf{C}}$, with values in $(-\pi/2, \pi/2]$, by*

$$\theta(z) = \tfrac{1}{2}(\arg(q_1(z)) - \arg(q_2(z))),$$
$$\theta'(z) = \tfrac{1}{2}(\arg(p_1(z)) - \arg(p_2(z))).$$

Then for C depending only on M, and $j = 1$, 2,

$$\int |\theta|^2 |q_j| < C\varepsilon.$$

Now let ε_0 be $\leqslant$ the Margulis constant. Let S be a component of $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{\geqslant \varepsilon_0}$ or $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{< \varepsilon_0}$. Take a suitable normalisation as follows. Let $\{0, 1, \infty\} \subset \psi_1(Y)$ with $\psi_1(0) = \psi_2(0)$, $\psi_1(1) = \psi_2(1)$, $\psi_1(\infty) = \psi_2(\infty)$. If S is a component of $\overline{\mathbf{C}} \setminus \psi_1(Y))_{\geqslant \varepsilon_0}$, let these three points all be in different components of $\overline{\mathbf{C}} \setminus S$. If S is a component of $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{< \varepsilon_0}$, let T be an adjacent component of $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{\geqslant \varepsilon_0}$ and let the three points all be in different components of $\overline{\mathbf{C}} \setminus (S \cup T)$. Then if $\varepsilon > 0$ is sufficiently small given ε_0, and $j = 1$ or 2,

$$\int_S |\theta'|^2 |p_j| \leqslant C\varepsilon.$$

Proof. — Let χ_j be the quasi-conformal homeomorphism of minimal distortion with $[\chi_j \circ \varphi] = [\psi_j]$. Let η minimize distortion with $[\eta \circ \psi_1] = [\psi_2]$. Then θ is the angle between the directions of maximal dilatation of χ_1, χ_2. Then $\theta + O(\sqrt{\varepsilon})$ is the angle between the directions of maximal dilation of χ_1 and $\eta^{-1} \circ \chi_2$, or of χ_2 and $\eta \circ \chi_1$. To see this, let v be a unit tangent vector at z in the direction of maximum dilatation for χ_2, let $v^\perp$ be a unit vector, perpendicular to v and let

$$v' = (\cos \alpha) v + (\sin \alpha) v^\perp$$

with $|\alpha| \leqslant \delta$. Then for a constant C_1 depending only on M,

$$\|D(\chi_2)_z(v')\| \leqslant \|D(\chi_2)_z(v)\|(1 - C_1\delta^2).$$

So there is a constant C_2 such that, if v' makes angle $\geqslant C_2\sqrt{\varepsilon}$ with $\pm v$, then v' cannot be in the direction of maximum dilatation of $\eta^{-1} \circ \chi_2$. The bounds on the integrals of θ^2 then follow directly from 8.3.

Now let S be a component of $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{\geqslant \varepsilon_0}$ or $(\overline{\mathbf{C}} \setminus \psi_1(Y))_{< \varepsilon_0}$, normalised as explained. Isotoping η if necessary, but keeping the distortion $1 + O(\varepsilon)$, and keeping $[\eta \circ \psi_1] = [\psi_2]$, we can ensure that η is within $O(\varepsilon)$ of the identity in the C^1 norm. If S has only three complementary components, this is trivial. Otherwise, we use 8.7. Then $\theta' + O(\sqrt{\varepsilon})$ is the angle between the directions of maximum dilatation of χ_1^{-1}, $\chi_2^{-1} \circ \eta$, and $\theta' \circ \eta + O(\sqrt{\varepsilon})$ is the angle between the directions of maximal dilatation of χ_2^{-1} and $\chi_1^{-1} \circ \eta^{-1}$. Then the results follow directly from 8.3. □

8.9. Triangular Lemma. — *Let φ, ψ_1, ψ_2, θ, q_1, p_1, q_2, p_2 be as in 8.8, but this time, let both $d([\varphi], [\psi_1])$ and $d([\psi_1], [\psi_2]) \geqslant 1/M$. Let χ_i $(1 \leqslant i \leqslant 3)$ minimize distortion with $[\chi_i \circ \varphi] = [\psi_i]$, $i = 1$, 2, and $[\chi_3 \circ \psi_1] = [\psi_2]$. Write*

$$\varepsilon = d([\varphi], [\psi_1]) + d([\psi_1], [\psi_2]) - d([\varphi], [\psi_2]).$$

Let $q_3(z)dz^2$ be the quadratic differential for $d([\psi_1], [\psi_2])$ at $[\psi_1]$. Let

$$\theta' = \tfrac{1}{2}(\arg(q_3) - \arg(p_1)).$$

Then

$$(1) \qquad \int |\theta|^2 |q_2| \leqslant C\varepsilon,$$

$$\int |\theta' \circ \chi_1|^2 |q_2| \leqslant C\varepsilon, \tag{2}$$

$$\int |\theta' \circ \chi_3^{-1}|^2 |p_2| \leqslant C\varepsilon. \tag{3}$$

Proof. — Write $K_j = K(\chi_j)$, $j = 1$, 2 or 3. These are all constants. We shall obtain (2) first, by an application of 8.2. Note that $K(\chi_3 \circ \chi_1)(z)$ is λ^2 where $\lambda^2 \geqslant \lambda^{-2} > 0$ are the eigenvalues of $A^t A$ where $\theta'' = \theta' \circ \chi_1$ and

$$A = \begin{pmatrix} \sqrt{K_3} & 0 \\ 0 & 1/\sqrt{K_3} \end{pmatrix} \begin{pmatrix} \cos\theta'' & -\sin\theta'' \\ \sin\theta'' & \cos\theta'' \end{pmatrix} \begin{pmatrix} \sqrt{K_1} & 0 \\ 0 & 1/\sqrt{K_1} \end{pmatrix}.$$

Then

$$\lambda^2 + \lambda^{-2} = \text{trace}\, A^t A = (K_1 K_3 + (K_1 K_3)^{-1}) \cos^2\theta'' + (K_1 K_3^{-1} + K_3 K_1^{-1}) \sin^2\theta''.$$

So

$$\lambda^2 + \lambda^{-2} \leqslant K_1 K_3 + (K_1 K_3)^{-1} - C_0 \theta''^2 K_1 (K_3 - K_3^{-1}).$$

So

$$K(\chi_1 \circ \chi_3) = \lambda^2 \leqslant K_1 K_3 (1 - C_1 \theta''^2).$$

By 8.2 we have

$$K_2 \leqslant \int K(\chi_3 \circ \chi_1)|q_2| \leqslant K_1 K_3 \left(1 - C_1 \int \theta'(\chi_1(z))^2 |q_2|\right).$$

Since

$$d([\varphi], [\psi_1]) = \tfrac{1}{2}\log K_1,\ d([\varphi], [\psi_2]) = \tfrac{1}{2}\log K_2,\ d([\psi_1], [\psi_2]) = \tfrac{1}{2}\log K_3,$$

(2) follows. Then (3) is exactly similar, with $[\varphi]$ and $[\psi_2]$ interchanged.

We shall obtain (1) by an application of 8.3. Let θ_1 be the angle between the directions of maximal distortion of χ_2 and $\chi_3 \circ \chi_1$. To obtain (1) by an application of 8.3, it suffices to show that

$$\theta_1 = \theta + O(\theta''). \tag{3}$$

This seems natural, since $\theta_1 = \theta$ if $\theta'' = 0$. Arguing much as in 8.8, let v be a unit tangent vector at z in the direction of maximum dilatation for χ_1. Let $v^\perp$ be a unit vector, perpendicular to v and let

$$v' = (\cos\alpha)v + (\sin\alpha)v^\perp$$

with $|\alpha| \leqslant \delta$. Then for a constant C_1 depending only on M,

$$\|D(\chi_1)_z(v')\| \leqslant \|D(\chi_1)_z(v)\|(1 - C_1\delta^2).$$

Then if $K_1\delta \geqslant C_2|\theta''|$, we obtain

$$\|D(\chi_3 \circ \chi_1)_z(v')\| < \|D(\chi_3 \circ \chi_1)_z(v)\|.$$

So if v' is in the direction of maximum dilation of $\chi_3 \circ \chi_1$, we must have $\delta = O(\theta''/K_1)$, which gives (3), as required. □

8.10. Pullback and Pushforward of Quadratic Differentials. — Let $q(z)dz^2$ be a quadratic differential on $(\overline{\mathbb{C}}, Y)$. This means that any poles of q are in Y, and are at most simple. Let $s : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be holomorphic. Then $s^*q(z)dz^2$, the *pullback* of $q(z)dz^2$ under s, is defined by

$$s^*q(z)dz^2 = (s'(z))^2 q(z)dz^2.$$

Then any poles of $s^*q(z) = (s'(z))^2 q(s(z))$ are in $s^{-1}(Y)$, and are at most simple. So $s^*q(z)dz^2$ is a quadratic differential on $(\overline{\mathbb{C}}, s^{-1}(Y))$.

We define the *pushforward* $s_*q(z)dz^2$ of $q(z)dz^2$ by

$$s_*q(z) = \sum_{s(w)=z} \frac{q(w)}{(s'(w))^2}$$

if z is not a critical value. Then s_*q is meromorphic with poles at most at the critical values of s and the images under s of the poles of q. It is easily checked, that, since q has at most simple poles, so does s_*q. In fact, if $\delta > 0$ is sufficiently small,

$$\int_{|z-b|=\delta} (z-b)s_*q(z)dz = \sum_{s(a)=b} \int_{|w-a|=\delta} \frac{(s(w)-b)q(w)}{s'(w)} dw = 0.$$

Therefore $s_*q(z)dz^2$ is a quadratic differential on $(\overline{\mathbb{C}}, s(Y))$. Furthermore:

$$\mathrm{Res}(s^*q, a) = \mathrm{Res}(q, s(a)) \cdot s'(a),$$

$$\mathrm{Res}(s_*q, b) = \sum_{s(a)=b} \mathrm{Res}(q/s', a).$$

Both formulae are pretty easy. The second needs a change of variable, like the one above. Thus

$$\int_{|z-b|=\delta} \sum_{s(w)=z} \frac{q(w)}{(s'(w))^2} dz = \sum_{s(a)=b} \int_{|z-a|=\delta} \frac{q(w)}{s'(w)} dw,$$

which gives the result. Note also that if s has degree d, then

$$\int |s_*q| \leqslant \int |q| = \frac{1}{d} \int |s^*q|.$$

Equality holds in the lefthand inequality if and only if all terms $q(w)/(s'(w))^2$ in the sum defining $s_*q(z)$ have the same argument. In this case, they must also have the same modulus. So lefthand equality holds if and only if $q = (1/d)s^*(s_*q)$. The equality $q = (1/d)s_*(s^*q)$ always holds.

8.11. The derivative formula for $d(x, \tau x)$. — Let $\tau : \mathcal{T}(Y) \to \mathcal{T}(Y)$ be a pullback map for a space $B(Y, f_0)$, as defined in 6.7. We have seen that, roughly speaking,

$$\tau([\varphi]) = [s^{-1} \circ \varphi \circ f_0]$$

for a suitable holomorphic function s, given φ. Then we have the following. We use local coordinates as in 8.4.

The Derivative Formula for $d(x, \tau x)$. *Let $q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], \tau([\varphi])$ and let $p(z)dz^2$ be the stretch of $q(z)dz^2$ at $\tau([\varphi])$. Then*

$$d([\varphi + \underline{h}], \tau([\varphi + \underline{h}])) = d([\varphi], \tau([\varphi + \underline{h}]) + 2\pi \operatorname{Re}\Big(\sum_{y \in Y} \operatorname{Res}(q - s_* p, \varphi(y)h(y)\Big) + o(\underline{h})$$

Proof. It suffices to prove this formula for $\underline{h} = (h(y'))$ where, for just one $y \in Y$, $h(y) \neq 0$. Of course, we normalise so that, for this y, $\varphi(y) \neq \infty$. Then write $h(y) = h$. Write $[\psi] = \tau([\varphi])$. Let s be the holomorphic branched covering with $[\psi] = [s^{-1} \circ \varphi \circ f_0]$ (by abuse of notation). Then there is a holomorphic branched covering s_h, depending on h and the choice of $y \in Y$, such that

$$\tau([\varphi + \underline{h}]) = [s_h^{-1} \circ \varphi \circ f_0].$$

If y is not a critical value of f_0 then $s_h = s$. In any case, $h \mapsto s_h$ and $h \mapsto \underline{k}(h)$ are holomorphic functions in h near 0. We normalise again so that, for $\underline{k} = (k(y'))$, $k(y') = 0$ if $\psi(y') = \infty$. This essentially follows from the Riemann Mapping Theorem. There is a quasiconformal homeomorphism χ of distortion $1 + O(h)$ mapping $\varphi(y)$ to $\varphi(y) + h$ and fixing $\varphi(y')$ for $y' \in Y \setminus \{y\}$. Then we consider the covering space of $\overline{\mathbf{C}}$ using the branched covering $\chi \circ s$, with a complex manifold structure with respect to which $\chi \circ s$ is holomorphic. The covering space is conformally equivalent to the sphere, with holomorphic covering map s_h. Clearly, s_h is unique up to right composition by a Möbius transformation. Then $s_h^{-1} \circ \chi \circ s$ also has distortion $1 + O(h)$. It follows by 8.7 that the critical points of s_h are within $O(h)$ of those of s, after right composition by a suitable Möbius transformation. So $h \mapsto s_h$ and $h \mapsto \underline{k}(h)$ are Lipschitz. But s_h and $\underline{k}(h)$ are implicitly defined by a set of holomorphic functions. So $h \mapsto s_h$ and $h \mapsto \underline{k}(h)$ are holomorphic near 0.

We obviously want to apply the general derivative formula 8.4. This means that if $\tau([\varphi + \underline{h}]) = [\psi + \underline{k}]$, we want to show that

$$\sum_{y' \in Y} \operatorname{Res}(p, \psi(y'))k(y') = \operatorname{Res}(s_* p, \varphi(y))h.$$

But

$$s_* p(z) = \sum_{y' \in Y,\ \psi(y') \neq \infty} \frac{\operatorname{Res}(p, \psi(y'))}{z - \psi(y')}.$$

So by the formula of 8.10 for residues of $s_* p$, it suffices to show that for $y' \in Y$, $\psi(y') \neq \infty$,

(1) $$k(y') = \sum_{f_0(w) = y} h \operatorname{Res}((s'(z)(z - \psi(y'))^{-1}, \psi(w)).$$

We now know that $\underline{k} = O(h)$. Since $h \mapsto s_h$ is holomorphic, we can find polynomials r, t, α, β, where $\deg(\beta) < \deg(t)$ and r and t have no common factors such that

$$s = \frac{r}{t}, \quad s_h = \frac{r + h\alpha + O(h^2)}{t + h\beta + O(h^2)} = s + h\frac{\alpha t - r\beta}{t^2} + O(h^2).$$

Of course, $\alpha = \beta = 0$ if y is not a critical value of f_0. Now for $y' \in Y$,

$$s_h(\psi(y') + k(y')) = \varphi(f_0(y')) + h\delta_{y,f_0(y')}, \tag{2}$$

where $\delta_{y,w}$ denotes the Kronecker δ. Write

$$\gamma = \frac{\alpha t - \beta r}{t^2}.$$

Then expanding (2), and using $s(\psi(y')) = \varphi(f_0(y'))$, we obtain

$$s'(\psi(y'))k(y') + h\gamma(\psi(y')) = h\delta_{y,f_0(y')} + O(h^2) \tag{3}$$

Suppose that y' is a critical point of f_0 of multiplicity m. Here, multiplicity $m = 0$ means that y' is not a critical point. Then

$$s^{(i)}(\psi(y')) = s_h^{(i)}(\psi(y') + k(y')) = 0, \quad 0 < i \leqslant m, \ s^{(m+1)}(\psi(y')) \neq 0.$$

Expanding, and using $k(y') = O(h)$, we obtain

$$\gamma^{(i)}(\psi(y')) = 0, \quad 0 < i < m,$$
$$k(y')s^{(m+1)}(\psi(y')) + h\gamma^{(m)}(\psi(y')) = 0.$$

Thus, (1) reduces to: if y' is critical of multiplicity m (including $m = 0$),

$$\sum_{f_0(w)=y} \operatorname{Res}((s'(z)(z - \psi(y')))^{-1}, \psi(w)) + \frac{\gamma^{(m)}(\psi(y'))}{s^{(m+1)}(\psi(y'))} = 0. \tag{4}$$

For this, we consider the function

$$\frac{\gamma(z)}{s'(z)(z - \psi(y'))}.$$

By the choice of γ, this function is meromorphic with poles only at points $\psi(w)$ for $f_0(w) = y$ and at $\psi(y)$, and is $O(1/z^2)$ for large z. So the residue sum is 0. Moreover, $(\gamma - 1)/s'$ is holomorphic at points $\psi(w)$ with $f_0(w) = y$. So the lefthand side of (4) is the residue sum, as required. □

CHAPTER 9

PRODUCT STRUCTURE IN THE THIN PART OF TEICHMÜLLER SPACE AND TEICHMÜLLER DISTANCE

9.1. Product Structure in the Thin Part. — We continue to use $\mathcal{T}(Y)$ to denote the Teichmüller space of a sphere $\overline{\mathbf{C}}$ with finite set Y of marked points. Let Γ be a set of disjoint simple nontrivial loops in $\overline{\mathbf{C}} \setminus Y$. As in 6.5, we write

$$\mathcal{T}(\Gamma, \varepsilon_0) = \{[\varphi] : \varphi(\gamma) \text{ has length } < \varepsilon_0 \text{ for } \gamma \in \Gamma\}.$$

Here, length is with respect to the Poincaré metric on $\overline{\mathbf{C}} \setminus \varphi(Y)$. Having fixed a suitable ε_0 ($\leqslant$ the Margulis constant), the *thin part* of $\mathcal{T}(Y)$ is the union of all such sets $\mathcal{T}(\Gamma, \varepsilon_0)$. There is a particularly simple way to give a component of the thin part of $\mathcal{T}(Y)$ a product structure, as follows.

Let Σ denote the set of all loops of Γ, and of all components α of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ such that $\alpha \setminus Y$ is not an annulus. We call such components *gaps*, as in 2.13. We choose a set $A(\alpha) \subset Y$ such that each component of $\overline{\mathbf{C}} \setminus \alpha$ contains exactly one point of $A(\alpha)$. For $\gamma \in \Sigma \cap \Gamma$, we choose $A(\gamma) \subset Y$ containing exactly two points on each side of γ. If $S \in \Sigma$ is nearest to γ on one side, the two points of $A(\gamma)$ on that side are in different components of $\overline{\mathbf{C}} \setminus S$. If $\gamma \subset \partial\alpha$, we also choose $A(\alpha)$ and $A(\gamma)$ to intersect in three points. This is possible, by starting with some S or $\gamma \in \Sigma$ and then working outwards from this base. Let $[\varphi] \in \mathcal{T}(Y)$. Then we have a map

$$\mathcal{T}(Y) \longrightarrow \prod_{\alpha \in \Sigma} \mathcal{T}(A(\alpha))$$

given by

$$[\varphi] \longmapsto ([\varphi_\alpha]),$$

where $[\varphi_\alpha]$ is the class of $[\varphi]$ in $\mathcal{T}(A(\alpha))$. We write π_α for the projection into $\mathcal{T}(A(\alpha))$.

It is not clear that the map (π_α) is a homeomorphism. However, if we now restrict (π_α) to $\mathcal{T}(\Gamma, \delta)$ for δ sufficiently small, it becomes injective (and hence open). The reason is roughly as follows. Each point of Y occurs in at least one $A(\alpha)$. If α and β are an adjacent loop and gap, then $A(\alpha) \cap A(\beta)$ contains 3 points. Hence, if $[\varphi_\alpha] = [\psi_\alpha]$ for all α, we can normalise so that $\varphi(Y) = \psi(Y)$, and if $\gamma \in \Gamma$, then

$\varphi(\gamma) = \psi(\gamma)$ up to isotopy. So $\psi^{-1} \circ \varphi$ is defined, and fixes all loops and gaps of Γ, up to isotopy. Then $\psi_\alpha^{-1} \circ \varphi_\alpha$ is isotopic to the identity via an isotopy preserving $A(\alpha)$ for each gap or loop α. This means, in particular, that $\psi_\alpha^{-1} \circ \varphi_\alpha$ has zero Dehn twist round each loop α of Γ. So $[\varphi] = [\psi]$.

If α is a nonperipheral loop, then $A(\alpha)$ contains precisely four points and $\mathcal{T}(A(\alpha))$ identifies isometrically with the upper half plane. To make an identification we need to choose two simple nonperipheral nonisotopic loops α_1, α_2 in $\overline{\mathbf{C}} \setminus A(\alpha)$ which intersect precisely twice. It is natural to choose $\alpha_1 = \alpha$. There is some choice for α_2, but we choose it so that intersections between α and α_2 are positively oriented. Having chosen these loops, we can define a torus branched cover of each $[\varphi] \in \mathcal{T}(A(\alpha))$ such that the covering map is a homeomorphism from each preimage loop onto α and a double covering of a single preimage loop $\widetilde{\alpha}_2$ onto α_2. Let the covering transformations determined by α and $\widetilde{\alpha}_2$ be $z \mapsto z + \frac{1}{2}$ and $z \mapsto z + a$. Then $\mathrm{Im}(a) > 0$. Then $[\varphi] \mapsto -\overline{a}$ is our chosen identification of $\mathcal{T}(A(\alpha))$ with the upper half plane. It may seem perverse to take $-\overline{a}$ rather than a, but this ensures that right action on $\mathcal{T}(A(\alpha))$ by the mapping class group of the torus translates to left action on the upper half plane by $\mathrm{SL}(2, \mathbf{Z})$. So we can identify π_α with a map from $\mathcal{T}(Y)$ to the upper half plane. If we restrict to $\mathcal{T}(Y)(\Gamma, \delta)$ then the image is open, and contains a half plane of the form

$$\{z : \mathrm{Im}(z) > (2\pi^2/\delta) - C\}$$

for a suitable $C > 0$ independent of δ. We have chosen this normalisation because, if $[\varphi] \in \mathcal{T}(Y)$ and the geodesic homotopic to $\varphi(\alpha)$ has length δ_1 for $\delta_1^{-1} - \varepsilon_0^{-1} > 0$ and bounded from 0, then $2\pi^2/\delta_1 + O(1/\varepsilon_0)$ is also the modulus of the component of $(\overline{\mathbf{C}} \setminus \varphi(A(\alpha)))_{<\varepsilon_0}$ homotopic to $\varphi(\alpha)$. In future, especially from Chapter 16 onwards, we shall often use the quantity

$$m_\alpha([\varphi]) = \log \mathrm{Im}(\pi_\alpha([\varphi])).$$

Let $C > 0$ be suitably chosen. Let $1/\delta' = 1/\delta - C$. We see that a point $([\varphi_\alpha])$ in the image of $\mathcal{T}(\Gamma, \delta)$ is in the interior of the image if $[\varphi_\alpha] \in (\mathcal{T}(A(\alpha))_{\geqslant \delta'}$ for all gaps α and $[\varphi_\alpha] \in \{z : \mathrm{Im}(z) > (2\pi^2/\delta) - C\}$. Each such set is connected. It follows that the image of $\mathcal{T}(\Gamma, \delta)$ contains

$$\prod_{\alpha \text{ a gap}} (\mathcal{T}(A(\alpha))_{\geqslant \delta'} \times \prod_{\alpha \text{ a loop}} \{z \in \mathbf{C} : \mathrm{Im}(z) > (2\pi^2/\delta) - C\}.$$

9.2 The Teichmüller metric as a maximal metric. — Let d_α denote the Teichmüller metric on $\mathcal{T}(A(\alpha))$, and d_Y the Teichmüller metric on $\mathcal{T}(Y)$. Then d_α is a semimetric on $\mathcal{T}(Y)$, that is, nonnegative, symmetric and satisfies the triangular inequality, but it is possible to have $d_\alpha(x, y) = 0$ for $x \neq y$. Then, clearly,

$$d_\alpha \leqslant d_Y.$$

This chapter is concerned with refinements of this. We consider how to estimate Teichmüller distance in the thin part of Teichmüller space using the natural product structure there. The very rough idea is that the distance d_Y is approximately the maximum of the distances d_α on $\mathcal{T}(\Gamma, \varepsilon_0)$. We give a procedure for estimating distance depending on precisely what coordinates in the product give dominant distance. These results are summarised in the Same Shape and Maximal Distance Lemmas.

9.3. Subsurfaces, semimetrics and associated quadratic differentials

Definition of Subsurfaces. Let $x = [\varphi] \in \mathcal{T}(\Gamma, \varepsilon)$, with $\varepsilon \leqslant$ the Margulis constant. Let Σ denote the set of loops and gaps (as in 9.1). We shall usually write $S = \overline{\mathbf{C}} \setminus \varphi(Y)$. Then $S_{<\varepsilon}$ denotes the set of points through which it is possible to draw a nontrivial *and nonperipheral* closed loop of Poincaré length $< \varepsilon$, and $S_{\geqslant\varepsilon} = S \setminus S_{<\varepsilon}$. Then for $\alpha \in \Sigma$, we define $S(\alpha, [\varphi], \varepsilon)$- also called $S(A(\alpha), [\varphi], \varepsilon)$ – as follows. If α is a loop, then $S(\alpha, [\varphi], \varepsilon)$ is the component of $S_{<\varepsilon}$ homotopic to $\varphi(\alpha)$. If α is a gap, then $S(\alpha, [\varphi], \varepsilon)$ is the union U of components of $S_{\geqslant\varepsilon}$ and $S_{<\varepsilon}$ homotopic to $\varphi(\alpha)$, with components of $S_{\geqslant\varepsilon}$ adjacent to ∂U. In particular, if $\gamma \subset \partial\alpha$ is peripheral, then $S(\alpha, [\varphi], \varepsilon)$ includes the component of $S_{<\varepsilon}$ homotopic to $\varphi(\gamma)$. We may write $S(\alpha)$ if ε and $[\varphi]$ are fixed.

We can also define $S(A, \varepsilon)$ for any $A \subset Y$: we take $S(A)$ to be the union of all components T of $S_{\geqslant\varepsilon}$ and $S_{<\varepsilon}$ such that each component of $\overline{\mathbf{C}} \setminus T$ contains a point of A, and at least two points if T is a component of $S_{<\varepsilon_0}$. If $A = A(\alpha)$, this definition agrees with the previous one. Of course, if $A' \subset A \subset Y$, then $S(A') \subset S(A)$.

Let U be any connected subsurface of S such that $\partial U \subset S_{<\varepsilon}$ consists of round circles up to bounded distortion (under the natural normalisation of $S_{<\varepsilon}$). We shall sometimes choose $A(U) \subset Y$ to intersect each component of $\overline{\mathbf{C}} \setminus U$ in precisely either one or two points. If U is an annulus $A(U)$ will intersect each component in two points. If U is not an annulus, then $A(U)$ will intersect each component in one point. Then

$$S(A(U)) \subset U.$$

Semimetrics and Other Metrics. For any $Y' \subset Y$, we have a natural projection $\mathcal{T}(Y) \to \mathcal{T}(Y')$, and consequently a semimetric $d_{Y'}$ on $\mathcal{T}(Y)$, which also identifies with the Teichmüller metric on $\mathcal{T}(Y')$. If $Y' = A(\alpha)$, we shall also call this semimetric d_α. If $U = S(Y')$ (as above), we shall also call this semimetric d_U. More generally, for any U for which we can define $A(U)$ (as above) we shall write $d_U = d_{A(U)}$. Of course, this depends on the choice of $A(U)$, but in fact the different choices do not change d_U much.

Quadratic Differentials. For any $Y' \subset Y$, let $\pi_{Y'} : \mathcal{T}(Y) \to \mathcal{T}(Y')$ denote the natural projection. For any $[\varphi]$, $[\psi] \in \mathcal{T} = \mathcal{T}(Y)$, there is a quadratic differential $q_{Y'}(z)dz^2$ at $\pi_{Y'}([\varphi])$ for $d_{Y'}([\varphi], [\psi])$. (See 8.1.) This can also be regarded as a quadratic differential on $\overline{\mathbf{C}}, \varphi(Y))$, since poles occur at most at the points $\varphi(Y') \subset$

$\varphi(Y)$. One of the subjects of this chapter is the relationship between $q_{Y'}(z)dz^2$ and the quadratic differential $q(z)dz^2$ at $[\varphi]$ for $d([\varphi],[\psi])$.

9.4. Good Boundary, the Pole-Zero Condition, and Dominant Area. — We need to make some definitions which we shall use several times. Let $x = [\varphi] \in \mathcal{T}(\Gamma, \varepsilon)$ and $S = \overline{\mathbf{C}} \setminus \varphi(Y)$.

We say that a subsurface U has C_0-*good boundary* if $\partial U \subset S_{<\varepsilon_0}$ and all components are images of planar round circles under maps with derivative of modulus between C_0 and C_0^{-1}. We say that an annulus $T \subset S$ is a C_0- *good shape annulus* if T is an annulus with good boundary of modulus $\geqslant C_0^{-1}$. There is C_0 such that, whenever ε_0 is sufficiently small, there are C_0-good shape annuli adjacent to each boundary component in every component of $S_{<\varepsilon_0}$. The constant C_0 is independent of $[\varphi]$, and even of Y. This follows from standard distortion results for univalent functions. See, for example, section 2 of [**R2**]. We say that a subsurface $U \subset S$ has C_0-*good shape* if U contains disjoint embedded C_0-good shape annuli adjacent to each component of ∂U.

From now on in this chapter, we fix C_0 and $\varepsilon_0 > 0$ such that C_0-good shape annuli exist in $S_{<\varepsilon_0}$ adjacent to boundary components, and we shall simply talk of *good shape*.

Areas. Let $q(z)dz^2$ be a quadratic differential at $[\varphi]$. If α is a gap, of Γ, and $U \subset S$ is a subsurface, let $a(\alpha, U, q)$ denote the q-area of U, that is,

$$\int_U |q|.$$

We write $a(\alpha, U)$ if it is clear which quadratic differential is being used. If α is a gap, write $a(\alpha) = a(\alpha, S(\alpha, [\varphi], \varepsilon_0), q)$ if ε_0, $[\varphi]$ and q are fixed. We may also write $a(\alpha, q)$, and so on. If α is a loop, let $a(\alpha) = a(\alpha, q) = a(\alpha, \varepsilon, q)$ denote the minimal area in $S(\alpha, [\varphi], \varepsilon)$ of a good boundary annulus of modulus 1.

Pole-Zero Condition. If $D \subset S$, let p_D, z_D be the numbers of poles and zeros of q in D. Then we say that $U \subset S$ satisfies the *Pole-Zero Condition* (for q) if, for every component D of $\overline{\mathbf{C}} \setminus U$, if U is not an annulus,

$$p_D - z_D \leqslant 1,$$

and if U is an annulus,

$$p_D - z_D = 2.$$

Let $U \subset S$ with good boundary. First suppose U contains at least one component of $S_{\geqslant\varepsilon_0}$. We say that U has C_1-*dominant area* for q if the following holds for any good shape annulus $T \subset U$ adjacent to the boundary of modulus $\leqslant C_0$ and any good shape $T' \subset Q$ separated from ∂U by at least some points of $S_{\geqslant\varepsilon_0}$:

$$\frac{C_1 a(T, q)}{a(T', q)} \leqslant 1.$$

Now let $U \subset S(\alpha, \varepsilon_0)$ for some loop α, and be homotopic to $S(\alpha, \varepsilon_0)$. Then U has C_1-*dominant area* for q if U satisfies the Pole Zero Condition for q and has modulus $\geqslant m_\alpha(1 - 1/C_1)$. Here, $m_\alpha = m_\alpha([\varphi])$ is as in 9.1.

Boundedly Proportional. Again, fix a suitable $C_0 > 0$. We say that q and q_U are *boundedly proportional on* U if there is a constant $C > 0$ such that the following holds for any good shape annuli T, T':

$$\frac{a(T', q)}{Ca(T, q)} \leqslant \frac{a(T', q_U)}{a(T, q_U)} \leqslant \frac{Ca(T', q)}{a(T, q)}.$$

We are now ready to start stating and proving results about the closeness of quadratic differentials for d and d_α.

9.5. Same Shape Lemma. *Let C_0 be given. Let C_1 and C_1' be sufficiently large given C_0. Let $[\varphi]$, $[\psi] \in \mathcal{T}(Y)$, and $M = \frac{1}{2}\log K = d([\varphi], [\psi])$. Let $q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi])$. Let $S = \overline{\mathbf{C}} \setminus \varphi(Y)$. Take any connected good shape $U \subset S$, satisfying the Pole-Zero Condition, such that each component of $U \cap S_{<\varepsilon_0}$ which adjoins ∂U has modulus $\geqslant C_1(M+1)$. Let Q be U minus a good shape annulus of modulus $C_1(M+1)$ adjacent to each boundary component. Let Q have C_1-dominant area for either q or q_1.*

Let $q_1(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d_U([\varphi], [\psi]) = \frac{1}{2}\log K_1$. Let $\theta_1(z)$ denote the angle between $\sqrt{q(z)}$ and $\sqrt{q_1(z)}$. Let b be the maximum of the $a(T, q)$, and b_1 the maximum of the $a(T, q_1)$, for T running over the annulus components of $U \setminus Q$. Then

$$K_1 + C_1'^{-1}(K-1)\int |\theta_1|^2 |q_1| \leqslant K, \tag{1}$$

$$K \leqslant K_1 + C_1' b_1(K-1) - C_1'^{-1}(K-1)\int |\theta_1|^2 |q_1|. \tag{2}$$

Also, q and q_1 are boundedly proportional on Q. Let $T \subset Q$ be of good shape, and of modulus $\leqslant C_0$, if an annulus. Normalise so that T is bounded, and at least two boundary components have diameters bounded from 0. Let $a(T, q_1) = a_1$. Then any zeros of q in T are distance $\leqslant C_1'\sqrt{b_1/a_1}$ from the same number of zeros (up to multiplicity) of q_1, and similarly with q and q_1 interchanged.

Proof. (1) is immediate from 8.3 (although the rôles of q and q_1 have been interchanged). In fact, to get (2), we shall also use 8.3, this time with q playing the same rôle here as there. Let $B = U \setminus Q$. Let χ, χ_1 be the quasi-conformal homeomorphisms with $[\chi \circ \varphi]_Y = [\psi]_Y$, $[\chi_1 \circ \varphi]_{A(U)} = [\psi]_{A(U)}$. Then we claim that we can construct a quasi-conformal homeomorphism χ_2 which is χ_1 on Q, and χ on $\overline{\mathbf{C}} \setminus U$, and with $K(\chi_2) = 1 + O(K-1)$ on B. We shall do this in the lemma below. Then 8.3 gives

$$K \leqslant \int_{\overline{\mathbf{C}} \setminus U} K|q| + \int_Q K_1|q| + a(U \setminus Q) + C_3(K-1)b - C_2(K-1)\int_Q |\theta_1|^2 |q|.$$

This gives

$$(3) \qquad \int_U K|q| \leqslant \int_Q K_1|q| - C_2(K-1)\int_Q |\theta_1|^2|q| + b + C_3(K-1)b.$$

Note that $a(\overline{\mathbf{C}} \setminus Q, q_1) = O(b_1)$. So we can derive (2) from (3) if, for a suitable constant C_4,

$$(4) \qquad \frac{b_1}{C_4 b} \leqslant \frac{a(Q,q_1)}{a(Q,q)} \leqslant \frac{C_4 b_1}{b}.$$

So it remains to show that q and q_1 are boundedly proportional on Q (see 9.4). If U is an annulus, this is automatic, by the Pole Zero Condition. So now suppose that U is not an annulus. Then we shall prove (4) by proving (4) for good shape surfaces $T \subset Q$, by induction on the distance of T from ∂U. If T is an annulus, we take it of bounded modulus. The idea is that if q and q_{U} are boundedly proportional on T' adjacent to T, then they are also boundedly proportional on T with a somewhat worse bound, which we can then improve. Let $T \subset Q$ be any good shape surface which is either a component of $S_{\geqslant \varepsilon_0}$ or bounded modulus annulus in $S_{<\varepsilon_0} \cap Q$. Normalise so that T is bounded, and at least two components of $\overline{\mathbf{C}} \setminus T$ are bounded apart, with diameters bounded from 0. Then by (3)

$$(5) \qquad \int_T |\theta_1|^2|q| \leqslant C_3 b/C_2.$$

If T' is a component of $U \setminus Q$, and $T'' \subset Q$ is a good shape annulus of modulus C_0 adjacent to T', then $a(T',q) = O(a(T'',Q)$ by the Pole-Zero Condition for q. By the inductive hypothesis, and the C_1-dominant area condition for q or q_1, we deduce from (5) that $C_3 b/C_2 = o(a(q,T))$ if C_1 is large enough. So any zeros of q in T are close to zeros of q_1 in the normal plane metric, and vice versa. It then follows immediately that q and q_1 are boundedly proportional on T, with a better bound, and we can continue the induction. Let $a_1 = a(T, q_1)$, as in the statement of the Lemma. We also obtain, by Cauchy-Schwartz on (5), and using bounded proportionality of q, q_1:

$$(6) \qquad \int |\theta_1||q_1| \leqslant C_4\sqrt{b_1 a_1}.$$

The bound on the distance between zeros follows immediately. So the last statement of the lemma is also proved. □

9.6. Lemma. — *χ_2 can be constructed as in 9.5.*

Proof. — Continue writing $S = \overline{\mathbf{C}} \setminus \varphi(Y)$, and also write $S' = \overline{\mathbf{C}} \setminus \psi(Y)$. We need to construct χ_2 on each component T of $U \setminus Q$, to be χ_1 on one boundary component, and χ on the other. Normalise φ and ψ so that the boundary component ∂ of $S_{<\varepsilon_0}$ homotopic to T satisfies

$$\partial \subset \{z : C_2 e^{-2\pi^2/\varepsilon_0} \leqslant |z| \leqslant C_3 e^{-2\pi^2/\varepsilon_0}\}$$

with T nearer 0. Take the same normalisation for the component of $S'_{<\varepsilon_0}$ homotopic to $\chi(\partial)$. Now we can find annuli T_1, $T_2 \subset T$ of comparable modulus which contain no zeros of q or q_1, with T_1 separating T_2 from U.

There are $\widetilde{\chi}_1$ and holomorphic coverings $s_1(z) = \lambda_1 z^p$, $s_2(z) = \lambda_2 z^2$ for an integer $p \geqslant 1$ and $|\lambda_1| = |\lambda_2| = 1$ such that if $\widetilde{T}_1 = s_2^{-1}(s_1(T_1))$, then

$$\widetilde{\chi}_1(x+iy) = (\sqrt{K_1}x + y/\sqrt{K_1})(1+o(1)),$$

$$D\widetilde{\chi}_1 = \begin{pmatrix} \sqrt{K_1} & 0 \\ 0 & 1/\sqrt{K_1} \end{pmatrix} + o(1/\sqrt{K_1}).$$

We can further reduce T_1, again to an annulus of comparable modulus, and assume that

$$\widetilde{T}_1 \subset \{z : e^{-b} < |z| < e^{-a}\}$$

for $K_1 e^{-a} = o(1)$ and $b - a \geqslant \Delta$, if C_1 is large enough. Now we claim that we can construct χ_2 which is χ_1 on the outer boundary of T_1 and the identity on the inner boundary. We do this by constructing $\widetilde{\chi}_2$ on $\widetilde{T}_1$.

We take $\rho = \sqrt{x^2+y^2}$, $s = \log(1/(\log \rho))$ and $t : [-b, -a] \to [0,1]$ to be a function of bounded derivative with $t = 0$ near b and 1 near a. This is possible if $b - a$ is large enough, that is, if C_1 is large enough. Then we define

$$\widetilde{\chi}_2(x,y) = (1 - t(s(\rho)))\widetilde{\chi}_1(x,y) + t(s(\rho)) \begin{pmatrix} x \\ y \end{pmatrix}.$$

Then $s'(\rho) = -1/(\rho \log \rho)$, so $\rho s'(\rho) \to 0$ as $\rho \to 0$. Note that, writing $\widetilde{\chi}_1(x,y) = (x', y')$,

$$\begin{aligned} D\widetilde{\chi} = (1-t) \begin{pmatrix} \sqrt{K_1} & 0 \\ 0 & (1/\sqrt{K_1}) \end{pmatrix} &+ tI + o(1/\sqrt{K_1}) \\ &+ t'(s(\rho))s'(\rho)\Big(\begin{pmatrix} x \\ y \end{pmatrix} D_{\rho(x,y)} - \widetilde{\chi}_1(x,y) D\rho_{(x,y)} \Big). \end{aligned}$$

It follows that $D\widetilde{\chi}_2$ is invertible and

$$K(D\widetilde{\chi}_2) = K((1-t)D\widetilde{\chi}_1 + tI) + o(1),$$

and we get the required estimates.

Similarly we can construct χ_2 on T_2 to be χ on component of ∂T_2 furthest from U. Then the construction of χ_2 is complete. □

9.7. Maximum Distance Lemma. — *The following holds for a suitable constant C_2 depending only on $\#(Y)$. Let $[\varphi]$, $[\psi] \in \mathcal{T}(\Gamma, \varepsilon) \subset \mathcal{T}(Y)$ with $M = d([\varphi],[\psi])$ and $C_2 \varepsilon M < 1$. Let $[\varphi] \notin \mathcal{T}(\gamma, C_2\varepsilon)$ for $\gamma \notin \Gamma$. Let Σ_1 denote the set of gaps of Γ. (See 9.1.)*

There is at least one $\alpha \in \Sigma_1 \cup \Gamma$ of C_1-dominant area. For any such α,

$$d([\varphi],[\psi]) \leqslant \begin{cases} d_\alpha([\varphi],[\psi]) + e^{-1/C_2\varepsilon} & \text{if } \alpha \in \Sigma_1, \\ d_\alpha([\varphi],[\psi]) + C_2 M \varepsilon & \text{if } \alpha \in \Gamma. \end{cases}$$

Remark. — The conclusion of this lemma means that the Same Shape Lemma can then be applied. As we shall see, we also use the Same Shape Lemma to prove this lemma.

Proof. — We need to find Q and U as in the Same Shape Lemma 9.5, such that either $U \subset S(\gamma)$ for some $\gamma \in \Gamma$, or $U \setminus S_{<\varepsilon_0} = S(\alpha)$ for some gap α of Γ. (See 9.3.) Let b_1, C_1 be as in the Same Shape Lemma. For some C_2, we also need $C_1 b_1 \leqslant C_2 M\varepsilon$ if $U \subset S_{<\varepsilon_0}$, and $C_1 b_1 \leqslant e^{-C_2/\varepsilon}$ if $U \setminus S_{<\varepsilon_0}$ is $S(\alpha)$ for some gap α. If C_2 is sufficiently large given C_1 of 9.5, Q will have C_1-dominated area, from the condition $[\varphi] \notin \mathcal{T}(\gamma, C_2\varepsilon)$ for $\gamma \notin \Gamma$. So it remains to find such a U

Suppose there is a good shape annulus $A \subset S(\gamma)$ of modulus $\geqslant \pi^2/\varepsilon$ satisfying the Pole-Zero Condition. Then $b_1 \leqslant 2C_1 M\varepsilon$. Then if C_2/C_1 is large enough, we can take $U = A$. So now suppose that no such A exists. Then we can construct a sequence of good shape surfaces S_i ($0 \leqslant i \leqslant N$, $N \leqslant \#(Y)$) such that S_N is our required surface U, as follows. Take $S_0 = S$. Suppose S_i has been constructed. Inductively we assume that S_i satisfies the Pole-Zero Condition, and that an annulus of modulus $\geqslant \pi^2/3\varepsilon$ in $S_{<\varepsilon_0} \cap S_i$ adjoins each component of ∂S_i. We take $i = N$ if S_i is homotopic to $S(\alpha)$ for some α. Suppose not. Then there is $\gamma \in \Gamma$ and a $S(\gamma) \subset S_i$ which is not homotopic to any boundary component. At least one component of $S_i \setminus S(\gamma)$ can be extended by adding an annulus of modulus $\geqslant (\pi^2/2\varepsilon) - O(1/\varepsilon_0)$ in $S(\gamma)$, so that the Pole Zero Condition is satisfied. Then assuming C_2 is sufficiently large (again depending only on C_1), we simply take S_{i+1} to be this extended component of $S_i \setminus S(\gamma)$. Then S_{i+1} satisfies the inductive hypotheses. The condition $C_1 b_1 \leqslant e^{-C_2/\varepsilon}$ then follows simply for C_2 large enough depending on C_1. □

9.8. Another Maximum Distance Lemma. — *The following holds for some $C_3 > 0$ depending only on $\#(Y)$, $M_0 > 0$, and for $\varepsilon > 0$ sufficiently small given $\#(Y)$, $M_0 > 0$. Let $d([\varphi], [\psi]) = M \leqslant M_0$. Let $q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi])$. Let $[\chi \circ \varphi] = [\psi]$ with χ minimizing distortion. Let $[\varphi']$, $[\psi'] \in \mathcal{T}(\Gamma, \varepsilon)$. Let ξ_t, ξ'_t ($t \in [0, 1]$) be isotopies with $[\varphi'] = [\xi_0 \circ \varphi']$, $[\psi'] = [\xi'_0 \circ \psi']$, $[\varphi] = [\xi_1 \circ \varphi']$, $[\psi'] = [\xi'_1 \circ \psi]$. For all t, let*

$$\int (K(\xi_t) + K(\xi'_t) \circ \chi^{-1})|q| = \eta_t, \quad \eta_t \leqslant C_3^{-2} \mathrm{Min}_T a(T, q),$$

where the minimum of T is taken over all C_1-good shape (see 9.3) subsurfaces of $\overline{\mathbf{C}} \setminus \varphi(Y)$ Then

$$d([\varphi'], [\psi']) \leqslant d([\varphi], [\psi]) + C_3 \eta. \tag{1}$$

Proof. — This is rather similar to 9.5. We shall prove more generally that

$$d([\varphi_t], [\psi_t]) \leqslant d([\varphi], [\psi]) + C_3 \eta_t. \tag{2}$$

Write $q_t(z)dz^2$ for the quadratic differential for $d([\varphi_t],[\psi_t])$. By 8.2 (as in 9.5) it suffices to prove that

$$\int K(\xi_t' \circ \chi \circ \xi_t)|q_t| \leqslant e^{2M}(1 + C_3\eta_t). \tag{3}$$

Clearly, (3) will hold so long as q_t and q are boundedly proportional (see 9.4), and this has to be true for t sufficiently small. Let χ_t minimize distortion with $[\chi_t \circ \xi_t \circ \varphi] = [\xi_t' \circ \psi]$. Then $[\xi_t'^{-1} \circ \chi_t \circ \xi_t \circ \varphi] = [\psi]$. As in 8.8-9, we see that the angle θ_t between the directions of maximal distortion of χ and $\xi_t'^{-1} \circ \chi_t \circ \xi_t$ at z is controlled by $\theta(q, q_t)$ (the angle between q and q_t) and $K(\xi_t) + K(\xi_t' \circ \chi^{-1})$.

Then by 8.3 if θ_t denotes the angle between q and q_t we obtain

$$\begin{aligned} d([\varphi],[\psi]) &\leqslant \int K(\xi_t'^{-1} \circ \chi_t \circ \xi_t)|q| - C_4 \int \theta_t^2 |q| \\ &\leqslant d([\xi_t \circ \varphi],[\xi_t' \circ \psi]) + C_5 \int (K(\xi_t) \circ \chi^{-1} + K(\xi_t'))|q| - C_4 \int \theta_t^2 |q| \end{aligned}$$

which yields

$$\int \theta(q_t, q)^2 |q| \leqslant C_6 C_3 \eta_t \leqslant C_3^{-1} C_6 \min_T a(T, q).$$

So then, if C_3 is sufficiently large, we obtain that the angle is very small and (3) will hold for a $t' > t$, if $t < 1$. □

CHAPTER 10

THE FORMULA FOR THE SECOND DERIVATIVE OF TEICHMÜLLER DISTANCE

10.1. The purpose of this chapter is to give a formula for the second derivative of the Teichmüller distance function d for the Teichmüller space of marked spheres of a set of real codimension 2 in $\mathcal{T}(Y) \times \mathcal{T}(Y)$. We start with a simple proof that it is real analytic off this set. This first proof is based on the methods of Chapter 8. The explicit formula 10.16 for the second derivative fails to make sense at precisely the points identified in 10.2. We shall see, however, in Chapter 12, that the Teichmüller distance function is C^2 on $\{([\varphi],[\psi]) \in \mathcal{T}(Y) \times \mathcal{T}(Y) : [\varphi] \neq [\psi]\}$.

10.2. Lemma. *Let $[\varphi] \neq [\psi]$. Let $q(z)dz^2$ be the quadratic differential for $d([\varphi],[\psi])$ at $[\varphi]$ (see 8.1). Suppose that q has only simple zeros, and nonzero residues at all points of $\varphi(Y)$. Then d is real analytic in a neighbourhood of $([\varphi],[\psi])$.*

Proof. We assume without loss of generality that $0,\ 1,\ \infty \in Y$ and, for all $[\varphi_1] \in \mathcal{T}(Y)$, we choose φ_1 to fix $0,\ 1,\ \infty$. A quadratic differential $q_1(z)dz^2$ near $q(z)dz^2$, with poles at φ_1 for $[\varphi_1]$ near $[\varphi]$, is determined up to scale by its poles and zeros, and by the residue at one pole, up to positive scaling. A Beltrami differential is then determined by q_1 (up to scale) and by a distortion K_1 near K, where $\frac{1}{2}\log K = d([\varphi],[\psi])$. We parameterize the set of K_1 near K and q_1 (up to scale) near q by Λ where Λ is an open subset in $\mathbf{R}^{4n-12}$. For convenience, we take $\underline{0} \in \Lambda$ corresponding to (K,q). Write (K_λ, q_λ) for the distortion and quadratic differential up to scale. The corresponding Beltrami differential ν_λ then satisfies

$$\nu_\lambda(z) = (K_\lambda - 1)\overline{q_\lambda(z)}/((K_\lambda + 1)|q_\lambda(z)|).$$

This is clearly well-defined for all z apart from the poles and zeros of q_λ.

Let $p(z)dz^2$ be the stretch of $q(z)dz^2$ by factor $\sqrt{K}$, and let $p_\lambda(z)dz^2$ be the stretch of $q_\lambda(z)dz^2$ by factor K_λ. (See 8.1.) Parametrize quadratic differentials near p similarly by an open set $\Lambda' \subset \mathbf{R}^{4n-12}$. We are now going to show that the map

$$F : (K_\lambda, q_\lambda) \longmapsto (K_\lambda, q_\lambda, K_\lambda, p_\lambda) : \Lambda \longrightarrow \Lambda \times \Lambda'$$

is real-analytic. There is a similar real-analytic map

$$G : \Lambda''(\subset \Lambda') \longrightarrow \Lambda \times \Lambda'$$

and we shall also be interested in the map

$$F_1 : \Lambda \longrightarrow \mathcal{T}(Y) \times \mathcal{T}(Y)$$

obtained by projecting Λ, Λ' to neighbourhoods of $[\varphi]$, $[\psi] \in \mathcal{T}(Y)$ by using the standard projection of quadratic differentials to $\mathcal{T}(Y)$. This map is well-known to be injective [**Abi**], and hence a homeomorphism onto its image.

We want to use the real analytic version of the Measurable Riemann Mapping Theorem [**A-B**]. We cannot apply the theorem to the family $\lambda \mapsto \nu_\lambda : \Lambda \to L^\infty(\overline{\mathbf{C}})$ because this is not even continuous, let alone real analytic. However, we can find a family of homeomorphisms ζ_λ of $\overline{\mathbf{C}}$ such that ζ_λ maps the zeros of poles of q to those of q_λ, ζ_λ is pure translation in a neighbourhood of the zeros and poles of q, $\zeta_{\underline{0}}$ = identity, $(\lambda, z) \mapsto \zeta_\lambda(z) : \Lambda \times \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ is C^∞, and real analytic in λ. Then let $\mu_\lambda = \zeta_\lambda^* \nu_\lambda$, that is, identifying ν_λ with the corresponding Riemannian metric,

$$\mu_\lambda(u, v) = \nu_\lambda(D\zeta_\lambda u, D\zeta_\lambda v).$$

Then the singularities of μ_λ are the poles and zeros of q, and $\lambda \mapsto \mu_\lambda(z)$ is constant for z near the poles and zeros of q, and $\lambda \mapsto \mu_\lambda : \Lambda \to L^\infty(\overline{\mathbf{C}})$ is real analytic. Let σ_0 be the standard Beltrami differential (that is, corresponding to the standard metric). Let ξ_λ be the quasi-conformal homeomorphism, normalised so that ξ_λ fixes 0, 1, ∞, for example, with $(\xi_\lambda)_* \sigma_0 = \mu_\lambda$. We can normalise so that $\xi_{\underline{0}}(\varphi(Y)) = \psi(Y)$. By Theorem 3 of [**A-B**], it follows that $\lambda \mapsto \xi_\lambda(z)$ is real analytic for all z.

In particular, $\lambda \mapsto \xi_\lambda(z)$ is real analytic when z is a pole or zero of q. We note also that the points $\zeta_\lambda(\varphi(y))$ are the poles for q_λ, while the points $\xi_\lambda(\varphi(y))$ are the poles of p_λ, and the images under ξ_λ of the zeros of q are the zeros of p_λ. Moreover, $(\lambda, z) \mapsto \xi_\lambda \circ \zeta_\lambda^{-1}(z)$ is real analytic in (λ, z), and a diffeomorphism in z, away from the poles and zeros of q_λ. To see this, if we identify complex numbers with 2×2 real matrices in the usual way, we have

$$\begin{pmatrix} \sqrt{K_\lambda} & 0 \\ 0 & 1/\sqrt{K_\lambda} \end{pmatrix} \sqrt{q_\lambda} \circ \zeta_\lambda \circ \xi_\lambda^{-1} = \sqrt{p_\lambda} D(\xi_\lambda \circ \zeta_\lambda^{-1})_{\zeta_\lambda \circ \xi_\lambda^{-1}}.$$

So now $(z, \lambda) \mapsto \zeta_\lambda \circ \xi_\lambda^{-1}(z)$ is real analytic away from the poles and zeros of p_λ. So $(z, \lambda) \mapsto p_\lambda(z)$ is real analytic in annuli round the poles of p_λ, which are near the poles of p. Then let A_λ be an annulus of fixed internal and external radii centred on a pole z_λ of p_λ, bounded from all zeros and poles. Then for a constant a,

$$\mathrm{Res}(p_\lambda, z_\lambda) = a \int_{A_\lambda} p_\lambda.$$

It follows that the residue is real analytic in λ. This completes the proof that F is real analytic.

Now we consider the real analytic local homeomorphism F_1. The distance function is the log of the first coordinate of F_1^{-1}. So the distance function is real analytic at $([\varphi],[\psi])$ if the derivative of F_1 is invertible at $F_1^{-1}([\varphi],[\psi])$. We use the notation $[\varphi_1]=[\varphi+\underline{h}_1]$ of 8.4, where $\underline{h}_1\in\mathbf{C}^{n-3}$. Then it suffices that

(1) $F_1^{-1}([\varphi+\underline{h}_1],[\psi+\underline{k}_1])-F_1^{-1}([\varphi+\underline{h}_2],[\psi+\underline{k}_2])=o(\|\underline{h}_1-\underline{h}_2\|^{1/2})+o(\|\underline{k}_1-\underline{k}_2\|^{1/2})$

It suffices to prove this in each of the cases $\underline{k}_1=\underline{k}_2$ and $\underline{h}_1=\underline{h}_2$. We consider first the case $\underline{k}_1=\underline{k}_2$, and for simplicity of writing, we assume that $\underline{k}_1=\underline{k}_2=\underline{0}$ and $\underline{h}_1=\underline{0}$, $\underline{h}_2=\underline{h}$. If $[\varphi]$ and $[\varphi+\underline{h}]$ are obtained from $[\psi]$ by different stretches of the same quadratic differential with poles at $\psi(y)$ $(y\in Y)$, then $d([\varphi],[\psi])-d([\varphi+\underline{h}],[\psi])=O(\|\underline{h}\|)$ (which can be estimated straightforwardly, although of course it also follows from the Derivative Formula of 8.4). Then (1) follows because F is C^∞, and we get $O(\|\underline{h}\|)$ rather than $o(\|\underline{h}\|^{1/2})$. This means that it suffices to prove (1) with $d([\varphi],[\psi])=d([\varphi+\underline{h}],[\psi])$ and $\underline{k}_1=\underline{k}_2=0$, or similarly with $\underline{h}_1=\underline{h}_2=0$ and $d([\varphi],[\psi])=d([\varphi],[\psi+\underline{k}])$.

Now let χ be the quasi-conformal homeomorphism of minimal distortion such that $[\chi\circ\varphi]=[\psi]$, and let $\chi_{\underline{h}}$ be as in 8.4, so that $[\chi_h\circ(\varphi+\underline{h})]=[\psi]$. The term $I(\underline{h})$ of 8.4 is 0. So the calculation there shows that

$$K\leqslant\int K(\chi_{\underline{h}})|q|\frac{dz\wedge d\overline{z}}{2i}=K+o(\underline{h}).$$

Let θ denote the angle between the direction of maximum distortion for $\chi_{\underline{h}}$ at z, and $\sqrt{q_2(z)}$, where $q_2(z)dz^2$ is the quadratic differential at $[\varphi+\underline{h}]$ for $d([\varphi+\underline{h}],[\psi])$. Then it follows from 8.3 that

$$\int|\theta|^2|q|=o(\underline{h}).$$

But over all but small neighbourhoods of the points $\varphi(y)$ $(y\in Y)$, the direction of maximum distortion for $\chi_{\underline{h}}$ at z is the direction of $\sqrt{q(z)}$. It follows that q_2 is within $o(\|\underline{h}\|^{1/2})$ of q, and hence that

$$F^{-1}([\varphi],[\psi])-F^{-1}([\varphi+\underline{h}],[\psi])=o(\|\underline{h}\|^{1/2})$$

as required. In the case $\underline{h}_1=\underline{h}_2=\underline{k}_1=0$, $\underline{k}_2=\underline{k}$, let q, q_2 be the quadratic differentials for $d([\varphi],[\psi])$ and $d([\varphi],[\psi+\underline{k}]$ at $[\varphi]$ and let p, p_2 be the stretches at $[\psi]$, $[\psi+\underline{k}]$. We show in the same way that p and p_2 are $o(\|\underline{h}\|^{1/2})$ apart. Then because G is real analytic, q and q_2 are $o(\|\underline{h}\|^{1/2})$ apart, and the proof of (1) is finished, as required. □

10.3. Standing Assumption and Notation. — For the rest of this Chapter, we fix $[\varphi]$, $[\psi]\in\mathcal{T}(Y)$, $\#(Y)=n\geqslant 4$, and quasiconformal homeomorphism χ with $[\psi]=[\chi\circ\varphi]$ and $K(\chi)=K$ minimal, so that $\frac{1}{2}\log K=d([\varphi],[\psi])$. Thus, χ has a linear form with respect to singular coordinates given by the quadratic differential $q(z)dz^2$ for $d([\varphi],[\psi])$ at $[\varphi]$, and the stretch $p(z)dz^2$ of $q(z)dz^2$ at $[\psi]$, as in 8.1. We also suppose that q has only simple zeros and nonzero residues at all points $\varphi(y)$, $y \in Y$. (Thus, similar properties hold for p also.)

We assume without loss of generality that three fixed points of Y are mapped by both φ and ψ to 0, 1, ∞. We write b_j ($1 \leqslant j \leqslant n-3$) for the other points $\varphi(y)$, that is, the other poles of q, and b'_j for the other poles of p. We write

$$a_j = \operatorname{Res}(q, b_j) \text{ and } a'_j = \operatorname{Res}(p, b'_j).$$

10.4. The Second Derivative. — We have seen in 8.4 that a neighbourhood of $([\varphi], [\psi])$ identifies with a neighbourhood of $\underline{0}$ in $\mathbf{C}^{n-3} \times \mathbf{C}^{n-3} \cong \mathbf{C}^{2n-6} \cong \mathbf{R}^{4n-12}$. Therefore, the second derivative of the Teichmüller Distance Function d identifies with a real symmetric matrix of real dimension $4n-12$. The Derivative Formula of Chapter 8 tells us that the first derivative of d identifies with the row vector

$$2\pi(\overline{a}_1, \ldots, \overline{a}_{n-3}, -\overline{a}'_1, \ldots, -\overline{a}'_{n-3}),$$

where $\overline{a}_j$ denotes complex conjugate, and we identify a complex number $c + id$ with (c, d) in the usual way. Then the derivative $D\overline{a}_j$ of $\overline{a}_j$ is a $2 \times (4n-12)$ real matrix, and the second derivative of d (if it exists) is

$$2\pi(D\overline{a}_1, \ldots, D\overline{a}_{n-3}, -D\overline{a}'_1, \ldots, -D\overline{a}'_{n-3}).$$

Note that this matrix is not obviously symmetric. Before stating a more illuminating formula, we need to describe our framework and establish notation.

10.5. Holomorphic 1-Forms associated to q, p. — We can write

$$q = \sum_{j=1}^{n-3} a_j q_j, \; q_j = \frac{b_j(b_j - 1)}{z(z-1)(z-b_j)},$$

and similarly for p in terms of functions p_j. Let S be the closed nonsingular Riemann surface

$$\{(z, u) \in \overline{\mathbf{C}}^2 : q(z) = u^2\}.$$

Thus S has genus $n-3$. Let $\pi : S \to \overline{\mathbf{C}}$ be projection onto the first coordinate. Thus π is a 2-fold covering branched precisely at the zeros and poles of q (including ∞), that is, at $2n-4$ points. Let S', π' be similarly defined for p. Then

$$\pi^*((q_j/\sqrt{q})dz) = \frac{q_j(z)dz}{u}$$

is a (single-valued) holomorphic 1-form on S. Note that $q_j(z)dz/u$ is holomorphic at the preimages under π of 0, 1, ∞, b_j and the zeros of q, since $q'(z)dz - 2udu = 0$ on S. Similarly, $q_j(z)dz/u$ has zeros at the preimages under π of the other poles of q. So $\pi^*((q_j/\sqrt{q})dz) \neq 0$ at $\pi^{-1}(b_k)$ if and only if $j = k$. It follows that for

$1 \leqslant j \leqslant n-3$ these forms are linearly independent. In fact they form a basis for the space of holomorphic forms over $\mathbf{C}$, since this has dimension $n-3$ [**Gun**]. We have

$$\pi^*(\sqrt{q}dz) = \sum_{j=1}^{n-3} \pi^*\Big(\frac{q_j}{\sqrt{q}}dz\Big).$$

10.6. Harmonic 1-forms. — Now given any meromorphic 1-form on S, given locally by

$$f(z)dz = (u(x,y) + iv(x,y))(dx + idy),$$

we naturally define

$$\mathrm{Re}(f(z)dz) = udx - vdy, \quad \mathrm{Im}(f(z)dz) = \mathrm{Re}(if(z)dz) = vdx + udy.$$

We call the real part of a holomorphic 1-form *harmonic*. Equivalently, a C^1 form given locally by $vdx + udy$ is harmonic if and only if

$$\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = 0.$$

Note that a nonzero harmonic form on S cannot be a coboundary — for if it were, it would be dh for a harmonic function h — and since S is compact, it admits no nonconstant harmonic functions. Now let $f_1(z)dz$ and $f_2(z)dz$ be meromorphic 1-forms and define $\mathrm{Re}(f_j(z)dz)$, $\mathrm{Im}(f_j(z)dz)$ as above. Then we also have the following.

$$\mathrm{Re}(f_1(z)dz) \wedge \mathrm{Re}(f_2(z)dz) = \mathrm{Im}(f_1(z)dz) \wedge \mathrm{Im}(f_2(z)dz) = \mathrm{Im}(f_1(z)\overline{f_2(z)})dx \wedge dy,$$
$$\mathrm{Re}(f_1(z)dz) \wedge \mathrm{Im}(f_2(z)dz) = -\,\mathrm{Im}(f_1(z)dz) \wedge \mathrm{Re}(f_2(z)dz) = \mathrm{Re}(f_1(z)\overline{f_2(z)})dx \wedge dy.$$

10.7. Definition of v_j, v_j', w_j. — We define

$$v_{2j-1} = \mathrm{Re}\Big(\pi^*\Big(\Big(\frac{q_j}{\sqrt{q}}\Big)dz\Big)\Big), \quad v_{2j} = \mathrm{Im}\Big(\pi^*\Big(\Big(\frac{q_j}{\sqrt{q}}\Big)dz\Big)\Big), \quad 1 \leqslant j \leqslant n-3,$$

$$v'_{2j-1} = \chi^*\Big(\mathrm{Re}\Big(\pi'^*\Big(\big(p_j/\sqrt{p}\big)dz\Big)\Big)\Big),\ v'_{2j} = \chi^*\Big(\mathrm{Im}\Big(\pi'^*\Big(\Big(\frac{p_j}{\sqrt{p}}\Big)dz\Big)\Big)\Big),\ 1 \leqslant j \leqslant n-3,$$

where $\chi : S' \to S$ is a lift of $\chi : (\overline{\mathbf{C}}, \varphi(Y)) \to (\overline{\mathbf{C}}, \psi(Y))$ (by abuse of notation) chosen so that $\sqrt{K}v = v'$, (rather than $\sqrt{K}v = -v$), where

$$v = \mathrm{Re}(\pi^*(\sqrt{q}dz)), \quad v' = \chi^*(\mathrm{Re}\,\pi'^*(\sqrt{p}dz)).$$

Define

$$f_j(z) = \frac{a_j q_j(z)}{z - b_j},$$

$$w_{2j-1} = \mathrm{Re}\Big(\pi^*\Big(*\Big(\frac{f_j}{\sqrt{q}}\Big)dz\Big)\Big), \quad w_{2j} = \mathrm{Im}\Big(\pi^*\Big(\Big(\frac{f_j}{\sqrt{q}}\Big)dz\Big)\Big), \quad 1 \leqslant j \leqslant n-3.$$

Note that $\pi^*((f_j/\sqrt{q})dz)$ is a meromorphic form on S with just one singularity — a double pole at $\pi^{-1}(b_j)$.

10.8. The Cup Form. — The v_j $(1 \leqslant j \leqslant 2n-6)$ are linearly independent, since the set of $v_{2j-1} + iv_{2j}$, for $1 \leqslant j \leqslant n-3$, is linearly independent. They form a basis for the space of harmonic forms on S, which intersects the space of coboundary forms trivially. Hence they form a basis for $H^1(S, \mathbf{R})$. Similarly, the $\chi_* v_j'$ form a basis for the harmonic forms on S'. So the $v_j' = (\chi^{-1})^* \chi_* v_j'$ also form a basis for $H^1(S, \mathbf{R})$, although these forms are not harmonic. (They can be treated as differential forms even though χ is not differentiable at the preimages of zeros of q, because χ can be approximated by diffeomorphisms.) Since each w_j is the real part of a meromorphic form with zero residue, it also defines an element of $H^1(S, \mathbf{R})$. The cup form

$$\cup : H^1(\mathbf{R}) \times H^1(\mathbf{R}) \longrightarrow \mathbf{R}$$

is nondegenerate antisymmetric and for differential 1-forms u, u' we have

$$u \cup u' = \int_S u \wedge u'.$$

This also holds if u or $u' \in \{v_j' : 1 \leqslant j \leqslant 2n-6\}$. It follows that

$$v_{2j-1} \cup v_{2k-1} = v_{2j} \cup v_{2k} = 2 \int \operatorname{Im} \frac{q_j \overline{q}_k}{|q|},$$

$$v_{2j-1} \cup v_{2k} = -v_{2j} \cup v_{2k-1} = 2 \int \operatorname{Re} \frac{q_j \overline{q}_k}{|q|}.$$

We also have

$$v_j \cup v_k' = \chi_* v_j \cup \chi_* v_k', \quad v_j' \cup v_k' = \chi_* v_j' \cup \chi_* v_k'.$$

Note that there is no reason to suppose that $v_j \cup w_k = \int_S v_j \wedge w_k$, and in fact this is not true in general. I shall be doing the calculation in 11.10. However, we do have

$$\int_S w_{2j-1} \wedge v_{2k-1} = 2 \int \operatorname{Im} \frac{q_j \overline{q}_k}{(z - b_j)|q|},$$

and so on. If $j = k$, this has to be interpreted as an improper integral.

10.9. Definition of J, J'. — We define a linear map J on $H^1(S, \mathbf{R})$ by: $u + iJu$ is the unique holomorphic form with real part u. Then $J^2 = -I$ and J preserves $\cup$. Note that

$$Jv_{2j-1} = v_{2j}, \; Jv_{2j} = -v_{2j-1}.$$

We define J' by: if $u' = \chi^* u$ for u harmonic and $u + ix$ holomorphic on S', then $J'u' = \chi^* x$. Then, again, $J'^2 = -I$ and J' preserves $\cup$. We see that

$$v + iJV = \pi^*(\sqrt{q}dz), \quad \chi_* v' + i\chi_* J'v' = \pi'^*(\sqrt{p}dz) = \chi_*\Big(\sqrt{K}v + \frac{iJv}{\sqrt{K}}\Big),$$

so

$$v' = \sqrt{K}v, \quad J'v' = \frac{Jv}{\sqrt{K}}, \quad J'v = \frac{Jv}{K}, \quad J'Jv = -Kv.$$

Note also that

$$v \cup Jv = 2 \int |q| = 2.$$

10.10. Definitions of V, W, V', W', U, U', T. — Define

$$V = (v_{i,j}) = (v_j \cup v_i), \quad W = (w_{i,j}) = \Big(\int w_j \wedge v_i \Big), \quad T = (v'_j \cup v_i)$$

We define V', W' similarly using v'_j and w'_j. For any $X \subset H^1(S, \mathbf{R})$, let

$$X^\perp = \{u : u \cup x = 0 \text{ for all } x \in X\}.$$

Of course $x \in \{x\}^\perp$ for all x. Now let u_j (or u'_j) be the projection of v_j (or v'_j) onto $\{v, Jv\}^\perp = \{v', J'v'\}^\perp$ along $\mathrm{sp}(v, Jv) = \mathrm{sp}(v', J'v')$. Then let U, U', T_1 be obtained by deleting the last two rows and columns from the matrices with (i,j)-entries

$$u_j \cup u_i, \quad u'_j \cup u'_i, \quad u'_j \cup u_i.$$

It can be checked that

$$V' = -T^t V^{-1} T, \quad U' = -T_1^t U^{-1} T_1.$$

Since $\cup$ is nondegenerate, V is invertible. Also, for any j, $u_{2j-1} + iu_{2j}$ can be expressed as a complex linear combination of $u_{2k-1} + iu_{2k}$ for $k \neq j$. It follows that U is invertible. The space spanned by

$$\{u_k : k \neq 2j-1, 2j\} \quad \text{or} \quad \{u'_k : k \neq 2j-1, 2j\}$$

is $\{v, Jv\}^\perp$, (and the same as the space spanned by all u_k, or all u'_k). In particular, T_1 is invertible.

10.11. Definition of A, A', B, B', Π, Σ_1, E_1. — The matrix A is the $2n-6$-dimensional row vector

$$(\overline{a}_1, \ldots, \overline{a}_{n-3}),$$

where we identify the complex number $c + id$ with the row vector (c, d) in the usual way. The matrix A' is defined similarly. If we write $M(c+id)$ for the 2×2 real matrix

$$\begin{pmatrix} c & -d \\ d & c \end{pmatrix},$$

then B is given by the matrix which has the matrices

$$M\left(\frac{a_j(2b_j - 1)}{b_j(b_j - 1)} \right)$$

down the diagonal ($1 \leqslant j \leqslant n-3$) and zeros elsewhere. The matrix B' is defined similarly using a'_j and b'_j. The $(2n-6) \times (2n-6)$ matrices Σ, Π have the 2×2 matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = M(i), \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

respectively down the diagonal, and zeros elsewhere. The matrix $(2n-8) \times (2n-8)$ matrix Σ_1 is defined similarly to Σ. The matrix E_1 is given by deleting the last two columns of E, where

$$E = I - A^t AV\Sigma - \Sigma A^t AV.$$

The matrix E' is defined similarly.

10.12. Symmetric and Positive Matrices. — The matrix $-V\Sigma = (v_i \cup Jv_j)$ is symmetric and positive, since

$$\begin{pmatrix} v_{2i-1} \cup Jv_{2j-1} & v_{2i-1} \cup Jv_{2j} \\ v_{2i} \cup Jv_{2j-1} & v_{2i} \cup Jv_{2j} \end{pmatrix} = M\left(\int 2\frac{q_i\overline{q}_j}{\mid q \mid}\right).$$

In fact, $-V\Sigma$ is positive definite, since it is invertible (even though it is an integral of rank one matrices). Similarly, $-V'\Sigma$, $-U\Sigma_1$, $-U'\Sigma$ are symmetric and positive definite. So the same is true for the inverses $V^{-1}\Sigma$, $V'^{-1}\Sigma$, $U^{-1}\Sigma_1$, $U'^{-1}\Sigma_1$.

The matrices ΠB and $\Pi B'$ are symmetric and of zero trace (and hence not positive). We shall see in 11.11 that $-V^{-1}W\Pi$ and $V'^{-1}W'\Pi$ are also symmetric — that does not seem to be obvious, although it is clear that they have zero trace.

10.13. An Inner Product on Harmonic 1-forms on S. — We define a positive-definite inner product $\langle\, ,\, \rangle$ on harmonic 1-forms on S by

$$\langle x, y\rangle = x \cup Jy.$$

Of course, we can also regard this as an inner product on $H^1(S)$.Then $\{v, Jv\}^\perp$, defined as before, is also the orthogonal complement of $\{v, Jv\}$ with respect to $\langle\, ,\, \rangle$. Note that the linear maps JJ' and $J'J$ are symmetric with respect to this inner product, while J is skew-symmetric and J' has adjoint $-JJ'J$. Write

$$C = -JJ' \text{ on } \{v, Jv\}^\perp.$$

Then it is easily checked that C is positive with respect to this inner product. In fact we shall see later (in 11.8) that

$$\frac{I}{K} < C < KI.$$

If we defined C on the whole space $H^1(S)$ we would have to replace $<$ by $\leqslant$. Note that we have

$$(\langle u_i, u_j\rangle) = -U\Sigma_1, \quad (\langle u'_i, Cu'_j\rangle) = -U'\Sigma_1, \quad (\langle u'_i, u_j\rangle) = -T_1\Sigma_1,$$
$$(\langle u_i, Cu_j\rangle) = \Sigma_1 T_1 \Sigma_1 U T_1^{-1}.$$

10.14. Rules Relating Linear maps on $H^1(S)$ and Matrices of Integrals

Let

$$X : \mathbf{R}^{2n-8} \longrightarrow \{v, Jv\}^\perp$$

be defined by

$$X((x_i)) = u \text{ if } \langle u, u_i\rangle = x_i \text{ for all } i.$$

Then, if $P : \{v, Jv\}^\perp \to \{v, Jv\}^\perp$ is any linear map, the matrix of $X^{-1}PX : \mathbf{R}^{2n-8} \to \mathbf{R}^{2n-8}$ (with respect to the standard basis on $\mathbf{R}^{2n-8}$) is given by

$$(\langle u_i, Pu_j\rangle)U^{-1}\Sigma_1 = (u_i \cup JPu_j)U^{-1}\Sigma_1.$$

To see this, note that if $\langle u, u_i\rangle = x_i$ for all i, and

$$u = \sum_j b_j u_j,$$

then

$$\Big(\sum_j b_j u_j\Big) \cup Ju_i = x_i \iff (u_j \cup Ju_i)\underline{b} = \underline{x} \iff -\Sigma_1 U\underline{b} = \underline{x} \iff \underline{b} = U^{-1}\Sigma_1\underline{x},$$

where $\underline{x}$ and $\underline{b}$ denote the column vectors (x_i) and (b_i) respectively. The result then follows, since

$$X^{-1}PX\underline{x} = (\langle Pu, u_i\rangle) = (\langle u_i, Pu\rangle) = (\langle u_i, Pu_j\rangle)\underline{b} = (\langle u_i, Pu_j\rangle)U^{-1}\Sigma_1\underline{x}.$$

As a consequence, we have the following:

Rule 1. — $(\langle u_i, P_1P_2u_j\rangle) = (\langle u_i, P_1u_j\rangle)U^{-1}\Sigma_1(\langle u_i, P_2u_j\rangle)$.

Rule 2. — $I = (\langle u_i, Pu_j\rangle)U^{-1}\Sigma_1(\langle u_i, P^{-1}u_j\rangle)U^{-1}\Sigma_1$.

10.15. Some Symmetric Positive Linear Maps on $H^1(S)$. — We define the following linear maps, all of which are rational functions of C as in 10.13, and hence commute.

$$D = I - \frac{C}{K}, \quad F = D^{-1}\Big(I + \frac{C}{K}\Big), \quad G = \frac{-2CD^{-1}}{\sqrt{K}}, \quad H = CF.$$

Note that

$$D > 0, \quad FH - G^2 = D^{-2}C\Big(\Big(I + \frac{C}{K}\Big)^2 - 4\frac{C}{K}\Big) = C > 0,$$

assuming Lemma 11.8. Now we are ready to describe our formula for the second derivative of Teichmüller distance.

10.16. The Second Derivative Formula. — The second derivative formula $2\pi R$ is given by

$$R = R_1 + R_2 + R_3,$$

where

$$R_1 = \begin{pmatrix} -V^{-1}W\Pi - \Pi B & 0 \\ 0 & V'^{-1}W'\Pi + \Pi B' \end{pmatrix},$$

$$R_2 = \frac{2\pi}{K^2-1}P^t\begin{pmatrix} K^2+1 & 2K \\ 2K & K^2+1 \end{pmatrix}P, \quad \text{where} \quad P = \begin{pmatrix} A\Sigma & 0 \\ 0 & -A'\Sigma \end{pmatrix},$$

$$R_3 = 4\pi Q\begin{pmatrix} (\langle Fu_i, u_j\rangle) & (\langle Gu_i, u'_j\rangle) \\ (\langle Gu'_i, u_j\rangle) & (\langle Hu'_i, u'_j\rangle) \end{pmatrix}Q^t, \quad \text{where} \quad Q = \begin{pmatrix} E_1U^{-1}\Sigma_1 & 0 \\ 0 & E'_1U'^{-1}\Sigma_1 \end{pmatrix}.$$

The kernel $K_1 \oplus K_2$ in $\mathbf{R}^{2n-6} \oplus \mathbf{R}^{2n-6}$ of Q^t above is such that K_j has dimension 2.

10.17. Remark. — We have already noted symmetry of R_2 and R_3, and have indicated that symmetry of R_1 will be proved in 11.11. It is also clear that R_2 is positive (of rank two). Also, by 10.15, R_3 is positive, but has a kernel, as stated in 10.16. We shall see in 11.8 that the kernel is as claimed. We shall also see that the kernel of $R_2 + R_3$ has dimension 2. We expect something like this to be true, for the following reason. Let $[\chi]$ be a pseudo-Anosov isotopy class. Then we know that the critical points of $[\varphi] \mapsto d([\varphi], [\varphi \circ \chi])$ are minima, and lie on a geodesic in Teichmüller space. Therefore, we expect the second derivative of this map to be positive, with a nullspace of dimension one. This means that if $2\pi R$ is the second derivative matrix of $(x, y) \mapsto d(x, y)$ at $(x, y) = ([\varphi], [\varphi \circ \chi])$, then we expect

$$(\underline{h}^t\ \underline{h}^t)\, R \begin{pmatrix} \underline{h} \\ \underline{h} \end{pmatrix} \geqslant 0,$$

with equality for $\underline{h}$ in a one-dimensional subspace. At such a point we have $B = B'$, $V = V'$, $W = W'$. So the first term in the Second Derivative Formula vanishes.

10.18. Start of Proof: the Equation for the Second Derivative Matrix

Let $\underline{h} = (h_j)$ and $\underline{h}' = (h'_j)$ be small. Let χ_1 be the quasiconformal homeomorphism of minimal distortion with $\chi_1 \circ (\varphi + \underline{h})] = [\psi + \underline{h}']$. Let $q_1(z)dz^2$ be the quadratic differential at $[\varphi + \underline{h}]$ for $d([\varphi + \underline{h}], [\psi + \underline{h}'])$, and let $p_1(z)dz^2$ be the stretch of $q_1(z)dz^2$ at $[\psi + \underline{h}']$. Let the residues at $b_j + h_j$, $b'_j + h'_j$ for q_1, p_1 be $a_j + k_j$, $a'_j + k'_j$. Let $K(1 + \varepsilon) = K(\chi_1)$ (recalling $K = K(\chi)$, $\log K = d([\varphi], [\psi])$). Then if $\overline{\underline{k}} = (\overline{k}_j)$ (where $\overline{k}_j$ denotes complex conjugate), the second derivative matrix $2\pi R$ (if it exists) is given by

$$\text{(SDE1)} \qquad \begin{pmatrix} \overline{\underline{k}} \\ \overline{\underline{k}}' \end{pmatrix} = R \begin{pmatrix} \underline{h} \\ \underline{h}' \end{pmatrix} + o(\underline{h}) + o(\underline{h}').$$

Conversely, if R exists satisfying (SDE1), then $2\pi R$ is the second derivative. We claim that (SDE1) is implied by the existence of a unique solution (x, x', ε) to (SDE2) and similar equations involving h'_ℓ, where

$$x + iJx = \sum_{j=1}^{n-3} k_j(v_{2j-1} + iv_{2j}), \qquad x' + iJ'x' = \sum_{j=1}^{n-3} k'_j(v'_{2j-1} + iv'_{2j}),$$

$$w + it = \sum_{j=1}^{n-3} h_j(w_{2j-1} + iw_{2j}), \qquad b + iJb = \sum_{\ell=1}^{n-3} \frac{a_\ell(2b_\ell - 1)h_\ell}{b_\ell(b_\ell - 1)}(v_{2\ell-1} + iv_{2\ell}).$$

$$\text{(SDE2)} \qquad \begin{gathered} \begin{pmatrix} x \\ Jx \end{pmatrix} - \begin{pmatrix} \sqrt{1/K} & 0 \\ 0 & \sqrt{K} \end{pmatrix} \begin{pmatrix} x' \\ J'x' \end{pmatrix} + \varepsilon \begin{pmatrix} v \\ -Jv \end{pmatrix} = -\begin{pmatrix} b \\ Jb \end{pmatrix} - \begin{pmatrix} w \\ t \end{pmatrix}, \\ x \cup Jv = -b \cup Jv - \int w \wedge Jv. \end{gathered}$$

Here, (SDE2) is to be regarded as an equation for real and imaginary parts in $H^1(S)$: we are *not* requiring pointwise equality for the forms on each side of the equation, which is clearly impossible, since singularities occur on the righthand side only.

Proof. — Take $h'_m = 0$ for all m. Write

$$S_1 = \{(z,u) \in \overline{\mathbf{C}}^2 : q_1(z) = u^2\},$$

and let S'_1 be similarly defined using p_1. Let $\pi_1 : S_1 \to \overline{\mathbf{C}}$, $\pi'_1 : S'_1 \to \overline{\mathbf{C}}$ be the projections onto the first coordinate. By abuse of notation write $\chi_1 : S_1 \to S'_1$ for the lift through π_1, π'_1 of $\chi_1 : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$. Then the relationship between q_1 and p_1 can be expressed as

$$\pi_1^*(\sqrt{q_1}dz) = \begin{pmatrix} (1/\sqrt{K})(1-\varepsilon)^{1/2} & 0 \\ 0 & \sqrt{K}(1+\varepsilon)^{1/2} \end{pmatrix} \chi_1^* \pi_1'^*(\sqrt{p_1}dz),$$

where we identify complex-valued forms with column vectors of real-valued forms, just as we had similar equations involving q, p, χ, π, π'. This can be rewritten in the following way. Let γ be any loop on S_1, and let $\gamma' = \chi_1\gamma$. Then

$$(1) \qquad \int_\gamma \pi_1^*(\sqrt{q_1}dz) = \begin{pmatrix} (\sqrt{1/K})(1-\varepsilon)^{1/2} & 0 \\ 0 & \sqrt{K}(1+\varepsilon)^{1/2} \end{pmatrix} \int_{\gamma'} \pi_1'^*(\sqrt{p_1}dz).$$

We also have

$$(2) \qquad \int |q_1| = 1.$$

We can assume γ, γ' are bounded away from the zeros and poles of q, p (and hence also from the zeros and poles of q_1, p_1). Then, projecting, (1) gives

$$(3) \qquad \int_{\pi_1\gamma} \sqrt{q_1}dz = \begin{pmatrix} \sqrt{1/K}(1-\varepsilon)^{1/2} & 0 \\ 0 & \sqrt{K}(1+\varepsilon)^{1/2} \end{pmatrix} \int_{\pi'_1\gamma'} \sqrt{p_1}dz.$$

Now

$$(z - b_\ell - h_\ell)^{-1} = (z-b_\ell)^{-1} + h_\ell(z-b_\ell)^{-2} + o(h_\ell).$$

So

$$\frac{a_\ell((b_\ell+h_\ell)^2 - (b_\ell+h_\ell))}{z(z-1)(z-b_\ell-h_\ell)} = a_\ell q_\ell + h_\ell f_\ell + \frac{a_\ell(2b_\ell-1)h_\ell}{b_\ell(b_\ell-1)} q_\ell + o(h_\ell),$$

and

$$\sqrt{q_1(z)} = \sqrt{q(z) + \sum_{j=1}^{n-3} k_j q_j + \sum_{\ell=1}^{n-3}\left(h_\ell f_\ell + \frac{(2b_\ell-1)a_\ell h_\ell}{b_\ell(b_\ell-1)} q_\ell\right)} + o(\underline{h}) + o(\underline{k})$$

$$(4) \qquad = \sqrt{q(z)} + \sum_{j=1}^{n-3} \frac{k_j q_j}{2\sqrt{q(z)}} + \sum_{\ell=1}^{n-3}\left(\frac{h_\ell f_\ell}{2\sqrt{q(z)}} + \frac{(2b_\ell-1)a_\ell h_\ell q_\ell}{2b_\ell(b_\ell-1)\sqrt{q(z)}}\right) + o(\underline{h}) + o(\underline{k}),$$

$$(5)\quad |q_1| = |q| + |q|\operatorname{Re}\left(\sum_{j=1}^{n-3}\frac{k_jq_j}{q(z)} + \sum_{\ell=1}^{n-3}\left(\frac{h_\ell f_\ell}{q(z)} + \frac{(2b_\ell-1)a_\ell h_\ell q_\ell}{b_\ell(b_\ell-1)q(z)}\right)\right) + o(\underline{h}) + o(\underline{k}).$$

We have a similar expression for p_1 in terms of p — except that there is no h'_m term for any m, since $h'_m = 0$ for all m (and no h_m term for any m either). We also have

$$(6)\quad \begin{aligned}\sqrt{1/K}(1-\varepsilon)^{1/2} &= \sqrt{1/K} - \sqrt{1/K}(\varepsilon/2) + o(\varepsilon),\\ \sqrt{K}(1+\varepsilon)^{1/2} &= \sqrt{K} + \sqrt{K}(\varepsilon/2) + o(\varepsilon),\end{aligned}$$

$$(7)\quad \begin{pmatrix}\sqrt{1/K} & 0\\ 0 & \sqrt{K}\end{pmatrix}\overline{\int_{\pi'_1\gamma'}\sqrt{p(z)}dz} = \overline{\int_{\pi_1\gamma}\sqrt{q(z)}dz}.$$

Putting (4) to (7) in (3), (2), and then lifting up via π, π', we obtain (SDE2), up to the addition of a term $o(\underline{h}) + o(\underline{k}) + o(\underline{k}') + o(\varepsilon)$. The solution to (SDE2) is unique if and only if, for any solution (x, x', ε) which is correct to within $o(\underline{h}) + o(\underline{k}) + o(\underline{k}') + o(\varepsilon)$, x, x' and ε are all $O(\underline{h})$. So the existence of a unique solution to (SDE2) implies (SDE1). □

CHAPTER 11

SOLVING THE SECOND DERIVATIVE EQUATION

11.1. In this chapter, we shall prove the Second Derivative Formula of 10.16 by solving equation (SDE2) of 10.18. We also give an independent verification of the fact that the matrix R of 10.16 is symmetric. This essentially means showing that the matrix R_1 of 10.16 is symmetric.

Throughout this chapter, we shall use expressions such as (h_j) to denote column vectors. Row vectors will be denoted by expressions such as $(h_j)^t$.

11.2. Splitting up the Equation (SDE2). — To solve (SDE2) of 10.18, it clearly suffices to solve in the case when $h'_j = 0$ for all j, and the case when $h_j = 0$ for all j. Apart from briefly in 11.8, from now on we suppose that $h'_j = 0$ for all j. Write

$$h_j = \theta_{2j-1} + i\theta_{2j}.$$

Consider the righthand side of (SDE2). We are going to write it as a sum of 3 terms. Let $w^{(1)}$ be the unique element of $H^1(S)$ with

$$w^{(1)} \cup v_j = \int w \wedge v_j, \quad 1 \leqslant j \leqslant 2n-6.$$

Write $w^{(2)} = w - w^{(1)}$. Then the righthand side of (SDE2) can be written as

$$\begin{pmatrix} -b \\ -Jb \end{pmatrix} + \begin{pmatrix} -w^{(1)} \\ -Jw^{(1)} \end{pmatrix} + \begin{pmatrix} -w^{(2)} \\ -t + Jw^{(1)} \end{pmatrix}.$$

We let (SDE2.1), (SDE2.2), (SDE2.3) be the equations with the same lefthand side as (SDE2), and first, second and third terms, respectively, from the righthand side of (SDE2). Then it is clear that (SDE2.1) and (SDE2.2) can be solved immediately by taking $x' = 0$ and $x = -b$, $x = -w^{(1)}$ respectively. So to solve (SDE2), we only need to solve (SDE2.3) and add the solutions to (SDE2.1), (SDE2.2), (SDE2.3). Specifically, we write

$$\begin{pmatrix} \underline{k} \\ \underline{k}' \end{pmatrix} = \left(\begin{pmatrix} S^{(1,1)} \\ 0 \end{pmatrix} + \begin{pmatrix} S^{(2,1)} \\ 0 \end{pmatrix} + \begin{pmatrix} S^{(3,1)} \\ S^{(3,2)} \end{pmatrix} \right) (\theta_\ell),$$

(SDE2.1) $$(v_j)^t S^{(1,1)}(\theta_\ell) = -b,$$

(SDE2.2) $$(v_j)^t S^{(2,1)}(\theta_\ell) = -w^{(1)}.$$

We shall show in 11.5 and 11.9 that

$$t = Jw^{(1)} - Jw^{(2)}.$$

Then, ε and

$$\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} (v_j)^t S^{(3,1)} \\ (v_j')^t S^{(3,2)} \end{pmatrix} (\theta_i)$$

give a solution to

(SDE2.3) $$\begin{pmatrix} x \\ Jx \end{pmatrix} - \begin{pmatrix} \sqrt{1/K} & 0 \\ 0 & \sqrt{K} \end{pmatrix} \begin{pmatrix} x' \\ J'x' \end{pmatrix} + \varepsilon \begin{pmatrix} v \\ -Jv \end{pmatrix} = \begin{pmatrix} -w^{(2)} \\ Jw^{(2)} \end{pmatrix}, \quad x \cup Jv = 0.$$

We also need to show uniqueness of solutions. So to solve (SDE2), we need to solve (SDE2.3), as well as giving the form of $S^{(1,1)}$ and $S^{(2,1)}$.

11.3. The form of $S^{(1,1)}$, $S^{(2,1)}$. — Let B be as in 10.11, and V and W as in 10.12. Let

$$\frac{a_j(2b_j - 1)}{b_j(b_j - 1)} = \beta_{2j-1} + i\beta_{2j}, \quad 1 \leqslant j \leqslant n-3,$$

Then

$$(v_j)^t S^{(1,1)}(\theta_i) = -\sum_{k=1}^{n-3}((\beta_{2k-1}\theta_{2k-1} - \beta_{2k}\theta_{2k})v_{2k-1} - (\beta_{2k}\theta_{2k-1} + \beta_{2k-1}\theta_{2k})v_{2k}),$$

which gives

$$S^{(1,1)} = -\Pi B.$$

Taking cup product with (v_i) on the left gives

$$-VS^{(2,1)} = W\Pi, \quad S^{(2,1)} = -V^{-1}W\Pi.$$

It follows that $S^{(1,1)} + S^{(2,1)}$ is the top left quarter of R_1 in 10.16. So in solving (SDE2.3), we shall want to see the left half of the matrix $R_2 + R_3$ from the formula of 10.16.

11.4. First Solution of (SDE2.3). — To solve (SDE2.3), write

$$-w^{(2)} = \gamma Jv + \delta v + z,$$
$$x = \alpha Jv + m,$$
$$x' = \alpha' Jv' + \beta' v' + m' = \frac{\alpha'}{\sqrt{K}} Jv + \beta' \sqrt{K} v + m'$$

with z, m, $m' \in \{v, Jv\}^{\perp}$. (We have incorporated the information $x \cup Jv = 0$.) Then (SDE2.3) gives two equations, which become six when we equate the components in the span of each of v, Jv, and in $\{v, Jv\}^{\perp}$. Altogether, we obtain

$$\alpha - \frac{\alpha'}{K} = \gamma, \qquad -\alpha + K\alpha' = \gamma,$$
$$-\beta'\sqrt{K} + \varepsilon = \delta, \qquad -\frac{\beta'}{\sqrt{K}} - \varepsilon = -\delta,$$
$$m - \frac{m'}{\sqrt{K}} = z, \qquad Jm - \sqrt{K}J'm' = -Jz.$$

From these, we obtain

$$\alpha = \gamma\frac{K^2+1}{K^2-1}, \quad \alpha' = \frac{2\gamma K}{K^2-1}, \quad \beta' = 0, \quad \varepsilon = \delta.$$

The last two equations give

$$\Big(-\frac{1}{\sqrt{K}} - \sqrt{K}JJ'\Big)m' = 2z, \quad \Big(1 + \frac{J'J}{K}\Big)m = \Big(1 - \frac{J'J}{K}\Big)z.$$

Remembering $C = -JJ'$ and $-J'J = C^{-1}$ and using the notation of 10.15, we obtain

$$m = \Big(C - \frac{I}{K}\Big)^{-1}\Big(C + \frac{I}{K}\Big)z = Fz,$$
$$m' = \frac{2}{\sqrt{K}}\Big(C - \frac{I}{K}\Big)^{-1}z = \frac{2}{\sqrt{K}}D^{-1}z.$$

11.5. More Information about the Righthand side of (SDE2.3). — It follows from the lemma below that

$$w^{(2)}_{2j-1} \cup v_{2j} = -4\pi, \qquad w^{(2)}_{2j-1} \cup v_k = 0, \quad k \neq 2j,$$
$$w^{(2)}_{2j} \cup v_{2j-1} = -4\pi, \qquad w^{(2)}_{2j} \cup v_k = 0, \quad k \neq 2j-1.$$

This implies that, for all j, k,

$$w^{(2)}_{2j-1} \cup Jv_k = w^{(2)}_{2j} \cup v_k, \quad \text{so} \quad w^{(2)}_{2j} = -Jw^{(2)}_{2j-1}, \quad w^{(2)}_{2j-1} = Jw^{(2)}_{2j}.$$

Thus

$$\sum_{j=1}^{n-3}\theta_{2j-1}w^{(2)}_{2j} + \sum_{j=1}^{n-3}\theta_{2j}w^{(2)}_{2j-1} = -J\Big(\sum_{j=1}^{n-3}\theta_{2j-1}w^{(2)}_{2j-1} - \sum_{j=1}^{n-3}\theta_{2j}w^{(2)}_{2j}\Big),$$

which justifies the righthand side of (SDE2.3).

Lemma. — *Let μ be a meromorphic 1-form on a compact Riemann surface S with zero residues at all singularities and at worst double poles. Let ν be a holomorphic 1-form on S. Let $p_j \in S$ $(1 \leqslant j \leqslant r)$ be those singularities of μ which are not zeros* *of ν. At p_j, choose local coordinates so that $\nu = d\zeta$ and the ζ^{-2}- term in the Laurent expansion of μ is $\dfrac{a_j}{\zeta^2}d\zeta$. Write $\nu = \nu_1 + i\nu_2$, where ν_1, ν_2 are harmonic. Then*

$$\mu \cup \nu_1 = \int_S \mu \wedge \nu_1 - \pi i \sum_{j=1}^{r} a_j,\ \mu \cup \nu_2 = \int_S \mu \wedge \nu_2 - \pi \sum_{j=1}^{r} a_j.$$

We shall prove this lemma in 11.10. Here, we explain why it gives the required identities for $w_j^{(2)} \cup v_k$. We take $\mu = w_{2j-1} + iw_{2j}$ and $\nu = v_{2k-1} + iv_{2k}$. The only singularity of μ is a single double pole. If $k \neq j$ then this double pole coincides with a zero of ν, and the lemma gives

$$\mu \cup \nu_1 = \int_S \mu \wedge \nu_1,\ \mu \cup \nu_2 = \int_S \mu \wedge \nu_2.$$

This gives all the required identities involving $w_{2j-1}^{(2)}$, $w_{2j}^{(2)}$, v_{2k-1} and v_{2k} for $k \neq j$. So now suppose that $k = j$. Then at the double pole of μ, if ν is expressed in local coordinates as $d\zeta$, we see that μ takes the form $(4/\zeta^2)d\zeta$. So the lemma gives

$$\mu \cup \nu_1 - \int_S \mu \wedge \nu_1 = -4\pi i,\ \mu \cup \nu_2 - \int_S \mu \wedge \nu_2 = -4\pi,$$

or, as required,

$$w_{2j} \cup v_{2j-1} - \int_S w_{2j} \wedge v_{2j-1} = -4\pi = w_{2j-1} \cup v_{2j} - \int_S w_{2j-1} \wedge v_{2j}.$$

11.6. Interpreting the Solution of (SDE2.3). — We have obtained expressions for α, α' m, m' in terms of γ, δ, z. (We do not yet have expressions for γ, δ, z in terms of known quantities.) The relation to the matrix $S^{(3)}$ is given by

$$(1) \qquad \begin{pmatrix} \alpha J v + m \\ -\alpha' J' v' - m' \end{pmatrix} = \begin{pmatrix} (v_j)^t & 0 \\ 0 & (v_j')^t \end{pmatrix} S^{(3)}(\theta_\ell).$$

Let A be the matrix of 10.11, which is a row matrix. We can write

$$v + iJv = \sum_{j=1}^{n-3} a_j(v_j + iJv_j),\ v = (v_j)A^t,\ Jv = (v_j)\Sigma A^t,$$

and similarly for v', $J'v'$, using A' instead of A. Let U, U' be as in 10.12. Then for any $y \in \{v, Jv\}^\perp$, we have

$$y = (u_j)^t U^{-1}(y \cup u_j) = (u_j')^t U'^{-1}(y \cup u_j').$$

Here, we have $1 \leqslant j \leqslant 2n-8$, since U is a $(2n-8) \times (2n-8)$ matrix. So

$$\begin{aligned} m &= (u_j)U^{-1}(m \cup u_k) = (u_j)U^{-1}(Fz \cup u_j) = (u_j)U^{-1}(Fu_k \cup u_j)U^{-1}(z \cup u_k) \\ &= (u_j)U^{-1}\Sigma_1(\langle u_j, Fu_k\rangle)U^{-1}(z \cup u_k), \end{aligned}$$

$$
\begin{aligned}
m' &= (u'_j)U'^{-1}(m' \cup u'_j) = (u'_j)U'^{-1}\Big(\frac{2D^{-1}z}{\sqrt{K}} \cup u'_j\Big) \\
&= (u'_j)U'^{-1}\Big(\frac{2D^{-1}u_k}{\sqrt{K}} \cup u'_j\Big)U^{-1}(z \cup u_k) \\
&= (u'_j)U'^{-1}\Sigma_1\Big(\frac{2D^{-1}u_k}{\sqrt{K}} \cup (-J')u'_j\Big)U^{-1}(z \cup u_k) \\
&= (u'_j)U'^{-1}\Sigma_1\Big(u'_j \cup \frac{2}{\sqrt{K}}J(-JJ')D^{-1}u_k\Big)U^{-1}(z \cup u_k) \\
&= (u'_j)U'^{-1}\Sigma_1\Big\langle u'_j, \frac{2}{\sqrt{K}}CD^{-1}u_k\Big\rangle U^{-1}(z \cup u_k).
\end{aligned}
$$

To proceed further, we need more information about z, γ, δ.

11.7. We first note that

$$-(w^{(2)} \cup v_i) = 4\pi\Sigma(\theta_\ell).$$

Next we analyse E, defined by

$$(v_j)E = (u_1, \ldots, u_{2n-6}).$$

Now for all j,

$$u_j = v_j - \frac{v_j \cup Jv}{2}v + \frac{v_j \cup v}{2}Jv,$$

since $v \cup Jv = 2$. So

$$
\begin{aligned}
(v_j)E &= (v_j) - v\Big(\frac{v_j \cup Jv}{2}\Big) + Jv\Big(\frac{v_j \cup v}{2}\Big) \\
&= (v_j) - (v_\ell)A^t\Big(\frac{v_j \cup Jv}{2}\Big) + (v_\ell)\Sigma A^t\Big(\frac{v_j \cup v}{2}\Big),
\end{aligned}
$$

where (v_j), (v_ℓ) are row vectors. Applying $\cup v_i$ on the right, we obtain

$$VE = V - VA^t\Big(\frac{v_j \cup Jv}{2}\Big) + V\Sigma A^t\Big(\frac{v_j \cup v}{2}\Big),$$

or

$$E = I - \frac{A^tA\Sigma V}{2} + \frac{\Sigma A^tAV}{2} = I - \frac{A^tAV\Sigma}{2} + \frac{\Sigma A^tAV}{2}.$$

So E is as defined in 10.11. It follows that $\Sigma E = E\Sigma$, and $E(V\Sigma)^{-1}$ is symmetric, or

$$EV^{-1}\Sigma = V^{-1}\Sigma E^t.$$

So

$$\Sigma = VV^{-1}\Sigma = VEV^{-1}\Sigma + V(I-E)V^{-1}\Sigma,$$

where

$$
\begin{aligned}
VEV^{-1}\Sigma &= VV^{-1}\Sigma E^t = \Sigma E^t, \\
V(I-E)V^{-1}\Sigma &= \frac{1}{2}VA^tA - \frac{1}{2}V\Sigma A^tA\Sigma = \frac{1}{2}(v \cup v_i)A - \frac{1}{2}(Jv \cup v_i)A\Sigma.
\end{aligned}
$$

So the component of $-w^{(2)}$ in $\mathrm{sp}(v, Jv)$ is

$$2\pi(A(\theta_\ell))v - 2\pi(A\Sigma(\theta_\ell))Jv.$$

So

$$\gamma = -2\pi A\Sigma(\theta_\ell), \quad \delta = 2\pi A(\theta_\ell).$$

Since $\delta = \varepsilon$, we recover the First Derivative Formula 8.4. We also obtain, for $1 \leqslant i \leqslant 2n-8$,

$$(z \cup u_i) = (-w^{(2)} \cup u_i) = E_1^t(-w^{(2)} \cup v_i) = 4\pi E_1^t \Sigma(\theta_\ell) = 4\pi \Sigma_1 E_1^t(\theta_\ell).$$

So we obtain

$$\alpha J v = -2\pi \frac{K^2+1}{K^2-1}(v_j)^t \Sigma A^t A \Sigma(\theta_\ell) = 2\pi \frac{K^2+1}{K^2-1}(v_j)^t \Sigma^t A^t A \Sigma(\theta_\ell)$$

and similarly

$$\begin{aligned} \alpha' J v' &= 2\pi \frac{K}{K^2-1}(v_j')^t \Sigma^t A'^t A \Sigma(\theta_\ell), \\ m &= 4\pi (v_j)^t E_1 U^{-1} \Sigma_1 \langle u_j, F u_k \rangle U^{-1} \Sigma_1 E_1^t(\theta_\ell), \\ m' &= 4\pi (v_j')^t E_1' U'^{-1} \Sigma_1 (\langle u_j', -G u_k \rangle) U^{-1} \Sigma_1 E_1^t(\theta_\ell). \end{aligned}$$

Then (1) of 11.6 gives an expression for $S^{(3)}$. This agrees with the left half of the matrix $R_2 + R_3$ of the Second Derivative Formula 10.16. Thus, as far as the first half of the matrix R is concerned, the proof of the Second Derivative Formula is now reduced to proving the lemma of 11.5.

11.8. The Second Half. — Now we consider the changes needed to solve (SDE2.3) in the case when $h_j = 0$ for all j, but $h_j' \neq 0$. (The changes to (SDE2.1) and (SDE2.2) are clear.) These are obtained from the first half by interchanging J and J' and replacing u_i by $J'u_i'$, u_i' by Ju_i. Thus U is replaced by U' (since J' preserves $\cup$) and $\langle u_j, F u_k \rangle = u_j \cup JFu_k$ is replaced by

$$J'u_j' \cup J'(J'F(-J'))J'u_k' = u_j' \cup JCFu_k' = \langle u_j', Hu_k' \rangle,$$

since

$$J'FJ' = J'\Big(I - \frac{JJ'}{K}\Big)\Big(I + \frac{JJ'}{K}\Big)^{-1}(-J') = \Big(I + \frac{J'J}{K}\Big)\Big(I - \frac{J'J}{K}\Big)^{-1}.$$

Similarly

$$\langle u_j', G u_k \rangle = u_j' \cup JGu_k$$

is replaced by

$$Ju_j \cup J'J'G(-J')J'u_k' = u_j \cup JGu_k' = \langle u_j, Gu_k' \rangle$$

since

$$J'G(-J') = J'(-JJ')\Big(I + \frac{JJ'}{K}\Big)^{-1}(-J').$$

11.9. Lemma. — *Define an inner product on harmonic forms by*

$$\langle \omega, \omega' \rangle = \omega \cup J\omega'.$$

Then with respect to this inner product on $\{v, Jv\}^{\perp}$, *the symmetric linear operators* JJ' *and* $J'J$ *satisfy*

$$\frac{1}{K} I < -JJ', \quad -J'J < KI.$$

Proof. — It is easily checked that JJ' and $J'J$ are symmetric with respect to this inner product. It suffices to show that

$$0 \leqslant -JJ' < KI$$

since, after transferring to S' by χ_*, the corresponding inequalities for $J'J$ are proved similarly. The sharper lower bound on $-JJ'$ is obtained by using $-J'J = (-JJ')^{-1}$. So we need to show that for any harmonic form $\omega \neq 0$ in $\{v, Jv\}^{\perp}$,

$$0 < \omega \cup J(-JJ'\omega) = \omega \cup J'\omega < K(\omega \cup J\omega).$$

We have the lower bound, because if ω' is the harmonic form on S' which is cohomologous to $\chi^*\omega$, then $\chi^* J'\omega$ is cohomologous to the harmonic conjugate $J''\omega'$ of ω', and $\omega \cup J'\omega = \omega' \cup J''\omega' > 0$. We use $\langle\, ,\, \rangle$ to denote the inner product $\langle \mu, \nu \rangle = \mu \cup J''\nu$ on harmonic forms on S' also (by abuse of notation). Write $\xi + i\eta = \zeta = \zeta(\pi^{-1}z)$ for the coordinate for the singular Euclidean structure on S, S' given locally by

$$\zeta(\pi^{-1}z) = \int_{z_0}^{z} \sqrt{q(t)}dt,$$

and similarly for S', using p, so that the matrix of $D\chi$ with respect to these coordinates is $\left(\begin{smallmatrix} \sqrt{K} & 0 \\ 0 & 1/\sqrt{K} \end{smallmatrix}\right)$. Then for each of S, S', $\langle\, ,\, \rangle$ is the restriction to harmonic 1-forms of an inner product defined on C^0 forms by

$$\langle \mu, \nu \rangle = \int (a_1 b_1 + a_2 b_2) d\xi d\eta.$$

where μ, ν are given in local coordinates by $a_1 d\xi + a_2 d\eta$ and $b_1 d\xi + b_2 d\eta$ respectively. If μ is cohomologous to μ' and ν is harmonic, and ν' is the harmonic conjugate of ν, then

$$\langle \mu, \nu \rangle = \mu \cup \nu' = \mu' \cup \nu' = \langle \mu', \nu \rangle.$$

Now we have

$$\langle \chi_*\omega, \chi_*\omega \rangle < K\langle \omega, \omega \rangle,$$

with strict inequality since ω is not a multiple of $d\xi$ in local coordinates (because $\omega \in \{v, Jv\}^{\perp}$). By Cauchy Schwarz,

$$\langle \omega', \omega' \rangle^2 = \langle \chi_*\omega, \omega' \rangle^2 \leqslant \langle \chi_*\omega, \chi_*\omega \rangle \langle \omega', \omega' \rangle < K\langle \omega, \omega \rangle \langle \omega', \omega' \rangle.$$

So

$$\langle \omega', \omega' \rangle < K\langle \omega, \omega \rangle,$$

that is,

$$\omega \cup J'\omega = \omega' \cup J''\omega' < K(\omega \cup J\omega),$$

as required. □

11.10. Proof of Lemma 11.5. — Write $\mu = \mu_1 + i\mu_2$, where μ_1 and μ_2 are real and define elements of $H^1(S, \mathbf{R})$. So let ω_k be a harmonic 1-form with $\mu_k = \omega_k$ in $H^1(S, \mathbf{R})$. Fix any $x_0 \in S$. Then a function f_j is well-defined except at the singularities of μ by

$$f_k(x) = \int_{\gamma_x} (\mu_k - \omega_k),$$

where γ_x is any path from x_0 to x. Then we have

$$\mu_k = \omega_k + df_k,$$

and f_j is harmonic with singularities at the singularities of μ. It is important to note, however, that f_1 and f_2 are *not* usually conjugate harmonic functions. In fact, there is usually not a conjugate harmonic function to f_k defined globally on S. Since $df_k = 0$ in $H^1(S, \mathbf{R})$, we have $df_k \cup \nu = 0$. However,

$$\omega_k \cup \nu = \int_S \omega_k \wedge \nu = \lim_{\delta \to 0} \int_{S_\delta} \omega_k \wedge \nu,$$

where S_δ is a surface obtained by deleting δ-parametrised neighbourhoods of the singularities of μ, and these neighbourhoods converge to the singularities as $\delta \to 0$. So

$$\mu_k \cup \nu - \int_S \mu_k \wedge \nu = \mu_k \cup \nu - \lim_{\delta \to 0} \int_{S_\delta} \mu_k \wedge \nu = -\lim_{\delta \to 0} \int_{S_\delta} df_k \wedge \nu.$$

Now, on S_δ, $df_k \wedge \nu = d(f_k \nu)$. So, if we let $\gamma_j(\delta)$ be the boundary component of S_δ excluding the j'th singularity, and oriented anticlockwise around the singularity,

$$\xi_k \cup \nu - \int_S \xi_k \wedge \nu = \sum_j \lim_{\delta \to 0} \int_{\gamma_j(\delta)} f_k \nu.$$

If the j'th singularity is a zero of ν, take $a_j = 0$. So to prove the lemma, it suffices to prove, for each j,

$$-\pi i a_j = \lim_{\delta \to 0} \int_{\gamma_j(\delta)} (f_1 + i f_2)\nu_1, \quad -\pi a_j = \lim_{\delta \to 0} \int_{\gamma_j(\delta)} (f_1 + i f_2)\nu_2.$$

First, suppose that p_j is not a zero of ν, and choose a local coordinate with $\nu = d\zeta$. Write $\zeta = \xi + i\eta$. Then $\nu_1 = d\xi$, $\nu_2 = d\eta$. By linearity, we can assume that $a_j = 1$. Take $\gamma_j(\delta)$ to be a square of sidelength 2δ and centred on 0. Then the ζ^{-1} term in the Laurent expansion of $f_1 + i f_2$ about 0 is $-\zeta^{-1}$. This is the only term which gives a nonzero contribution in the limit. Then

$$-\int_{\gamma_j(\delta)} \zeta^{-1}\nu_1 = -2\int_{-\delta}^{\delta} \frac{(\xi + i\delta)d\xi}{\xi^2 + \delta^2} = -\pi i,$$

$$-\int_{\gamma_j(\delta)} \zeta^{-1}\nu_2 = -2\int_{-\delta}^{\delta} \frac{(\delta + i\eta)d\eta}{\eta^2+\delta^2} = -\pi,$$

as required.

If p_j is a zero of ν then for some $n > 0$, $\nu = (x+iy)^n(dx+idy)$ in a suitable local coordinate, while $f_1 + if_2$ locally has only a simple pole singularity. So the limits of the integrals round $\gamma_j(\delta)$ as $\delta \to 0$ are 0. □

11.11. Introduction of Singular Harmonic Functions to show that $V^{-1}W\Pi$ is symmetric. — Observe that $V^{-1}W$ is essentially a complex matrix, that is, $V^{-1}W$ is divided into 2×2 submatrices $C_{j,k}$ all of the form $M(\alpha - i\beta)$. Let $-4\pi V^{-1}\Sigma\Pi = D$ and $D = (D_{j,k})$ for 2×2 matrices $D_{j,k}$. Then $V^{-1}\Sigma$ is also essentially a complex matrix, and symmetric, with resulting consequences for the $D_{j,k}$.

Lemma. — *There are harmonic functions g_j, h_j with singularities only at $\pi^{-1}(b_j)$ such that if the local coordinate $\zeta = \xi + i\eta$ is given locally by*

$$\zeta(\pi^{-1}(z)) = \int_{b_k}^{z} \frac{q_k(t)dt}{\sqrt{q(t)}},$$

then $g_j(\zeta) + \mathrm{Re}(4/\zeta) = G_j(\zeta)$, $h_j(\zeta) + \mathrm{Im}(4/\zeta) = H_j(\zeta)$ are continuous near $\pi^{-1}(b_j)$. Let $k \neq j$. The (k,j) block $C_{k,j} + D_{k,j}$ of $V^{-1}W - 4\pi V^{-1}\Sigma\Pi$ is given by

$$C_{k,j} + D_{k,j} = -\begin{pmatrix} (g_j)_\xi(0,0) & (h_j)_\xi(0,0) \\ (g_j)_\eta(0,0) & (h_j)_\eta(0,0)\end{pmatrix}.$$

Let $k = j$ and let

$$r_j(z) = \frac{q(z)}{(q_j(z))^2(z-b_j)}.$$

Then

$$C_{j,j} + D_{j,j} = -\begin{pmatrix} (G_j)_\xi(0,0) & (H_j)_\xi(0,0) \\ (G_j)_\eta(0,0) & (H_j)_\eta(0,0)\end{pmatrix} - M(r_j'(b_j)/3).$$

Remark. — Using the notational convention $H_\zeta = (H_\xi - iH_\eta)/2$, $H_{\bar\zeta} = (H_\xi + iH_\eta)/2$, and if M is as in 10.11, this becomes

$$\begin{aligned} C_{j,k} &= -M((g_j + ih_j)_\zeta(0))^t \quad \text{if } j \neq k, \\ C_{j,j} &= -M((G_j + iH_j)_\zeta(0))^t - M(r_j'(b_j)/3). \\ D_{j,k}\Pi &= -M((g_j + ih_j)_{\bar\zeta}(0))^t \quad \text{if } j \neq k, \\ D_{j,j}\Pi &= -M((G_j + iH_j)_{\bar\zeta}(0))^t. \end{aligned}$$

Proof of the lemma. — It follows from 11.5 (proved in 11.10) that, in $H^1(S)$,

$$(w_j)^t = (v_j)^t(V^{-1}W - 4\pi V^{-1}\Sigma\Pi).$$

Then the following equation holds pointwise for harmonic functions g_j, h_j with singularities only at $\pi^{-1}(b_j)$, that is, only at the (common) singularity of w_{2j-1} and w_{2j}.

$$\big(w_{2j-1}\ w_{2j}\big) = \sum_\ell \big(v_{2\ell-1}\ v_{2\ell}\big)\, C_{\ell,j} + \sum_\ell \big(v_{2\ell-1}\ v_{2\ell}\big)\, D_{\ell,j} + \big(dg_j\ dh_j\big)\,.$$

First assume $k \neq j$, and evaluate both sides at $\pi^{-1}b_k$, using the local coordinate $\xi + i\eta$ with 0 corresponding to $\pi^{-1}(b_k)$. The lefthand side vanishes, and on the righthand side terms involving v_m for $m \neq 2k-1$ or $2k$ also vanish. The term $\big(v_{2k-1}\ v_{2k}\big)$ is simply $\big(d\xi\ d\eta\big)$ So we have the required formula.

Now let $k = j$. Again we want to evaluate at $(0,0)$. We have to subtract off the singularity $\big(\mathrm{Re}(4/(\xi+i\eta)^2)d\xi\ \mathrm{Im}(4/(\xi+i\eta)^2)d\eta\big)$ from both $\big(w_{2j-1}\ w_{2j}\big)$ and $\big(dg_j\ dh_j\big)$. The evaluation of the sum of the $\big(v_{2\ell-1}\ v_{2\ell}\big)$ is as before. This time there is a possibly nonzero constant term in $\big(w_{2j-1}\ w_{2j}\big)$, which we need to calculate. We have

$$\zeta^2 = \frac{4(z-b_j)}{a_j}(1+o(1)), \quad a_j = \mathrm{Res}(q, b_j) = \mathrm{Res}(q/q_j^2, b_j) = r_j(b_j).$$

We need to calculate ζ more accurately from the formula

$$\begin{aligned}\zeta = \int_{b_j}^z \frac{1}{\sqrt{(t-b_j)r_j(t)}}dt &= \int_{b_j}^z \frac{1}{\sqrt{a_j(t-b_j)}}\Big(1 - \frac{r_j'(b_j)(t-b_j)}{2a_j} + O((t-b_j)^2)\Big)dt \\ &= 2\sqrt{\frac{z-b_j}{a_j}}\Big(1 - \frac{r'(b_j)(z-b_j)}{6a_j} + O((z-b_j^2)\Big) \\ &= 2\sqrt{\frac{z-b_j}{a_j}}\Big(1 - \frac{r'(b_j)\zeta^2}{24} + O(\zeta^4)\Big).\end{aligned}$$

So

$$\begin{aligned}\frac{a_j}{z-b_j} &= \frac{4}{\zeta^2}\Big(1 - \frac{r'(b_j)\zeta^2}{24} + O(\zeta^4)\Big)^2 = \frac{4}{\zeta^2} - \frac{r_j'(b_j)}{3} + O(\zeta^2), \\ \frac{a_j q_j dz}{(z-b_j)\sqrt{q}} &= \frac{4d\zeta}{\zeta^2} - \frac{r_j'(b_j)d\zeta}{3} + O(\zeta^2)d\zeta,\end{aligned}$$

which gives the result. □

11.12. Introduction of Green's functions to prove that $V^{-1}W\Pi$ is Symmetric. — First, we recall some properties of Green's functions on Riemann surfaces and establish some notation. Fix j, k, $1 \leqslant j, k \leqslant n-3$. Choose charts

$$\begin{gathered}\varphi_0 : \{\zeta : |\zeta| < \delta_0\} \longrightarrow S,\ \varphi_1 : \{\zeta : |\zeta - 1| < \delta_0\} \longrightarrow S, \\ \varphi_0(0) = \pi^{-1}b_k, \quad \varphi_1(1) = \pi^{-1}b_j, \\ \varphi_0^*(v_{2k-1} + iv_{2k}) = d\zeta, \quad \varphi_1^*(v_{2j-1} + iv_{2j}) = d\zeta.\end{gathered}$$

Let

$$S_\delta = S \setminus (\varphi_0(\{\zeta : |\zeta| < \delta\}) \cup \varphi_1(\{\zeta : |\zeta - 1| < \delta\})),$$

and let G_δ be the Green's function of S_δ, so that G_δ is defined on $S_\delta \times S_\delta \setminus$ diagonal. If we use a holomorphic coordinate ζ on S_δ as above, then G_δ is real-analytic off the diagonal, and in a neighbourhood of the diagonal, $G_\delta(\zeta_1, \zeta_2) + \log|\zeta_1 - \zeta_2|$ is real-analytic. Locally, if $j \neq k$, in the product of the charts for φ_0 and φ_1, we can consider G_δ as a function of $(\zeta_1, 1 + \zeta_2)$ for

$$\delta \leqslant |\zeta_1|, \quad |\zeta_2| < \delta_0.$$

We also have

$$G_\delta(\zeta_1, 1 + \zeta_2) = 0 \quad \text{if } |\zeta_1| = \delta \text{ or } |\zeta_2| = \delta.$$

By considering Fourier series, we see that we have an expansion of G_δ of the form

$$G_\delta(\zeta_1, 1 + \zeta_2) = \operatorname{Re}\Big(\sum_{n,m=0}^{\infty} (c_{n,m}(\delta) f_n(\zeta_2) f_m(\zeta_1) + d_{n,m}(\delta) \overline{f}_n(\zeta_2) f_m(\zeta_1))\Big),$$

where

$$f_0(\zeta) = \log(|\zeta|/\delta), \quad f_n(\zeta) = \zeta^n - \delta^{2n}\overline{\zeta}^{-2n} \text{ for } n > 0.$$

The coefficients $c_{n,m}$ and $d_{n,m}$ are uniquely determined if we specify that $c_{n,0}$ and $c_{0,n}$ are real and $d_{n,0} = d_{0,n} = 0$, for all n.

If $j = k$, we take $\varphi_1(1 + \zeta) = \varphi_0(\zeta)$, and expand $G_\delta(\zeta_1, \zeta_2)$ and then we have a similar expansion except that we have in addition a term

$$-\log \frac{|\zeta_1 - \zeta_2|}{|\delta^2 - \overline{\zeta}_2\zeta_1|}.$$

11.13. Verification of symmetry of $V^{-1}W\Pi$

Lemma. — *Fix j and k. Then*

$$(\partial/\partial\zeta)(g_j + ih_j)(0) = \lim_{\delta\to 0} 2c_{1,1}(\delta),$$
$$(\partial/\partial\overline{\zeta})(g_j + ih_j)(0) = \lim_{\delta\to 0} 2\overline{d_{1,1}(\delta)}.$$

By 11.11, this is enough to prove symmetry of $V^{-1}W\Pi$ (and confirms symmetry of $V^{-1}\Sigma$, which, of course, we already know). To see this, note that because $G(1 + \zeta_2, \zeta_1) = G(\zeta_1, 1 + \zeta_2)$, the coefficient $c_{1,1}$ of $f_1(\zeta_1)f_1(\zeta_2)$ in the expansion of both functions is the same.

Proof of the lemma. — Now, g_j and h_j are harmonic functions on the surface S_δ, and hence they and their derivatives can be expressed in terms of the Green's function G_δ and their values on ∂S_δ. We use the local coordinates ζ_1, $1 + \zeta_2$ as already indicated. We are only interested in the values of the derivatives of g_j and h_j (and hence of the functions themselves) for small ζ, and need to consider $G_\delta(\zeta, \zeta')$ for small ζ, and ζ' near 0 or 1. If g is any harmonic function on S_δ, then Green's Theorem gives, for any $\zeta \in S_\delta$,

$$2\pi g(\zeta) = -\int_{\partial S_\delta} g * d_2 G_\delta(\zeta, .) \tag{1}$$

where $*d_2G_\delta(\zeta,.)$ is the conjugate harmonic form to d_2G_δ and d_2G denotes d of the function $\xi_2+i\eta_2\mapsto G(\zeta,\xi_2+i\eta_2)$, that is, if $\zeta_2=\xi_2+i\eta_2$, then

$$*d_2G_\delta(\zeta,\zeta_2))=-\frac{\partial G_\delta}{\partial\eta_2}(\zeta,\zeta_2)d\xi_2+\frac{\partial G_\delta}{\partial\xi_2}(\zeta,\zeta_2)d\eta_2=-i\frac{\partial G_\delta}{\partial\zeta_2}d\zeta_2+i\frac{\partial G_\delta}{\partial\overline{\zeta}_2}d\overline{\zeta}_2.$$

and similarly if $\zeta_2=1+\xi_2+i\eta_2$, and we orient ∂S_δ anticlockwise round 0, 1 (in the local coordinates). Now let g_0 be a continuous function defined on

$$\{\zeta:|\zeta|<\delta_0\}\cup\{\zeta:|\zeta-1|<\delta_0\}.$$

Then for any $\delta>0$ we can define a harmonic function on S_δ by the formula

$$-\int_{\partial S_\delta}g_0*d_2G_\delta(\zeta,.)$$

which is bounded between the maximum and minimum values of g_0. Taking limits as $\delta\to0$, we obtain a bounded harmonic function on S itself (since a bounded harmonic function cannot have isolated singularities), which must, therefore, be constant. Now $g_j(\zeta)$ and $h_j(\zeta)$ are bounded for $|\zeta|<\delta_0$, and $g_j(1+\zeta)$, $h_j(1+\zeta)$ differ by bounded harmonic functions from $-4\,\mathrm{Re}(1/\zeta)$, $-4\,\mathrm{Im}(1/\zeta)$ respectively for $|\zeta|<\delta_0$, since, in our local coordinate $1+\zeta$, $w_{2j-1}+iw_{2j}$ takes the form $(4/\zeta^2)d\zeta$. So taking limits as $\delta\to0$, we obtain, for $|\zeta|<\delta_0$, and constants M_j, N_j,

$$\pi g_j(\zeta)=2\lim_{\delta\to0}\int_{|\zeta_2|=\delta}\mathrm{Re}(1/\zeta_2)*d_2G_\delta(\zeta,1+\zeta_2)+M_j,$$

$$\pi h_j(\zeta)=2\lim_{\delta\to0}\int_{|\zeta_2|=\delta}\mathrm{Im}(1/\zeta_2)*d_2G_\delta(\zeta,1+\zeta_2)+N_j.$$

The convergence is locally uniform in ζ. So if $\zeta=\xi+i\eta$, we can differentiate the formulae with respect to ξ or η, and the constants will disappear. The derivatives of g_j and h_j are continuous for $|\zeta|<\delta_0$. So we obtain

$$\pi^2\frac{\partial g_j}{\partial\xi}(0,0)=\lim_{\delta'\to0}\lim_{\delta\to0}\int_{|\zeta|=\delta'}\frac{d\zeta}{i\zeta}\int_{|\zeta_2|=\delta}\mathrm{Re}(1/\zeta_2)\frac{\partial}{\partial\xi}(*d_2G_\delta(\zeta,1+\zeta_2)),$$

and similarly for the other terms. So

$$\pi^2\Big(\frac{\partial g_j}{\partial\xi}(0,0)+i\frac{\partial h_j}{\partial\xi}(0,0)\Big)=\lim_{\delta'\to0}\lim_{\delta\to0}\int_{|\zeta|=\delta'}\frac{d\zeta}{i\zeta}\int_{|\zeta_2|=\delta}\frac{1}{\zeta_2}\frac{\partial}{\partial\xi}(*d_2G_\delta(\zeta,1+\zeta_2)),$$

and similarly for $\dfrac{\partial}{\partial\eta}$. So

$$(2)\qquad \pi^2(\partial/\partial\zeta)(g_j+ih_j)=\lim_{\delta'\to0}\lim_{\delta\to0}\int_{|\zeta|=\delta'}\frac{d\zeta}{i\zeta}\int_{|\zeta_2|=\delta}\frac{1}{\zeta_2}\Big(\frac{\partial}{\partial\zeta}\Big)(*d_2G_\delta(\zeta,1+\zeta_2)).$$

Now

$$\int_{|\zeta|=\delta}\frac{\overline{\zeta}^n d\overline{\zeta}}{\zeta}=0\text{ for }n\neq-2,\qquad\int_{|\zeta|=\delta}\zeta^{n-1}d\zeta=0\text{ for }n\neq0.$$

So we only need to consider the $\overline{\zeta}_2^{-2}$-term in $\dfrac{\partial^2 G_\delta}{\partial\zeta\partial\overline{\zeta}_2}$ and the constant term in $\dfrac{\partial^2 G_\delta}{\partial\zeta\partial\zeta_2}$. Then we obtain

$$\pi^2(\partial/\partial\zeta)(g_j + ih_j)(0) = \lim_{\delta'\to 0}\lim_{\delta\to 0}\int_{|\zeta|=\delta'}\frac{d\zeta}{i\zeta}\int_{|\zeta_2|=\delta}\frac{-ic_{1,1}(d\zeta_2 - \delta^2\overline{\zeta}_2^{-2}d\overline{\zeta}_2)}{\zeta_2} = 4\pi^2 c_{1,1}.$$

Now suppose that $j = k$. Then our expressions for $\dfrac{\partial^2 G_\delta}{\partial\zeta\partial\overline{\zeta}_2}$, $\dfrac{\partial^2 G_\delta}{\partial\zeta\partial\zeta_2}$ are augmented by

$$-\frac{1+\overline{\zeta}_2\zeta}{(\delta^2 - \overline{\zeta}_2\zeta)^2}, \quad (\zeta_2 - \zeta)^{-2},$$

because of the log term. Both contribute nothing to the double integral. So in this case also we obtain the formula.

The case of $\partial/\partial\overline{\zeta}$ is similar.

CHAPTER 12

THE SECOND DERIVATIVE OF TEICHMÜLLER DISTANCE IS CONTINUOUS

12.1. The main purpose of this chapter is to prove the following theorem.

Theorem. — *The function*

$$d : ([\varphi], [\psi]) \longmapsto d([\varphi], [\psi]) : \mathcal{T}(Y) \times \mathcal{T}(Y) \longrightarrow [0, \infty)$$

is C^2 at any point $([\varphi], [\psi])$ with $d([\varphi], [\psi]) > 0$. Furthermore, if $\#(Y) = n$ and the multiplicities of zeros of q are k_i, $1 \leqslant i \leqslant s$, then the kernel of the term $R_2 + R_3$ of the second derivative (in the notation of 10.16) has dimension

$$2 + \sum_{i=1}^{s} \left[\frac{k_i}{2} \right].$$

We shall also obtain bounds on the second derivative $D^2 d$ in certain cases: of quadratic differentials with zero residues (in 12.9), as distance approaches 0 (in 12.11) and for quadratic differentials of a certain shape in the thin part of Teichmüller space (in 12.10).

12.2. Extending continuously is enough. — To show that Teichmüller distance is C^2 at $([\varphi], [\psi])$ with $d([\varphi], [\psi]) > 0$, we only have to show that

Proposition. — *$D^2 d$ extends continuously from the set of $([\varphi], [\psi])$ for which the quadratic differential $q(z)dz^2$ for $d([\varphi], [\psi])$ at $[\varphi]$ satisfies the Standing Assumption of 10.3.*

We now show why this is enough. We claim, first, that the set of $[\psi]$ for which $q(z)dz^2$ at $[\varphi]$ does not satisfy the Standing Assumption is a union of real analytic submanifolds, all of codimension at least 2. This is because the set of $q(z)dz^2$ with zero residues or multiple zeros is a union of real analytic submanifolds. Let M be one of these submanifolds. Let $\{(q_\lambda, K_\lambda) : \lambda \in \Lambda\}$ be a parametrisation of $M \times (1, \infty)$. Let $\{\mu_\lambda : \lambda \in \Lambda\}$ be the real analytic family of Beltrami differentials determined by $(q_\lambda, K_\lambda) : \lambda \in \Lambda\}$ as in 10.2. Then as in 10.2, Theorem 3 of [**A-B**] gives a real analytic submanifold of $[\psi_\lambda]$ such that $q_\lambda(z)dz^2$ is the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi_\lambda])$, with $\frac{1}{2}\log K_\lambda = d([\varphi], [\psi_\lambda])$. Then for any $([\varphi], [\psi])$ with $d([\varphi], [\psi]) > 0$, for almost all vectors $(\underline{h}, \underline{k})$ of length $\leqslant \delta$ for δ sufficiently small, $D^2 d$ is defined at $([\varphi + t\underline{h}], [\psi + t\underline{k}])$ for almost all $t \in [0, 1]$. Now if g is a function for which $g''(t)$ is defined almost everywhere on $[0, 1]$ and extends continuously then $g''(0+)$ exists and is defined by the formula

$$g''(0+) = \lim_{\delta \to 0} \delta^{-1} \int_0^\delta g''(t)dt,$$

and thus coincides with the continuous extension. Thus (for suitable choice of directions), all second order partial derivatives exist and coincide with the continuous extension. So $D^2 d$ exists and is continuous.

12.3. The Main Reduction. — Let $q(z)dz^2$, be the quadratic differential at $[\varphi]$, and $p(z)dz^2$ the stretch at $[\psi]$. Let q satisfy the Standing Assumption of 10.3. Let v, v' be as in 10.7, J, J' as in 10.9, V, W, T as in 10.10. The following is the main step needed to prove that Teichmüller distance is C^2. It will be proved in 12.7.

Proposition. — *Let $[\varphi] \in (\mathcal{T}(Y))_{\geqslant \varepsilon_0}$. The functions V^{-1}, $V^{-1}W$ extend continuously to all $([\varphi], q)$, and are bounded by a constant $C(\varepsilon_0)$. If $a_j = \mathrm{Res}(q, b_j) = 0$, then the j'th row and column of V^{-1} are 0.*

12.4. We shall also need the following.

Lemma. — *Let $[\varphi]$, $[\psi] \in (\mathcal{T}(Y))_{\geqslant \varepsilon_0}$ with $1/M \leqslant d([\varphi], [\psi]) = \frac{1}{2} \log K \leqslant M$. The linear map $(K + J'J)^{-1}$ on $\{v, Jv\}^\perp$ is bounded by $C(\varepsilon_0, M)$, using the cup form inner product of 10.8.*

Proof. — We examine the proof of 11.9. We see that it suffices to show that for any harmonic $\omega \in \{v, Jv\}^\perp$,

$$\chi_* \omega \cup J'' \chi_* \omega \leqslant (K - \beta)(\omega \cup J\omega).$$

As in 11.9, J'' denotes harmonic conjugate on S'. This is proved similarly to 11.9, but we just check the conditions. We assume without loss of generality that $\omega \cup J\omega = 1$, and that ω is the real part of $r(z)dz/\sqrt{q(z)}$ in local coordinates. Then the condition $\omega \in \{v, Jv\}^\perp$ gives

$$\int \frac{r\overline{q}}{|q|} = 0,$$

and hence

$$\int \frac{|r \pm q|^2}{|q|} = 2.$$

The functions r^2/q and q have a bounded number of poles and zeros, and the poles are bounded apart and simple. The integral of modulus of each function is 1. It follows that we can find a bounded disc of size bounded from 0 on which $\arg(r/\sqrt{q})$ is bounded from $\arg \sqrt{q}$. Then we can complete the proof as in 11.9. $\square$

12.5. Proof of 12.1 given 12.3. — By 12.2, we only need to show that the second derivative extends continuously. We consider the terms R_1, R_2, R_3 in the Second Derivative Formula 10.16. Clearly R_1 extends continuously, and by 12.3 R_2 also extends continuously. It remains to consider R_3. The projections E and E' of 10.11 are continuous because the entries of AV are

$$M\Big(\int \frac{q\overline{q}_j}{|q|}\Big),$$

and similarly for $A'V'$. The top left quarter of R_3 is $U^{-1}\Sigma_1(Fu_i \cup u_j)U^{-1}\Sigma_1$, with others similarly defined (see 10.15). We claim that these extend continuously because V^{-1} and V'^{-1} do.

First, we show continuity of U^{-1}, and, similarly, of U'^{-1} (see 10.9). We assume, renumbering if necessary, that $a_{n-3} \neq 0$. Let S be the $(2n-6) \times (2n-8)$ matrix which has the $(2n-8) \times (2n-8)$ identity matrix in the first $2n-8$ rows and the matrices

$$M\Big(\frac{-a_j}{a_{n-3}}\Big) \quad (1 \leqslant j \leqslant n-4)$$

in the last two rows. We claim that

$$U^{-1}\Sigma_1 = S^t V^{-1}\Sigma S, \text{ or, equivalently, } I = -S^t V^{-1}\Sigma S U \Sigma_1. \tag{1}$$

This is because $S^t E_1$ is the $(2n-8) \times (2n-8)$ identity matrix, and

$$V\Sigma E_1 = (Jv_j \cup v_i)E_1 = (Jv_j \cup u_i) = (Ju_j \cup u_i) = SU\Sigma_1.$$

Multiplying on the left by $-S^t V^{-1}\Sigma$ gives (1). We also have

$$V^{-1}TV'^{-1}\Sigma T^t V^{-1} = V^{-1}(JJ'v_i \cup v_j)V^{-1},$$

$$U^{-1}(JJ'u_i \cup u_j)U^{-1} = S^t E V^{-1}(JJ'v_i \cup v_j)V^{-1}E^t S.$$

Now we consider the continuity of $U^{-1}\Sigma_1(Fu_i \cup u_j)U^{-1}\Sigma_1$, and the other quarters of R_3. We consider these as functions of q. Note that, when the Standing Assumption is satisfied,

$$U^{-1}\Sigma_1 = (\langle x_i, x_j\rangle) = (x_i \cup Jx_j),$$

for harmonic 1-forms $x_i \in \{v, Jv\}^\perp$ with $x_{2i} = Jx_{2i-1}$. Since $U^{-1}\Sigma_1$ extends continuously to the set of all q, $x_i \cup Jx_i$ remains bounded for all i. So x_i must be a harmonic 1-form on the corresponding (possibly degenerate) surface $S = S_q$ (see 10.5), even if q does not satisfy the Standing Assumption. Moreover, $Jx_{2i-1} = x_{2i}$. Now let $q(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi])$, with stretch $p(z)dz^2$ at $[\psi]$. Let $S' = S_p$. Then by 12.4, $C = -JJ'$ and $D^{-1} = (I + JJ'/K)^{-1}$ are defined and bounded on the span of the x_i, even in the absence of the Standing Assumption, and are continuous in (q, p), that is, in $([\varphi], [\psi])$. Let x_i' be similarly defined, using U'. It follows that F, G, H (see 10.15) are defined on the span of the x_i or x_i', and the matrices $(\langle Fx_i, x_j\rangle)$, $(\langle Gx_i', x_j\rangle)$, $(\langle Hx_i', x_j'\rangle)$ are defined and continuous in q, p. These are the four quarters of R_3. So R_3 extends continuously, as required. □

12.6. How genus reduces. — Let $q(z)dz^2$ be any quadratic differential at any point of $\mathcal{T}(Y)$. Let S be obtained by removing singularities and then filling in punctures on

$$\{(z,u) \in \overline{\mathbf{C}}^2 : q(z) = u^2\}.$$

If $q_m dz^2$ are quadratic differentials satisfying the Standing Assumption with $q_m \to q$, then the corresponding surfaces S_m degenerate as $m \to \infty$, reflecting the fact that the genus of S is lower than that of S_m. In fact, the genus S is described by the following formula.

Lemma. — *Let n' be the number of poles of q (all simple, possibly $< n$ in general). Let z_i $(1 \leqslant i \leqslant s)$ be the zeros of q, with z_i of multiplicity k_i. Let $s = s_1 + s_2$, where s_1, s_2 are respectively the numbers of zeros of odd and even multiplicities. Then the genus of S is*

$$\frac{n' + s_1}{2} - 1 = n' - 3 - \sum_{i=1}^{s} \left[\frac{k_i}{2}\right].$$

Proof. — We have

$$\sum_{i=1}^{s} k_i = n' - 4.$$

So

$$\sum_{i=1}^{s} \left[\frac{k_i}{2}\right] = \sum_{i=1}^{s} \frac{k_i}{2} - \frac{s_1}{2} = \frac{n' - s_1}{2} - 2.$$

So the two different claimed expressions for the genus coincide. As usual, define $\pi(z,u) = z$. Let A denote the set of zeros and poles of q. Then

$$\pi : S \smallsetminus \pi^{-1}(A) \longrightarrow \overline{\mathbf{C}} \smallsetminus A$$

is a covering. So the Euler characteristic of S is $2(2 - \#(A)) + \#(\pi^{-1}(A))$, and the genus is $\#(A) - 1 - \frac{1}{2}\#(\pi^{-1}(A))$. The crucial point of the calculation is that $\pi^{-1}(z_i)$ has one point if k_i is odd, and two if k_i is even. □

12.7. Proof of 12.3. — Let $q_m \to q$. We need to show convergence of V_m^{-1} and $W_m V_m^{-1}$. Recall that the matrices $W_m V_m^{-1}$ and V_m^{-1} give the coefficients of the $v_{i,m}$ in the expressions of the singular forms $w_{j,m}^{(1)}$ and $w_{j,m}^{(2)}$. In turn, we showed in 11.11 that these coefficients are given by the *linear terms* in the expansions of $g_{j,m}$, $h_{j,m}$ in local coordinates given at $\pi_m^{-1}(b_i)$ by $q_{i,m}/\sqrt{q_m}$, with an extra term $r'_{j,m}(b_{j,m})/12$ giving 2×2 matrices down the diagonal of $W_m V_m^{-1}$, where

$$r_{j,m}(z) = \frac{q_m(z)}{(q_{j,m}(z))^2(z - b_{j,m})}.$$

The term $r'_{j,m}(b_{j,m})$ is easily seen to converge: to 0 if b_j is a zero of q. Here, $g_{j,m}$ $h_{j,m}$ are harmonic functions with singularities at $\pi_m^{-1}(b_{j,m})$ of a certain form with respect to local coordinates given by $q_{j,m}/\sqrt{q_m}$. Thus, in order to prove 12.3, it suffices to prove the following.

Lemma. — *Let $q_m \to q$ where q_m has only simple zeros and all residues non-zero. Let m be a real parameter. Then the linear terms of $g_{j,m}$, $h_{j,m}$ in expansions in the appropriate local coordinates at $\pi_m^{-1}(b_{i,m})$ do converge. If the limit b_i of $b_{i,m}$ is a zero of q then these limits are 0 for all j, and for all $j \neq i$ if b_i is not a pole of q. If b_i is not a pole of q and not a zero of q, then the limit for $j = i$ is $-r_i'(b_i)/3$.*

Proof. — The idea is to produce a nonconstant harmonic function on a compact surface without boundary, if the linear terms do not converge, and thus to obtain a contradiction. Let S_m be the surface which is the usual double branched cover of

$$\{(z, w) : q_m(z) = w^2\}$$

with branched cover π_m. Fix i, j, that is, fix convergent sequences $b_{i,m}$, $b_{j,m}$ of poles of q_m. We may as well take $i = 1$. Let b_1, b_j be the limits, which may or may not be poles of q.

Now use the local coordinate ζ near $\pi_m^{-1}(b_{1,m})$, where $\zeta(\pi_m^{-1}(b_{1,m})) = 0$ and

$$\zeta(\pi_m^{-1}(z)) = \int_{b_{1,m}}^{z} (q_{1,m}/\sqrt{q_m})(t)dt.$$

Then there is an embedded disc $K_m = \{\zeta : |\zeta| \leqslant r_m\}$ in S_m with $r_m \geqslant r_0 > 0$ with r_0 depending only on q. In fact, if b_1 is a zero of q, we can take $r_m \to \infty$, because we can take $r_m \geqslant C\sqrt{|b_{1,m} - z_{2,m}|}^{-1}$, where $z_{2,m}$ is the *second* nearest zero of q_m to $b_{1,m}$. Thus, $g_{1,m}$ and $h_{1,m}$ can be regarded as functions on $K_m = \{\zeta : |\zeta| \leqslant r_m\} \supset K_0$, such that (by 11.11)

$$g_{1,m}(x, y) + 4\,\mathrm{Re}(1/(x + iy)), \quad h_{1,m}(x, y) + 4\,\mathrm{Im}(1/(x + iy))$$

are harmonic functions on K_m.

The general Fourier series expansion of a harmonic function g in $\{re^{i\theta} : r \in (0, r_0)\}$ (as used in 11.12) is of the form

$$\sum_{n\neq 0,\ n\in\mathbf{Z}} r^n(\alpha_n e^{in\theta} + \overline{\alpha}_n e^{-in\theta}) + \alpha_0 + \alpha_0' \log r,$$

where

$$2\pi(\alpha_n r^n + \overline{\alpha}_n r^{-n}) = \int_0^{2\pi} g(re^{i\theta})e^{-in\theta}d\theta \text{ if } n \neq 0,$$

$$2\pi(\alpha_0 + \alpha_0' \log r) = \int_0^{2\pi} g(re^{i\theta})d\theta.$$

So the coefficients can be computed from integrals for just two values of r, say $r = r_0/2$ and $r = r_0/4$. Let $\alpha_{n,1,m}$ be the coefficients for $g_{1,m}$. We have $\alpha_{n,1,m} = 0$ for $n \leqslant -2$, $\alpha_{-1,1,m} = -2$, $\alpha_{0,1,m}' = 0$. We need to show that $\lim_{m\to\infty} \alpha_{1,1,m}$ exists: and $= 0$ if b_1 is a zero for q, and similarly for the coefficients for $h_{1,m}$. We also need to prove convergence of corresponding coefficients in Fourier expansions of $g_{1,m}$, $h_{1,m}$ in discs round points $\pi_m^{-1}(b_{j,m})$ for $j \neq 1$.

For the moment we continue to consider $g_{1,m}$ and $0 \in K_0$ (corresponding to $\pi_m^{-1}(b_{1,m})$). We only consider $g_{1,m}$ from now on, since the arguments for $h_{1,m}$ are exactly analogous. Choose real $0 < \lambda_{1,m} \leqslant 1$ and $c_{1,m}$ such that $\lambda_{1,m} g_{1,m} + c_{1,m}$ is bounded on ∂K_0 -and bounded from constant if we have to take $\lim_{m\to\infty} \lambda_{1,m} = 0$. Then $\lambda_{1,m} g_{1,m} + c_{1,m}$ is harmonic on $S_m \setminus \{\pi_m^{-1}(b_m)\}$ and bounded on $S_m \setminus K_0$. Take a convergent subsequence. The limit function g_1 can be regarded as a function on: S (if b_1 is a pole of q); or $\mathbf{C}$ if b_1 is not a pole of q. If b_1 is a zero of q, then

$$\mathbf{C} = \lim_{m\to\infty} K_m,$$

and

$$g_1 + \lim_{m\to\infty} 4\lambda_{1,m} \operatorname{Re}(1/(x+iy))$$

is also a well-defined bounded harmonic function on $\mathbf{C}$. Thus in all cases, $K_0 = \{\zeta : |\zeta| \leqslant r_0\}$ can be regarded as a subset of the domain of g_1. In all cases, the function is harmonic in the complement of $0 \in K_0$, and bounded in the complement of K_0. Moreover the Fourier series coefficients satisfy the same restrictions as those of the $g_{1,m}$, because we simply take limits in the integral expressions for the coefficients. Then if $\lim_{m\to\infty} \lambda_{1,m} = 0$, we see that g_1 is nonconstant and extends continuously to 0, giving a nonconstant bounded harmonic function on $\mathbf{C}$ or S, which is impossible. Therefore $\{\lambda_{1,m}\}$ is bounded from 0, and we can take a subsequence so that

$$g_1 = \lim_{k\to\infty} g_{1,m_k} + c_{1,m_k}.$$

If g_1 is a function on $\mathbf{C}$ (that is, if b_1 is a zero of q), we deduce that $g_1(x+iy) + 4\operatorname{Re}(1/(x+iy))$ is constant, and thus $\lim_{k\to\infty} \alpha_{1,1,m_k} = 0$. Because there are no nonconstant bounded harmonic functions on $\mathbf{C}$, or S, we see that all subsequences of $g_{1,m} + c_{1,m}$ have the same limit by suitable choice of $\{c_{1,m}\}$. Hence we can choose $c_{1,m}$ so that $\lim_{m\to\infty} g_{1,m} + c_{1,m}$ exists, and thus $\lim_{m\to\infty} \alpha_{1,1,m}$ exists, as required.

We compute $\lim_{m\to\infty} \alpha_{1,1,m}$ when b_1 is neither a pole nor a zero of q. Then $b_1 = \lim_{m\to\infty} b_{1,m} = \lim_{m\to\infty} z_{1,m}$ where $z_{1,m}$ is a zero of q_m, but all other zeros of q_m are bounded from b_m. Write

$$r'_m = r'_{1,m}(b_{1,m}), \quad a_{1,m} = \operatorname{Res}(q_m, b_{1,m}) = r_{1,m}(b_{1,m}).$$

(This is the usual definition of $a_{1,m}$.) We have

$$\frac{q_m(z)}{(q_{1,m}(z))^2} = r_{1,m}(z)(z - b_{1,m}) = \frac{z^2(z-1)^2}{b_{1,m}^2(b_{1,m}-1)^2}(z - z_{1,m})(A + O(z - b_{1,m}))$$

for a constant A. We deduce that

$$r_{1,m}(b_{1,m}) = A(b_{1,m} - z_{1,m}) = a_{1,m},\ r'_m = r'_{1,m}(b_{1,m}) = A + O(b_{1,m} - z_{1,m}).$$

So

$$r_{1,m}(z)(z - b_{,m}) = r'_m(z - b_{1,m})(z - z_{1,m})(1 + O(z - b_{1,m}) + O(z_{1,m} - b_{1,m})).$$

Note that r'_m is bounded from 0 (since the limit is nonzero), and that $z_{1,m} - b_{1,m} = O(a_{1,m})$. Then define

$$(b_{1,m} - z_{1,m})u/2 = z - (z_{1,m} + b_{1,m})/2 + \sqrt{(z - z_{1,m})(z - b_{1,m})}.$$

Then the u coordinate describes a branched cover of the z coordinate. The branch-points are $u = \pm 1$, and for $|z - b_{1,m}| \leqslant C\delta$, for a suitable constant C, u takes all values in

$$\{u : |a_{1,m}|/\delta \leqslant |u| \leqslant \delta/|a_{1,m}|\}.$$

Also, we have

$$\frac{du}{u} = \frac{dz}{\sqrt{(z - b_{1,m})(z - z_{1,m})}}.$$

So

$$\begin{aligned}\zeta &= \int_{b_{1,m}}^{z} \frac{q_{1,m}(t)dt}{\sqrt{q_m(t)}} \\ &= \int_{b_{1,m}}^{z} \frac{dt}{\sqrt{r'_m(t - b_{1,m})(t - z_{1,m})(1 + O(t - b_{1,m}) + O(z_{1,m} - b_{1,m}))}} \\ &= \frac{1}{\sqrt{r'_m}} \log u(z)(1 + O(z_{1,m} - b_{1,m})) + O(z - b_{1,m}).\end{aligned}$$

So

$$u = \exp(\sqrt{r'_m}\zeta(1 + O(z_{1,m} - b_{1,m})) + O(z - b_{1,m}))$$

So now in the u coordinate, $g_{1,m}$ and $h_{1,m}$ must converge, respectively, to the real and imaginary parts of $-4\sqrt{r'_m}/(u-1)$. We have $z - b_{1,m} = O(a_{1,m}\zeta^2)$ for small $z - b_{1,m}$. So $g_{1,m}$, $h_{1,m}$ are given to first approximation by the ζ term in

$$\begin{aligned}\frac{-4\sqrt{r'_m}}{\exp(\sqrt{r'_m}\zeta) - 1} &= \frac{-4}{\zeta(1 + \sqrt{r'_m}\zeta/2 + r'_m\zeta^2/6)} \\ &= (-4/\zeta)(1 - \sqrt{r'_m}\zeta/2 - r'_m\zeta^2/6 + r'_m\zeta^2/4 + O(\zeta^3)) \\ &= -4/\zeta + 2\sqrt{r'_m} - r'_m\zeta/3 + O(\zeta^2).\end{aligned}$$

So the limit of the ζ-term is as required.

Now we consider the Fourier series expansion of $g_{1,m}$ in a disc round $\pi_m^{-1}(b_{j,m})$ for $j \neq 1$. Taking limits of $\lambda_{j,m}g_{1,m} + c_{j,m}$ for suitable $\lambda_{j,m} \leqslant 1$, $c_{j,m}$ as before, we obtain a harmonic function on a surface which is a compact minus finitely many points, and bounded on the complement of K_0. As in the previous case, we can deduce that $\{\lambda_{j,m}\}$ is bounded and therefore take $\lambda_{j,m} = 1$ for all m. The surface is disjoint from K_0 if b_1 is not a pole of q, or if b_j is not a pole of q. Therefore in these cases the limit function is bounded and for the Fourier coefficients, $\lim_{m\to\infty} \alpha_{1,j,m} = 0$. If b_1 and b_j are both poles of q, then $\lim_{m\to\infty} \alpha_{1,j,m}$ exists, just as for $\{\alpha_{1,1,m}\}$. □

12.8. The Kernel. — The following lemma gives the dimension of the $R_2 + R_3$ term of the second derivative of the Teichmüller distance function that was claimed in 12.1. This will complete the proof of 12.1.

Lemma. — *Write*

$$X_1 = \lim_{m\to\infty} V_m^{-1} W_m, \quad -4\pi \lim_{m\to\infty} V_m^{-1}\Sigma\Pi = X_2.$$

Then, using the notation of 12.6,

$$2 \times \text{genus}(S) = \text{rank}(X_2).$$

Proof. — The substance of 12.7 is that linear terms in the expansions of the harmonic functions $g_{j,m}$, $h_{j,m}$ converge to linear terms in the expansions of limiting harmonic functions on S. We have seen that if b_i is not a pole of q, then $C_{i,j} = D_{i,j} = 0 = C_{j,i} = D_{j,i}$ for all i.

So now we can assume, after deleting some points, that the residue of q at each point b_j is nonzero. We use 11.11 and take limits. So then if we write $X_1 = (C_{i,j})$, $X_2 = (D_{i,j})$, we can interpret the nonzero 2×2 matrices $C_{i,j}$, $D_{i,j}$ as follows:

(1) $$\eta = \omega(C_{i,j} + D_{i,j}) + \omega' + \omega'' \text{ in } H^1,$$

where η, ω, ω' and ω'' have the following properties. Let ζ be the local coordinate at $\pi^{-1}(b_j)$ with $\zeta = 0$ at $\pi^{-1}(b_j)$ and $d\zeta = \pi^*(q_j dz/\sqrt{q})$. Then η is any meromorphic 1-form on S given in this local coordinate by $d\zeta(4\zeta^{-2} + O(1))$ (which is the same form as $\pi^*(f_j dz/\sqrt{q})$ — see 11.5.) Now let ζ be the local coordinate at $\pi^{-1}(b_i)$ with $\zeta = 0$ at $\pi^{-1}(b_i)$ and $d\zeta = \pi^*(q_i dz/\sqrt{q})$. Then ω and ω' are holomorphic 1-forms, ω'' is antiholomorphic vanishing at $\pi^{-1}(b_i)$, $\omega = d\zeta(1 + O(\zeta))$ near $\pi^{-1}(b_i)$, and $\omega' - \eta = O(\zeta)d\zeta$ near $\pi^{-1}(b_i)$. Although ω, ω' and ω'' are not uniquely determined, the 2×2 matrices $C_{i,j}$ and $D_{i,j}$ are.

Let S have genus $N + 3$. Fix a rational function Q whose poles are 0, 1, b_j, $1 \leqslant j \leqslant N$, and whose zeros are the other poles of q and the distinct odd order zeros of q. Then S is biholomorphic to the surface $\{(z, u) : Q(z) = u^2\}$. Then as in 10.5, $\pi^*(q_j/\sqrt{Q})$ $(1 \leqslant j \leqslant N)$ are a basis for the holomorphic 1-forms on S, and thus are linearly independent in $H^1(S, \mathbf{C})$. For all $1 \leqslant j \leqslant n-3$, let ω_j be a holomorphic 1-form such that in the local coordinate ζ at $\pi^{-1}(b_j)$, $\omega_j = (1 + O(\zeta))d\zeta$. For $1 \leqslant j \leqslant N$, we can choose ω_j to be a (nonzero) multiple of $\pi^*(q_j dz/\sqrt{Q})$. For $1 \leqslant j \leqslant n - 3$ $(n - 3 \geqslant N)$, let η_j be a multiple of $\pi^*(f_j dz/\sqrt{Q})$ such that in the local coordinate ζ at $\pi^{-1}(b_j)$, $\pi^*(f_j dz/\sqrt{Q}) = (4\zeta^{-2} + O(1))d\zeta$, again agreeing to first order with $\pi^*(f_j dz/\sqrt{q})$. Then by 11.5 (applied to Q),

$$\eta_j \cup \omega_k = -8\pi i \delta_{j,k}, \quad 1 \leqslant j, k \leqslant N.$$

So $\{\omega_k, \eta_j : 1 \leqslant j, k \leqslant N\}$ form a basis of $H^1(S, \mathbf{C})$ (over $\mathbf{C}$). So any meromorphic 1-form on S with zero residues at any poles can be written as a complex linear combination of these, as an element of $H^1(S, \mathbf{C})$. In particular, this can be done for η_ℓ for any $n-3 \geqslant \ell > N$. So

$$\eta_\ell = \sum_{j=1}^{N} t_{j,\ell}\eta_j + \xi,$$

for some $t_{j,\ell} \in \mathbf{C}$ and holomorphic 1-form ξ. For any $1 \leqslant k \leqslant n-3$, we also have, in H^1

$$\begin{aligned} \eta_\ell &= \omega_k(C_{k,\ell} + D_{k,\ell}) + \omega'_{k,\ell} + \omega''_{k,\ell}, \\ \eta_j &= \omega_k(C_{k,j} + D_{k,j}) + \omega'_{k,j} + \omega''_{k,j}, \quad 1 \leqslant j \leqslant N \end{aligned}$$

where these expressions satisfy the same conditions as in (1). So then, we have

$$\omega_k(D_{k,\ell} - \sum_{j=1}^{N} t_{j,\ell} D_{k,j}) = 0 \quad \text{at } \pi^{-1}(b_i),$$

and hence

$$D_{k,\ell} = \sum_{j=1}^{N} t_{j,\ell} D_{k,j}, \quad 1 \leqslant k \leqslant n-3,$$

and this gives the bound on the rank of X_2. The lower bound on the rank is clear from the linear independence of $\{\eta_j : 1 \leqslant j \leqslant N\}$.

12.9. When X_1 can be bounded by X_2. — We saw in 12.8 that, in the presence of multiple zeros or zero residues of q, X_2 has a kernel, which can be computed. It does not appear to be the case, however, that X_1 also has a kernel in general. If one examines the proof in 12.8, then it is clear that $X_1(\mathrm{Ker}(X_2))$ can be computed, and that one can find a basis of this subspace, such that coordinates of vectors in this basis are rational functions of the a_i and b_j. It appears to be possible for $X_1 \mid \mathrm{Ker}(X_2)$ to have maximal rank. However, 12.8 shows that a reduction in the rank of both X_1 and X_2 is associated with zero residues of q. We now give a more precise estimate. We write $X_1 = (C_{i,j})$ and $X_2 = (D_{i,j})$ as before. The result has content when a residue is small, when the result of 12.8 implies that for some i the i'th row and column of both X_1 and X_2 are small, for some i.

Lemma. — *As usual, let $a_j = \mathrm{Res}(q, b_j)$ with all poles bounded and bounded apart. If at most one zero of q is close to b_j, then $C_{j,k}$ and $D_{j,k}$ are $O(a_j)$ for all $k \neq j$. In general, $C_{j,k}$ and $D_{j,k}$ are $O(\sqrt{a_j})$ for all $j \neq k$ and $C_{j,j}$ and $D_{j,j}$ are $O(a_j/\varepsilon_j)$, where ε_j is the distance from b_j of the closest zero.*

Remark. — The proof will show that better estimates are possible, but it does not seem worth pursuing them.

Proof. — There is nothing to prove if a_j is bounded from 0. To simplify notation we take $j = 1$. We assume that q satisfies the Standing Assumption 10.3, since the general result follows by the continuity arguments of 12.7. We need to consider the harmonic functions g_1 and h_1 which have singularities $4\operatorname{Re}(1/\zeta)$, $4\operatorname{Im}(1/\zeta)$ for a local coordinate

$$\zeta = \int_{b_1} \sqrt{1/a_1(t-b_1)}(1+O(t-b_1))dt = O(\sqrt{(z-b_1)/a_1})(1+O(z-b_1)).$$

As usual we take the surface S branched over the zeros and poles of q. If b_1 is approximated by at least one zero of q then S is close to a degenerate surface S_1 which might be disconnected. If ε is the distance of the nearest zero of q from b_1, then $\zeta_1 = \sqrt{(z-b_1)/\varepsilon_1} = \sqrt{a_1/\varepsilon_1}\zeta(1+O(\zeta))$ is a local coordinate which is close to a coordinate bounded and bounded from 0 on S_1. It follows from the type of continuity arguments used in 12.7 that $\sqrt{\varepsilon_1/a_1}(g_1+ih_1)$ is bounded with respect to coordinate ζ_1, away from the singularity $\pi^{-1}(b_1)$. So $g_1 + ih_1 = O(\sqrt{a_1/\varepsilon_1})$ away from the singularity, which means that g_1 and h_1 have expansions in the ζ_1 coordinate about $\zeta_1 = 0$ in which the coefficients of ζ_1 and $\overline{\zeta}_1$ are $O(\sqrt{a_1/\varepsilon_1})$. So if we expand in terms of ζ, the coefficients of ζ and $\overline{\zeta}$ are $O(a_1/\varepsilon_1)$, that is, $C_{1,1}$ and $D_{1,1}$ are $O(a_1/\varepsilon_1)$. Of course, this has no content unless b_1 is approximated by at least two zeros of q. Now we consider $k \neq 1$. There is a finite union of disjoint annuli in S separating $\pi^{-1}b_k$ from $\pi^{-1}b_1$ such that any path from $\pi^{-1}b_1$ to $\pi^{-1}b_k$ crosses annuli with sum of moduli $\geqslant \frac{1}{2}\log(1/\varepsilon_1) - O(1)$. If b_1 is approximated by just one zero of q, then there are just two annuli, both of modulus $O\log(1/\varepsilon_1) - O(1)$. The functions g_1 and h_1 are bounded on all of these annuli. Moreover, annuli which are the same distance from $\pi^{-1}b_1$ — that is, separated from $\pi^{-1}b_1$ by the same number of other annuli — are preimages of the same annulus under π_1. So it suffices to prove the following. Let S_2 be a surface with finitely many boundary components, and let there be an annulus embedded in S_2 of modulus m homotopic to and adjacent to each boundary component. Let $S_2' \subset S_2$ be the complement of the annuli. Let g_1 be harmonic on S_2 and varying by at most L on ∂S_2, where any boundary components of ∂S_2 are preimages under π of a the same annulus. Then g_1 varies by at most $O(Le^{-m})$ on S_2. We shall then apply this to a decreasing family of surfaces S_2 containing $\pi^{-1}b_k$ but not containing $\pi^{-1}b_1$. This will complete the proof.

We consider the lift of g_1 to the universal cover of S_2, taking this to be the unit disc D. We continue to call the lifted harmonic function g_1. Then

$$g_1(re^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} g_1(e^{it})\frac{1-r^2}{|re^{i\theta}-e^{it}|^2}dt \tag{1}$$

where $g_1 \mid S^1$ is invariant under the action of Fuchsian group Γ which has zero measure limit set and there is a fundamental domain for the action of Γ on the domain of discontinuity in S^1 which is a finite union of at most two intervals of length $O(e^{-m})$.

If there are two intervals I_1 and I_2 then

$$\frac{\int_{I_1} g_1}{|I_1|} = \frac{\int_{I_2} g_1}{|I_2|}(1 + O(e^{-m})).$$

Also

$$\frac{\int_{\gamma I_1} g_1}{|\gamma I_1|} = \frac{\int_{I_1} g_1}{|I_1|}(1 + O(e^{-m})).$$

This means that we can write the integral (1) up as a countable sum of integrals over the orbits of one or two intervals, and we obtain

$$g_1(re^{i\theta}) = \frac{\int_{I_1} g_1}{|I_1|}(1 + O(e^{-m}))$$

for any such interval, as required. □

12.10. Another continuity result. — Let $q_m(z)dz^2$, $p_m(z)dz^2$ be quadratic differentials with $p_m(z)dz^2$ a stretch of $q_m(z)dz^2$. Assume for the moment that they satisfy the Standing Assumption with corresponding bases of harmonic 1-forms $v_{j,m}$, $(\chi_m)_*(v'_{j,m})$, as in Chapter 10, and $w_{j,m}$, $w'_{j,m}$ for real and imaginary parts of meromorphic 1-forms. Let $T_m = (v'_{j,m} \cup v_{i,m})$, like T in 10.10. We use matrices V_m^{-1}, $V_m'^{-1}$, $V_m^{-1}W_m$, $V_m'^{-1}W'_m$, as usual.

Lemma. — *Let $q_m \to q$, $p_m \to p$ with $p \neq q$. Then $\{V_m^{-1}T_mV_m'^{-1}\}$ is precompact, the kernel of any limit contains the kernel of $V'^{-1}\Sigma$ and the image is orthogonal to* $\mathrm{Ker}(V^{-1}\Sigma)$.

Proof. — First we claim that, for a constant C, for all vectors $\underline{x} = (x_j)$, $\underline{y} = (y_j)$,

$$|\langle \Sigma T_m\underline{x}, \underline{y}\rangle|^2 \leqslant C\langle \Sigma^t V'_m\underline{x}, \underline{x}\rangle\langle \Sigma^t V_m\underline{y}, \underline{y}\rangle.$$

To see this, write

$$\xi = \sum_j x_jv'_j,\ \eta = \sum_j y_jv_j.$$

Then, by Cauchy Schwartz,

$$\begin{aligned}\langle \Sigma T_m\underline{x}, \underline{y}\rangle^2 &= (\xi \cup J_m\eta)^2 \leqslant (\xi \cup J_m\xi)(\eta \cup J_m\eta) \leqslant C(\xi \cup J_m(-J_mJ'_m\xi))(\eta \cup J_m\eta)\\ &= C(\xi \cup J'_m\xi)(\eta \cup J_m\eta) = C\langle \Sigma^t V'_m\underline{x}, \underline{x}\rangle\langle \Sigma^t V_m\underline{y}, \underline{y}\rangle.\end{aligned}$$

Then replacing $\underline{x}$ by $\Sigma V_m'^{-1}\underline{x}$ and $\underline{y}$ by $\Sigma V_m^{-1}\underline{y}$ yields

$$|\langle V_m^{-1}T_mV_m'^{-1}\Sigma\underline{x}, \underline{y}\rangle|^2 \leqslant C\langle \Sigma V_m'^{-1}\underline{x}, \underline{x}\rangle\langle \Sigma V_m^{-1}\underline{y}, \underline{y}\rangle.$$

This gives the result. □

12.11. We use V^{-1} and $V^{-1}W$, $V^{-1}TV'^{-1}$ even in degenerate cases, when V and W are not actually defined (because, as we have seen, V^{-1} and $V^{-1}W$ still make sense, by taking limits).

Lemma. — *Let $q_\delta(z)dz^2$ be the quadratic differential at distance δ along the geodesic given by $q(z)dz^2 = q_0(z)dz^2$, with corresponding V_δ^{-1} and $V_\delta^{-1}W_\delta$ and $T_\delta = (v_{j,\delta} \cup v_i)$. Then*

$$\lim_{\delta\to 0} V^{-1} - V_\delta^{-1} = 0, \quad \lim_{\delta\to 0} V^{-1}W - V_\delta^{-1}W_\delta = 0, \quad \lim_{\delta\to 0} V^{-1}T_\delta V_\delta^{-1} - V^{-1} = 0.$$

Proof. — The first two limits follow from 12.7. To get the other limit, we use the fact that the matrix $V^{-1}T_\delta V_\delta^{-1}$ gives the coefficients of $w_{j,\delta}^{(2)}$ in terms of the v_k. As $\delta \to 0$, $w_{j,\delta}^{(2)} \to w_j^{(2)}$ in homology (because of the formulae 11.5 for $w_{j,\delta}^{(2)} \cup v_{k,\delta}$, and because $v_{k,\delta} \to v_k$ in homology). This, together with the bound on image from the previous lemma, gives the result. □

12.12. Another result about bounds on entries of X_1, X_2. — The following lemma (as we shall see) is proved by a very similar method to 12.7. It will be needed towards the end of the proof of Descending Points in Chapter 22.

Lemma

(1) *Given $C_1 > 0$, there is a constant $C_2 > 0$ such that the following holds. Let the poles $\{b_j : 1 \leqslant j \leqslant n-3\}$, and zeros $\{z_j : 1 \leqslant j \leqslant n-4\}$ of q be bounded, and bounded from the poles 0, 1, and let*

$$e^{-m}/C_1 \leqslant |b_j - b_k| \leqslant C_1 e^{-m}, \quad |b_j - z_k| \leqslant C_1 e^{-m} \quad \text{for all } j,\ k.$$

Then all entries in V^{-1}, WV^{-1} are $\leqslant C_2 e^m$.

(2) *Let q be as in 1, and let $q' = q_\delta$ be similarly defined, with residues within δ of those of q, and poles within $e^{-m}\delta$ of those of q, and with corresponding matrices V_δ^{-1}, $V_\delta^{-1}W_\delta$, $V^{-1}T_\delta V_\delta^{-1}$. Then there is $\varepsilon(\delta)$ with $\lim_{\delta\to 0}\varepsilon(\delta) = 0$ such that*

$$\|e^{-m}V^{-1} - e^{-m}V_\delta^{-1}\| \leqslant \varepsilon(\delta),$$
$$\|e^{-m}V^{-1}W - e^{-m}V_\delta^{-1}W_\delta\| \leqslant \varepsilon(\delta),$$
$$\|e^{-m}V^{-1} - e^{-m}V^{-1}T_\delta V_\delta^{-1}\| \leqslant \varepsilon(\delta).$$

Proof

(1) It suffices to look at an arbitrary sequence $\{q_m\}$ satisfying the conditions of q, and to show that all entries in the corresponding matrices V_m^{-1}, $W_m V_m^{-1}$ are $\leqslant C_2 e^m$ for some C_2. We use much of the notation of 12.7. Again, the terms $r'_{j,m}(b_{j,m})$ are easily seen to be $O(e^m)$: we are considering the $z - b_{j,m}$ term in the Taylor series expansion of $q_m(z)/(z-b_{j,m})(q_{m,j}(z))^2$. Expand out only one term at a time. There are the same number of terms within $O(e^{-m})$ of $b_{j,m}$ on top and bottom, all terms on the bottom are e^{-m}/C at $b_{j,m}$ and those on top are closer to $b_{j,m}$, if anything.

So now we only need to consider the coefficients in the harmonic functions. We take an embedded disc in S_m round $\pi_m^{-1}(b_{j,m})$. In the ζ-coordinate we only know, now, that $K_{j,m}$ can be chosen to have radius $\geqslant r_0 e^{-m/2}$. Let $K_0 = \{\zeta : |\zeta| \leqslant r_0\}$. Then we consider the harmonic functions

$$G_{1,j,m}(\zeta) = e^{-m/2} g_{1,m}(e^{-m/2}\zeta), \quad \zeta \in K_0 \setminus \{0\},$$

and similarly for $h_{1,m}$. If $j \neq 1$, these functions are harmonic on all of K_0. If $j = 1$, and $4\,\mathrm{Re}(1/\zeta)$, $4\,\mathrm{Im}(1/\zeta)$ are added to these functions, we obtain harmonic functions on all of K_0. The Fourier coefficient expansions for these functions are $\{e^{-m} a_{n,j,m}\}$ $(n \geqslant -1)$ where $\{a_{n,j,m}\}$ are the coefficients in $K_{j,m}$ for $g_{1,m}$. We can then take limits along subsequences of the functions $\lambda_{j,m} e^{-m/2} g_{1,m} + c_{j,m}$ for suitable constants $c_{j,m}$ on limits of the S_m which contain the disc K_0, which is a local chart at $\lim_{m\to\infty} \pi_m^{-1}(b_{j,m})$. Arguing as in 12.7, we can take $\lambda_{j,m} = 1$ for all m. In local coordinates the limit function is the limit of $G_{1,m,j}$, and thus we obtain that $\lim_{m\to\infty} e^{-m} a_{1,j,m}$ exists. Then for the entire sequence, $\{a_{1,j,m}\}$ is bounded, as required.

(2) For the first two inequalities, it suffices to take a sequence q_m as above for which poles and residues converge. The S_m converge, and the harmonic functions $\lambda_{j,m} e^{-m/2} g_{1,m} + c_{j,m}$ must converge, because otherwise we obtain nonconstant harmonic functions (without singularities) in the limit. So $\lim_{m\to\infty} e^{-m} a_{1,j,m}$ exists for all j. The limit of the $z - b_{j,m}$ term in the Taylor series expansion of $q_m(z)/(z - b_{j,m})(q_{m,j}(z))^2$ exists. So $\lim_{m\to\infty} e^{-m} V_m^{-1}$, $\lim_{m\to\infty} e^{-m} V_m^{-1} W_m$ exist, and this is enough for the first two inequalities.

The last inequality is proved similarly to the last limit in 12.11. For this we do need to take q_m, q'_m with the same limiting residues and corresponding V_m, V'_m, T_m. By 12.10 we know that $\{e^{-m} V_m^{-1} T_m V_m'^{-1}\}$ is precompact, with bounds on image space of any limit.Then we obtain the result as in 12.11, using convergence in homology on the appropriate subsurface.

CHAPTER 13

THE SECOND DERIVATIVE AND THE SOLUTION OF A DIFFERENTIAL EQUATION

13.1. In this chapter, we give yet another interpretation of terms in the second derivative of the Teichmüller distance function for marked spheres. Continuing with the notation of Chapters 10-12, let $q(z)dz^2$ be a quadratic differential on $\overline{\mathbf{C}}$ with at worst simple poles at the points 0, 1, ∞ and b_j, $1 \leqslant j \leqslant n-3$, and let $a_j = \text{Res}(q, b_j)$. Assume in this chapter that $a_j \neq 0$ for all j, that residues at 0, 1, ∞ are also $\neq 0$ and that all zeros are simple. We again call this the *Standing Assumption*, as in 10.3. Let S be as in 10.5, with genus $n-3$, v_j, w_j as in 10.7, V, W as in 10.10, Π, Σ as in 10.11 and $\Omega = \Sigma\Pi$. Then let

$$X = X(q) = V^{-1}W - 4\pi V^{-1}\Omega.$$

We have seen in 11.11 that, if (v_j) and (w_j) denote row vectors, then

$$(w_j) = (v_j)X.$$

We showed that $X\Pi$ is symmetric, and, indeed, that both of $V^{-1}W\Pi$ and $V^{-1}\Sigma = V^{-1}\Omega\Pi$ are symmetric (although the second is obvious by inspection). In this chapter, we shall exhibit X as a solution of a differential equation. One of the consequences is another proof of the symmetry of $X\Pi$. Recall that the proof in Chapter 11 used Green's functions. The proof given here is more elementary, but heavier computationally. It is the proof that one expects to find first, because the only technique used is integration by parts. In fact, it was not the first proof to emerge. The reason is that, although the forms v_j form a basis for $H^1(S)$ over $\mathbf{R}$, in order to get a basis of meromorphic 1-forms over $\mathbf{C}$, we have to include all of $v_{2j-1} + iv_{2j}$ and $w_{2j-1} + iw_{2j}$, $1 \leqslant j \leqslant n-3$. This is related to the fact that X is not a complex-linear matrix. Instead, it occurs as the solution of a differential equation in which the coefficients are complex-matrix-valued functions.

13.2. A notation convention. — Let $P = (p_{i,j})$ be an $m \times m$ matrix with complex coefficients. We shall denote by P^{ct} the matrix $(p_{j,i})$. Note that P can also be regarded as a real $2m \times 2m$ matrix, by replacing each entry $a + ib$ (as usual) by $M(a + ib) = \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$. By abuse of notation, we shall write products such as PY, where Y is a real $2m \times 2m$ matrix, not necessarily associated with a complex $m \times m$ matrix. Note, for example that

$$\Pi P \Pi = \overline{P},$$

where $\overline{P} = (\overline{p}_{i,j})$ as a complex matrix. If we regard P as a $2m \times 2m$ real matrix, and then P^t denotes the real transpose. Thus, $P^t = (\overline{P})^{ct}$, by abuse of notation. So $P = P^{ct}$ if and only if $P\Pi = (P\Pi)^t$.

13.3. The form of the differential equation

Quadratic Differential Equation Theorem. — *Let λ be a real variable, and let $q_\lambda dz^2$ be any C^1 family of quadratic differentials on $\overline{\mathbf{C}}$ satisfying the Standing Assumption. Write $X(\lambda) = X(q_\lambda)$. Then $X(\lambda)$ is a solution of a differential equation*

$$\frac{dY}{d\lambda} = YPY + QY + YQ^{ct} + R \tag{1}$$

where P, Q, R are complex-linear-matrix-valued functions of λ with $P = P^{ct}$, $R = R^{ct}$.

13.4. Remark. — It follows from (1) that

$$\frac{d(X\Pi)}{d\lambda} = (X\Pi)(\Pi P)(X\Pi) + Q(X\Pi) + X\Pi(\Pi Q^{ct}\Pi) + R\Pi, \tag{2}$$

where $(\Pi P)^t = \Pi P$, $(R\Pi)^t = R\Pi$ and $\Pi Q^{ct}\Pi = Q^t$. It follows immediately that if $X(\lambda_0)\Pi = (X(\lambda_0)\Pi)^t$ for some λ_0 then $X(\lambda)\Pi = (X(\lambda)\Pi)^t$ for all λ, because both $X(\lambda)\Pi$ and $(X(\lambda)\Pi)^t$ solve (1). Note also that if Y is a solution of (1) on a connected open set of λ and $Y(\lambda_0)$ is complex linear and $Y(\lambda_0) = Y(\lambda_0)^{ct}$, then uniqueness implies the same two properties hold for $Y(\lambda)$ for all λ. The following lemma shows that, given the theorem, $X(q)\Pi$ is symmetric for all q satisfying the Standing Assumption. Hence, by continuity, $X(q)\Pi$ is symmetric for all q.

13.5. Lemma. — *There exists $q(z)dz^2$ such that*

$$X(q)\Pi = (X(q)\Pi)^t.$$

Proof. — Choose q so that all poles b_j $(1 \leqslant j \leqslant n-3)$ and all zeros z_j $(1 \leqslant j \leqslant n-4)$ lie on the real line, and so that the following hold, where

$$T_1 = \{0,1\} \cup \{z_j : 1 \leqslant j \leqslant n-4\}, \quad T_2 = \{\infty\} \cup \{b_j : 1 \leqslant j \leqslant n-3\}.$$

(1) Points from T_1, T_2 alternate, so that any two adjacent points of T_1 are separated by precisely one point of T_2, and vice versa.

(2) Under some Möbius transformation of $\mathbf{R} \cup \{\infty\}$ to $\{z :\mid z \mid= 1\}$, points of $T_1 \cup T_2$ are mapped to points equally spaced around the circle.

It follows from 1 and 2 that there is a group of Möbius transformations of $T_1 \cup T_2$ which acts transitively, whose action is isomorphic to that of the dihedral group on $2n-4$ letters, and any transformation in the group which maps one point of T_2 into T_2 leaves both T_1 and T_2 invariant. Write $X = (E_{j,k})$ for the decomposition of X into 2×2 submatrices. Thus, $E_{j,k}$ represents the coefficients of v_{2j-1} and v_{2j} in w_{2k-1} and w_{2k}. We need to show that for all $j \neq k$,

$$E_{j,k} = E_{k,j}.$$

So fix $k \neq j$. By our choice of q, there is a Möbius involution σ which interchanges b_j and b_k and leaves each of the sets T_1, T_2 invariant. Let b_p be such that σ interchanges ∞ and b_p. Write σ also for the lifted involution to the surface S. Then for all $\ell \leqslant 2n-6$, $\sigma_* w_\ell$ can be written as a real linear combination of $\sigma_* v_m$ $(1 \leqslant m \leqslant 2n-6)$ and $E_{j,k}$ represents the coefficients of $\sigma_* v_{2j-1}$ and $\sigma_* v_{2j}$ in $\sigma_* w_{2k-1}$ and $\sigma_* w_{2k}$. Let $\pi : S \to \overline{\mathbf{C}}$ denote the natural projection (as before). Then the following hold.

a) For any $m \leqslant n-3$, $v_{2m-1} + iv_{2m}$ vanishes at $\pi^{-1}(b_\ell)$ except for $\ell = m$.

b) For any $r \neq p$, $\sigma_*(v_{2r-1} + iv_{2r})$ vanishes at $\pi^{-1}(b_\ell)$ except for $\ell = p$, t, where $\sigma(b_r) = b_t$ for $r \neq p$ and $t = p$ if $r = p$. In particular, $t = k$ if $r = j$.

c) The meromorphic form $\sigma_*(w_{2k-1} + iw_{2k})$ has a single double pole singularity at $\pi^{-1}(b_j)$ and vanishes at $\pi^{-1}(b_\ell)$ for $\ell \neq j$, p.

Therefore, $\sigma_*(v_{2r-1} + iv_{2r})$ can be written as a complex linear combination of $v_{2t-1} + iv_{2t}$ and $v_{2p-1} + iv_{2p}$. Similarly, $\sigma_*(w_{2k-1} + iw_{2k})$ can be written as a complex linear combination of $w_{2j-1} + iw_{2j}$, $v_{2j-1} + iv_{2j}$ and $v_{2p-1} + iv_{2p}$. It remains to show that the coefficient of $v_{2k-1} + iv_{2k}$ in $\sigma_*(v_{2j-1} + iv_{2j})$ is the same as the coefficient of $w_{2j-1} + iw_{2j}$ in $\sigma_*(w_{2k-1} + iw_{2k})$.

Because $\sigma^2 =$ identity on S, the coefficients of $v_{2k-1} + iv_{2k}$ in $\sigma_*(v_{2j-1} + iv_{2j})$ and of $v_{2j-1} + iv_{2j}$ in $\sigma_*(v_{2k-1} + iv_{2k})$ multiply to 1. In fact, these coefficients (which depend, up to sign, on the choice of the lift σ on S) are given respectively by

$$\sqrt{\frac{\sigma'(b_k)a_j}{a_k}}, \quad \sqrt{\frac{\sigma'(b_j)a_k}{a_j}}, \tag{1}$$

where the signs of the square root are such that the product is 1. (Note that $\sigma'(b_j)\sigma'(b_k) = 1$.) Let $v_{2j-1} + iv_{2j}$ be of the form $d\xi$ for a local coordinate ξ near $\pi^{-1}(b_j)$. We saw in 11.10 (and can check again) that in this coordinate, $w_{2j-1} + iw_{2j}$ is of the form

$$\Big(\frac{4}{\xi^2} + O(1)\Big)d\xi. \tag{2}$$

We could also use ξ as a local coordinate for $\sigma_*(v_{2j-1} + iv_{2j})$ near $\pi^{-1}(b_k)$. Then $\sigma_*(w_{2j-1} + iw_{2j})$ is also given by (2) in this coordinate. The same holds with j

replaced by k, and a local coordinate ζ near $\pi^{-1}(b_k)$. It follows from (1) that

$$d\zeta = \left(\sqrt{\frac{\sigma'(b_k)a_j}{a_k}} + o(1)\right)d\xi, \quad \zeta = \left(\sqrt{\frac{\sigma'(b_k)a_j}{a_k}} + o(1)\right)\xi,$$

$$\frac{4}{\xi^2}d\xi = \left(\sqrt{\frac{\sigma'(b_k)a_j}{a_k}} + o(1)\right)\frac{4}{\zeta^2}d\zeta,$$

and hence the coefficient of $w_{2j-1} + iw_{2j}$ in $\sigma_*(w_{2k-1} + iw_{2k})$ is

$$\sqrt{\frac{\sigma'(b_k)a_j}{a_k}}$$

with the same choice of sign as in (1), as required. □

13.6. First Reduction in 13.3

Lemma. — *To prove Theorem 13.3 it suffices to show that we have complex linear equations (using column vectors)*

(1) $$\left(\frac{d(v_{2k-1} + iv_{2k})}{d\lambda}\right) = C(v_{2k-1} + iv_{2k}) + D(w_{2k-1} + iw_{2k}),$$

(2) $$\left(\frac{d(w_{2k-1} + iw_{2k})}{d\lambda}\right) = G(v_{2k-1} + iv_{2k}) + H(w_{2k-1} + iw_{2k})$$

with

$$C^{ct} = -H, \quad D = D^{ct}, \quad G = G^{ct}.$$

Proof. — Assume that (1) and (2) hold. We have

$$(w_k) = (v_k)X.$$

Then

(3) $$\left(\frac{dw_k}{d\lambda}\right) = \left(\frac{dv_k}{d\lambda}\right)X + (v_k)\frac{dX}{d\lambda}.$$

Then reverting to real row vectors and using the abuse of notation explained in 13.2 we have

(4) $$\left(\frac{dv_k}{d\lambda}\right) = (v_k)C^t + (w_k)D^t = (v_k)(C^t + XD^t),$$

(5) $$\left(\frac{dw_k}{d\lambda}\right) = (v_k)G^t + (w_k)H^t = (v_k)(G^t + XH^t).$$

Then from (3), (4) and (5) we obtain

$$(v_k)(G^t + XH^t) = (v_k)(C^t + XD^t)X + (v_k)\frac{dX}{d\lambda}.$$

This gives

$$\frac{dX}{d\lambda} = -XD^tX - C^tX + XH^t + G^t.$$

Write $P = -D^t = -\overline{D}^{ct}$ (see 13.2), $Q = -C^t = -\overline{C}^{ct}$ and $R = \overline{G}^{ct}$. Then $P = P^{ct}$, $H^t = Q^{ct}$ and $R = R^{ct}$ as required for 13.3. □

13.7. Second Reduction: in Choice of Family. — Let $q(z)dz^2$ be any quadratic differential, as usual with poles at 0, 1, ∞, b_j $(1 \leqslant j \leqslant n-3)$ and zeros at z_j $(1 \leqslant j \leqslant n-4)$. Thus, for some $\alpha \neq 0$,

$$q(z) = \alpha \frac{\prod_{j=1}^{n-4}(z-z_j)}{z(z-1)\prod_{j=1}^{n-3}(z-b_j)}.$$

If q_λ is any C^1 family satisfying the Standing Assumption, then $\alpha(\lambda)$, $b_j(\lambda)$ and $z_j(\lambda)$ are all C^1 functions of λ and

$$\frac{d}{d\lambda} = \frac{\partial}{\partial\alpha}\frac{d\alpha}{d\lambda} + \sum_{j=1}^{n-3}\frac{\partial}{\partial b_j}\frac{db_j}{d\lambda} + \sum_{j=1}^{n-4}\frac{\partial}{\partial z_j}\frac{dz_j}{d\lambda}.$$

It follows that we can restrict to considering complex parameter families in which only one of the functions α, b_j, z_j is nonconstant. For each such parameter we need to satisfy (1) of 13.6. So we need to prove the following Reduced Theorem.

13.8. Reduced Theorem. — *There are complex-matrix-valued functions C_j, D_j, E_j, F_j, G_j, J_j, such that, in $H^1(S)$, if $(v_{2k-1}+iv_{2k})$ and $(w_{2k-1}+iw_{2k})$ denote column vectors,*

$$\frac{\partial}{\partial z_j}(v_{2k-1}+iv_{2k}) = C_j(v_{2k-1}+iv_{2k}) + D_j(w_{2k-1}+iw_{2k}),$$
$$\frac{\partial}{\partial z_j}(w_{2k-1}+iw_{2k}) = G_j(v_{2k-1}+iv_{2k}) + H_j(w_{2k-1}+iw_{2k}),$$

where

$$C_j^{ct} = -H_j, \quad D_j = D_j^{ct}, \quad G_j = G_j^{ct}.$$

Similarly,

$$\frac{\partial}{\partial b_j}(v_{2k-1}+iv_{2k}) = E_j(v_{2k-1}+iv_{2k}) + F_j(w_{2k-1}+iw_{2k}),$$
$$\frac{\partial}{\partial b_j}(w_{2k-1}+iw_{2k}) = J_j(v_{2k-1}+iv_{2k}) + K_j(w_{2k-1}+iw_{2k}),$$

where

$$E_j^{ct} = -K_j, \quad F_j = F_j^{ct}, \quad J_j = J_j^{ct}.$$

Furthermore,

$$\frac{\partial}{\partial\alpha}(v_{2k-1}+iv_{2k}) = \frac{-1}{2\alpha}(v_{2k-1}+iv_{2k}),$$
$$\frac{\partial}{\partial\alpha}(w_{2k-1}+iw_{2k}) = \frac{1}{2\alpha}(w_{2k-1}+iw_{2k}).$$

The specific formulae are as follows.

$$C_j(k,\ell) = \frac{a_\ell b_k(b_k-1)b_\ell(b_\ell-1)}{2(b_k-z_j)(b_\ell-z_j)^2 z_j^2(1-z_j)^2 q'(z_j)} \quad (k \neq \ell),$$

$$C_j(k,k) = \frac{a_k b_k^2(b_k-1)^2}{2(b_k-z_j)^3 z_j^2(1-z_j)^2 q'(z_j)} + \frac{1}{2(b_k-z_j)},$$

$$D_j(k,\ell) = \frac{b_k(b_k-1)b_\ell(b_\ell-1)}{2(b_k-z_j)(b_\ell-z_j)z_j^2(1-z_j)^2 q'(z_j)},$$

$$E_j(k,k) = \frac{-1}{2(b_k-b_j)},\ k\neq j, \quad E_j(j,j) = \frac{2b_j-1}{b_j(b_j-1)},$$

$$E_j(j,k) = \frac{b_k(b_k-1)}{2(b_k-b_j)b_j(b_j-1)},\ k\neq j, \quad E_j(k,\ell)=0 \text{ if } j\neq k \text{ and } k\neq \ell,$$

$$F_j(k,k) = \frac{1}{2a_k}, \quad F_j(k,\ell) = 0 \text{ if } k\neq\ell,$$

$$G_j(k,\ell) = \frac{-a_k a_\ell b_k(b_k-1)b_\ell(b_\ell-1)}{2(b_k-z_j)^2(b_\ell-z_j)^2 z_j^2(1-z_j)^2 q'(z_j)}, \quad k\neq\ell,$$

$$G_j(k,k) = \frac{-a_k^2 b_k^2(b_k-1)^2}{2(b_k-z_j)^4 z_j^2(1-z_j)^2 q'(z_j)} - \frac{a_k}{2(b_k-z_j)^2},$$

$$J_j(k,k) = \frac{a_k}{2(b_k-b_j)^2},\ j\neq k, \quad J_j(k,j) = J_j(j,k) = \frac{-a_k b_k(b_k-1)}{2b_j(b_j-1)(b_k-b_j)^2},\ j\neq k,$$

$$J_j(j,j) = \frac{1}{2}\Big(\sum_{\substack{k=1\\k\neq j}}^{n-3} \frac{a_k b_k(b_k-1)}{b_j(b_j-1)(b_j-b_k)}\Big(\frac{1}{b_j-b_k} - \frac{1}{b_j} - \frac{1}{b_j-1}\Big) - \frac{a_j}{b_j(b_j-1)}\Big).$$

13.9. Remark. — As one would expect, this is an extremely tedious calculation, and the reader will probably want to skip the rest of this chapter, at least in the first instance.

13.10. Calculation of C_j, D_j, E_j, F_j. — Recall that

$$v_{2k-1} + iv_{2k} = \pi^*\left(\frac{q_k dz}{\sqrt{q}}\right).$$

Now

$$\frac{\partial}{\partial z_j}\left(\frac{q_k}{\sqrt{q}}\right) = \frac{1}{\sqrt{q}}\frac{\partial q_k}{\partial z_j} - \frac{q_k}{2\sqrt{q}}\frac{1}{q}\frac{\partial q}{\partial z_j}$$

Using

$$q = \alpha\frac{\prod_{m=1}^{n-4}(z-z_m)}{z(z-1)\prod_{p=1}^{n-3}(z-b_p)},$$

we have

$$\frac{\partial q_k}{\partial z_j} = 0, \quad \frac{1}{q}\frac{\partial q}{\partial z_j} = \frac{-1}{z-z_j}.$$

So

$$\frac{\partial}{\partial z_j}\left(\frac{q_k}{\sqrt{q}}\right) = \frac{q_k}{2\sqrt{q}(z-z_j)} = \frac{1}{2(b_k-z_j)}\frac{q_k}{\sqrt{q}} - \frac{b_k(b_k-1)}{2(b_k-z_j)}\frac{g_j}{\sqrt{q}}$$

where

$$g_j(z) = \frac{1}{(z-z_j)z(z-1)}.$$

So the formulae for C_j and D_j then follow from Lemma 13.14 below. Now remember that

$$q_k(z) = \frac{b_k(b_k-1)}{z(z-1)(z-b_k)}, \quad f_k(z) = \frac{a_k b_k(b_k-1)}{z(z-1)(z-b_k)^2}$$

So if $k \neq j$, we have

$$\frac{\partial}{\partial b_j}\frac{q_k}{\sqrt{q}} = \frac{-q_k}{2\sqrt{q}}\frac{1}{q}\frac{\partial q}{\partial b_j} = \frac{-q_k}{2(z-b_j)\sqrt{q}} = \frac{b_k(b_k-1)q_j}{2(b_k-b_j)b_j(b_j-1)\sqrt{q}} - \frac{q_k}{2(b_k-b_j)\sqrt{q}},$$

and

$$\frac{\partial}{\partial b_j}\frac{q_j}{\sqrt{q}} = \frac{(2b_j-1)q_j}{b_j(b_j-1)\sqrt{q}} + \frac{f_j}{a_j\sqrt{q}} - \frac{f_j}{2a_j\sqrt{q}} = \frac{(2b_j-1)q_j}{b_j(b_j-1)\sqrt{q}} + \frac{f_j}{2a_j\sqrt{q}}.$$

This immediately gives the formulae for E_j and F_j.

13.11. The $\partial/\partial\alpha$ Calculations. — We have

$$\frac{\partial q_k}{\partial \alpha} = 0, \quad \frac{1}{q}\frac{\partial q}{\partial \alpha} = \frac{1}{\alpha}.$$

Recall that

$$a_k = \alpha\frac{\prod_{m=1}^{n-4}(b_k - z_m)}{b_k(b_k-1)\prod_{p\neq k}(b_k-b_p)},$$

and hence

$$\frac{\partial a_k}{\partial \alpha} = \frac{a_k}{\alpha}.$$

So

$$\frac{\partial}{\partial\alpha}\frac{q_k}{\sqrt{q}} = \frac{-q_k}{2q\sqrt{q}}\frac{\partial q}{\partial \alpha} = \frac{-q_k}{2\alpha\sqrt{q}}, \quad \frac{\partial}{\partial\alpha}\frac{f_k}{\sqrt{q}} = \frac{f_k}{\alpha\sqrt{q}} - \frac{f_k}{2\alpha\sqrt{q}} = \frac{f_k}{2\alpha\sqrt{q}},$$

which gives the $\partial/\partial\alpha$ calculations required for 13.8.

13.12. Calculation of G_j, H_j, J_j, K_j. — From the formula for a_k given in 13.11, we have

$$\frac{\partial a_k}{\partial z_j} = \frac{-a_k}{b_k - z_j}, \quad \frac{\partial a_k}{\partial b_j} = \frac{a_k}{b_k-b_j} \text{ if } k \neq j,$$

$$\frac{\partial}{\partial b_j}(a_j b_j(b_j-1)) = -a_j b_j(b_j-1)\Big(\frac{1}{b_j} + \frac{1}{b_j-1} + \sum_{k\neq j}\frac{1}{b_j-b_k}\Big).$$

So

$$\frac{\partial}{\partial z_j}\frac{f_k}{\sqrt{q}} = \frac{-f_k}{(b_k - z_j)\sqrt{q}} + \frac{f_k}{2(z - z_j)\sqrt{q}}$$
$$= \frac{-f_k}{2(b_k - z_j)\sqrt{q}} - \frac{a_k q_k}{2(b_k - z_j)^2\sqrt{q}} + \frac{a_k b_k (b_k - 1) g_j}{2(b_k - z_j)^2\sqrt{q}},$$

using

(1) $$\frac{1}{(z - b_k)^2(z - z_j)} = \frac{1}{(b_k - z_j)(z - b_k)^2} + \frac{1}{(b_k - z_j)^2(z - z_j)} - \frac{1}{(b_k - z_j)^2(z - b_k)}.$$

The calculation of G_j and H_j then follows from 13.14. If $j \neq k$, we have

$$\frac{\partial}{\partial b_j}\frac{f_k}{\sqrt{q}} = \frac{f_k}{(b_k - b_j)\sqrt{q}} - \frac{f_k}{2(z - b_j)\sqrt{q}}$$
$$= \frac{f_k}{2(b_k - b_j)\sqrt{q}} + \frac{a_k q_k}{2(b_k - b_j)^2\sqrt{q}} - \frac{a_k b_k (b_k - 1) q_j}{2(b_k - b_j)^2 b_j (b_j - 1)\sqrt{q}},$$

using

(2) $$\frac{1}{(z - b_k)^2(z - b_j)} = \frac{1}{(b_k - b_j)(z - b_k)^2} + \frac{1}{(b_k - b_j)^2(z - b_j)} - \frac{1}{(b_k - b_j)^2(z - b_k)}.$$

This gives the formulae of 13.8 for $J_j(k, \ell)$ and $K_j(k, \ell)$ if $k \neq j$. Finally, we have

$$\frac{\partial}{\partial b_j}\frac{f_j}{\sqrt{q}} = \Big(\sum_{m=1}^{n-4}\frac{1}{b_j - z_j} - \sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{1}{b_j - b_k} + \frac{2b_j - 1}{b_j(b_j - 1)}\Big)\frac{f_j}{\sqrt{q}} + \frac{3a_j b_j (b_j - 1)}{2z(z - 1)(z - b_j)^3\sqrt{q}}.$$

Then the formulae of 13.8 for $J_j(j, \ell)$ and $K_j(j, \ell)$ follow from 13.15 and 13.16.

13.13. Lemma. — *In $H^1(S)$,*

(1) $$\Big(\sum_{k=1}^{n-3} a_k b_k\Big)\pi^*\Big(\frac{dz}{z(z - 1)^2\sqrt{q}}\Big) = \sum_{k=1}^{n-3}\pi^*\Big(\frac{a_k b_k}{b_k - 1}\frac{q_k dz}{\sqrt{q}} + b_k\frac{f_k dz}{\sqrt{q}}\Big),$$

(2)
$$\Big(\sum_{k=1}^{n-3} a_k (b_k - 1)\Big)\pi^*\Big(\frac{dz}{z^2(z - 1)\sqrt{q}}\Big) = \sum_{k=1}^{n-3}\pi^*\Big(-(b_k - 1)\frac{f_k dz}{\sqrt{q}} + \frac{a_k(1 - b_k)}{b_k}\frac{q_k}{\sqrt{q}}\Big).$$

Proof. — We have

$$\frac{d}{dz}(z\sqrt{q}) = \sqrt{q} + \frac{zq'}{2\sqrt{q}} = \sqrt{q} + \frac{z}{2\sqrt{q}}\sum_{k=1}^{n-3} a_k q_k\Big(\frac{-1}{z} - \frac{1}{z - 1} - \frac{1}{z - b_k}\Big)$$
$$= \frac{-\sqrt{q}}{2} - \frac{1}{2\sqrt{q}}\sum_{k=1}^{n-3}\Big(\frac{a_k q_k}{z - 1} + \frac{a_k b_k q_k}{z - b_k}\Big).$$

Then using

$$\frac{1}{(z - 1)(z - b_k)} = \frac{1}{b_k - 1}\Big(\frac{1}{z - b_k} - \frac{1}{z - 1}\Big),$$

and rearranging, we obtain (1). Similarly,

$$\frac{d}{dz}((z-1)\sqrt{q}) = \sqrt{q} + \frac{(z-1)q'}{2\sqrt{q}} = \sqrt{q} - \frac{1}{2\sqrt{q}}\sum_{k=1}^{n-3} a_k q_k \Big(\frac{z-1}{z} + 1 + \frac{z-1}{z-b_k}\Big)$$
$$= \frac{-\sqrt{q}}{2} + \frac{1}{2\sqrt{q}}\sum_{k=1}^{n-3}\Big(\frac{a_k q_k}{z} - \frac{(b_k-1)a_k q_k}{z-b_k}\Big).$$

Using

$$\frac{1}{z(z-b_k)} = \frac{1}{b_k}\Big(\frac{1}{z-b_k} - \frac{1}{z}\Big),$$

and rearranging, we obtain (2). □

13.14. Lemma. — *In* $H^1(S)$,

$$z_j^2(z_j-1)^2 q'(z_j)\pi^*\Big(\frac{g_j dz}{\sqrt{q}}\Big) = -\sum_{k=1}^{n-3}\frac{b_k(b_k-1)}{b_k-z_j}\pi^*\Big(\frac{f_k dz}{\sqrt{q}}\Big) - \sum_{k=1}^{n-3}\frac{a_k b_k(b_k-1)}{(b_k-z_j)^2}\pi^*\Big(\frac{q_k dz}{\sqrt{q}}\Big).$$

Proof

(1)
$$\frac{d}{dz}\frac{\sqrt{q}}{z-z_j} = \frac{-q}{(z-z_j)^2\sqrt{q}} + \frac{q'}{2(z-z_j)\sqrt{q}}.$$

Now we write

(2)
$$\frac{z(z-1)q}{(z-z_j)^2} = \frac{c}{z-z_j} + \sum_{k=1}^{n-3}\frac{c_k}{z-b_k}.$$

Then

$$c_k = \lim_{z\to b_k}\frac{z(z-1)q(z)(z-b_k)}{(z-z_j)^2} = \frac{b_k(b_k-1)a_k}{(b_k-z_j)^2},$$
$$c = \lim_{z\to z_j}\frac{z(z-1)(q(z)-q(z_j))}{z-z_j} = z_j(z_j-1)q'(z_j).$$

Now write

(3)
$$\frac{z(z-1)q'}{z-z_j} = \frac{d}{z-z_j} + \frac{e}{z} + \frac{f}{z-1} + \sum_{k=1}^{n-3}\frac{s_k}{z-b_k} + \sum_{k=1}^{n-3}\frac{t_k}{(z-b_k)^2}.$$

Then

$$d = \lim_{z\to z_j} z(z-1)q'(z) = z_j(z_j-1)q'(z_j).$$

Since

$$q_k'(z) = \frac{b_k(b_k-1)}{z(z-1)(z-b_k)}\Big(\frac{-1}{z} - \frac{1}{z-1} - \frac{1}{z-b_k}\Big),$$

we have

$$\begin{aligned} e &= \lim_{z\to 0} \frac{z^2(z-1)q'(z)}{z-z_j} = \lim_{z\to 0} \sum_{k=1}^{n-3} \frac{a_k z^2(z-1)q_k'(z)}{z-z_j} \\ &= \lim_{z\to 0} -\sum_{k=1}^{n-3} \frac{a_k b_k(b_k-1)}{(z-z_j)(z-b_k)} = \frac{-1}{z_j}\sum_{k=1}^{n-3} a_k(b_k-1). \end{aligned}$$

Similarly,

$$f = \lim_{z\to 1} -\sum_{k=1}^{n-3} \frac{a_k b_k(b_k-1)}{(z-b_k)(z-z_j)} = \sum_{k=1}^{n-3} \frac{a_k b_k}{1-z_j}.$$

Also,

$$t_k = \lim_{z\to b_k} \frac{z(z-1)(z-b_k)^2 q'(z)}{z-z_j} = -\lim_{z\to b_k} \frac{a_k b_k(b_k-1)}{z-z_j} = \frac{-a_k b_k(b_k-1)}{b_k-z_j}.$$

Finally, s_k is the coefficient of $1/(z-b_k)$ in the partial fraction expansion of

$$\frac{a_k z(z-1)q_k'}{z-z_j} = \frac{-a_k b_k(b_k-1)}{(z-z_j)(z-b_k)}\Big(\frac{1}{z-b_k} + \frac{1}{z} + \frac{1}{z-1}\Big).$$

So

$$s_k = \frac{a_k b_k(b_k-1)}{b_k-z_j}\Big(\frac{1}{b_k-z_j} - \frac{1}{b_k} - \frac{1}{b_k-1}\Big).$$

Now substituting (2) and (3) in (1), we obtain, in $H^1(S)$,

$$\begin{aligned} (4) \quad \frac{z_j(z_j-1)q'(z_j)}{2}\pi^*\Big(\frac{g_j dz}{\sqrt{q}}\Big) &= \sum_{k=1}^{n-3} \frac{\frac{1}{2}s_k - c_k}{b_k(b_k-1)}\pi^*\Big(\frac{q_k dz}{\sqrt{q}}\Big) \\ -\sum_{k=1}^{n-3} \frac{a_k(b_k-1)}{2z_j}\pi^*\Big(&\frac{dz}{z^2(z-1)\sqrt{q}}\Big) + \sum_{k=1}^{n-3} \frac{a_k b_k}{2(1-z_j)}\pi^*\Big(\frac{dz}{(z-1)^2 z\sqrt{q}}\Big) \\ &- \sum_{k=1}^{n-3} \frac{1}{2(b_k-z_j)}\pi^*\Big(\frac{f_k dz}{\sqrt{q}}\Big). \end{aligned}$$

Then using (1) and (2) of 13.13, we obtain

$$\begin{aligned} z_j(z_j-1)q'(z_j)\pi^*\Big(\frac{g_j dz}{\sqrt{q}}\Big) &= \sum_{k=1}^{n-3} \frac{s_k - 2c_k}{b_k(b_k-1)}\pi^*\Big(\frac{q_k dz}{\sqrt{q}}\Big) \\ &+ \frac{1}{1-z_j}\Big(\sum_{k=1}^{n-3} \frac{a_k b_k}{b_k-1}\pi^*\Big(\frac{q_k dz}{\sqrt{q}}\Big) + \sum_{k=1}^{n-3} b_k\pi^*\Big(\frac{f_k dz}{\sqrt{q}}\Big)\Big) \\ + \frac{1}{z_j}\Big(\sum_{k=1}^{n-3}(b_k-1)\pi^*\Big(\frac{f_k dz}{\sqrt{q}}\Big) - \sum_{k=1}^{n-3} &\frac{a_k(1-b_k)}{b_k}\pi^*\Big(\frac{q_k dz}{\sqrt{q}}\Big)\Big) - \sum_{k=1}^{n-3} \frac{1}{b_k-z_j}\pi^*\Big(\frac{f_k dz}{\sqrt{q}}\Big). \end{aligned}$$

Then the coefficient of $\pi^*\Big(\dfrac{f_k dz}{\sqrt{q}}\Big)$ is

$$\frac{b_k}{1-z_j}+\frac{b_k-1}{z_j}-\frac{1}{b_k-z_j}=\frac{-b_k(b_k-1)}{z_j(z_j-1)(b_k-z_j)}.$$

The coefficient of $\pi^*\Big(\dfrac{q_k dz}{\sqrt{q}}\Big)$ is

$$\begin{aligned}
&\frac{s_k-2c_k}{b_k(b_k-1)}+\frac{a_kb_k}{(b_k-1)(1-z_j)}+\frac{a_k(b_k-1)}{z_jb_k}\\
&\qquad=\frac{a_k}{b_k-z_j}\Big(\frac{-1}{b_k-z_j}-\frac{1}{b_k}-\frac{1}{b_k-1}\Big)+\frac{a_kb_k}{(b_k-1)(1-z_j)}+\frac{a_k(b_k-1)}{z_jb_k}\\
&\qquad=\frac{-a_kb_k(b_k-1)}{(b_k-z_j)^2z_j(z_j-1)}.
\end{aligned}$$

□

13.15. Lemma

$$\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{a_jb_j(b_j-1)(b_j-b_k)}=\sum_{\ell=1}^{n-4}\frac{1}{b_j-z_\ell}-\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{1}{b_j-b_k}. \tag{1}$$

Proof. — We claim that both sides of (1) represent

$$\lim_{z\to b_j}\frac{\dfrac{d}{dz}(z(z-1)(z-b_j)q(z))}{z(z-1)(z-b_j)q(z)}.$$

This is clearly true of the righthand side of (1). For the lefthand side, note that

$$\frac{\dfrac{d}{dz}(z(z-1)(z-b_j)q(z))}{z(z-1)(z-b_j)q(z)}=\frac{q'(z)+\Big(\dfrac{1}{z}+\dfrac{1}{z-1}+\dfrac{1}{z-b_j}\Big)q(z)}{q(z)}. \tag{2}$$

Then the limit of (2) as $z\to b_j$ is

$$\begin{aligned}
&\frac{1}{a_j}\lim_{z\to b_j}(z-b_j)\Big(q'(z)+\Big(\frac{1}{z}+\frac{1}{z-1}+\frac{1}{z-b_j}\Big)q(z)\Big)\\
&\qquad=\frac{1}{a_j}\lim_{z\to b_j}\sum_{\substack{k=1\\k\neq j}}^{n-3}a_kq_k(z)=\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{a_jb_j(b_j-1)(b_j-b_k)},
\end{aligned}$$

as required. □

13.16. Lemma. — *In $H^1(S)$,*

$$(1)\quad \pi^*\Big(\frac{3a_jb_j(b_j-1)dz}{2z(z-1)(z-b_j)^3\sqrt{q}}\Big)$$
$$= -\Big(\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{a_jb_j(b_j-1)(b_j-b_k)}+\frac{2b_j-1}{b_j(b_j-1)}\Big)\pi^*\Big(\frac{f_jdz}{\sqrt{q}}\Big)$$
$$+\Big(\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{b_j-b_k}\Big(\frac{1}{b_j-b_k}-\frac{1}{b_j}-\frac{1}{b_j-1}\Big)-\frac{a_j}{b_j(b_j-1)}\Big)\pi^*\Big(\frac{q_jdz}{2\sqrt{q}}\Big)$$
$$-\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{b_k(b_k-1)}{2(b_k-b_j)b_j(b_j-1)}\pi^*\Big(\frac{f_kdz}{\sqrt{q}}\Big)-\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{2(b_k-b_j)^2b_j(b_j-1)}\pi^*\Big(\frac{q_kdz}{\sqrt{q}}\Big).$$

Proof. — We have

$$\frac{d}{dz}\Big(\frac{\sqrt{q}}{z-b_j}\Big)=\frac{-q}{(z-b_j)^2\sqrt{q}}+\frac{q'}{2(z-b_j)\sqrt{q}}$$
$$=-\sum_{k=1}^{n-3}\frac{a_kq_k}{(z-b_j)^2\sqrt{q}}+\sum_{k=1}^{n-3}\frac{a_kq_k'}{2(z-b_j)\sqrt{q}}.$$

So

$$(2)\quad \frac{d}{dz}\Big(\frac{\sqrt{q}}{z-b_j}\Big)+\frac{3a_jb_j(b_j-1)}{2(z-b_j)^3z(z-1)\sqrt{q}}$$
$$=-\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{z(z-1)(z-b_k)(z-b_j)\sqrt{q}}\Big(\frac{1}{z-b_j}+\frac{1}{2(z-b_k)}\Big)$$
$$-\sum_{k=1}^{n-3}\frac{a_kb_k(b_k-1)}{2z(z-1)(z-b_j)(z-b_k)}\Big(\frac{1}{z}+\frac{1}{z-1}\Big).$$

Now, if $k\neq j$,

$$(3)\quad \frac{1}{z(z-b_k)(z-b_j)}=\frac{1}{b_kb_jz}+\frac{1}{b_k(b_k-b_j)(z-b_k)}+\frac{1}{b_j(b_j-b_k)(z-b_j)},$$

$$(4)\quad \frac{1}{(z-1)(z-b_k)(z-b_j)}=\frac{1}{(b_k-1)(b_j-1)(z-1)}+\frac{1}{(b_k-1)(b_k-b_j)(z-b_k)}+\frac{1}{(b_j-1)(b_j-b_k)(z-b_j)},$$

while

$$(5)\quad \frac{1}{z(z-b_j)^2}=\frac{1}{b_j(z-b_j)^2}+\frac{1}{b_j^2z}-\frac{1}{b_j^2(z-b_j)},$$

$$(6)\quad \frac{1}{(z-1)(z-b_j)^2} = \frac{1}{(b_j-1)(z-b_j)^2} + \frac{1}{(b_j-1)^2(z-1)} - \frac{1}{(b_j-1)^2(z-b_j)}.$$

Using (3) to (6), and (2) of 13.12, the righthand side of (2) above becomes

$$\begin{aligned}(7)\quad & \frac{f_j}{\sqrt{q}}\Big(-\sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kb_k(b_k-1)}{a_jb_j(b_j-1)(b_j-b_k)} - \frac{1}{2b_j} - \frac{1}{2(b_j-1)}\Big)\\ & + \frac{q_j}{2\sqrt{q}}\Big(\sum_{k=1}^{n-3}\frac{a_kb_k(b_k-1)}{b_j(b_j-1)(b_j-b_k)}\Big(\frac{1}{b_j-b_k} - \frac{1}{b_j} - \frac{1}{b_j-1}\Big) + \frac{a_j}{b_j^2} + \frac{a_j}{(b_j-1)^2}\Big)\\ & - \sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{a_kq_k}{2(b_k-b_j)\sqrt{q}}\Big(\frac{1}{b_k-b_j} + \frac{1}{b_k} + \frac{1}{b_k-1}\Big) - \sum_{\substack{k=1\\k\neq j}}^{n-3}\frac{f_k}{2(b_k-b_j)\sqrt{q}}\\ & - \sum_{k=1}^{n-3}\frac{a_k}{2z(z-1)\sqrt{q}}\Big(\frac{b_k-1}{b_jz} + \frac{b_k}{(b_j-1)(z-1)}\Big).\end{aligned}$$

Now we use Lemma 13.13 to substitute for the last two terms of (7). The f_k and q_k contributions to the last two terms are, respectively,

$$\Big(\frac{b_k-1}{2b_j} - \frac{b_k}{2(b_j-1)}\Big)\frac{f_k}{\sqrt{q}},\qquad \Big(-\frac{a_k(1-b_k)}{2b_jb_k} - \frac{a_kb_k}{2(b_j-1)(b_k-1)}\Big)\frac{q_k}{\sqrt{q}}.$$

The last two terms give

$$\begin{aligned}& \frac{1}{2b_j}\sum_{k=1}^{n-3}\frac{(b_k-1)f_k}{\sqrt{q}} - \frac{1}{2(b_j-1)}\sum_{k=1}^{n-3}\frac{b_kf_k}{\sqrt{q}}\\ & \qquad - \frac{1}{2b_j}\sum_{k=1}^{n-3}\frac{a_k(1-b_k)}{b_k}\frac{q_k}{\sqrt{q}} - \frac{1}{2(b_j-1)}\sum_{k=1}^{n-3}\frac{a_kb_k}{b_k-1}\frac{q_k}{\sqrt{q}}.\end{aligned}$$

So the coefficients of $\dfrac{f_k}{\sqrt{q}}$ and $\dfrac{q_k}{\sqrt{q}}$ are given respectively by

$$\frac{b_k-1}{2b_j} - \frac{b_k}{2(b_j-1)},\qquad \frac{-a_k(1-b_k)}{b_jb_k} - \frac{a_kb_k}{2(b_j-1)(b_k-1)}.$$

□

CHAPTER 14

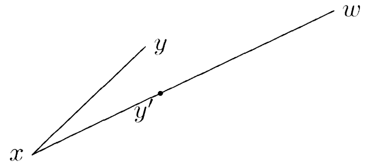
DISTANCE BETWEEN GEODESICS

14.1. Let $\ell_1 = \{x_t : t \in [0, T]\}$ and $\ell_2 = \{y_t : t \in [0, T]\}$ be two geodesic segments in the hyperbolic plane parametrised by length, with $d(x_0, y_0) \leqslant \delta$, $d(x_T, y_T) \leqslant \delta$. Then

$$d(x_t, y_t) \leqslant C\delta e^{-\min(t, T-t)}.$$

Of course, a similar statement holds in hyperbolic space of any dimension. However, we cannot prove such a strong statement in Teichmüller space with the Teichmüller distance d. This is related to the other well-known ways in which Teichmüller space differs from hyperbolic space, such as, for example, there being several different boundaries of Teichmüller space, each of which inherits some, but not all, of the characteristics of the boundary of hyperbolic space. We do, however, have the following. We use the notation $S(\alpha, [\varphi])$, $a(\alpha, q)$, m_α of 9.3, 9.4, 9.1 for quadratic-differential areas and moduli of the appropriate subsurfaces.

14.2. We shall prove the following in 14.9. The diagram shows the situation under consideration.



The geodesic and y.

Proposition. — *Given $\varepsilon_0 > 0$, $M > 0$, there is $M_1 > 0$ such that the following hold. Let $[x, w]$ be a geodesic segment in the Teichmüller space $\mathcal{T}(Y)$, with respect to the Teichmüller metric d. Let $y \in \mathcal{T}(Y)$. Suppose that $y' = [\varphi'] \in [x, w]$ with $d(x, y') = d(x, y)$. Let $q(z)dz^2$ denote the quadratic differential at y' for $d(y', w)$. Let*

$$d(x, y) + d(y, w) - d(x, w) \leqslant M.$$

(1) *Let* $y' \in \mathcal{T}_{\geqslant \varepsilon_0}$. *Then* $d(y, y') \leqslant M_1$.

(2) *Given* $\varepsilon > 0$ *there is* $\varepsilon' > 0$ *such that if* $y \in \mathcal{T}_{<\varepsilon'}$ *then* $y' \in \mathcal{T}_{<\varepsilon}$.

(3) *The following holds given* $C > 0$, *for* M_1 *depending on* ε_0, M *and* C. *Suppose that* $y' \in \mathcal{T}_{<\varepsilon_0}$, Γ *is the set of loops* γ *with* $\varphi'(\gamma)$ *of length* $\leqslant \varepsilon_0$, *and* α *is a loop or gap of* Γ *with* $a(\alpha, q) \geqslant C$ *if* α *is a gap, and* $a(\alpha, q) \geqslant C/m_\alpha(y')$ *if* α *is a loop. Then all geodesics homotopic to* $\varphi'(\partial\alpha)$ *have length* $\leqslant M_1$ *at* y, *and*

$$d_\alpha(y, y') \leqslant M_1.$$

Remark. — It can happen that $y \in \mathcal{T}_{\geqslant \varepsilon_0}$ but $y' \in \mathcal{T}_{<\varepsilon}$, with ε arbitrarily small if $d(x, y)$ is large enough.

14.3. Poincaré Length and a Modification. — Let γ be a loop in $\overline{\mathbf{C}} \setminus Y$. Then we denote by $|\varphi(\gamma)|$ the length of the geodesic homotopic to $\varphi(\gamma)$, with respect to the Poincaré metric on $\overline{\mathbf{C}} \setminus \varphi(Y)$. We normalise so that $\overline{\mathbf{C}} \setminus \varphi(Y)$ has area 1. Strictly speaking, we should use a notation of the form $|\varphi(\gamma)|_{[\varphi]}$, but it should be clear from the context that we are using the Poincaré metric on $\overline{\mathbf{C}} \setminus \varphi(Y)$. Let ρ denote the Poincaré metric. We need to consider another length with respect to another measurable Riemannian metric ρ' on $\overline{\mathbf{C}} \setminus \varphi(Y)$ when studying Teichmüller distance. We fix $\varepsilon_0 \leqslant$ the Margulis constant. We define ρ' to be ρ on $(\overline{\mathbf{C}} \setminus \varphi(Y))_{\geqslant \varepsilon_0}$ and peripheral components of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$. So it remains to define ρ' on a nonperipheral component A of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$. Let $\gamma' \subset \overline{\mathbf{C}} \setminus Y$ be the nontrivial nonperipheral simple loop such that A is homotopic in $\overline{\mathbf{C}} \setminus \varphi(Y)$ to $\varphi(\gamma')$. We map A biholomorphically to an annulus of the form

$$\{x + iy : 0 < y < 1/\sqrt{\varepsilon}\}/(x + iy \sim x + iy + \sqrt{\varepsilon}).$$

Here, ε is, of course, uniquely determined, and is boundedly proportional to the Poincaré length of the geodesic homotopic to $\varphi(\gamma')$. Then we take ρ' to be the image of the standard Euclidean metric $dx^2 + dy^2$ in A.

Let γ be any simple geodesic. Then we define $|\varphi(\gamma)|'$ to be the ρ'-length of the Poincaré geodesic homotopic to $\varphi(\gamma)$. Note that $\varphi(\gamma)$ is bounded from $\varphi(Y)$, so there is no problem with the definition of ρ' near $\varphi(Y)$. Let α be a loop or subsurface such that $\varphi(\alpha)$ is homotopic to a component $S(\alpha)$ of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{\geqslant \varepsilon_0}$ or $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$. We also define $|\varphi(\gamma \cap \alpha)|'$ to be the ρ' length of the geodesic homotopic to $\varphi(\gamma)$ with $S(\alpha)$.

We have

$$\rho \leqslant C\rho', \quad |\varphi(\gamma)| \leqslant C|\varphi(\gamma)|', \quad |\varphi(\gamma \cap \alpha)| \leqslant C|\varphi(\gamma \cap \alpha)|' \tag{1}$$

for a constant C. The reverse inequality does not hold in general. This is because the shortest Poincaré length of a path between the components of ∂A is $O(\log(1/\varepsilon)$, while for ρ' it is $O(1/\sqrt{\varepsilon})$. However, we obviously have $\rho' = \rho$ if $[\varphi] \in \mathcal{T}_{\geqslant \varepsilon_0}$. We have $|\varphi(\gamma)| = |\varphi(\gamma)|'$ whenever either of these quantities is bounded.

14.4. The metric ρ' and Teichmüller distance. — We can use $|\cdot|'$ to estimate Teichmüller distance. Let $\gamma \subset \overline{\mathbf{C}} \setminus Y$ be any simple loop. This ensures that, for a suitable $\varepsilon_0 > 0$, and any $[\varphi]$, the geodesic homotopic to $\varphi(\gamma)$ does not intersect peripheral components of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$. Note that $1/|\varphi(\gamma)|'$ is, to within a bounded constant depending on ε_0, the maximum modulus of an annulus in $\overline{\mathbf{C}} \setminus \varphi(Y)$ homotopic to $\varphi(\gamma)$. Then there is a constant $C > 0$ (depending on ε_0) such that the following holds. Let $\chi : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ with $\|\chi\|_{\mathrm{qc}} = K$ (8.1), and let $[\varphi] \in \mathcal{T}(Y)$. Then

$$|\chi \circ \varphi(\gamma)|' \leqslant C\sqrt{K}|\varphi(\gamma)|'.$$

So if χ minimizes distortion up to isotopy constant on $\varphi(Y)$, we obtain

$$\frac{|\chi \circ \varphi(\gamma)|'}{|\varphi(\gamma)|'} \leqslant Ce^{d([\varphi],[\chi\circ\varphi])}. \tag{1}$$

We shall see below (in 14.7) that we have a reverse inequality also. In fact, we shall see that, for a suitable constant C we can find γ such that

$$\frac{e^{d([\varphi],[\chi\circ\varphi])}}{C} \leqslant \frac{|\chi \circ \varphi(\gamma)|'}{|\varphi(\gamma)|'}. \tag{2}$$

14.5. Q-d length. — We need to recall some facts about length of loops with respect to the singular metrics induced by quadratic differentials. We consider $\mathcal{T}(Y)$. Let $q(z)dz^2$ be a quadratic differential at $[\varphi]$, let $\{[\chi_t \circ \varphi] : t \in [0,\infty)\}$ be the half-geodesic determined by $q(z)dz^2$, and let $\mathcal{F}_\pm$ be the corresponding foliations, so that χ_t multiplies length on leaves of $\mathcal{F}_\pm$ by $e^{\pm t}$. Let γ be any simple nontrivial nonperipheral loop in $\overline{\mathbf{C}} \setminus Y$. Then $\varphi(\gamma)$ can be chosen up to a homotopy which is a limit of isotopies to be a union of segments which are each at constant angle to $\mathcal{F}_\pm$. If there is more than one segment, they can be chosen so that any two are between zeros of q or points of $\varphi(Y)$, with angle $\geqslant \pi$ at any zero outside $\varphi(Y)$. We shall then say that $\varphi(\gamma)$ is *in good position*. Good position is unique, up to reparametrisation. Moreover, no two segments intersect transversally, and the number of isotopically distinct segments is $\leqslant \#(Y) - 3$. We then define $|\varphi(\gamma)|_q$ to be the sum of the lengths of the segments with respect to the singular metric $\rho_q = |q(x+iy)|(dx^2+dy^2)$. Then $|\varphi(\gamma)|_q$ depends only on the isotopy class of γ. Then for a constant C depending only on ε_0,

$$\rho_q \leqslant C\rho', \quad |\varphi(\gamma)|_q \leqslant C|\varphi(\gamma)|'. \tag{1}$$

This follows from simple analysis of the metric ρ_q. For suitable C depending on ε_0, we have $\rho_q \leqslant C\rho'$ (equivalently $\rho_q \leqslant C\rho$) on any component of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{\geqslant\varepsilon_0}$. For a nonperipheral component A of $(\overline{\mathbf{C}} \setminus \varphi(Y))_{<\varepsilon_0}$, if we use the biholomorphic equivalence

$$\varphi : A \longrightarrow \{x+iy : 0 < y < 1/\sqrt{\varepsilon}\}/(x+iy \sim x+iy+\sqrt{\varepsilon}\}$$

then we see that $\varphi_*\rho_q \leqslant dx^2+dy^2$, and thus $\rho_q \leqslant \rho_2$ on A, which yields $\rho_q \leqslant C\rho'$. We are using the fact that the ρ_q-area of $\overline{\mathbf{C}} \setminus \varphi(Y)$ is 1, that is, $\int |q| = 1$.

Let $[\varphi]$, q, α be such that $a(\alpha,q) \geqslant C_0$ if α is a gap, and $a(\alpha,q)m_\alpha([\varphi]) \geqslant C_0$ if α is a loop. Let $[\varphi] \in (\mathcal{T}(A(\alpha))_{\geqslant\varepsilon_0}$ if α is a gap. We can always find at least one α

such that these conditions hold, for ε_0 and C_0 sufficiently small (depending only on $\#(Y)$). Then for a constant $C_1 = C_1(C_0, \varepsilon_0)$, and any loop γ in α

$$|\varphi(\gamma)| = |\varphi(\gamma)|' \leqslant C_1|\varphi(\gamma)|_q. \tag{2}$$

This holds whether α is a loop or a gap. It follows immediately from the definition of ρ_q. To see this in the case when α is a loop, note that the Pole-Zero Condition of 9.4 must be satisfied for $S(\alpha, [\varphi], \varepsilon)$ for all sufficiently small ε, depending on C_0. By (1) and (2), ρ_q is then boundedly proportional to ρ' on $S(\alpha, [\varphi], \varepsilon_0)$, with bound depending on C_0.

If γ is a closed loop in $\overline{\mathbf{C}} \setminus Y$ and α is a subsurface of $\overline{\mathbf{C}} \setminus Y$, then we also define $|\varphi(\gamma \cap \alpha)|_q$. To do this, we simply put $\varphi(\alpha)$ and $\varphi(\gamma)$ in good position to compute $|\varphi(\gamma \cap \alpha)|_q$.

14.6. The Q-d Length function along a geodesic. — Let $\{[\chi_t \circ \varphi] : t \in [0, T]\}$ denote the geodesic in $\mathcal{T}(Y)$ joining $[\varphi]$ and $[\chi \circ \varphi]$. Let $\chi_0 =$ identity, $\chi_T = \chi$, $d([\varphi], [\chi_t \circ \varphi]) = t$. Let $q_t(z)dz^2$ be the quadratic differential for this geodesic at $[\chi_t \circ \varphi]$, with $q_0 = q$. Let $\gamma \subset \overline{\mathbf{C}} \setminus Y$ be any simple loop. Then $|\chi_t \circ \varphi(\gamma)|_{q_t}$ is relatively easily computed, because if γ_1 is a segment at angle θ with $\mathcal{F}_+$ and with $|\gamma_1|_q = L$,

$$|\chi_t \circ \varphi(\gamma_1)|_{q_t} = L\sqrt{e^{2t}\cos^2\theta + e^{-2t}\sin^2\theta}.$$

So

$$|\chi_t \circ \varphi(\gamma)| = \sum_{j=1}^{r} \sqrt{a_j^2 e^{2t} + b_j^2 e^{-2t}}$$

for suitable a_j, $b_j \geqslant 0$ and r. In fact, $r \leqslant \#(Y) - 3$, but this does not matter. Write

$$a = \sum_{j=1}^{r} a_j, \quad b = \sum_{j=1}^{r} b_j.$$

Then we have

$$\frac{ae^t + be^{-t}}{2} \leqslant |\chi_t \circ \varphi(\gamma)|_{q_t} \leqslant \sqrt{2}(ae^t + be^{-t}).$$

It follows that $t \mapsto |\chi_t \circ \varphi(\gamma)|_{q_t}$ has at most one essential minimum in the following sense. There is a constant C such that the following holds. Given γ, there is $s \in [0, T]$ such that

$$e^{|t-s|}|\chi_s \circ \psi(\gamma)|_{q_s} \leqslant |\chi_t \circ \varphi(\gamma)|_{q_t} \leqslant Ce^{|t-s|}|\chi_s \circ \varphi(\gamma)|_{q_s} \quad \text{for } t \in [0, T]. \tag{1}$$

14.7. The upper bound on Teichmüller distance in terms of Q-d length and modified Poincaré length. — We are going to use (1) of 14.6 and (1) of 14.5 to obtain the upper bound (2) of 14.4 on Teichmüller distance. If γ is such that the essential minimum $s = 0$ or s is bounded, then we have, for suitable C',

$$e^{d([\varphi], [\chi \circ \varphi])} \leqslant C' \frac{|\chi \circ \varphi(\gamma)|_{q_T}}{|\varphi(\gamma)|_q}.$$

If γ also satisfies (1) of 14.5, then we obtain

$$e^{d([\varphi],[\chi\circ\varphi])} \leqslant C'' \frac{|\chi\circ\varphi(\gamma)|'}{|\varphi(\gamma)|'},$$

which gives (2) of 14.4 (with C'' replacing C) as claimed. So it remains to find such a γ. As we have seen, if S is a component of $(\overline{\mathbf{C}} \smallsetminus \varphi(Y))_{\geqslant\varepsilon_0}$ or a nonperipheral component of $(\overline{\mathbf{C}} \smallsetminus \varphi(Y))_{<\varepsilon_0}$, of q-area bounded from 0, and satisfying the Pole-Zero Condition of 9.4 if S is an annulus, then (1) of 14.5 holds for any simple loop γ with $|\varphi(\gamma)|'/|\varphi(\gamma)\cap S|'$ bounded. So we need to find such a loop for which the essential minimum $s=0$.

First, let S be a component of $(\overline{\mathbf{C}} \smallsetminus \varphi(Y))_{\geqslant\varepsilon_0}$ homotopic to $\varphi(\alpha)$ for a subsurface α of $\overline{\mathbf{C}} \smallsetminus Y$. Consider a set Γ of $\gamma \subset \alpha$ such that $\alpha \smallsetminus (Y \cup (\cup\Gamma)$ is a union of topological discs with at most one puncture and annuli parallel to $\partial\alpha$, and such that $|\varphi(\gamma)|$ is bounded for $\gamma\in\Gamma$. If $|\chi_t\circ\varphi(\gamma)|_{q_t} < \varepsilon_1$ for some t, then this happens for some $t\leqslant T_1$, for T_1 depending on ε_1. Then $|\chi_t\circ\varphi(\gamma)| < C\varepsilon_1$. If this is true for all loops adjacent to a component of $\alpha \smallsetminus (Y\cup(\cup\Gamma))$, then the area of that component is $O(\varepsilon_1^2)$. So if ε_1 is sufficiently small, given the q-area of α, for each t, there is at least one loop $\gamma\in\Gamma$ with $|\varphi_t(\gamma)|_{q_t} \geqslant \varepsilon_1$. So there is ε_1' (depending on ε_1) and $\gamma\in\Gamma$ such that $|\varphi_t(\gamma)|_{q_t} \geqslant \varepsilon_1'$ for all t. So then the essential minimum for γ is within a bounded distance — depending on ε_0 — of 0.

Now let S be a nonperipheral component of $(\overline{\mathbf{C}} \smallsetminus \varphi(Y))_{<\varepsilon_0}$. Let S be homotopic to $\varphi(\gamma')$. Fix γ_0 such that $|\varphi(\gamma_0)\smallsetminus S|$ is bounded and γ_0 intersects γ' exactly twice. Let γ_n ($n\in\mathbf{Z}$) be obtained from by Dehn twist n times round γ'. Then we have the following formula for $|\chi_t\circ\varphi(\gamma_n)|_{q_t}$.

$$|\chi_t\circ\varphi(\gamma_n)|_{q_t} = \sqrt{(a_1+nc)^2e^{2t}+(b_1+nd)e^{-2t}} + \sum_{j=2}^{r}\sqrt{a_j^2e^{2t}+b_j^2e^{-2t}}$$

where

$$|\chi_t\circ\varphi(\gamma')|_{q_t} = \sqrt{c^2e^{2t}+d^2e^{-2t}},$$
$$|a_j|+|b_j| \leqslant C(|a_1|+|b_1|),\quad |c|+|d|\leqslant C,\quad 1/C\leqslant a_1d-b_1c\leqslant 1.$$

So either $|c|>|d|C^3$ or $|c|\leqslant |d|/C^3$, and we can choose n so that $|b_1+nd|\leqslant |d|/2$. Then

$$|(b_1+nd)c| \leqslant \frac{d^2}{2C^3} \leqslant \frac{1}{2C}$$
$$|a_1+nd| \geqslant \frac{1}{2C|d|} \geqslant \frac{1}{2C^2} \geqslant \frac{|b_1+nd|}{C}.$$

It follows that 0 is an essential minimum for either γ' or for γ_n, for some n.

14.8 Stable and Unstable lengths. — Let $[\varphi]\in\mathcal{T}(Y)$ and let $q(z)dz^2$ be a quadratic differential at $[\varphi]$, with corresponding expanding and contracting foliations $\mathcal{F}_+$, $\mathcal{F}_-$. At each nonsingular point, the tangent space of the surface $\overline{\mathbf{C}}\smallsetminus\varphi(Y)$ is the direct sum of the tangent spaces of $\mathcal{F}_{\pm}$, and we have corresponding seminorms $\|\cdot\|_{\pm}$ on tangent vectors. Again, let $\varphi(\gamma)$ be in good position. Let $D\varphi(\gamma)(t)$ define the derivative of $\varphi(\gamma)$ at t. Then we define

$$|\varphi(\gamma)|_{+}=\int_0^t \|D\varphi(\gamma)(u)\|_{+}du,$$

and similarly for $|\varphi(\gamma)|_{-}$. Again, this definition depends only on the isotopy class of γ, and good position gives the minimum value of the integral, in the homotopy class of $\varphi(\gamma)$. We shall write $|\varphi(\gamma)|_{q,\pm}$ if we are considering more than one quadratic differential.

In the notation of [**F-L-P**], we have

$$|\varphi(\gamma)|_{+}=i(\gamma,\mathcal{F}_{-}),$$

where i is the *intersection function*. We shall need, in Chapter 25, to use the intersection number between two measured foliations. If $\mathcal{F}$ and $\mathcal{G}$ are measured foliations on surfaces S and S', and $\chi: S\to S'$ is a homeomorphism, then $i(\chi\mathcal{F},\mathcal{G})$ is a continuous function. Indeed, it is Lipschitz [**R1**] with respect to a natural metric. It can be defined as $\lim_{n\to\infty} i(\mathcal{F}_n,\mathcal{G})$ by taking any limit of a sequence $\{\mathcal{F}_n\}$ of measured foliations supported on simple loops [**F-L-P**], [**R1**].

As for $|\cdot|_q$, we can define $|\varphi(\gamma\cap\alpha)|_{\pm}$ if α is a subsurface of $\overline{\mathbf{C}}\setminus Y$.

14.9. Proof of 14.2. — Let $y'\in[x,w]$ with $d(x,y')=d(x,y)=t$. Write

$$\{[\chi_s\circ\varphi]: s\in[0,T]\}=[x,w],\ x=[\varphi],\ y'=[\chi_t\circ\varphi],\ w=[\chi_T\circ\varphi]=[\xi],\ y=[\psi],$$

and let $q_s(z)dz^2$, be the quadratic differential at $[\chi_s\circ\varphi]$ for $d([\chi_s\circ\varphi],[\chi_T\circ\varphi])$.

Let $y'\in\mathcal{T}_{\geqslant\varepsilon_0}$, in which case take $\Gamma=\varnothing$, or let Γ be the set of loops such that $\psi(\gamma)$ has length $<\varepsilon_0$. If $\Gamma=\varnothing$, let $\alpha=\overline{\mathbf{C}}$. If $\Gamma\neq\varnothing$, let α be a loop or gap of Γ with $a(\alpha,q_t)\geqslant C_1$ if α is a gap, or $a(\alpha,q_t)\geqslant C_1/m_\alpha(y')$ if α is a loop. We call such an α a *good component*. Then we claim that, for any nontrivial nonperipheral loop γ with nontrivial intersection with α, either

$$\frac{e^{d(x,y')}}{C}\leqslant\frac{|\varphi(\gamma)|'}{|\chi_t\circ\varphi(\gamma\cap\alpha)|'},\tag{1}$$

or similarly for w, $\chi_T\circ\varphi$ replacing x, φ. If $\alpha\in\Gamma$ we we claim that, further, either (1) or its equivalent for w holds for $y'=[\chi_t\circ\varphi]$ replaced by any $y_1=[\chi_s\circ\varphi]$ in $[x,w]\cap\mathcal{T}(\alpha,\varepsilon_0)$. To see this, we consider the function $s\mapsto|\chi_s\circ\varphi(\gamma)|_{q_s}$ for any simple loop γ. Then because this function has just one essential minimum (14.6) we have either

$$\frac{e^t}{C}\leqslant\frac{|\varphi(\gamma)|_{q_0}}{|\chi_t\circ\varphi(\gamma)|_{q_t}}$$

or

$$\frac{e^{2(T-t)}}{C}\leqslant\frac{|\chi_T\circ\varphi(\gamma)|_{q_T}}{|\chi_t\circ\varphi(\gamma)|_{q_t}}.$$

But we also have, by (1) of 14.5,

$$|\varphi(\gamma)|_{q_0} \leqslant C|\varphi(\gamma)|', \quad |\chi_T \circ \varphi(\gamma)| \leqslant C|\chi_T \circ \varphi(\gamma)|'.$$

This allows us to replace $|\cdot|_{q_0}$, $|\cdot|_{q_T}$ by $|\cdot|'$. The area lower bound gives that ρ_{q_t} is boundedly proportional to ρ' on $S(\alpha)$. If α is a loop, then ρ_{q_s} is boundedly proportional to ρ' on $S(\alpha)$ for all $[\chi_s \circ \varphi] \in \mathcal{T}(\alpha, \varepsilon_0)$. This allows us to replace $|\chi_t \circ \varphi(\gamma)|_{q_t}$ by $|\chi_t \circ \varphi(\gamma \cap \alpha)|'$

We claim that, for a constant C_1 depending on M, and any good component α, and any nontrivial nonperipheral any loop γ intersecting α nontrivially,

$$|\chi_t \circ \varphi(\gamma \cap \alpha)|' \leqslant C_1|\psi(\gamma)|'. \tag{2}$$

This would be enough to prove 1 of 14.2 (with $\alpha = \overline{\mathbf{C}}$) and most of 3 of 14.2. To see (2): by (1) of 14.4,

$$\frac{|\varphi(\gamma)|'}{|\psi(\gamma)|'} \leqslant Ce^{d([\varphi],[\psi])} = Ce^t = Ce^{d(x,y')},$$

$$\frac{|\chi_T \circ \varphi(\gamma)|'}{|\psi(\gamma)|'} \leqslant Ce^{d([\psi],\chi_T\circ\varphi])} \leqslant Ce^M e^{(T-t)} = Ce^M e^{d(x,y')}.$$

Then (2) follows from (1) above.

If $|\psi(\gamma)|'$ (equivalently $|\psi(\gamma)|$) is sufficiently small then we deduce from (2) that γ has no essential intersections with α. This proves 2 of Proposition 14.2, namely, that for suitable ε_0', if $y = [\psi] \in \mathcal{T}_{<\varepsilon_0'}$ then $y' = [\chi_t \circ \varphi] \in \mathcal{T}_{<\varepsilon_0}$. We also see from (2) that, if $\psi(\gamma)$ is bounded, then $|\chi_t \circ \varphi(\gamma \cap \alpha)|$ is bounded and hence so is $\#(\partial\alpha \cap \gamma)$. Then we can homotope $\partial\alpha$ into a finite union of segments of loops γ such that $|\psi(\gamma)|$ is bounded. Hence $|\psi(\partial\alpha)|$ is bounded. This gives part of 3 of Proposition 14.2.

Now we need to consider two cases of α. First, suppose that α is not an annulus. Then we can find a set of loops Γ in α such that all components of $\alpha \smallsetminus (\cup\Gamma \cup Y)$ are at-most-once-punctured discs or annuli parallel to α, and with $|\psi(\gamma)|$ bounded for $\gamma \in \Gamma$ (but also, as we have seen, bounded from 0). Then we also have, by (2), that $|\chi_t \circ \varphi(\gamma)|$ is bounded for all $\gamma \in \Gamma$, which gives the bound on $d_\alpha(y, y')$. Now let α be a loop. Let $[y_1', y_2']$ be the maximal interval on $[x, w]$ such that $z \in \mathcal{T}(\alpha, \varepsilon_0)$ for $z \in [y_1', y_2']$. Let $y_1 \in [x, y]$ and $y_2 \in [y, w]$ with $d(x, y_i) = d(x, y_i')$. What we have proved above for y and y' holds equally well for y_i and y_i', because the triangular equality still holds. Write $y_i = [\psi_i]$. If $y_i = x$ or w, then $d_\alpha(y_i, y_i')$ is bounded. Then suppose that $y_i \neq x$ or w. Then $|\psi_i(\alpha)|$ is bounded above. Then there is a subsurface $\beta_i \supset \alpha$ such that $d_{\beta_i}(y_i, y_i')$ is bounded (in terms of M). It follows that $d_\alpha(y_i, y_i')$ is bounded for $i = 1$, 2. Since $(\mathcal{T}(A(\alpha)), d_\alpha)$ is isometric to the upper half plane with the Poincaré metric, we deduce that $d_\alpha(y, y')$ is bounded. □

14.10. The function d_α'. — This is a concept which will be needed in Chapters 25-31, and also in 14.13 below.

Let $x = [\varphi] \in \mathcal{T}(\partial\alpha, \varepsilon_0)$ and $y = [\psi] \in \mathcal{T}(Z)$. Let $x \in (\mathcal{T}(A(\alpha))_{\geqslant\varepsilon_0}$ if α is a gap. If α is a gap, we choose a set of loops Γ_α, as follows. Take Γ_α to be a set of nontrivial nonperipheral loops in α with $|\varphi(\Gamma_\alpha)| \leqslant C(\varepsilon_0)$ (for a suitable constant $C(\varepsilon_0)$) and such that every component of $\overline{\mathbf{C}} \smallsetminus (\cup\Gamma_\alpha \cup Z)$ is a topological disc with at most one puncture or an annulus parallel to the boundary. We then define

$$d'_\alpha(x,y) = \sup\{|\log(|\psi(\gamma)|' + 1) - \log(|\varphi(\gamma)|'| + 1)| : \gamma \in \Gamma_\alpha\}.$$

Of course, this definition depends on the choice of Γ_α, but only to within a bounded constant, depending on ε_0. If α is a loop, we choose x' nearest to y on the geodesic $[x,y]$ with $x' \in \mathcal{T}(\alpha, \varepsilon_0)$. Then we define

$$d'_\alpha(x,y) = d_\alpha(x,x') + |\log(|\psi(\alpha)|' + 1) - \log(|\varphi'(\alpha)|'| + 1)|.$$

By 14.4 and 14.5 (applied to $\mathcal{T}(A(\alpha))$), this differs by a bounded constant from $d_\alpha(x,y)$, if $y \in \mathcal{T}(\partial\alpha, \varepsilon_0)$ also. Now let β be a loop or gap, and let $S(\beta)$ be as in 9.3. If $y \in \mathcal{T}(\partial\beta, \varepsilon_0)$, we can also define, if α is a gap,

$$d'_{\alpha,\beta}(x,y) = \sup\{|\log(|\psi(\gamma \cap S(\beta))|' + 1) - \log(|\varphi(\gamma)|' + 1)| : \gamma \in \Gamma_\alpha\}.$$

To measure $|\psi(\gamma\cap S(\beta))|'$, we choose ψ so that any component of $\psi(\gamma)\cap(\overline{\mathbf{C}}\smallsetminus\psi(Z))_{\geqslant\varepsilon_0}$ intersects the boundary of $(\overline{\mathbf{C}} \smallsetminus \psi(Z))_{\geqslant\varepsilon_0}$ perpendicularly. If α is a loop, let x' be as above. We define

$$d'_{\alpha,\beta}(x,y) = d_\alpha(x,x') + |\log(|\psi(\alpha \cap S(\beta))|' + 1) - \log(|\varphi'(\alpha)|' + 1)|.$$

Thus,

$$d'_\alpha(x,y) = d'_{\alpha,\overline{\mathbf{C}}}(x,y).$$

14.11. Lemma. — *For a constant C,*

$$|d'_{\alpha,\beta}(x,y) - d'_{\beta,\alpha}(y,x)| \leqslant C.$$

Remark. — This can be applied, in particular, with $\beta = \overline{\mathbf{C}}$.

Proof. — If α and β are gaps then this follows immediately from the fact that $\exp(d'_{\alpha,\beta}(x,y))$ and $\exp(d'_{\beta,\alpha}(y,x))$ are both boundedly proportional to $\#(\Gamma_\alpha\cap\Gamma_\beta)$. If α is a loop, and β is not, then both $\exp(d'_{\alpha,\beta}(x,y))$ and $\exp(d'_{\beta,\alpha}(y,x))$ are boundedly proportional to $\#(\alpha \cap \Gamma_\beta)\exp d_\alpha(x,x')$ for x' as defined above. If both α and β are loops then both $\exp(d'_{\alpha,\beta}(x,y))$ and $\exp(d'_{\beta,\alpha}(y,x))$ are boundedly proportional to

$$(\#(\alpha \cap \beta) + 1)\exp(d_\alpha(x,x') + d_\beta(y,y'))$$

for y' similarly defined. □

14.12. Comparing Geodesics in Two Teichmüller Spaces. — Throughout this work, we are interested in the relationship between Teichmüller spaces $\mathcal{T}(Y)$, $\mathcal{T}(Z)$ with $Z \subset Y$, $\#(Y \setminus Z) = 1$. The following lemma is concerned with this set-up.

Lemma. — *Let $Y = Z \cup \{v\}$. Given $\varepsilon_0 > 0$ there is C such that the following holds. Let $[\varphi_1]$, $[\varphi_2] \in \mathcal{T}(Y)$. Let $q(z)dz^2$ be the quadratic differential for $d_Z([\varphi_1], [\varphi_2])$ at $[\varphi_1]$, and $|\cdot|_+$ the corresponding measure of unstable length. Let $[\varphi_2]_Z = [\chi \circ \varphi_1]_Z$, where χ minimizes distortion up to isotopy constant on $\varphi_1(Z)$. Then*

$$[\varphi_2]_Y = [\sigma_{\gamma_1} \circ \sigma_{\gamma_2} \circ \sigma_{\gamma_3} \circ \chi \circ \varphi_1]_Y = [\chi \circ \sigma_\beta \circ \varphi_1]_Y,$$

where the paths γ_j have the following properties:

$$|\gamma_1|' \leqslant Ce^{d_Y([\varphi_1],[\varphi_2])}, \quad |\gamma_3|' \leqslant Ce^{d_Z([\varphi_1],[\varphi_2])}.$$

The path γ_2 is trivial unless $[\varphi_1]$, $[\varphi_2] \in \mathcal{T}(\eta, \varepsilon_0)$ for a loop η such that $\varphi_1(v) \in S(\eta, [\varphi_1], \varepsilon_0)$. In that case, γ_2 is a single segment in $S(\eta, [\varphi_2], \varepsilon_0)$.

Proof. — We consider the surface $\overline{\mathbf{C}} \setminus \varphi_1(Y)$. If $[\varphi_1] \in (\mathcal{T}(Z))_{\geqslant \varepsilon_0}$, let α be a path joining v to z for some $z \in Z$ such that $|\varphi_1(\alpha)|$ is bounded. Now suppose that $[\varphi_1] \in (\mathcal{T}(Z))_{<\varepsilon_0}$ and that $\varphi_1(v) \in S(\zeta, [\varphi_1], \varepsilon_0) \cup S(\eta, [\varphi_1], \varepsilon_0)$ for a gap ζ, and a loop η adjacent to ζ. Suppose that ζ contains at least one point z of Z. Then we choose a path $\alpha = \alpha_\zeta$ joining v and z such that

$$\varphi_1(\alpha) \subset S(\zeta, [\varphi_1], \varepsilon_0) \cup S(\alpha, [\varphi_1], \varepsilon_0),$$

$|\varphi_1(\alpha) \cap S(\zeta, [\varphi_1], \varepsilon_0)|$ is bounded, $\varphi_1(\alpha)$ has at most one component of intersection with $S(\eta, [\varphi_1], \varepsilon_0)$, and then only if $\varphi_1(v) \in S(\eta, [\varphi_1], \varepsilon_0)$. Now suppose that ζ does not contain a point of Z. Let N be the modulus of the Euler characteristic of ζ. Decompose ζ into $\leqslant 2N$ cells, using a set Γ of arcs in ζ with endpoints on $\partial\zeta$, and such that each component of $\partial\zeta$ is some $\gamma \in \Gamma$, up to homotopy. Choose these so that $|\varphi_1(\Gamma)|$, is bounded, and $\varphi_1(\gamma)$ is in good position with respect to the quadratic differential for $d_Z([\varphi_1], [\varphi_2])$ for all $\gamma \in \Gamma$. Then choose α with one endpoint at v and the other in some disc bounded by $\cup\Gamma$, and otherwise to have the same properties as before.

First, we assume that either $[\varphi_1] \in \mathcal{T}_{\geqslant \varepsilon_0}$ or $v \in \zeta$ for a gap ζ. By 14.4, both

$$|\varphi_2(\alpha)|' = O(e^{d_Y([\varphi_1],[\varphi_2])}), \quad |\chi \circ \varphi_1(\alpha)|' = O(e^{d_Z([\varphi_1],[\varphi_2])}),$$

and similarly for Γ, where defined. If α ends at a point of Z, we deduce that $\gamma = \varphi_2(\alpha) * \overline{\chi \circ \varphi_1(\alpha)}$ also satisfies

$$|\gamma|' = O(e^{d_Y([\varphi_1],[\varphi_2])}),$$

as required. If Γ is defined then the sets $\varphi_2(\Gamma)$ and $\chi \circ \varphi_1(\Gamma)$ intersect. It follows that there is an isotopy mapping $\varphi_2(\Gamma)$ to $\chi \circ \varphi_1(\Gamma)$, and $\varphi_2(\partial\zeta)$ to $\chi \circ \varphi_1(\partial\zeta)$, which moves all points distance $O(e^{d_Y([\varphi_1],[\varphi_2])})$. We can then improve this isotopy further to map $\varphi_2(v)$ to $\chi \circ \varphi_1(v)$, and the distance moved is again $O(e^{d_Y([\varphi_1],[\varphi_2])})$. So then take $\gamma_3 = \chi \circ \varphi_1(\alpha)$ and γ_1 to be the union of $\varphi_2(\alpha)$ and a homotopy arc between $\varphi_2(\Gamma)$ and $\chi \circ \varphi_1(\Gamma)$. Then γ_1 and γ_3 have the required properties.

Now we suppose that $\varphi_1(v) \in S(\eta, [\varphi_1], \varepsilon_0)$ for a loop η. Let η_1, η_2 be the loops in $\overline{\mathbf{C}} \setminus Y$ such that $[\varphi_1] \in \mathcal{T}(\eta_1, \varepsilon_0') \cap \mathcal{T}(\eta_2, \varepsilon_0')$: this is, in fact, true for $\varepsilon_0' = 2\varepsilon_0(1+o(1))$. First, suppose that $[\varphi_2]_Z \notin \mathcal{T}(\eta, \varepsilon_0)$. Then, by numbering η_1 and η_2 suitably, there is a first point $[\varphi_1']$ on the geodesic in $\mathcal{T}(Y)$ joining $[\varphi_1]$ to $[\varphi_2]$, such that $[\varphi_1'] \notin \mathcal{T}(\eta_1, \varepsilon_0')$. Assume without loss of generality that ζ (used in the definition of α) is adjacent to η_1. Then

$$|\varphi_1'(\alpha)|' \leqslant C_0 e^{d_Y([\varphi_1], \varphi_1'])}.$$

Let $[\varphi_1'']$ be the first point on the geodesic joining $[\varphi_1]$ and $[\varphi_2]$in $\mathcal{T}(Z)$ with $[\varphi_1''] \notin \mathcal{T}(\eta, \varepsilon_0)$. Then

$$|\varphi_1''(\alpha)|' \leqslant C_0 e^{d_Z([\varphi_1], [\varphi_1''])}.$$

and hence

$$|\varphi_2(\alpha)|' \leqslant C_1 e^{d_Y([\varphi_1], [\varphi_2])}, \quad |\chi \circ \varphi_1(\alpha)|' \leqslant C_1 e^{d_Z([\varphi_1], [\varphi_2])}.$$

Then again let $\gamma_3 = \chi \circ \varphi_1(\alpha)$ and let γ_1 be the union of $\varphi_2(\alpha)$ and an arc joining $\varphi_2(\Gamma)$ and $\chi \circ \varphi_1(\Gamma)$ if necessary. Then γ_1 and γ_3 have the required properties.

Now let $[\varphi_2] \in \mathcal{T}(\eta, \varepsilon_0)$. A point $[\varphi_1']$ on the geodesic in $\mathcal{T}(Y)$ joining $[\varphi_1]$ and $[\varphi_2]$ as above might still exist, in which case we still have a bound as above on $|\varphi_2(\alpha)|'$. Let $[\varphi_1'']$ be the first point on the geodesic joining $[\varphi_1]$ and $[\varphi_2]$ in $\mathcal{T}(Z)$, if it exists, with $\varphi_1''(v) \in S(\zeta', [\varphi_1''], \varepsilon_0)$ for a gap ζ'. If $\zeta' = \zeta$, then we can bound $|\varphi_1''(\alpha)|'$ and $|\chi \circ \varphi_1(\alpha)|'$ as before, and can define γ_1 and γ_3 as before. If $\zeta' \neq \zeta$, then ζ' is the other gap adjacent to η, then we write $\alpha = \alpha_1' \cup \alpha_2' \cup \alpha_3'$ where we have

$$|\varphi_1''(\alpha_j')| \leqslant C e^{d_Z([\varphi_1], [\varphi_1''])}$$

for $j = 1$, 3, and $\varphi_1''(\alpha)$ is a single arc in the component of $(\overline{\mathbf{C}} \setminus \varphi_1''(Z))_{<\varepsilon_0}$ homotopic to $\varphi_1''(\eta)$. Then up to homotopy we can assume that $\chi \circ \varphi_1(\alpha_2')$ is a single arc in the component of $(\overline{\mathbf{C}} \setminus \varphi_2(Z))_{<\varepsilon_0}$ homotopic to $\varphi_2(\eta)$ and for $j = 1$, 3,

$$|\varphi_2(\alpha_j')|' \leqslant C e^{d_Z([\varphi_1], [\varphi_2])}.$$

This gives the required estimates. □

14.13. Another Comparison. — Here is another lemma similar to the last, which is, again, a result about comparing certain paths in $\mathcal{T}(Y)$ with geodesics in $\mathcal{T}(Z)$ with the same endpoints. This will be needed in Chapters 25-30.

Lemma. — *Let $[\varphi_0] \in \mathcal{T}(Y)$ and let $[\varphi_1] = [\chi_1 \circ \varphi_0]$, $[\varphi_2] = [\chi_2 \circ \varphi_1] \in \mathcal{T}(Y)$, where χ_1, χ_2 minimize distortion up to isotopies constant on $\varphi_0(Z)$, $\varphi_1(Z)$ respectively. Let $[\varphi_2] = [\chi \circ \varphi_0]$, where χ minimizes distortion up to isotopy constant on $\varphi_0(Z)$. Then*

$$[\varphi_2] = [\sigma_\gamma \circ \chi \circ \varphi_0]_Y$$

where $\gamma = \gamma_5 * \cdots \gamma_1$ *and the following properties hold for the paths* γ_i*:*

$$|\gamma_1|' \leqslant Ce^{d_Z([\varphi_0],[\varphi_2])},$$
$$|\gamma_3|' \leqslant C(e^{d_Z([\varphi_0],[\varphi_2])} + e^{d_Z([\varphi_1],[\varphi_2])}),$$
$$|\gamma_5|' \leqslant Ce^{d_Z([\varphi_1],[\varphi_2])}.$$

The path γ_2 *is trivial unless* $\varphi_0(v) \in S(\eta_0, [\varphi_0], \varepsilon_0)$ *for a loop* η_0 *with* $[\varphi_0]_Z$, $[\varphi_2]_Z \in \mathcal{T}(\eta_0, \varepsilon_0)$, *in which case* γ_2 *is at most a single segment in* $S(\eta_0, [\varphi_2], \varepsilon_0)$. *The path* γ_4 *is trivial unless* $\varphi_1(v) \in S(\eta_1, [\varphi_1], \varepsilon_0)$ *for a loop* η_1, *with* $[\varphi_1]_Z$, $[\varphi_2]_Z \in \mathcal{T}(\eta_1, \varepsilon_0)$ *in which case* γ_4 *is at most a single arc in* $S(\eta_1, [\varphi_2], \varepsilon_0)$

More generally let β *be any loop or gap with* $[\varphi_2] \in \mathcal{T}(\partial\beta, \varepsilon_0)$. *Write* $S(\beta) = S(\beta, [\varphi_2], \varepsilon_0)$. *Then*

$$|\gamma_1 \cap S(\beta)|' \leqslant Ce^{d'_\beta([\varphi_0],[\varphi_2])},$$
$$|\gamma_3 \cap S(\beta)|' \leqslant C(e^{d'_\beta([\varphi_0],[\varphi_2])} + e^{d'_\beta([\varphi_1],[\varphi_2])}),$$
$$|\gamma_5 \cap S(\beta)|' \leqslant Ce^{d'_\beta([\varphi_1],[\varphi_2])}.$$

Here, $|\gamma_1 \cap S(\beta)|'$ *is measured as in 14.10.*

Proof. — This is a very similar method to 14.12. This time we need to define paths α_0 and α_1, and a loop set Γ_0. We start with the loop set Γ_0. This is like the loop set Γ of 14.12 (which is not always defined). Choose a gap ζ_0 and adjacent loop η_0 so that

$$\varphi_0(v) \in S(\zeta_0, [\varphi_0], \varepsilon_0) \cup S(\eta_0, [\varphi_0], \varepsilon_0).$$

In addition we can choose ζ_0 so that the following holds, if $\varphi_0(v) \in S(\eta_0, [\varphi_0], \varepsilon_0)$. Let η_0', η_0'' be the loops in $\overline{\mathbf{C}} \smallsetminus Y$ homotopic to η_0 in $\overline{\mathbf{C}} \smallsetminus Z$ such that $[\varphi_0]_Y \in \mathcal{T}(\eta_0', \varepsilon_0') \cap \mathcal{T}(\eta_0'', \varepsilon_0')$ for $\varepsilon_0' = 2\varepsilon_0(1 + o(1))$. Let ℓ be the union of the geodesic in $\mathcal{T}(Z)$ joining $[\varphi_0]_Z$ and $[\varphi_1]_Z$, and the geodesic in $\mathcal{T}(Z)$ joining $[\varphi_1]_Z$ and $[\varphi_2]_Z$. We can regard this as a path in $\mathcal{T}(Y)$, and we can regard points in ℓ as points in $\mathcal{T}(Y)$. Let $[\varphi']$ be the first point on ℓ such that $[\varphi'] \notin \mathcal{T}(\eta_0', \varepsilon_0') \cap \mathcal{T}(\eta_0'', \varepsilon_0')$, if this exists. Number so that $[\varphi'] \notin \mathcal{T}(\eta_0', \varepsilon_0')$. We can choose ζ_0 so that ζ_0 is adjacent to η_0'.

We choose a loop set Γ_0 in ζ_0 such that $|\varphi_0(\Gamma_0)|$ is bounded, and each component of $\zeta_0 \smallsetminus (\cup\varphi_0(\Gamma_0) \cup \varphi_0(Z))$ is either a disc or punctured disc round a point of $\varphi_0(Z)$ or parallel to the boundary. We choose Γ_0 so that $\varphi_0(\Gamma_0)$ is in good position (14.5) with respect to the quadratic differential at $[\varphi_0]$ for $d_Z([\varphi_0], [\varphi_1])$. As in 14.12, we choose α_0 ending at v so that

$$\varphi_0(\alpha_0) \subset S(\zeta_0, [\varphi_0], \varepsilon_0) \cup S(\eta_0, [\varphi_0], \varepsilon_0),$$

$|\varphi_0(\alpha_0) \cap S(\zeta_0, [\varphi_0], \varepsilon_0)|$ is bounded, $\varphi_0(\alpha_0) \cap S(\eta_0, [\varphi_0], \varepsilon_0)$ has at most one component, and then only if $\varphi_0(v) \in S(\eta_0, [\varphi_0], \varepsilon_0)$.

Then $\varphi_2(\alpha_0)$ has the same properties as $\chi\circ\varphi_1(\alpha)$ in 14.12, with the same differences depending on whether or not $\varphi_0(v) \in S(\eta_0, [\varphi_0], \varepsilon_0)$. The paths γ_1, γ_2 (which may be trivial) and part of γ_3 come from $\varphi_2(\alpha_0)$.

Now we choose α_1. Let Γ_0' be obtained from $\varphi_1(\Gamma_0)$ by a bounded homotopy, so that $\varphi_1(\Gamma_0')$ is in good position with respect to the quadratic differential $q_1(z)dz^2$ at $[\varphi_1]$ for $d_Z([\varphi_1], [\varphi_2])$. Then choose α_1 ending at v so that:

a) $\varphi_1(\alpha_0') = \varphi_1(\alpha_0) * \overline{\varphi_1(\alpha_1)}$ has a bounded number of intersections with $\varphi_1(\Gamma_0')$ and is in good position with respect to $q_1(z)dz^2$, and $\varphi_1(\alpha_0')$ ends in $(\overline{\mathbf{C}} \setminus \varphi_1(Z))_{\geqslant\varepsilon_0}$,

b) $|\varphi_1(\alpha_1)|_{\geqslant\varepsilon_0}$ is bounded and $\varphi_1(\alpha_1)$ has at most one intersection with a component of $(\overline{\mathbf{C}} \setminus \varphi_1(Z))_{<\varepsilon_0}$ — and only if $\varphi_1(v) \in S(\eta_1, [\varphi_1], \varepsilon_0)$ for some loop η_1.

All this is possible, because $\varphi_1(\alpha_0)$ has only boundedly many intersections with $\varphi_1(\Gamma_0)$. Since $|\varphi_2(\alpha_0')|'$ is bounded by $|\varphi_2(\Gamma_0)|'$, we obtain

$$|\varphi_2(\alpha_0')|' \leqslant Ce^{d_Z([\varphi_0],[\varphi_2])},$$

Arguing as in 14.12, either

$$|\varphi_2(\alpha_1)|' \leqslant Ce^{d_Z([\varphi_1],[\varphi_2])},$$

or $\alpha_1 = \alpha_{1,1} * \alpha_{1,2} * \alpha_{1,3}$ where $\varphi_1(\alpha_{1,2})$ is a single segment in $S(\eta_1, [\varphi_1], \varepsilon_0)$ for some loop η_1 and for $j = 1, 3$,

$$|\varphi_2(\alpha_{1,j})|' \leqslant Ce^{d_Z([\varphi_1],[\varphi_2])}.$$

The path γ_3 then comes from $\varphi_2(\alpha_0')$, parts of $\varphi_2(\alpha_{1,3})$ and $\varphi_2(\alpha_0)$ and their joins along $\varphi_2(\Gamma_0)$ to $\varphi_2(\alpha_0')$. The path γ_5 comes from $\varphi_2(\alpha_{1,1})$.

The estimates on $|\gamma_j \cap S(\beta)|'$ are proved in exactly the same way, by bounding $|\varphi_2(\alpha_0' \cap S(\beta))|'$, $|\varphi_2(\alpha_1 \cap S(\beta))|'$ and $|\varphi_2(\alpha_{1,j} \cap S(\beta))|'$ □

CHAPTER 15

TRIANGLES OF GEODESICS

15.1. In this chapter, we continue the study of geodesics in Teichmüller space. The results include: a way of recognizing bounded loops at some point of a given geodesic ℓ, in 15.9, and a theorem about triangles of geodesics in 15.8. The latter generalizes the following result about triangles of geodesics in the hyperbolic plane H for a suitable constant C independent of the geodesics being considered. Given geodesics between any two of the points x_1, x_2, $x_3 \in H$, there is a point $y \in H$ which is distance $\leqslant C$ from a point on each of the three geodesics. As usual, any difficulties in generalizing this result concern the thin part of Teichmüller space. The first lemma gives the generalization in a relatively easy case, using the results of Chapter 14. Throughout this chapter, we continue to use the notation m_α, $S(\alpha)$, $a(\alpha)$ of 9.1, 9.3, 9.4.

15.2. Lemma. — *Let $[x_1, x_3]$ be a geodesic segment in $\mathcal{T}(Z)$, and let*

$$x_2 \in [x_1, x_3] \cap \mathcal{T}_{\geqslant \varepsilon_0}.$$

Let $x_4 \in \mathcal{T}(Z)$. Then for a suitable constant $C > 0$, at least one of the following holds:

$$d(x_1, x_2) + d(x_2, x_4) \leqslant d(x_1, x_4) + C,$$
$$d(x_3, x_2) + d(x_2, x_4) \leqslant d(x_3, x_4) + C.$$

Proof. — Write $x_i = [\varphi_i]$. Let $|\cdot|_{q,+}$ denote the unstable length for the quadratic differential $q(z)dz^2$ for $d(x_2, x_4)$ at x_2. Let $|\cdot|_{p,\pm}$ denote the unstable and stable lengths for the quadratic differential $p(z)dz^2$ for $d(x_2, x_3)$ at x_2. Let $r(z)dz^2$ be the stretch of $p(z)dz^2$ at x_3. By 14.7, there is γ such that, for a constant C_1,

$$|\varphi_2(\gamma)|' \geqslant C_1 e^{d(x_4, x_2)} |\varphi_4(\gamma)|'.$$

But since $[\varphi_2] \in \mathcal{T}_{\geqslant \varepsilon_0}$, $|\varphi_2(\gamma)| = |\varphi_2(\gamma)|'$ is boundedly proportional to $|\varphi_2(\gamma)|_p$ (see 14.5), and hence to one of $|\varphi_2(\gamma)|_{p,+}$ or $|\varphi_2(\gamma)|_{p,-}$. Assume without loss of generality that it is the former. Then

$$|\varphi_3(\gamma)|_r \geqslant C_2 e^{(d(x_3,x_2)+d(x_2,x_4))} |\varphi_4(\gamma)|'.$$

Then by 14.4, 14.5, we have

$$e^{2d(x_3,x_4)} \geqslant C_3 \frac{|\varphi_3(\gamma)|'}{\varphi_4(\gamma)|'} \geqslant C_4 e^{d(x_3,x_2)+d(x_2,x_4)},$$

which gives the result. □

15.3. Thick and dominant over long distance. — The result 15.2 does not quite imply that x_2 in 15.2 is a bounded distance from one of the geodesics $[x_1, x_3]$, $[x_3, x_4]$, even if $x_2 \in \mathcal{T}_{\geqslant \varepsilon_0}$. But we want to head towards such a result. If $x_2 \in \mathcal{T}_{<\varepsilon_0}$ in the lemma above, then the approximate triangular equality of 15.2 is clearly not likely to hold, and a suitable analogue needs to be sought. For this, we need the concept of *long*, (ν)-*thick and dominant* gaps. This is more special than the concept of dominant area introduced in 9.4, and also more special than a common type of restriction that we make: $x \in \mathcal{T}(\partial\alpha, \varepsilon) \cap (\mathcal{T}(A(\alpha))_{\geqslant \nu}$ for $\varepsilon << \nu$.

Fix a geodesic segment ℓ in $\mathcal{T}(Z)$, a point $x \in \ell$, and a gap α with $x \in \mathcal{T}(\partial\alpha, \varepsilon_0)$, for $\varepsilon_0 \leqslant$ the Margulis constant. Fix functions $r : (0,1) \to (0,1)$, $L : (1,\infty) \to (0,\infty)$ and $s : (0,1) \to (0,1)$. Let α be a gap. Let $\pm q(z)dz^2$ be the quadratic differential for $d(x,x')$ at x, any $x' \in \ell$. Write $a(\alpha, q) = a(\alpha) = a(\alpha, x)$, $a(\partial\alpha, q) = a(\partial\alpha) = a(\partial\alpha, x)$. Note that this is boundedly proportional to the *square* of $|\varphi(\partial\alpha)|_q$ if $x = [\varphi]$. Then α is *long* (ν)-*thick and dominant* along a segment ℓ_1 of ℓ (for given functions r, Δ, s) if $\ell_1 \subset \mathcal{T}(\partial\alpha, r(\nu)) \cap (\mathcal{T}(A(\alpha))_{\geqslant \nu}$ and $a(\partial\alpha, x) \leqslant s(\nu)a(\alpha, x)$ for $x \in \ell_1$, and ℓ_1 has length $\geqslant 2\Delta(\nu)$. If x is the midpoint of some such segment ℓ_1, we say that α is long, ν-thick and dominant at x.

It was shown in 9.7 that there is always at least one gap or loop of dominant area. The following is similar.

15.4. Lemma. — *Let functions r, Δ s, and a constant D_1 be given. Then the following holds for suitable functions r' and Δ' and $\nu(r, \Delta, s) > 0$. Let $x' \in \ell' \cap \mathcal{T}(\partial\beta, r'(\nu')) \cap (\mathcal{T}(A(\beta))_{\geqslant \nu'}$ with $a(\partial\beta) \leqslant r'(\nu')$, and ℓ' of length $\geqslant 2\Delta'(\nu')$. Then either β contains a loop of D_1-dominant area, or x' is within $\Delta'(\nu')$ of a point x, such that some gap α is long distance thick and dominant for x, ℓ r, Δ, s, and for $\nu \geqslant \nu(r, \Delta, s)$.*

Remark. — This lemma can always be applied with $\beta = \overline{\mathbf{C}}$, no matter what r' is.

Proof. — Let

$$r_1(\nu) = r(\nu)e^{-C\Delta(\nu)}$$

for a suitable constant C. We can assume without loss of generality that $r_1(\nu) \leqslant \nu$. Put $N = \#(Z)$ and choose r' so that $r'(\eta) \leqslant r_1^N(\eta)$ for all η, where this denotes N-fold iteration. Take $\nu_0 = \varepsilon_0$ and $\nu_{i+1} = r_1(\nu_i)$. Assume without loss of generality that $\nu' \leqslant \varepsilon_0$. Then $r'(\nu') \leqslant r_1(\nu_{N-1})$. If either segment of length $\Delta(\nu_{N-1})$ starting from x' in ℓ' is in $(\mathcal{T}(A(\beta))_{\geqslant \nu_{N-1}}$, then we can take $x = x'$ and $\alpha = \beta$. If not, there is a nonempty loop set Γ_{N-1} in β, and x_{N-1} within $\Delta(\nu_{N-1})$ of x' such that $x_{N-1} \in \mathcal{T}(\Gamma_{N-1}, \nu_{N-1})$. Then inductively we can construct an m, $1 \leqslant m \leqslant N$, points x_i and loop sets Γ_i, for $m \leqslant i \leqslant N$, with $\Gamma_N = \varnothing$, $\Gamma_{i+1} \subset \Gamma_i$, $\Gamma_i \neq \Gamma_{i+1}$, x_i is within $L(\nu_i)$ of x_{i+1}, $x_i \in \mathcal{T}(\Gamma_i, \nu_i)$ and x_m is the centre of a segment of ℓ' of length $2\Delta(\nu_{m+1})$ in

$$\mathcal{T}(\Gamma_m, r(\nu_{m+1})) \setminus \bigcup\{\mathcal{T}(\gamma, \nu_{m+1} : \gamma \notin \Gamma_m\}.$$

If r is sufficiently fast decreasing given s and D_1 (as we can assume) then we can either find a loop α of D_1-dominant area, or a gap α with $a(\partial\alpha) \leqslant s(\nu_{m+1})a(\alpha)$. □

15.5. Subdominant Area. — There is another measurement of area which is sometimes more useful in measuring area that $a(\alpha)$. Let α be a gap. Let $\{w_t : t \in [0, T]\}$ be a geodesic segment in $\mathcal{T}(\partial\alpha, \varepsilon_0) \subset \mathcal{T}(Z)$ with length as the parameter t. Write π_α for the projection to $\mathcal{T}(A(\alpha))$. Write $q_t(z)dz^2$ for the quadratic differential for the geodesic at $w_t = [\chi_t \circ \varphi_0]$, where χ_t is the quasi-conformal homeomorphism minimising distortion. Choose a subsurface $S_t(\alpha) = \chi_t(S_0(\alpha))$ of $\overline{\mathbf{C}} \setminus \varphi_t(Z)$ which is homotopic to $\varphi_t(\alpha)$ such that the boundary components $\chi_t(\partial\alpha)$ are in good position (14.5). The surface $S_t(\alpha)$ may, however, be degenerate, in that it may not be the closure of its interior. Write $s_t(\alpha)$ for the q_t-area of $S_t(\alpha)$. Let $\ell_t(\partial\alpha)$ be the q_t-length of $\partial\alpha$. By 14.6 this has an *essential minimum*. So there is t_0 such that

$$C_1 e^{t_0 - t} \ell_{t_0}(\partial\alpha) \leqslant \ell_t(\partial\alpha) \leqslant C_2 e^{t_0 - t} \ell_{t_0}(\partial\alpha) \text{ for } t \leqslant t_0,$$
$$C_1 e^{t - t_0} \ell_{t_0}(\partial\alpha) \leqslant \ell_t(\partial\alpha) \leqslant C_2 e^{t - t_0} \ell_{t_0}(\partial\alpha) \text{ for } t \geqslant t_0.$$

In contrast, the function $t \mapsto s_t(\alpha)$ is constant. So, given constants $D_2 < D_1$ with D_1/D_2 bounded,there are $0 \leqslant u_1 \leqslant u_2$ with $u_2 - u_1$ bounded such that

$$s_t(\alpha) \leqslant D_2(\ell_t(\partial\alpha))^2 \quad \text{for } t \notin [t_0 - u_2, t_0 + u_2],$$
$$s_t(\alpha) \geqslant D_1(\ell_t(\partial\alpha))^2 \quad \text{for } t \in (t_0 - u_1, t_0 + u_1).$$

The case $u_1 = 0$ is allowed. In $(t_0 - u_2, t_0 + u_2)$, $s_t(\alpha)$ is boundedly proportional to $a(\alpha, q_t)$ of 9.4. If D_1 is sufficiently small and $[\chi_t \circ \varphi] \in (\mathcal{T}(A(\alpha))_{\geqslant \varepsilon_0}$ then α is of D_1-dominant area for $t \in (t_0 - u_1, t_0 + u_1)$. We shall refer to α being of D_2- *subdominant area* for $t \notin (t_0 - u_2, t_0 + u_2)$. Then we have the following.

15.6. Corollary. — *Take the same hypotheses as in 15.4, and in addition a constant* $D_2 > 0$. *Then a segment* ℓ *of length* $2\Delta(\nu)$ *starts within* $\Delta'(\nu')$ *of* x', *and a loop set* Γ, *such that the following hold.*

$$\ell \subset \mathcal{T}(\Gamma, r(\nu)) \setminus \bigcup_{\gamma \notin \Gamma} \mathcal{T}(\gamma, \nu).$$

For every loop of Γ, *either* γ *is* D_1*-dominant at all points of* ℓ, *or* D_2*-subdominant at all points of* ℓ. *For each gap* α *of* Γ, *either* $a(\partial\alpha, x) \leqslant s(\nu)a(\alpha, x)$ *for all* $x \in \ell$, *or* $a(\alpha, x) \leqslant D_2 a(\partial\alpha, x)$ *for all* $x \in \ell$. *If no loop is* D_1*-dominant, then at least one gap satisfies* $a(\partial\alpha, x) \leqslant s(\nu)a(\alpha, x)$ *for all* $x \in \ell$. *Put briefly, every gap is thick, and either dominant or subdominant along the entire length of* ℓ

Proof. — Arguing as in 15.4, find a segment ℓ'' of length $\geqslant \Delta_1(\nu)$ in $\mathcal{T}(\Gamma, r(\nu)) \setminus \cup_{\gamma\notin\Gamma}\mathcal{T}(\gamma,\nu)$, for a suitable $\Delta_1 = \Delta_1(\nu) \geqslant \Delta(\nu)$. Define the quantities $s_t(\alpha)$, $\ell_t(\partial\alpha)$ along this segment, as in 15.5. Take $t = 0$ as the centre of this segment. There are only finitely many gaps. We consider the functions $(\ell_t(\alpha))^2/s_t(\alpha)$, as α varies over the gaps. As we have already noted, $s_t(\alpha)$ is constant and each function behaves like $c_\alpha e^{-2|t-t_\alpha|}$ for different constants c_α and t_α. So each function takes values in the interval $[s(\nu), 1/D_2]]$ only on at most a union of two intervals, each of length $\log(1/D_2 s(\nu))$. So if $\Delta_1(\nu)$ is sufficiently large given D_2 and $s(\nu)$, we can find $\ell \subset \ell''$ on which all functions take values outside $[s(\nu), 1/D_2]$. We know that either there is a dominant loop, or at least one function must take a value $< 1/D_2$, and hence $< s(\nu)$. □

15.7. Projections look like geodesics. — We saw in 9.5 that projections of geodesics to dominant area gaps look like geodesics. But in fact this is true more generally. We have the following.

Lemma. — *Let α be a gap of Γ, and let $[w, u]$ be a geodesic segment in $\mathcal{T}(\Gamma, \varepsilon_0)$. Let the projection to $\mathcal{T}(A(\alpha))$ be in $(\mathcal{T}(A(\alpha)))_{\geqslant\varepsilon_0}$. Then the projection of $[w, u]$ to $\mathcal{T}(A(\alpha))$ is within a bounded distance of a geodesic segment.*

Remark. — We are not assuming that $d_\alpha(w, u)$ is close to maximal.

Proof. — Write
$$[w, u] = \{w_t : t \in [0, T]\}$$
with $w = w_0$, $u = w_T$ and with length as the parameter. Let u_1 and u_2 be as in 15.5. We claim that, for suitable choice of D_1 and D_2 the projection of $\{w_t : t \in [t_0 - u_1, t_0 + u_1]\}$ to $(\mathcal{T}(A(\alpha))_{\geqslant\varepsilon_0}$ is within a bounded distance of a geodesic. This follows from 9.5, because for $t \in [t_0 - u_1, t_0 + u_1]$, $s_t(\alpha) = a_t(\alpha)(1 + o(1)$ and $a_t(\partial\alpha) = O((\ell_t(\partial\alpha))^2)$, so the Dominant Area Condition of 9.4 holds. Then we claim that the projection of each component of $\{w_t : t \notin [t_0 - u_1, t_0 + u_1]\}$ is bounded. It suffices to prove that the projection of each component of $\{w_t : t \notin [t_0 - u_2, t_0 + u_2]\}$ is bounded.

It is obviously sufficient to consider the case $t \in [0, t_0 - u_2]$, since the case $t \in [t_0 + u_2, T]$ is exactly similar. It suffices to find a set of loops Γ_0 in S_0 such that every component of $S_0 \setminus (\cup\Gamma_0)$ is a disc or an annulus parallel to ∂S_0, and such that $|\chi_t(\Gamma_0)|$ is bounded for all $t \in [0, t_0 - u_2]$. Let $|\chi_t(\gamma)|_{q_t}$ denote the q-d length of $\chi_t(\gamma)$ for the quadratic differential $q_t(z)dz^2$. Let S_t' be the component of $(\overline{\mathbf{C}} \setminus \varphi_t(Z))_{\geqslant\varepsilon_0}$ homotopic to $\varphi_t(\alpha)$. Then S_t is a limit of isotopies in $\overline{\mathbf{C}} \setminus \varphi_t(Z)$ from S_t'. Let D_t be a union of discs in the interior of S_t' round the zeros of q_t. Then for any loops γ_1, γ_2 with $\gamma_j \subset S_t \setminus D_t$,
$$\frac{|\gamma_1|}{C_1|\gamma_2|} \leqslant \frac{|\gamma_1|_{q_t}}{|\gamma_2|_{q_t}} \leqslant \frac{C_1|\gamma_1|}{|\gamma_2|}.$$

Note that the loops of $\partial S'_t$ are not particularly short in the Poincaré metric, although they might be homotoped — outside S'_t — to be of very short Poincaré length. For any loop γ_1 in S'_t, the minimum of $|\gamma'_1|_{q_t}$, for γ'_1 homotopic to γ_1, is achieved on S_t, But the minimum can be achieved to within a bounded proportion on S'_t. So for geodesics in the interior of S'_t, their Poincaré lengths are boundedly proportional to the q_t q-d lengths of good position homotopy representatives in S_t. The loops of $\partial S'_t$ are bounded in the Poincaré metric. So bounded loops in S'_t are those whose homotopy representatives in S_t have q_t-lengths bounded by a multiple of the q_t-length of ∂S_t. So it suffices to find a constant C and Γ_0 such that

$$|\chi_t(\Gamma_0)|_{q_t} \leqslant C|\partial S_t|_{q_t}$$

for all $t \in [0, t_0 - u_2]$. There are unstable foliation arcs connecting the components of ∂S_0 cutting S_0 into cells and of lengths bounded by the lengths of ∂S_0. So we simply take Γ_0 to be any set of bounded loops in S_0 cutting S_0 into cells. Such loops can be arranged up to homotopy along finitely many segments of ∂S_0 and transverse unstable segments, Then the required bound on $|\chi_t(\Gamma_0)|_{q_t}$ follows.

15.8. Triangles of Geodesics. — We need a little notation to explain the following statement. A loop γ is a Δ_0-*Pole-Zero loop* along a geodesic segment ℓ_1 if for all $[\varphi] \in \ell_1$ there is an annulus in $S([\varphi], \gamma, \varepsilon_0)$ of modulus $\geqslant \Delta_0$ satisfying the Pole-Zero Condition 9.4. Also, if γ is a nontrivial nonperipheral loop in $\overline{\mathbf{C}} \setminus Z$, we need to be more precise than in 9.1 about the identification of $\mathcal{T}(A(\gamma))$ with the upper half plane. We need to fix a loop $\zeta \subset \overline{\mathbf{C}} \setminus Z$ which intersects γ exactly twice and such that each component of $\overline{\mathbf{C}} \setminus (\gamma \cup \zeta)$ contains exactly one point of $A(\gamma)$. Then we take $\mathrm{Re}(\pi_\gamma([\varphi]))$ to be within a bounded distance of n such that $|\sigma_\gamma^n \circ \varphi(\zeta)|$ is minimal. Up to bounded distance, this normalisation is independent of the choice of ζ.

Theorem. — *Let y_0, y_1, $y_2 \in \mathcal{T}(Z)$ with $y_j = [\varphi_j]$. Take any $y = [\varphi] \in [y_0, y_1]$.*

(1) *There exist $K_0 > 0$, $m_0 > 0$ such that the following hold. Let α be a loop which is m_0-Pole-Zero on $\ell \subset [y_0, y_1] \subset \mathcal{T}$. Then ℓ is a union of two segments ℓ_0 and ℓ_1, such that for all $y \in \ell_j$ there are y', ℓ'_j with $y' \in \ell'_j \subset [y_j, y_2]$ such that $y' \in \mathcal{T}(\alpha, \varepsilon_0)$, and*

$$|\,\mathrm{Re}(\pi_\alpha(y)) - \mathrm{Re}(\pi_\alpha(y'))| \leqslant K_0.$$

In addition, given $m'_1 > 0$, there exist $K_1 > 0$ and $m_1 > 0$ such that if α is m_1-Pole-Zero along ℓ then either α is m'_1-Pole-Zero along ℓ'_j, or ℓ_j and ℓ'_j have length $\leqslant K_1$.

(2) *There exist parameter functions r, s, Δ such that the following hold. Let α be a long ν- thick and dominant gap along $\ell \subset [y_0, y_1] \subset \mathcal{T}$ for parameter functions r, s, Δ. Then ℓ is a union of two segments ℓ_0 and ℓ_1, uniquely determined up to moving the endpoints a bounded distance, such that for all $y \in \ell_j$ there are y', ℓ'_j with*

$y' \in \ell'_j \subset [y_j, y_2]$ such that $y' \in \mathcal{T}(\partial\alpha, \varepsilon_0)$, and

$$d_\alpha(y, y') \leqslant C(\nu).$$

In addition, if parameter functions r', s', Δ' are given then for suitable choice of Δ, r, s given these, and two more functions ν_1, $\Delta_1 : (0,\infty) \to (0,\infty)$, either α is long $\nu_1(\nu)$-thick and dominant along ℓ'_j for parameter functions Δ', r', s', or ℓ_j and ℓ'_j have length $\leqslant \Delta_1(\nu)$.

(3) If $y \in \ell_j$ and y' are as in either 1 or 2 above and $w \in [y, y_j]$ with w, β, μ satisfying the conditions of y, α, μ in 1 or 2 above, and $\alpha \cap \beta \neq \varnothing$, then $w \in \mu_j$, and $w' \in [y', y_j]$, where μ_j are defined relative to μ as the ℓ_j to ℓ.

Idea of the proof. — The key is a characterisation of loops γ such that $|\varphi(\gamma)|$ (equivalently $|\varphi(\gamma)|'$) is bounded, because in order to find y' with $d_\alpha(y, y') \leqslant C(\nu)$, we need to find y' such that $|\varphi'(\gamma)|$ is bounded whenever $\gamma \subset \alpha$ and $|\varphi(\gamma)|$ is bounded. We need the following definition for the proof of 15.8.

15.9. Almost-bounded loops, and a lemma. — For a loop γ and $[\varphi] \in \mathcal{T}(Z)$, define

$$|\varphi(\gamma)|'' = |\varphi(\gamma) \cap (\overline{\mathbf{C}} \smallsetminus \varphi(Z))_{\geqslant \varepsilon_0}|' + \sum_\beta n_\beta(\gamma),$$

where the summation is over loops β with $[\varphi] \in \mathcal{T}(\beta, \varepsilon_0)$ such that $\beta \cap \gamma \neq \varnothing$ and $n_\beta(\gamma) = \#(\varphi(\gamma) \cap \ell(\beta))$, where $\ell(\beta)$ is an arc between the components of $\partial S(\beta, [\varphi], \varepsilon_0)$ which meets these components at rightangles (with respect to the Poincaré metric. We say that $\varphi(\gamma)$ is *almost bounded* if $|\varphi(\gamma)|''$ is bounded. Roughly speaking, this means that the intersection of $\varphi(\gamma)$ with $(\overline{\mathbf{C}} \smallsetminus \varphi(Z))_{\geqslant \varepsilon_0}$ is bounded — but there is also a bound on twists round short loops.

We shall use the following criterion to prove 15.8.

Lemma. — *Let $[y_0, y_1] \subset \mathcal{T}(Z)$ be any geodesic, and write $y_j = [\varphi_j]$. Let $[\varphi] = y \in [y_0, y_1]$. Let $q(z)dz^2$ be the quadratic differential at y for $d(y, y_1)$.*

The following holds for a sufficiently large constant $D_0 = D_0(\delta) > 0$. Let α be a D_0M-Pole Zero loop at y (15.8). Let γ' intersect α essentially. Let $|\varphi(\gamma' \cap \alpha)|_q \geqslant \delta^{-1}|\varphi(\gamma'\alpha)|_{q,-}$. (See the end of 14.5 for this notation.) Then

(1) $$M|\varphi_0(\alpha)|' \leqslant |\varphi_0(\gamma')|'.$$

A similar statement holds for φ_1 replacing φ_0 if $|\varphi(\gamma' \cap \alpha)|_q \geqslant \delta^{-1}|\varphi(\gamma' \cap \alpha)|_{q,+}$.

Given a function $C_0 : (0,1) \to (1,\infty)$, there are functions r, Δ, s as in 15.8, but this time of variables $(\nu, \delta) \in (0,1)^2$, and there is a function $C_1 : (0,1)^2 \to (1,\infty)$ and a constant D_0 such that the following hold. Let α be a long, ν-thick and dominant gap for y, $[y_0, y_1]$, r, Δ, s. Let $\gamma \subset \alpha$ with $|\varphi(\gamma)| \leqslant C_0(\nu)$. Let $\gamma' \cap \alpha \neq \varnothing$. Let $|\varphi(\gamma')|_q \leqslant \delta^{-1}|\varphi(\gamma')|_{q,-}$. Then

(2) $$|\varphi_0(\gamma)|'|\varphi(\gamma' \cap S(\alpha))|' \leqslant C_1(\nu, \delta)|\varphi_0(\gamma')|'.$$

A similar statement holds with φ_1 replacing φ_0, if $|\varphi(\gamma')|_q \leqslant \delta^{-1}|\varphi(\gamma')|_{q,+}$.

Conversely, suppose that $|\varphi(\gamma)|''$ is sufficiently large given M, for all $[\varphi] \in [y_0, y_1]$. Then there are γ_0' and γ_1' both intersecting γ transversally, with no transversal intersections between γ_0' and γ_1', such that, for both $j = 0$ or 1,

$$|\varphi_j(\gamma_j')|'M \leqslant |\varphi_j(\gamma)|'. \tag{3}$$

Remark. — It would be preferable to have a necessary and sufficient condition for $|\varphi(\gamma)|$ to be bounded. The above falls somewhat short of this, but can be improved upon in some circumstances, as we shall see.

Before we prove this, we need the following.

15.10. Lemma. — *Let $\mathcal{G}_\pm$ be the expanding and contracting foliations of a quadratic differential $q(z)dz^2$ at $[\varphi] \in (\mathcal{T}(Z))_{\geqslant \varepsilon_0}$. Let a decreasing function $\varepsilon : (0, \infty) \to (0, \infty)$ be given. Then there is a decreasing function $L : (0, \infty) \to (0, \infty)$ such that, given $\delta > 0$, one of the following holds.*

(1) *There are a nontrivial nonperipheral loop γ, and $L \leqslant L(\delta)$, such that $|\varphi(\gamma)|_q = L$ and $|\varphi(\gamma)|_+ \leqslant \varepsilon(L)$.*

(2) *We have $\ell_+ \cap \ell_- \neq \varnothing$ for any segments $\ell_\pm$ of $\mathcal{G}_\pm$ with ℓ_+ and $|\ell_-|_- \geqslant L(\delta)$, $|\ell_+|_+ \geqslant \delta$.*

Proof. — Let $N = 4\#(Z)$ and let the function g be defined by

$$g(L') = \frac{2N}{\varepsilon(2NL')}.$$

Let g^j denote the j'th iterate of g. Now define

$$L(\delta) = \sum_{j=0}^{N-1} g^j(2/\delta).$$

Fix a segment ℓ_+ of $\mathcal{G}_+$ of length δ. Suppose that 2 does not hold for at least one segment of length $L(\delta)$ of $\mathcal{G}_-$. Lift to the surface supporting the orientable cover of the foliation $\mathcal{G}_+$, which also supports the orientable cover of $\mathcal{G}_-$. Let λ denote a lift of the $\mathcal{G}_+$-leaf segment ℓ_+. So now we can assume that $\mathcal{G}_\pm$ are orientable foliations on an orientable surface, and it suffices to find a closed loop γ on this surface of length $L \leqslant L(\delta)$ with $|\gamma|_+ \leqslant \varepsilon(L)$. Such a loop with automatically project to a nontrivial nonperipheral loop.

Take the union of all $\mathcal{G}_-$-leaves of length $\leqslant L(\delta)$ which start on λ and end the first return to λ, or at length $L(\delta)$, whichever comes first. Then this union of leaves is a surface S with boundary (by our assumptions), which is also a union of $\leqslant N$ rectangles, with base on λ, with opposite side also on λ if the height is $< L(\delta)$. The boundary of the surface is contained in the union of the $\mathcal{G}_-$ sides of the rectangles, and $\leqslant 2N$ $\mathcal{G}_+$ segments, the sum of whose widths is $\leqslant 2\delta$, since the area is $\leqslant 2$. At least one of the rectangles has width $\geqslant \delta/N$. Order the rectangles as R_i, $1 \leqslant i \leqslant m$

such that R_i has width λ_i, height L_i, and $\lambda_{i+1} \leqslant \lambda_i$, $\lambda_1 \geqslant \delta/N$. Then $L_i \leqslant 2/\lambda_i$, since the total area is 2. (We took a double cover of a surface of area 1.) Then either there must be $1 \leqslant i < m$ with $\lambda_{i+1} < \varepsilon(2NL_i)/N$ and $L_j \leqslant g^{j-1}(2/\delta)$ for $j \leqslant i$, or $L_m \leqslant g^{m-1}(2/\delta)$ and S has purely leaf boundary. We put $i = m$ in the second case. Then consider the surface formed by $\cup_{j\leqslant i} R_j$, which again has boundary. In its boundary, we obtain a loop $\varphi(\gamma)$ with

$$|\varphi(\gamma)|_- \leqslant 2\sum_{j\leqslant i} L_j = L \leqslant \sum_{j=0}^{m-1} g^j(2/\delta) \leqslant L(\delta),$$
$$|\varphi(\gamma)|_+ \leqslant \sum_{j>i} \lambda_j \leqslant \varepsilon(L),$$

where the last sum in the above is interpreted as 0 if $i = m$. □

15.11. Adapting to thin part. — Lemma 15.10 implicitly makes sense if $[\varphi] \in \mathcal{T}_{<\varepsilon_0}$. However, the following version for $[\varphi] \in \mathcal{T}_{<\varepsilon_0}$ is proved in exactly the same way, and will be useful. We shall always apply it when 3 is assumed *not* to hold.

Lemma. — *Let $\mathcal{G}_\pm$ be the expanding and contracting foliations for a quadratic differential $q(z)dz^2$ at $[\varphi] \in \mathcal{T}(Z)$. Let $S \subset \overline{\mathbf{C}} \smallsetminus \varphi(Z)$ have good boundary (9.4), and q-area a. Let a decreasing function $\varepsilon : (0,\infty) \to (0,\infty)$ be given. Then there is a decreasing function $L : (0,\infty) \to (0,\infty)$ and a constant $C > 0$ such that, given $\delta > 0$, one of the following holds.*

(1) *There is a loop γ with $\varphi(\gamma) \subset S$ such that $|\varphi(\gamma)|_q = L\sqrt{a} \leqslant L(\delta)\sqrt{a}$ and $|\varphi(\gamma)|_+ \leqslant \varepsilon(L)\sqrt{a}$.*

(2) *We have $\ell_+ \cap \ell_- \neq \varnothing$ for any segments $\ell_\pm$ of $\mathcal{G}_\pm$ in $\varphi(S)$ with $|\ell_-|_- \geqslant L\sqrt{a}$, $|\ell_+|_+ \geqslant \delta\sqrt{a}$.*

(3) $|(\partial S)|_q \geqslant \varepsilon(L(\delta))\sqrt{a}/C$.

Proof. — This is exactly the same as 15.10, except for a couple of relatively minor points. The first is that the total area is a, so all length measurements are multiplied by $\sqrt{a}$. The second is that $\mathcal{G}_-$-leaf segments can leave S, so that we take leaf segments starting from λ up until first return to or exit from $\widetilde{S}$ (the lift of S to the orientable cover), whichever comes first. If we assume that 2 does not hold, we can then deduce 1 or 3. The closed loop as in 1 might be formed by taking a segment between two different components of ∂S and doubling back along this. □

15.12. Proof of 15.9. — Let $q(z)dz^2$ be the quadratic differential for $d(y, y_1)$ at $y = [\varphi]$, with stretch $q_1(z)dz^2$ at $y_1 = [\varphi_1]$.

We consider the case when α is a loop first, because then the idea is particularly simple. Let T be the annulus in $S(\alpha, [\varphi])$ satisfying the Pole-Zero Condition and of modulus $\geqslant D_0M$. Let

(1) $$|\varphi(\gamma' \cap \alpha)|_{q,+} \geqslant \delta|\varphi(\gamma' \cap \alpha)|_q.$$

Then $\varphi(\gamma')$ passes through this annulus in a direction bounded from stable. Then consider stable leaves starting from $\varphi(\gamma)$. Any one will cross $\varphi(\gamma')$ at least $D_0'M$ times in T, for D_0' arbitrarily large if $D_0 = D_0(\delta)$ is large enough. As we move along the geodesic towards $[\varphi_1]$, the stable leaf segments contract. Take any $[\psi]$ on the geodesic between $[\varphi]$ and $[\varphi_1]$ with corresponding quadratic differential $p(z)dz^2$. Every point on $\psi(\gamma)$ in any $S(\beta, [\psi])$ is matched with K_0' points of $\psi(\gamma')$. The matching preserves $|\cdot|_{p,+}$-length, which is boundedly proportional to $|\cdot|'$ on any bounded type subsurface of $\overline{\mathbf{C}} \setminus \psi(Z)$. It follows that, assuming $D_0 = D_0(\delta)$ is large enough,

$$|\varphi_1(\gamma')|' \geqslant M|\varphi_1(\alpha)|',$$

as required.

Now let α be a gap which is long distance thick and dominant at $[\varphi]$ for $[y_0, y_1]$, and functions r, Δ, s, and ν, that is, $[\varphi]$ is in the centre of a segment of length $\geqslant 2\Delta(\nu)$ of $[y_0, y_1] \cap \mathcal{T}(\partial\alpha, r(\nu)) \cap (\mathcal{T}(A(\alpha))_{\geqslant \nu}$ with $a(\partial\alpha, [\varphi]) \leqslant s(\nu)a(\alpha, [\varphi])$. As before, assume without loss of generality that (1) holds. In what follows, constants C_i depend on ν. Let R be a rectangle in S of q-area $\geqslant a(\alpha, q)/C_2$, bounded by two unstable and two stable leaf segments, and such that there are n segments of $\varphi(\gamma')$ crossing between the stable leaf segment $\partial_- R$ in ∂R, with $n \geqslant |\varphi(\gamma' \cap \alpha)|'/C_3$ and $n \geqslant 1$. Then by 15.11, one of two things happens, for a function $L_0(\nu)$ given $\varepsilon(L, \nu)$.

a) A stable leaf segment of length $\leqslant L_0 = L_0(\nu)$ starting from any point of $\varphi(\gamma)$ crosses an unstable side of R.

b) For some $L \leqslant L_0$ there is a nontrivial nonperipheral loop $\gamma'' \subset \alpha$ such that

$$|\varphi(\gamma'')|_q \leqslant L\sqrt{a(\alpha, q)}, \quad |\varphi(\gamma'')|_{q,+} \leqslant \varepsilon(L, \nu)\sqrt{a(\alpha, q)}.$$

Now take $\varepsilon(L, \nu) = L^{-1}e^{-8\pi^2 \#(Z)/\nu}$. The large distance dominant area condition for suitable functions r, Δ, s means that b) does not hold, because b) implies an entry into $\mathcal{T}(\gamma'', \nu)$ for some $\gamma'' \subset \mathrm{Int}(\alpha)$ within time $\frac{1}{2}\log(L_0(\nu) + 2\pi^2 \#(Z)/\nu)$. So every point on $\varphi(\gamma)$ can be joined by a stable segment of length $\leqslant L_0$ to n points on $\varphi(\gamma' \cap \alpha)$ in the rectangle R. At most boundedly finitely many points on $\varphi(\gamma)$ are matched with each set of points on $\varphi(\gamma' \cap \alpha)$. As before, the matching preserves $|\cdot|_{q,+}$ length, and persists as we move along the geodesic. Again, $|\cdot|_{q_1,+}$ is boundedly proportional to $|\cdot|_{q_1}$, and to $|\cdot|'$ restricted to any bounded type subsurface of $\overline{\mathbf{C}} \setminus \varphi_1(Z)$, on each of which every point of $\varphi(\gamma)$ is matched with n points of $\varphi(\gamma')$. So we have

$$n|\varphi_1(\gamma)|' \leqslant C_4|\varphi_1(\gamma')|',$$

that is

$$|\varphi_1(\gamma)|'|\varphi(\gamma' \cap S(\alpha))|' \leqslant C_1|\varphi_1(\gamma')|', \tag{2}$$

as required.

Now we consider the converse. So suppose that $|\varphi(\gamma)|''$ is sufficiently large for all $[\varphi] \in [y_0, y_1]$. We can assume $\varphi(\gamma)$ is in good position with respect to the quadratic differential $q(z)dz^2$ at any point $[\varphi] \in [[\varphi_0], [\varphi_1]]$. Let $q_0(z)dz^2$ be the quadratic differential of $d([\varphi_0],[\varphi_1])$ at $[\varphi_0]$, and let $q_1(z)dz^2$ be the stretch at $[\varphi_1]$. We can find $[\varphi]\in[[\varphi_0],[\varphi_1]]$ such that one of 1-3 holds. By 15.6, given D_1, we can assume that every gap at $[\varphi]$ is either long distance thick and dominant, or D_1-subdominant, and that every loop is either D_1^{-1}-dominant or D_1-subdominant. The constant C_5, as before, depends on ν, for suitable chosen functions r, s, L.

1. There is a union ℓ of segments of $\varphi(\gamma)$, with all segments of $|\cdot|'$-length $\geqslant 1/C_5$, with $|\ell|'\geqslant|\varphi(\gamma)|'/C_5$, such that, for any $\ell_1\subset\ell$,

$$|\ell_1|_{q,+}\geqslant\frac{|\ell_1|_q}{C_5}, \tag{3}$$

and $d([\varphi],[\varphi_0])\leqslant C_5$.

2. This is similar, but with $|\cdot|_{q,+}$ replaced by $|\cdot|_{q,-}$ and $d([\varphi],[\varphi_1])\leqslant C_5$.

3. There are unions ℓ and ℓ' such that 1 holds for ℓ and 2 holds for ℓ'. There is no restriction on the position of $[\varphi]$ on $[[\varphi_0],[\varphi_1]]$. It is possible, but not inevitable, in this case, that $\ell=\ell'$.

Now let $[\varphi]\in[[\varphi_0],[\varphi_1]]$ be such that 3 holds. Then there is a gap or loop α such that

$$|\ell\cap S(\alpha,[\varphi])|'\geqslant\frac{|\varphi(\gamma)|''}{C_5\,\#(Z)}.$$

If α is long distant thick and dominant, take any $\gamma_1'\subset\alpha$ with $|\varphi(\gamma_1')|'$ bounded and such that γ and γ_1' have at least one essential intersection. Replacing γ by γ_1' and γ' by γ in (2), we have

$$|\varphi_1(\gamma_1')|'|\varphi(\gamma\cap\alpha)|'\leqslant C_1|\varphi_1(\gamma)|'. \tag{4}$$

This then gives (3) of 15.9 if $|\varphi'(\gamma)|''$ is sufficiently large for all $[\varphi']\in[y_0,y_1]$ given M. If α is a subdominant gap, then we can still find $\gamma_1'\subset\alpha$ such that γ and γ_1' have at least one essential intersection, each point on $\varphi(\gamma_1')$ locks with $\geqslant M\varphi(\gamma\cap\alpha)/C_6$ segments of $\varphi(\gamma\cap S(\alpha)$ along short stable segments and $|\varphi(\gamma_1')|'$ is bounded. We can do this, because every stable segment in $\varphi(\alpha)$ (taking $\varphi(\partial\alpha)$ in good position) has length $o(|\varphi(\partial\alpha)|_q)$. Then, again, we obtain (4) above, for suitable C_1. Similarly, we can find γ_0'. If γ_0' and γ_1' are in different gaps then they are obviously disjoint. If they are in the same gap α, we can clearly take them to be disjoint if α is subdominant. If α is long distance thick and dominant, we claim that we can take $\gamma_0'=\gamma_1'$. Note that if $\gamma'\subset\alpha$ for α large distance dominant at φ and $|\varphi(\gamma')|$ is bounded, then $|\varphi(\gamma')|_q$ is boundedly proportional to both of $|\varphi(\gamma')|_{q,\pm}$. So the claim holds. So in all cases we obtain (3) of 15.9, as required.

If 1 holds, then we only need to construct γ_1', and can then take $\gamma_0'=\gamma_1'$, and similarly for 2. □

15.13. Proof of 1 of 15.8. — Now we consider the case of 15.8 when α is a m_0-Pole-Zero loop (for suitable m_0), which is the simplest case. Fix a loop β which intersects α just twice. Consider, for varying $n\in\mathbf{Z}$ and $[\psi]\in[y_i,y_j]$ (i, $j\in\{0,1,2\}$), the function $h_{[\psi]}(n) = |\psi(\sigma_\alpha^n(\beta))|'$, and let $n([\psi])$ denote the value for which this is minimal. Let $n_i = n([\varphi_i])$, $i = 0, 1, 2$. Let $M_{i,j}$ be the maximum modulus of an annulus homotopic to $\varphi(\alpha)$ for $[\varphi] \in [y_i, y_j]$. As usual let $m_\alpha(y_i)$ denote the modulus of $S(y_i, \alpha, \varepsilon_0)$ — zero if $S(y_i, \alpha, 2\varepsilon_0) = \varnothing$. We shall show the following, which is not needed for 15.8, but will be needed later:

$$(1) \qquad |n_i - n_j| + \tfrac{1}{2}(m_\alpha(y_i) + m_\alpha(y_j)) \leqslant CM_{i,j}.$$

Fix i and j. Let y_i' be the point nearest y_i on $[y_i, y_j]$ such that $y_i' \in \mathcal{T}(\alpha, 2\varepsilon_0)$ and $S(\alpha, \varepsilon_0)$ contains an annulus of modulus $\geqslant 1$ satisfying the Pole-Zero Condition. Let y_j' be similarly defined. If there are no such points we can simply define $y_i' = y_j' = y_i$. We also claim that $n([\psi])$ is constant between y_i and y_i', and similarly between y_j and y_j', and varies monotonically between y_i' and y_j'. This will suffice to prove 1 of 15.8, since

$$n([\psi]) = \mathrm{Re}(\pi_\alpha([\psi]) + O(1)$$

for $[\psi] \in [y_i', y_j']$.

The n for which the minimum of $h_{[\psi]}$ occurs can be recognized from the good position of the loops $\psi(\sigma_\alpha^m(\beta))$ with respect to the quadratic differential at $[\psi]$ for $d([\psi], y_j)$. Write $\beta' = \sigma_\alpha^{n_i}(\beta)$. For any $[\psi] \in [y_i, y_j]$, the good position of $\psi(\sigma_\alpha^m(\beta')$ is a union of up to four segments, at most two of which are close to segments of $\psi(\beta')$ and the others to segments of $\psi(\alpha)$ (twisted round $O(m)$ times). If the good position of $\psi(\alpha)$ consists of arcs at more than one angle, then the good position of $\psi(\sigma_\alpha^m(\beta'))$ is exactly on top of $\psi(\alpha) \cup \psi(\beta')$ all $[\psi] \in [y_i, y_j]$. If this is true then the minimum of $h_{[\psi]}$ is constant for $[\psi] \in [y_i, y_j]$, that is, $n_i = n_j$. If $\psi(\alpha)$ is at constant angle with the foliations of the quadratic differential for $d([\psi], y_j)$, then the minimum of $h_{[\psi]}$ can only change if relative lengths change, that is, the segments of stable (or unstable) foliation along which $\psi(\sigma_\alpha^m(\beta'))$ is locked with $\psi(\alpha) \cup \psi(\beta')$ grow proportionally much longer than $\psi(\alpha) \cup \psi(\beta')$. This is only possible if there is $[\psi] \in [y_i, y_j]$ with $[\psi] \in \mathcal{T}(\alpha, \varepsilon_0)$ and α is a m_1'-Pole-Zero loop on at $[\psi]$ (15.8) for m_1' bounded from 0. The more $h_{[\psi]}$ needs to change, the larger m_1' needs to be. So $n([\psi])$ remains constant on $[y_i, y_i'] \cup [y_j', y_j]$. It also varies monotonically on $[y_i', y_j']$ to within $O(1)$, that is, if $[\psi]$, $[\psi']$, $[\psi'']$ are successive points on $[y_i, y_j']$ then $n([\psi'']) - n([\psi'])$, $n([\psi']) - n([\psi])$ have the same sign unless one of these is $O(1)$.

Fix a set $A(\alpha)$ such that all points of $A(\alpha)$ are in different components of $\overline{\mathbf{C}} \smallsetminus (\alpha \cup \beta')$. We now regard α, β as loops in $\overline{\mathbf{C}} \smallsetminus A(\alpha)$. Write $|\cdot|_1$ for length in surfaces $\overline{\mathbf{C}} \smallsetminus \psi(A(\alpha))$. Then for $[\psi] \in [y_i', y_j']$, $n([\psi]) = n'(\pi_\alpha([\psi]) + O(1)$ where $n'([\psi_1])$ is the value of m for which $|\psi_1(\sigma_\alpha^m(\beta'))|_1$ is minimal. Then $n'(\pi_\alpha([\psi]) = \mathrm{Re}(\pi_\alpha([\psi]) + O(1)$. The path in the upper half-plane

$$\{\pi_\alpha([\psi]) : [\psi] \in [y_i', y_j']\}$$

is within bounded distance the arc of a circle with centre on a horizontal line on or above the real line, but below both $\pi_\alpha(y_i')$ and $\pi_\alpha(y_j')$. We have $\mathrm{Im}(\pi_\alpha(y_i')) = m_\alpha(y_i')$ and similarly for y_j'. The maximum of $m_\alpha(y_i')$, $m_\alpha(y_j')$ is $\geqslant$ the maximum of $m_\alpha(y_i)$,

$m_\alpha(y_j)$. The highest point on this circle arc is $M_{i,j} + O(1)$, and satisfies the bound (1).

15.14. Before starting to prove 2 of 15.8, we need to bound growth in gaps disjoint from the long thick and dominant.

Lemma. — *Fix long thick and dominant parameter function Δ, r, s, and $\Delta_0 > 0$. Given these there exists M such that the following holds. Let a geodesic segment ℓ and $\alpha \subset \overline{\mathbf{C}}$ be a maximal subsurface up to homotopy relative to Z with the property that α is disjoint from all (Δ, r, s) long thick and dominant gaps, and Δ_0-Pole-Zero loops, along segments $\ell_1 \subset \ell$. Suppose also that that $\alpha \setminus Z$ is nontrivial nonperipheral. Then*

$$|\varphi(\partial\alpha)| \leqslant M \quad \text{for all } [\varphi] \in \ell, \tag{1}$$

$$M^{-1} \leqslant \frac{|\varphi(\gamma)|'}{|\psi(\gamma)|'} \leqslant M \quad \text{for all } [\varphi],\ [\psi] \in \ell,\ \gamma \subset \mathrm{Int}(\alpha). \tag{2}$$

Proof. — First we show that (1) implies (2) for some $M_1 = M_1(M)$ replacing M. It will then follow that if (1) holds then (1) and (2) hold for some possibly larger constant M. We proceed by induction on $-\chi(\alpha \setminus Z)$, where χ denotes Euler characteristic. If $\alpha \setminus Z$ is an annulus there is nothing to prove. By hypothesis, α intersects no long thick and dominant gaps (for parameter functions Δ, r, s) or Δ_0-Pole-Zero loops. It follows that, except on a segment of ℓ of bounded length, $a(\alpha, [\varphi]) = o(a(\partial\alpha, [\varphi]))$ for $[\varphi]$ along ℓ, because otherwise we would have ℓ long and $a(\partial\alpha, [\varphi]) = o(a(\alpha, [\varphi])$ along all but bounded length of ℓ and we could find a long thick and dominant inside α. So: $\ell = \ell_1 \cup \ell_2 \cup \ell_3$ where ℓ_2 is bounded and for $[\varphi] \in \ell_1$ $\varphi(\partial\alpha)$ is mostly in the stable direction and there are finitely many homotopy classes of arcs ζ crossing α between boundary components such that points on most of the length of $\varphi(\partial\alpha)$ are joined to other point on $\varphi(\partial\alpha)$ by arcs in the homotopy classes $\varphi(\zeta)$ of comparatively short unstable length. A similar statement holds for ℓ_3, with stable and unstable interchanged. If the arcs ζ cut α into cells then we have (2) with ℓ_1 replacing ℓ. If not, then for any complementary component β of the arcs ζ in α, we have (1) with β replacing α and ℓ_1 replacing ℓ. Then we also have (2) for β for a suitable $M_1(M)$, by the inductive hypothesis. So we have (2) for suitable $M_2(M)$ for α with ℓ_1 replacing ℓ. Similarly we have it for ℓ_3 replacing ℓ, and since ℓ_2 is bounded, we have (2) for suitable $M_3(M)$.

So now it remains to prove (1). Along ℓ we have a finite (but arbitrarily large) N, successive segments $\ell_{n,1}$, $1 \leqslant n \leqslant N$, subsurfaces $\alpha_{n,1} \supset \alpha$, $1 \leqslant n \leqslant N$, such that $\ell_{n,1}$ and $\ell_{n+1,1}$ are adjacent for all n, $\alpha_{n,1}$ is a union of gaps and loops with $|\varphi(\partial\alpha_{n,1})|$ bounded along ℓ_n and $\beta \cap \alpha_{n,1} = \varnothing$ for any (Δ, r, s) long thick and dominant gap β or Δ_0-Pole-Zero loop β along a subsegment of $\ell_{n,1}$. We can obviously arrange that $\alpha_n \neq \alpha_{n+1}$ for all n. Then the conclusion of the lemma holds if we replace ℓ by ℓ_n. Now we amalgamate segments, and take intersections of gaps, as follows. For all

$n \leqslant N_2$, some N_2, we shall have $\alpha_{n,2} = \alpha_{m,1} \cap \alpha_{m+1,1}$ for some m. We start with $\alpha_{1,2} = \alpha_{1,1} \cap \alpha_{2,1}$. If $\alpha_{n,2} = \alpha_{m,1} \cap \alpha_{m+1,1}$ then we take $\alpha_{n+1,2} = \alpha_{m+2p,1} \cap \alpha_{m+2p+1,1}$ for the least $p \geqslant 1$ such that $\alpha_{m+2p,1} \cap \alpha_{m+2p+1,1} \neq \alpha_{n,2}$. If there is no such p then we take $n = N_2$ if N is even, and$N_2 = n+1$, $\alpha_{N_2,2} = \alpha_{N,1}$ if N is odd. We take $\ell_{n,2}$ to be the union of $\ell_{m+r,1}$ for $0 \leqslant r < 2p$. Then $|\varphi(\partial\alpha_{n,2})|$ is bounded for $[\varphi] \in \ell_{n,2}$. To see this, it suffices to see that $|\varphi(\partial(\alpha_{m,1} \cap \alpha_{m+1,1}))|$ is bounded (up to homotopy) for $[\varphi] \in \ell_{m,1} \cup \ell_{m+1,1}$, for each m as above. This follows since $|\varphi(\partial\alpha_{m+1,1} \cap \alpha_{m,1})|)$ is bounded for $[\varphi] \in \ell_{m,1}$ up to homotopy — since it is bounded at the right endpoint and (1) implies (2) — and hence is bounded along $\ell_{m,1} \cup \ell_{m+1,2}$, since $|\varphi(\partial\alpha_{m+1,1})|$ is bounded for $[\varphi] \in \ell_{m+1,1}$. Similarly $|\varphi(\partial\alpha_{m,1} \cap \alpha_{m+1,1})|$ is bounded for $[\varphi] \in \ell_{m,1} \cup \ell_{m+1,1}$. Since we have similar bounds for $m+2i$, $i \leqslant p$, we get the required bound on $|\varphi(\partial\alpha_{n,2})|$ for $[\varphi] \in \ell_{n,2}$. We continue the process, defining segments $\ell_{n,k}$ and $\alpha_{n,k}$. The process terminates for some k bounded in terms of $\#(Z)$, because $\alpha_{n,k}$ is properly contained in some $\alpha_{m,k-1}$, except for $n = N_k$, when we simply have $\alpha_{N_k,k} \subset \alpha_{N_{k-1},k-1}$. In fact, for k bounded in terms of $\#(Z)$ we have $N_k = 1$ and we must have $\alpha_{1,k} = \alpha$ and result is proved. □

15.15. Proof of 2 of 15.8: finding one almost-bounded loop. — Take any $[\varphi] \in [y_0, y_1]$. First, we consider a long, ν-thick and dominant α at $[\varphi]$ for $[y_0, y_1]$, and suitable functions r, Δ, s. Write $q(z)dz^2$ for the quadratic differential for $d([\varphi], y_1)$ at $[\varphi]$. Let $\gamma \subset \alpha$ with $|\varphi(\gamma)| \leqslant K$. We start by showing that there is $[\varphi'] \in [y_0, y_2] \cup [y_1, y_2]$ such that

$$|\varphi'(\gamma)|'' \leqslant C(K, \nu). \tag{1}$$

Suppose for contradiction that $|\varphi'(\gamma)|'' \geqslant C(K,\nu)$ for all such $[\varphi']$. Then by (3) of 15.9, assuming that $C(K,\nu)$ is sufficiently large given a constant M, there are loops γ_0, γ_2, γ_1', γ_2', all intersecting γ essentially, and so that γ_0, γ_2 have no transversal intersections, neither do γ_1' and γ_2', and

$$(|\varphi_2(\gamma_2)|' + |\varphi_2(\gamma_2')|')M \leqslant |\varphi_2(\gamma)|', \tag{2}$$

$$|\varphi_0(\gamma_0)|'M \leqslant |\varphi_0(\gamma)|', \quad |\varphi_1(\gamma_1')|'M \leqslant |\varphi_1(\gamma)|'.$$

Assume as usual that the images under φ of all loops are in good position with respect to $q(z)dz^2$. By (2) of 15.9, it follows that, if M is sufficiently large given δ,

$$|\varphi(\gamma_0 \cap \alpha)|_{q,-} \leqslant \delta|\varphi(\gamma_0 \cap \alpha)|_q, \tag{3}$$

$$|\varphi(\gamma_1' \cap \alpha)|_{q,+} \leqslant \delta|\varphi(\gamma_1' \cap \alpha)|_q. \tag{4}$$

Since $\gamma_0 \cap \gamma_2 = \varnothing$, we deduce from 15.11 that, if M is sufficiently large given δ, (3) also holds for γ_2, and (4) also holds for γ_2'. It follows from 15.11 that, if M is sufficiently large given N, then γ_2 and γ_2' have $\geqslant N$ intersections in α.

Now we claim that the method of 15.9 implies that, for a suitable $C_1 = C_1(\nu)$,

$$|\varphi_2(\gamma)|' \leqslant C_1(|\varphi_2(\gamma_2)|' + |\varphi_2(\gamma_2')|'), \tag{5}$$

which will contradict (2), and hence give (1). To see this, let $p(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], y_2)$. Then we only need to lock all of $\varphi(\gamma)$ to segments of $\varphi(\gamma_2 \cup \gamma_2')$ by stable foliation segments (for $p(z)dz^2$) of length $O(|\varphi(\gamma)|_p)$. We see this as follows. Continue to take $\varphi(\gamma_2)$, $\varphi(\gamma_2')$ in good position with respect to $q(z)dz^2$. Now $|\varphi(\gamma)|$ is bounded. So $\varphi(\gamma)$ is in a subsurface S' of $S(\alpha)$ bounded by segments of leaves of the stable and unstable foliations of $q(z)dz^2$, and we can ensure that $\varphi(\gamma)$ is bounded from the boundary of S'. Write $\partial S' = \partial_+ S' \cup \partial_- S'$, the partition into unstable and stable leaf segments. Since there are long segments of $\varphi(\gamma_2)$, $\varphi(\gamma_2')$ close to stable and unstable foliations respectively, by 15.11, we can find segments ℓ_2, ℓ_2' on $\varphi(\gamma_2)$, $\varphi(\gamma_2')$ in $S(\alpha)$ such that every component of $\partial_+ S'$ is close to a subsegment of ℓ_2, and every component of $\partial_- S'$ is close to a subsegment of ℓ_2', and such that, for a suitable constant $C_2 = C_2(\nu)$,

$$|\ell_2|' + |\ell_2'|' \leqslant C_2|\varphi(\gamma)|.$$

Then we can find a surface S'', with $S' \subset S'' \subset S(\alpha)$, with boundary a subset of $\ell_2 \cup \ell_2'$, such that $S'' \smallsetminus (\ell_2 \cup \ell_2' \cup \varphi(Z))$ is a union of at-most-once-punctured topological discs and containing $\varphi(\gamma)$ in its interior.

Then (S'', ℓ_2, ℓ_2') maps under a bounded distortion homeomorphism τ isotopic to the identity to good position with respect to $p(z)dz^2$. It follows that every point of $\tau \circ \varphi(\gamma)$ can be locked by a bounded segment of the stable foliation of $p(z)dz^2$ to $\tau \circ \varphi(\gamma_2 \cup \gamma_2')$. Therefore, by the method of 15.9, (5) holds contradicting (2) and hence (1) holds for some $[\varphi'] \in [y_0, y_2] \cup [y_1, y_2]$. □

15.16. Proof of 2 and 3 of 15.8: finding other almost-bounded loops

Assume without loss of generality that $[\varphi'] \in [y_0, y_2]$. Take any $[\psi] \in [y_0, [\varphi]]$ and γ' such that $\gamma' \subset \beta$ where β is m_0-Zero-Pole or long thick and dominant at β, $\beta \cap \alpha \neq \varnothing$ and $|\psi(\gamma')|$ is bounded. We are interested in the case $[\psi] = [\varphi]$, $\alpha = \beta$ to prove 2 of 15.8 and $[\psi] \subset [y_0, [\varphi]]$, $\alpha \cap \beta \neq \varnothing$ for 3 of 15.8. We no longer need α to be a gap — only that γ does not satisfy the criterion of 15.9 to be not almost bounded. We claim that $|\varphi''(\gamma')|$ is almost bounded for some $[\varphi''] \in [y_0, [\varphi']]$. We shall show this by a method similar to 15.15. We can assume that either $\alpha \subset \beta$ or $d([\psi], [\varphi])$ is large enough to ensure that (by 15.11 and the property $\alpha \cap \beta \neq \varnothing$) $\psi(\gamma) \cap \psi(\beta) \neq \varnothing$, that is, $\beta \cap \gamma \neq \varnothing$. Suppose, for contradiction, that $|\varphi''(\gamma')|'' \geqslant M_1$ for all $[\varphi''] \in [y_0, [\varphi']]$. Then by 15.9, if M_1 is sufficiently large given M, there are disjoint loops γ_3 and γ_4 such that

(1) $$|\varphi_0(\gamma_3)|' M \leqslant |\varphi_0(\gamma')|',$$

(2) $$(|\varphi'(\gamma_4)|' + |\varphi'(\gamma)|')M \leqslant |\varphi'(\gamma')|'.$$

Then $|\psi(\gamma_3 \cap \beta)|_q$ must be dominated by $|\varphi(\gamma_3 \cap \beta)|_{q,-}$. So $\varphi(\gamma_3)$ has long segments almost tangent to the stable foliation. Hence so does $\psi(\gamma_4)$, by 15.11, simply because γ_3 and γ_4 are disjoint. But $|\psi(\gamma \cap \beta)|_{q,+}$ is boundedly proportional to $|\varphi(\gamma)|_q$. Assume as usual that $\psi(\gamma')$, $\psi(\gamma)$ are in good position with respect to $q(z)dz^2$. We can find a union S_1 of topological discs in $S(\beta, [\psi])$ such that each topological disc in the union is bounded by stable segments and segments of $\psi(\gamma \cap \beta)$, and $\psi(\gamma') \subset S_1$. The stable segments can be taken as subsegments of a sufficiently long stable segment starting from any $x \in S(\beta, [\psi])$ by 15.11. This means that we can replace the stable segments by segments of $\psi(\gamma_4)$. So we have a union S_2 of topological discs bounded by segments of $\psi(\gamma)$ and $\psi(\gamma_4)$ with $\psi(\gamma') \subset S_2$. The topological discs in the union have bounded diameter. Then, as before, we can transfer the whole configuration to have good position with respect to the quadratic differential at $[\varphi]$ for $d([\psi], [\varphi'])$. Working as before we obtain, for a suitable C_1,

$$|\varphi'(\gamma')|' \leqslant C_1(|\varphi'(\gamma)|' + |\varphi'(\gamma_4)|'), \tag{3}$$

contradicting (2) if M is large enough. □

We are now in a position to apply the following.

15.17. Lemma. — *The following holds for a suitable function $C_2 : (0,1) \to (1,\infty)$, and for suitable long thick and dominant functions r, s, Δ, given functions C_0, C_1 : $(0,1) \to (1,\infty)$ and an integer N. Let Γ, Γ'_0, Γ'_1 be loop sets in a subsurface α such that:*

a) *$\alpha \setminus (\cup\Gamma \cup Z)$ is a union of discs, once-punctured discs and annuli parallel to $\partial\alpha$, and similarly for Γ'_0, Γ'_1, and similarly for $\{\gamma'_0, \gamma'_1\}$ where γ'_j is any loop in Γ'_j, $j = 0$, 1 and $\#(\gamma \cap \gamma') \leqslant N$ for any γ, $\gamma' \in \Gamma \cup \Gamma'_0 \cup \Gamma'_1$;*

b) *there are points $[\varphi'_0]$, $[\varphi]$, $[\varphi'_1]$ with $[\varphi] \in [[\varphi'_0], [\varphi'_1]$ such that α is long, ν-thick and dominant (for r, s, Δ) at all points of the geodesic segment $[[\varphi'_0], [\varphi'_1]]$ and such that*

$$|\varphi(\Gamma)| \leqslant C_0(\nu), \quad |\varphi'_j(\Gamma'_j)| \leqslant C_0(\nu),\ j = 0, 1;$$

c) *there is a geodesic segment $[y_0, y_2]$ such that, for any $\gamma \in \Gamma \cup \Gamma'_0 \cup \Gamma'_1$, there is some $[\varphi_\gamma] \in [y_0, y_2]$ such that $|\varphi_\gamma(\gamma)|'' \leqslant C_1(\nu)$.*

Then there is $[\varphi'] \in [y_0, y_2]$ such that

$$|\varphi'(\Gamma)|'' \leqslant C_2(\nu). \tag{1}$$

Proof. — It suffices to prove (1) for some $[\varphi']$, for Γ'_j replacing Γ, for one of $j = 0, 1$, because any loop of Γ is a union of finitely many bounded segments of Γ'_j, or for $\{\gamma'_0, \gamma'_j\}$ replacing Γ for one loop γ'_j from each Γ'_j, because any loop of Γ is a finite union of segments of γ'_0, γ'_1. So now suppose for contradiction that none of these possibilities occurs. Then we can choose $[\varphi'] \in [y_0, y_2]$, and loops $\gamma'_j \in \Gamma'_j$, such that $|\varphi'(\gamma'_j)|''$ are large for $j = 0$, 1, and such that there are points $[\varphi'_j]$ on opposite sides of $[\varphi']$ in $[y_0, y_2]$ such that $|\varphi'_j(\gamma'_j)|''$ are bounded. Presumably $[\varphi'_0] \in [y_0, [\varphi']]$ and $[\varphi'_1] \in [[\varphi'], y_2]$- and for convenience we assume this — but it does not matter if the opposite is the case. Let $q(z)dz^2$ be the quadratic differential for $d([\varphi'], y_2)$. By considering $\sigma^n_{\gamma'_1} \circ \sigma^m_{\gamma'_2}(\gamma'_1)$ for varying (but bounded) n and m, we can make a loop

$\gamma \subset \gamma_0' \cup \gamma_1'$ with bounded intersections with each such that if $\varphi'(\gamma)$ is taken in good position with respect to $q(z)dz^2$, then $\varphi'((\gamma_0' \cup \gamma_1') \cap \beta) \subset \varphi'(\gamma)$ for any subdominant β, and $\varphi'(\beta \cap \gamma) \neq \varnothing$ whenever $\varphi'(\beta \cap (\gamma_0' \cup \gamma_1')) \neq \varnothing$, if β is dominant. Every point of $\varphi'(\gamma_0')$ can be joined along a bounded stable segment to a different point of $\varphi'(\gamma)$ and every point of $\varphi'(\gamma_1')$ can be joined along a bounded unstable segment to a point of $\varphi'(\gamma)$. It follows that $|\varphi''(\gamma)|''$ cannot be bounded in $[[\varphi'], y_2]$, because $|\varphi''(\gamma_0'])|''$ is not bounded there, and similarly cannot be bounded in $[y_0, [\varphi']]$. Yet $|\varphi(\gamma)|$ is bounded. So we have a contradiction to 15.15.

15.18. Lemma. — *Under the same hypotheses as in 15.17, $|\varphi'(\Gamma)| \leqslant C(\nu)$. Moreover, we can assume that α is long thick and dominant at $[\varphi']$.*

Proof. — Since $\partial\alpha$ is in the convex hull of Γ, $|\varphi'(\partial\alpha)|'' \leqslant C_3(\nu)$ for a suitable function C_3. There is $[\varphi_0''] \in [y_0, [\varphi']]$ with $|\varphi_0''(\Gamma_0')|'' \leqslant C_2(\nu)$. So we also have $|\varphi_0''(\partial\alpha)|'' \leqslant C_2(\nu)$. Similarly, given an integer N_1, assuming the long thick and dominant parameters are suitably chosen, for $0 \leqslant j \leqslant N_1$, we can construct loop sets Γ_{-j}, $[\varphi'_{-(j+1)}] \in [[\varphi_0'], [\varphi'_{-j}]]$ and $[\varphi''_{-(j+1)}] \in [y_0, [\varphi''_{-j}]]$ such that $|\varphi'_{-j}(\Gamma_{-j})| \leqslant C_0(\nu)$, $|\varphi''_{-j}(\Gamma_{-j})|'' \leqslant C_2(\nu)$, and property a) of 15.17 holds. So then $|\varphi''_{-j}(\partial\alpha)|'' \leqslant C_3(\nu)$. By 15.9, $\partial\alpha$ does not intersect any long, thick and dominant gap or loop along $[[\varphi''_{-N_1}], [\varphi_0'']]$. By 15.14, there is a not necessarily connected union α_0 of gaps and loops, $\alpha_0 \supset \partial\alpha$, such that $\partial\alpha_0$ is bounded along $[[\varphi''_{-N_1}], [\varphi_0'']]$ and (2) of 15.14 holds for all $\gamma \subset \mathrm{Int}(\alpha_0)$. Then all Γ_k have boundedly many intersections with α_0. This is impossible unless $\partial\alpha \subset \partial\alpha_0$ up to homotopy. So $\partial\alpha$ is bounded along $[[\varphi''_{-N_1}], [\varphi_0'']]$. Then any long thick and dominant gap intersecting α along a segment of $[[\varphi''_{-N_1}], [\varphi_0'']]$ must be contained in α. Suppose that such a gap β is properly contained in α. Then it must be long, thick and dominant along $[[\varphi''_{-(k+N_2)}], [\varphi''_{-k}]]$ for N_2 arbitrarily large by suitable choice of the long, thick and dominant functions. Then $\#(\partial\beta \cap \Gamma_i)$ is bounded for $k \leqslant i \leqslant k + N_3$, which is impossible. So $\beta = \alpha$. There must be at least one long thick and dominant gap along $[[\varphi''_{-N_1}], [\varphi_0'']]$ by 15.14, if N_1 is sufficiently large. The only gap which can be long thick and dominant is α. So α is long thick and dominant all along $[[\varphi''_{-N_1}], [\varphi_0'']]$, and the loops of $\partial\alpha$ are bounded. In particular, we have $\varphi''(\Gamma'_{-j}) \leqslant C_4(N_1, \nu)$ for a suitable constant C_4. Since N_1 only has to be chosen suitably depending on ν, we obtain $|\varphi''_{-j}(\Gamma'_{-j})| \leqslant C(\nu)$ and $|\varphi'(\Gamma)| \leqslant C(\nu)$ for suitable choice of function C. □

15.19. Essential disjointness of ℓ_0, ℓ_1. — Suppose that $y = [\varphi]$, ℓ, α are as in 2 of 5.8 and we also have $w = [\psi] \in \ell$ and $w, y \in \ell_0 \cap \ell_1$, with $w \in [y_0, y]$. Let $y' = [\varphi']$, $w' = [\psi']$ be the corresponding points on $[y_0, y_2]$ and $y'' = [\varphi'']$, $w'' = [\psi'']$ the points similarly defined on $[y_1, y_2]$. Then we have $w' \in [y', y_2]$ and $y'' \in [w'', y_2]$, since we have already proved 3 of 15.8 in the gap case in 15.16. Fix loops $\gamma \subset \alpha$ and $\zeta \subset \beta$ such that $|\varphi'(\gamma)| \leqslant M$, $|\psi'(\zeta)| \leqslant M$ and similarly with φ'', ψ'' replacing φ', ψ'. By

(2) of 15.9 applied to $[y_0, y_1] = [y', y_2]$ we have

$$|\varphi_2(\zeta)|' \leqslant C_1(\nu)|\varphi_2(\gamma)| \frac{M}{|\psi'(\gamma)|'}$$

and applying to $[y_0, y_1] = [w'', y_2]$ we have

$$|\varphi_2(\gamma)|' \leqslant C_1(\nu)|\varphi_2(\zeta)| \frac{M}{|\varphi''(\zeta)|'}.$$

If $d(y, w)$ is sufficiently large we obtain a contradiction. □

15.20. Proof of 3 of 15.8 when β is a loop. — Let $y = [\varphi]$, $y' = [\varphi']$ α, w, β be as in 3 of 15.8. Fix γ intersecting β exactly twice, and $n([\psi])$ denote n which minimises $|\psi(\sigma_\gamma^n(\beta))|$. Then since $w \in [y_0, y]$ we have $n(w) \in [n(y_0), n(y)]$. Since $|\varphi(\partial\alpha)|$ is bounded, $n(y)$ is also the n such that the number of essential intersections of $\sigma_\gamma^n(\beta)$ with $\partial\alpha$ is minimal, and this is also $n(y')$ since $|\varphi'(\partial\alpha)|$ is bounded. So $n(w) \in [n(y_0), n(y')] \subset n(y_0, n(y_2)]$ and $w \in \mu_0$.

15.21. How to get short loops in the boundary. — We have the following extension of 15.14.

Lemma. — *There exist long thick and dominant parameter functions Δ, r, s, and $\Delta_0 > 0$ and an integer N_0 such that the following hold. Given $\varepsilon_1 > 0$, there exist $\Delta_1 > 0$ and an integer $N_1 > 0$ such that the following hold. Let $\ell \subset \mathcal{T}$ be a geodesic segment. Let γ be a loop which is disjoint from all (Δ, r, s)-long thick and dominant gaps and Δ_0-Pole-Zero loops along ℓ. Let γ be in the convex hull boundary of a set of N (Δ, r, s) long thick and dominant gaps and Δ_0-Pole-Zero loops β, where 1a) or 1b) holds, and 2a) or 2b) holds*

1a) $N \geqslant N_0$.

1b) *Any two of the β intersect.*

2a) *Each gap β is (Δ, r, s) long thick and dominant along a segment of ℓ of length $\geqslant \Delta_1$, and each loop β is a Δ_1-Pole-Zero loop along a segment of ℓ.*

2b) $N \geqslant N_1$.

Then $|\varphi(\gamma)| < \varepsilon_1$ for at least one $[\varphi] \in \ell$.

Proof. — Write $\ell = \{[\varphi_t] : t \in [0, T]\}$ where $d([\varphi_0], [\varphi_t]) = t$ and let $q_t(z)dz^2$ be the quadratic differential for ℓ at $[\varphi_t]$. For any loop γ', write

$$F(t, \gamma') = |\varphi_t(\gamma')|_{q_t}.$$

We are going to prove the Lemma by induction on the topological type of the convex hull of the β as in the statement of the lemma.

By 15.14, we know that there is $M > 0$ such that $|\varphi(\gamma)| \leqslant M$ along ℓ. This implies that, for any γ' which intersects γ or is not separated from γ by any loop of Poincaré length $< \varepsilon_0$,

$$(1) \qquad F(t,\gamma) \leqslant C_1 M F(t,\gamma'),$$

where $C_1 = C_1(\varepsilon_0)$ depends only on ε_0. To prove the lemma, we only need to strengthen this, and in fact, in one respect it suffices to prove a weaker statement. It suffices to show that, for suitable $\delta_1 > 0$ given $\varepsilon_1 > 0$ (in fact, $\delta_1 = e^{-k/\varepsilon_1}$ for some integer k will do), there is t such that

$$(2) \qquad F(t,\gamma) \leqslant \delta_1 F(t,\gamma')$$

for at least one γ' such that $\varphi_t(\gamma')$ has bounded Poincaré length and γ' is not separated from γ by any γ'' such that $\varphi_t(\gamma'')$ has short Poincaré length. Let t_0 be the essential minimum of $F(t,\gamma)$. We know that, for suitable C_1 (independent even of ε_0),

$$(3) \qquad \frac{ce^{|t-t_0|}}{C_1} \leqslant F(t,\gamma) \leqslant C_1 c e^{|t-t_0|}$$

for some c and t_0.

We can find m and $a_i < b_i$, $1 \leqslant i \leqslant m$, such that $a_{i+1} = b_i$ for $i < m$, $a_0 = 0$, $b_m < t_0$, and loops γ_i such that $|\varphi_t(\gamma_i) \leqslant C_1$ for $t \in I_i = [a_i, b_i]$ and γ_i is not separated from γ by any loop γ' such that $|\varphi_t(\gamma')| < \varepsilon_0$ for some $t \in I_i$. Then for $t \in [0,T]$, for suitable c_i and $t_i \in I_i$,

$$(4) \qquad \frac{c_i e^{|t-t_i|}}{C_1} \leqslant F(t,\gamma_i) \leqslant C_1 c_i e^{|t-t_i|},$$

and, for all γ_k with $\gamma_k \cap \gamma_i \neq \varnothing$, or γ_k separating γ_i from γ and $t \in I_i$,

$$(5) \qquad C_1^{-1} F(\gamma_i, t) \leqslant F(\gamma_k, t).$$

So now for $t \in I_k$, with $t \leqslant t_0$, we have

$$F(\gamma,t) \leqslant C_1 e^{t_i - t} F(\gamma, t_i) \leqslant C_1^2 e^{t_i - t} F(\gamma_i, t) \leqslant C_1^3 \frac{F(\gamma_i, t_i)}{F(\gamma_k, t_i} F(\gamma_k, t).$$

But also we have

$$\begin{aligned} C_1^2 e^{t_i - t} F(\gamma_i, t_i) &\leqslant C_1^3 e^{t_i - t + t_i - b_i} F(\gamma_i, b_i) \\ &\leqslant C_1^4 e^{2(t_i - b_i) + b_i - t} F(\gamma_k, b_i) \leqslant C_1^5 e^{2(t_i - b_i)} F(\gamma_k, t). \end{aligned}$$

So we obtain (2) if either

$$(6) \qquad C_1^5 e^{2(t_i - b_i)} \leqslant \delta_1$$

or

$$(7) \qquad C_1^4 \frac{F(\gamma_i, t_i)}{F(\gamma_k, t_i)} \leqslant \delta_1.$$

Of course, we obtain (6) if $b_i - t_i$ is sufficiently large for some i for which $\gamma_i \cap \gamma_k \neq \varnothing$ or γ_k separates γ_i from γ for some $k > i$. We obtain (7) for suitable i and k if m is sufficiently large, simply because γ_i and γ_k must then have a large number of intersections for some $i \neq k$, and if γ_i and γ_k have a sufficiently large number of intersections given δ_1, (7) must hold. Similarly we obtain (2) from analogues of (6) and (7) if there exist such intervals $I_i \subset [t_0, T]$ and loops γ_i such that either m is sufficiently large or $t_i - a_i$ is sufficiently large and there exists $k < i$ such that $\gamma_i \cap \gamma_k \neq \varnothing$ or γ_k separates γ_i from γ. If either of these occurs we shall say that γ is a δ_1-*good convex hull* of loops γ_i. So if γ is a δ_1-good convex hull of loops γ_i, $|\varphi_t(\gamma)| < \varepsilon_1$ for some t: in fact we can take t to be the essential minimum of $F(\gamma, t)$, which is a long way from the ends of the interval of s along which $|\varphi_s(\gamma)|$ is bounded. If $|\varphi_s(\gamma)|$ is small for $s = 0$ or T, then this statement is still true if we extend ℓ to a maximal segment ℓ' with $|\varphi(\gamma)| \leqslant M$ for $\varphi \in \ell'$.

So now suppose that γ_i satisfy (4) and (5), but that m is bounded. The proof is completed if, for at least one $i < k$, $|\varphi_t(\gamma_i)| < \varepsilon_1'$ for ε_1' sufficiently small and some $t \in I_i$, and $\gamma_i \cap \gamma_k \neq \varnothing$ or γ_k separates γ_i from γ. If $m = 1$, then $\gamma = \gamma_1$ must be the boundary of a long thick and dominant along most of $[[\varphi_0], [\varphi_T]]$, and the proof is finished. By abuse of notation we shall write $[0, T]$ instead of $[[\varphi_0], [\varphi_T]]$. So now suppose that $m > 1$ and $\gamma_i \neq \gamma$. This means that for any i the convex hull of long thick and dominant gaps and Pole-Zero loops along I_i is strictly smaller than along $[0, T]$, because there are none between γ and γ_i. For N_0 sufficiently large, and N_1, Δ_1 sufficiently large the hypotheses of the lemma are then satisfied for this smaller convex hull and any given $\varepsilon_1' > 0$ replacing δ_1 for some i with k existing as claimed. So then $|\varphi_t(\gamma_i)| < \varepsilon_1'$ for some $t \in I_i$ by the inductive hypothesis, and the proof is completed. □

15.22. We finish with a lemma which shows that for $\pi_\beta(z)$ to be within a bounded distance of $[y_0', y_1'] \cap \mathcal{T}(\partial\beta, \varepsilon_0)$ for points y_0', y_1', 15.9 gives a criterion which depends only on gaps intersecting β.

Lemma. — *The following holds for a suitable function $C_2 : (0,1) \to (1, \infty)$, given a constant $M_2 > 0$, and long thick and dominant parameter functions Δ, r, s and $m_0 > 0$ given Δ', r', s' and $m_0' > 0$. Let α_j be gaps or loops at $y_j = [\varphi_j] \in \mathcal{T}(Z)$, $j = 0$, 1. Let $y_j' = [\varphi_j']$, $j = 0$, 1 be any other points such that α_j is a loop or gap at y_j' and let $d_{\alpha_j}(y_j, y_j') \leqslant M_2$. Let $z = [\psi] \in [y_0, y_1]$, let β be long ν-thick and dominant at z with respect to Δ, r, s, or m_0-zero-pole with $\beta \cap \alpha_0 \neq \varnothing$, $\beta \cap \alpha_1 \neq \varnothing$. Then there is $z' \in [y_0', y_1']$ such that $d_\beta(z, z') \leqslant C_2(\nu)$ which long thick and dominant with respect to Δ', r', s' or m_0'-Zero-Pole.*

Proof. — If β is a loop we use 5.19 above. Now let β be a gap. Take any loop $\gamma \subset \beta$ such that $|\psi(\gamma)|$ is bounded. Then we shall use criterion (3) of 15.9 to show that γ is almost bounded at some point on $[y_0', y_1']$. Let γ_0' γ_1' be two disjoint loops which both intersect γ, and hence intersect β. Then because $\gamma_0 \cap \gamma_1 = \varnothing$ and β is long thick and dominant, $\psi(\gamma_0')$, $\psi(\gamma_1')$ are simultaneously boundedly transverse to either the stable or unstable manifold for the quadratic differential for $d(z, y_1)$. (This uses 15.11 — a loop which was close to the stable manifold would have to be long and then by 15.11 would intersect any loop close to the unstable manifold.) Suppose without loss of generality that $\psi(\gamma_0')$, $\psi(\gamma_1')$ are simultaneously boundedly transverse to the stable manifold. Then they intersect $\psi(\partial\alpha_1)$ transversely. Then $\psi(\gamma)$ is a union of finitely many segments with endpoints in $\psi(\partial\alpha_1)$, each of which can be joined along bounded stable segments to $\psi(\gamma_0')$, and similarly for γ_1'. This means that at y_1, $\varphi_1(\gamma)$ is a union of boundedly many arcs with endpoints in $\varphi_1(\partial\alpha_1)$, each arc bounding a rectangle with parallel side in $\varphi_1(\gamma_0')$ and adjacent sides in $\varphi_1(\partial\alpha_1)$. The same will be true for γ_1' replacing γ_0', and the same will be true for φ_1' replacing φ_1, simply by applying the homeomorphism $\varphi_1' \circ \varphi_1^{-1}$, which moves $\varphi_1(\partial\alpha_1)$ a bounded distance. So then by the criterion of 15.9, γ is almost-bounded at some point on $[y_0', y_1']$ (because (3) of 15.9 fails for γ). By 15.16, we obtain a point $[\psi']$ on $[y_0', y_1']$ where Γ is almost bounded, for any set of loops Γ which is bounded at $[\psi]$ Then by 15.18, we find that $\psi'(\Gamma)$ is bounded and β is long thick and dominant at $z' = [\psi']$. $\square$

CHAPTER 16

HARD SAME SHAPE

16.1. The main purpose of this chapter is to prove the Hard Same Shape Theorem. As one might expect, is a harder theorem with a similar conclusion to that of 9.5. It gives conditions under which quadratic differentials have the "same shape" on subsurfaces. Here, hard is a comparative term only, and it is by no means the hardest possible result of this nature. But such results are very delicate, because shape of a quadratic differential on subsurface of nonmaximal distance often changes drastically under small perturbations.

In 16.6 we show that the Teichmüller map is Hölder, a result which is most certainly known, but would be hard to reference, given our context of marked spheres. This result is needed in a small way in the proof of Hard Same Shape, but is also used to prove a result (16.8) about triangles of geodesics: a result which I suspect can be improved using the calculus of Teichmüller distance, and in particular the result 12.2 that Teichmüller distance is C^2, but I cannot currently see how to implement this improvement. However, a related result is proved in 16.9, using the same technique as in the proof of Hard Same Shape. The last results 16.10-12 (about the minimum of the mapping class function) are placed in this chapter for similar reasons: I suspect better estimates are possible using the calculus of Teichmüller distance. The calculation in the present proof bears some resemblance to that in 16.6. 16.12 also uses a concept from Chapter 14: otherwise the flavour of this chapter is closer to that of 10-13.

16.2. The Hard Same Shape Theorem. — We use the projections π_α of 9.1, and the identification of 9.1, if α is a loop, of $\pi_\alpha(\mathcal{T}(\partial\alpha, \varepsilon_0))$ with a subset of the upper half plane.

Hard Same Shape Theorem. — *Given $M > 0$, there are constants C_1, $L > 0$ such that the following holds for all sufficiently small ε and $\delta > 0$. Let $\alpha \subset \overline{\mathbf{C}}$ be a closed disc. For $j = 1, 2$, let $Y_j \subset \overline{\mathbf{C}} \setminus (\partial\alpha)$ be finite, with $Y_1 \cap \alpha = Y_2 \cap \alpha = Y_0$, $\#(Y_0) \geqslant 2$. Let*

$$[\varphi_j],\ [\psi_j] \in \mathcal{T}(\partial\alpha, \varepsilon) \setminus \cup\{\mathcal{T}(\gamma, \nu) : \gamma \subset \mathrm{Int}(\alpha)\} \subset \mathcal{T}(Y_j, \varepsilon), \quad \nu \geqslant L\varepsilon.$$

Let $q_j(z)dz^2$ be the quadratic differential for $d([\varphi_j],[\psi_j])$ at $[\varphi_j]$, with stretch $p_j(z)dz^2$ at $[\psi_j]$. Let $[\psi_j]=[\chi_j\circ\varphi_j]$ with χ_j minimizing distortion. Choose a point $y_0\in Y_0$ and another point $y_j\in Y_j\smallsetminus Y_0$. Normalise so that $\varphi_j(y_0)=\psi_j(y_0)=0$, $\varphi_j(y_j)=\psi_j(y_j)=\infty$. Write $z_{k,j}$ for the zeros of q_j with $|z_{k,j}|\geqslant|z_{k+1,j}|$. Let $n_j=\#(Y_j\smallsetminus Y_0)-2$. Write

Case 1:
$$q_j(z)=\frac{\prod_{k=1}^{n_j}(1-z/z_{k,j})}{\prod_{y\in Y_j\smallsetminus Y_0}(1-z/\varphi_j(y))}\sum_{y\in Y_0}\frac{a_{y,j}}{z-\varphi_j(y)},$$

Case 2:
$$q_j(z)=\frac{\prod_{k=1}^{n_j+1}(1-z/z_{k,j})}{\prod_{y\in Y_j\smallsetminus Y_0}(1-z/\varphi_j(y))}z^{-1}\sum_{y\in Y_0\smallsetminus\{y_0\}}\frac{a_{y,j}}{z-\varphi_j(y)}.$$

Normalise further so that for some real λ, $0<\lambda\leqslant 1$,

Case 1:
$$\lambda\sum_{y\in Y_0}a_{y,1}=\sum_{y\in Y_0}a_{y,2},$$

and the points of $\varphi_1(Y_0)$ are bounded and bounded apart, and similarly for $Y_0\smallsetminus\{y_0\}$ replacing Y_0 in Case 2.

Let
$$|z_{k,j}|\geqslant M^{-1}e^{2\pi^2/\varepsilon},$$
for $1\leqslant k\leqslant n_j$ or $1\leqslant k\leqslant n_j+1$ depending on whether we are in Case 1 or 2. Write

Case 1:
$$Q_j(z)=\sum_{y\in Y_0}\frac{a_{j,y}}{z-\varphi_j(y)},$$

Case 2:
$$Q_j(z)=z^{-1}\sum_{y\in Y_0\smallsetminus\{y_0\}}\frac{a_{j,y}}{z-\varphi_j(y)}.$$

Let Q_1 have no zeros in $\{z: e^{2\pi^2(1-1/M)/\varepsilon}<|z|\}$. Let
$$|d([\varphi_2],[\psi_2])-d([\varphi_1],[\psi_1])|\leqslant\delta,\quad d([\varphi_1],[\psi_1])\leqslant M,$$
$$|\varphi_1(y)-\varphi_2(y)|\leqslant\delta,\quad |\psi_1(y)-\psi_2(y)|\leqslant\delta\text{ for all }y\in Y_0.$$
Let $\chi_2=\xi'\circ\chi_1\circ\xi$ on $\{z:|z|\leqslant e^{2\pi^2(1-1/2M)/\varepsilon}\}$, where the homeomorphisms ξ, ξ' are isotopic to the identity via isotopies almost constant on the boundary and on $\varphi_2(Y_0)$, $\psi_1(Y_0)$ respectively. Let
$$A=\sum_{y\in Y_0}|a_{y,1}|\ (=O(a(\alpha,q_1)).$$
Write $B=\log(|z_{n_j+1,1}|+2)$ in Case 1, and $B=1$ in Case 2. Then
$$a_{y,2}=\lambda a_{y,1}+O((\delta^{1/C_1}+e^{-1/C_1\varepsilon})\lambda AB)\quad\text{for all }y\in Y_0.$$
Similar results hold for p_1, p_2.

16.3. Remarks

(1) In applications, $\alpha \subset \overline{\mathbf{C}}$ will often be a subsurface with more than one boundary component — but an application of the easier Same Shape result 9.5 will allow us to reduce to the case of a disc.

(2) The condition on isotopy between χ_1 and χ_2 might seem awkward and unnatural, but it is the right condition for applications. Roughly speaking, it says that projections under π_α, $\pi_{\partial\alpha}$ are close — but since in some applications Y_1 and Y_2 are completely unrelated, there may in general be no other way to make sense of the statement.

(3) The difference $a_{y,2} - a_{y,1}$ is small if, for example, $\delta = O(e^{-2\pi^2/\varepsilon})$.

16.4. Outline of Proof of the Hard Same Shape Theorem. — By (easy) Same Shape 9.5, there is nothing to prove unless

$$d_\alpha([\varphi_1],[\psi_1]) \leqslant d([\varphi_1],[\psi_1]) - \delta_1 \tag{1}$$

for $\delta_1 \geqslant e^{-C_0/\varepsilon}$. So from now on we assume that (1) holds. Recall that $q_j(z)dz^2$, $p_j(z)dz^2$ and K_j with $\frac{1}{2}\log K_j = d([\varphi_j],[\psi_j])$, are uniquely determined by solving

$$\int_{\varphi_j(\gamma)} \sqrt{q_j}dz - \begin{pmatrix} 1/\sqrt{K_j} & 0 \\ 0 & \sqrt{K_j} \end{pmatrix} \int_{\psi_j(\gamma)} \sqrt{p_j}dz = 0, \tag{2}$$

for all nontrivial nonperipheral loops $\gamma \subset \overline{\mathbf{C}} \setminus Y_j$. This, of course, is an equation we have considered extensively in Chapters 10-13. For such loops, we can expand the integrals for $\sqrt{q_2}$, $\sqrt{p_2}$ in terms of $\sqrt{q_1}$, $\sqrt{p_1}$, and thus we shall see that q_2 must be close to q_1, and p_2 to p_1. Given the representation of q_1 in the statement of Hard Shape in 16.2, it is natural to consider the function Q_j of 16.2, and the surface

$$S_j = \{(z,w) : Q_j(z) = w^2\}.$$

By our assumptions in 16.2, the numerator of Q_j is of degree $\#(Y_0)-1$, $\#(Y_0)-2$ in Cases 1, 2 respectively, and thus there are $\#(Y_0)-1$, $\#(Y_0)-2$ zeros respectively up to multiplicity. If all of these zeros are simple, then S_j has genus $\#(Y_0)-1$, $\#(Y_0)-2$ respectively. The condition (1) above ensures that in Case 1 there is at most one zero z_0' of q_1 in

$$\{z : D_0 \leqslant |z| \leqslant M^{-1}e^{2\pi^2/\varepsilon}\}$$

for suitable D_0 — and none at all in Case 2 — because otherwise some set $\{z : |z| \leqslant D_0'\}$ satisfies the Dominant Area Condition 9.4: a calculation would show that $\{z : r/2 \leqslant |z| \leqslant r\}$ had area $O(Ar^{-2})$. Then Same Shape 9.5 would contradict (1). So the surface S_j might have unbounded geometry, but, as we shall see, this is somewhat controlled. There are $\#(Y_0)$ variables $a_{j,y}$ with one relation between them given in 16.2. We shall see that these are determined uniquely by (2) above.

16.5. Solving the Equations. — With the calculations of 10.18 onwards in mind, write

$$\varphi_2(y) = \varphi_1(y) + h(y), \quad a_{y,2} = \lambda(a_{y,1} + k(y)).$$

Then for any closed loop $\gamma \subset \alpha$,

$$\int_{\varphi_j(\gamma)} \sqrt{q_j(z)} = \int_{\varphi_j(\gamma)} \sqrt{Q_j(z)} + O(Ae^{-2\pi^2/\varepsilon}),$$

$$\int_{\varphi_2(\gamma)} \sqrt{Q_2(z)} = \sqrt{\lambda} \int_{\varphi_1(\gamma)} \Big(\sqrt{Q_1(z)} + \sum_{y \in Y_0} \frac{\frac{1}{2}k(y)}{(z - \varphi_1(y))\sqrt{Q_1(z)}} - \sum_{y \in Y_0} \frac{\frac{1}{2}h(y)a_{y,1}}{(z-\varphi_1(y))^2\sqrt{Q_1(z)}} + O(A\|\underline{h}\|^2) + O(\|\underline{k}\|^2)\Big),$$

and similarly for $p_j(z)dz^2$, with Y_0 replaced by $Y_0 \setminus \{y_0\}$ in Case 2. Assume for the moment that all zeros of Q_1 are simple and distance at least δ_1 apart, with residues of modulus at least δ_1 at points $\varphi_1(y)$ ($y \in Y_0$ or $y \in Y_0 \setminus \{y_0\}$ in cases 1 and 2 respectively). Using the method of computation of the Second Derivative Formula (from 10.18 onwards through Chapter 11) we choose a basis of holomorphic 1-forms on S_j. Letting $\pi(z, w) = z$ be the usual projection, the natural choice is

Case 1: $$\pi^*\Big(\frac{dz}{\sqrt{Q_j(z)}}\Big(\frac{1}{z - \varphi_j(y)} - \frac{1}{z}\Big)\Big) \quad (y \in Y_0 \setminus \{y_0\})$$

Case 2: $$\pi^*\Big(\frac{dz}{\sqrt{Q_j(z)}}\Big(\frac{1}{z - \varphi_j(y)} - \frac{1}{z - \varphi_j(y_1)}\Big)\Big) \quad (y \in Y_0 \setminus \{y_0, y_1)$$

These are indeed nonsingular at ∞, since Q_j has one more pole than zero. We also have meromorphic 1-forms

$$\pi^*\Big(\frac{a_{y,j}dz}{\sqrt{Q_j(z)}}\Big(\frac{1}{(z - \varphi_j(y))^2}\Big)\Big)$$

for $y \in Y_0$ in Case 1 and $y \in Y_0 \setminus \{y_0\}$ in Case 2. Write

$$Y_0 \setminus \{y_0\} = \{y_\ell : 1 \leqslant \ell \leqslant m = \#(Y_0) - 1\}.$$

Then let $v_{2\ell-1} + iv_{2\ell}$, $w_{2\ell-1} + iw_{2\ell}$ be the holomorphic and meromorphic 1-forms defined above for $y = y_\ell$, on S_1, with v_i harmonic and w_i harmonic singular. As in Chapter 11, let J, J' send a harmonic 1-form on S_1, S_2 to its conjugate. Write $h_\ell = h(y_\ell)$, $k_\ell = k(y_\ell)$. Again working in parallel with Chapter 11, write

$$x + iJx = \sum_{\ell=1}^{m} k_\ell(v_{2\ell-1} + iv_{2\ell}), \quad w + it = \sum_{\ell=1}^{m} h_\ell(w_{2\ell-1} + iw_{2\ell}),$$

with $\ell = 1$ replaced by $\ell = 2$ in the expression for $x + iJx$ in Case 2. Then define v_i', w_i', x', w', t' in analogy with Chapter 11, using the holomorphic and meromorphic

1-forms on S_2 described above. Then to show that $a_{j,2}$ is close to $a_{j,1}$, it suffices to show that we can solve the following equations in $H^1(S_1)$:

$$(1)\quad \begin{pmatrix} x \\ Jx \end{pmatrix} - \begin{pmatrix} K_1^{-1/2}x' \\ K_1^{1/2}x' \end{pmatrix} = -\begin{pmatrix} w \\ t \end{pmatrix} + \begin{pmatrix} K_1^{-1/2}w' \\ K_1^{1/2}t' \end{pmatrix} + O(\delta) + O(e^{-2\pi^2/\varepsilon}) + O(A\|\underline{h}\|^2).$$

This mimics the equation (SDE2) of 10.18, but there are some differences. The most important is that this is the only equation. There is no additional requirement that $|Q_2|$ have integral 1: of course $|q_2|$ has integral one, but we are only considering $q_2 \mid S(\alpha)$, so the mass condition does not appear. The method of solution of (1) still mimics the solution in 11.2. We write the righthand side of (1) as a sum of terms and solve each set of equations separately. Each time we have to show that we can solve with

$$\underline{k} = O((\|\underline{h}\| + \delta e^{-2\pi^2/\varepsilon})ABe^{C_0/\nu}\delta_1^{-C_0''})$$

for suitable C_0, C_0''. We write $w = w^{(1)} + w^{(2)}$ as in 11.2, and similarly for w'. Following the derivation in 11.4, we obtain equations such as

$$(K_1 + JJ')x = (K_1 - JJ')w^{(2)}.$$

We claim that, for suitable C_0',

$$(K_1 + JJ')^{-1} = O(B\delta_1^{-C_0'}).$$

We use a method based on that of 11.9. We need to show that for any harmonic 1-form on S_1,

$$\omega \cup J'\omega \leqslant (K_1 - CB^{-1}\delta_1^{C_0'})(\omega \cup J\omega).$$

As in 11.9, we consider $(\omega + iJ\omega)(z)$ and $\sqrt{Q_1(z)}dz$ pointwise, and use the complex numbers $\sqrt{Q_1(z)}$, $i\sqrt{Q_1(z)}$ as a basis of $\mathbf{C}$ over the reals. It suffices to show that $\omega + iJ\omega$ is bounded from being a real multiple of $\sqrt{Q_1(z)}dz$ on a set $E = E(\omega)$ with

$$(\omega \wedge J\omega)(E) \geqslant B^{-1}\delta_1^{C_0'}(\omega \wedge J\omega)(S_1).$$

Our assumptions give that all zeros of Q_1 are in $\{z : |z| \leqslant D_0\}$ for a suitable $D_0 = \delta_1^{-r}$, $r \leqslant \#(Y)$, and:

$$\textit{Case 1:}\quad \alpha = \sum_{y\in Y_0} a_{y,1}, \quad Q_1(z) = \begin{cases} \alpha z^{-1}(1 + O(e^B z^{-1})), & |z| \geqslant 2e^B \\ \alpha z_{n_1+1} z^{-2}(1 + O(D_0 z^{-1})), & 2D_0 \leqslant |z| \leqslant \frac{1}{2}e^B, \end{cases}$$

$$\textit{Case 2:}\qquad Q_1(z) = \alpha z^{-2}(1 + O(z^{-1})), \quad |z| \geqslant D_0, \quad \alpha = \sum_{y\in Y_0 \smallsetminus \{y_0\}} a_{y,1}.$$

Now let ω be given. Note that any $\omega + iJ\omega$ is $\pi^*(\sqrt{R(z)dz})$ where R is a rational function which has a pole at z_{n_j+1} in Case 1. Let D_1 be the maximum modulus of a zero of R. Take

$$E_1 = \{z : e^B \leqslant |z| \leqslant 2e^B\}, \quad E_2 = \{z : 2D_0 \leqslant |z| \leqslant \tfrac{1}{2}D_1\}, \quad E = E_1 \cup E_2.$$

We have $E_2 \neq \varnothing$, so $E \neq \varnothing$. For a constant c_R depending on R, we have

Case 1: $$R(z) = \alpha^{-1} c_R z^{-2}(z - z_{n_{j+1}})^{-1}, \quad |z| \geqslant 2D_1$$

Case 2: $$R(z) = \alpha^{-1} c_R z^{-3}, \quad |z| \geqslant 2D_1.$$

On E_1 (which may be empty) R is bounded from being a real multiple of Q_1: except on finitely bounded modulus subannuli in E_1, R is asymptotic to $\alpha^{-1} c_{R,j} z^{-j}$ for on different annuli, for varying $j \geqslant 3$. If E_1 is of large modulus, the integral of $|R|$ over E_1 is at least a bounded multiple of the integral over $\{z : \frac{1}{2}D_0 \leqslant |z| \leqslant D_0\}$ which is at least D_0^t ($t \leqslant \#(Y)$) times the integral over $\{z : |z| \leqslant D_0\}$ The integral of $|R|$ over any annulus

$$\{z : r_1 \leqslant |z| \leqslant r_2\} \quad \text{in} \quad 2D_1 \leqslant r_1 < r_2 - 1 \leqslant r_2 \leqslant e^B$$

is then boundedly proportional to $|\alpha^{-1} c_R| \log(r_2/r_1)$. The integral over $|z| \geqslant e^B$ is $O(\alpha^{-1} c_R)$. So E has the required properties.

Then to obtain the required bound on $\underline{k}$, we also need to bound the coefficients of the v_n when expressing any w_ℓ as a linear combination of the v_n: we can derive the coefficients of $w_{2\ell-1}^{(j)}$, $w_{2\ell}^{(j)}$ ($j = 1, 2$) from those of $w_{2\ell-1}$, $w_{2\ell}$. The simplest way to do this is probably to use the harmonic functions approach of 12.7. According to this approach, write $v_{2\ell-1} + iv_{2\ell} = \pi^*(\sqrt{R_\ell(z)}dz)$. Let r_0 (as in 12.7) be the radius of a maximal embedded disc in S_1 round $\pi^{-1}(\varphi_1(y))$ in the coordinate

$$\zeta(\pi^{-1}(z)) = \int_{\varphi_1(y)}^{z} \sqrt{R_\ell(t)}dt.$$

Then the bound on coefficients is $O(r_0^{-1}) = O(Ae^{C_0/\nu})$ for suitable $C_0 > 0$, taking into account the distance between the points of $\varphi_1(Y_0)$. If we allow zeros of Q_1 to approach within $O(\delta_2)$, or residues to be $O(\delta_2)$ then the method still works with an increased error term. We shall deal with this in 16.7 below.

16.6. The Teichmüller Map is Hölder with respect to the Teichmüller metric. — The *Teichmüller map* is the map which sends $(q(z)dz^2, K)$ to $[\psi] \in \mathcal{T}(Y)$, where $d([\varphi], [\psi]) = \frac{1}{2}\log K$ and $q(z)dz^2$ is the quadratic differential at $[\varphi]$ for $d([\varphi], [\psi])$. One of the consequences of Teichmüller distance d being C^2 (12.1) is that the *inverse* of the Teichmüller map is C^1. We also know (from 10.2) that the Teichmüller map is real analytic at a generic point. It is not globally C^1 because the inverse map is C^2. However, it is Hölder with respect to the Teichmüller metric on the range. Here is a proof.

Lemma. — *Let $q_1(z)dz^2$, $q_2(z)dz^2$ denote quadratic differentials at $[\varphi] \in (\mathcal{T}(Y))_{\geqslant \nu}$ for $d([\varphi], [\psi_1])$, $d([\varphi], [\psi_2])$ respectively, where $d([\varphi], [\psi_1]) = d([\varphi], [\psi_2]) \leqslant M$. Let the residues of q_1, q_2 agree to within $O(\delta)$. Then for a constant C_0 depending only on $\#(Y)$,*

$$d([\psi], [\psi']) = O(\delta^{1/C_0} e^{C_0/\nu}).$$

Proof. — Let $[\psi_j] = [\chi_j \circ \varphi]$, where χ_j minimizes distortion. By 8.3 we only need to show that

$$\int K(\chi_2 \circ \chi_1^{-1})|r| = O(\delta^{1/C_0} e^{C_0/\nu}), \tag{1}$$

where $r(z)dz^2$ is the quadratic differential for $d([\psi_1], [\psi_2])$ at $[\psi]$. A direct computation (similar to those employed, for instance, in 8.3, 8.9) shows that the distortion is

$$K(\chi_2 \circ \chi_1^{-1})(z) = 1 + O((\theta(\chi_1^{-1}(z)))^2,$$

where θ is the angle between $q_1(z)dz^2$ and $q_2(z)dz^2$. An application of Rouché's Theorem

$$\frac{1}{2\pi i} \int_\alpha \frac{q_j'}{q_j} = \#(\text{zeros inside } \Gamma)$$

gives that q_1 and q_2 have the same number of zeros inside any contour Γ on which $|q_1| \geqslant R\delta$ for R sufficiently large. In particular, there is a pairing of the zeros of q_1 and q_2 (up to multiplicity) such that any pair is $O(\delta^{1/k})$ apart, where $k = \#(Y) - 4$ is the total number of zeros of q_1 up to multiplicity. Outside discs of radius $R\delta^{1/k}$ round zeros of q_1,

$$(\theta(z))^2 \leqslant C|\theta(z)| \leqslant \frac{\delta}{|q_1(z)|}.$$

Changing coordinates if necessary, assume that $z_1 = 0$. We consider the integral near 0. Number the zeros z_j so that $|z_j| \leqslant |z_{j+1}|$. Fix $1 \leqslant m \leqslant k$ so that, for some ρ, $|z_m| \leqslant \rho < |z_{m+1}|$,

$$\rho^m \prod_{j=m+1}^{k} |z_j| = \delta.$$

Then $\rho \leqslant \delta^{1/k}$. Write

$$\beta = \prod_{j=m+1}^{k} |z_j|.$$

To show (1), we need to show that (remembering that r has at most simple poles)

$$\int_{\rho \leqslant |z|} \delta |z|^{-(m+1)} \beta^{-1} dxdy = O(\delta^{1/k}),$$

where $z = x + iy$ and $dxdy$ is the usual planar measure. But the integral for $\ell = m$ is $O(\delta \rho^{1-m} \beta^{-1}) = O(\delta^{1/k})$. So the estimate holds. □

16.7. The case of Multiple Zeros or Zero Residues. — We return to the proof of the Hard Same Shape Theorem. We consider the case of q_1 having approximately multiple zeros, or approximately zero residues at points of $\varphi_1(Y_0)$. It seems to be impossible to treat this case by the direct method above. Instead we need to consider the maps

$$([\varphi_j], [\psi_j]) \longmapsto (q_j, p_j), \quad j = 1, 2.$$

As pointed out in 16.5, these maps are C^1. Therefore, if P_j is defined similarly to Q_j in 16.2, the maps

$$\Theta_j : ([\varphi_j],[\psi_j]) \longrightarrow (Q_j, P_j)$$

($j = 1, 2$) are C^1. Consider the following statements (1) and (2)

$$(1) \qquad \Theta_2([\varphi_2],[\psi_2]) = \Theta_1([\varphi_1],[\psi_1]) + O((\delta + e^{-2\pi^2/\varepsilon})ABe^{C_0/\nu}\delta_1^{-C_0''}\delta_2^{-C_0}),$$

$$(2) \qquad \varphi_2(y) - \varphi_1(y) = O(\delta), \quad \psi_2(y) - \psi_1(y) = O(\delta) \text{ for all } y \in Y_0$$

(for a suitable normalisation). Then (1) holds for $([\varphi_2],[\psi_2])$ as in (2) and for $([\varphi_1],[\psi_1])$ varying over a set in which zeros of Q_1 are bounded apart by δ_2, residues at points of $\varphi_1(Y_0)$ have modulus at least δ_2, and $([\varphi_1],[\psi_1])$ satisfying the hypotheses of 16.2. Then by 16.6, this set of $([\varphi_1],[\psi_1])$ intersects a δ_2^{1/C_0}- neighbourhood of every point $([\varphi_1],[\psi_1])$ satisfying the hypotheses of 16.2: regarding this set as a subset of $\mathbf{C}^{Y_0\smallsetminus\{y_0\}} \times \mathbf{C}^{Y_0\smallsetminus\{y_0\}}$. But these maps are C^1 and the derivatives are $O(ABe^{\#(Y_0)/\nu})$. So on this set also we have

$$\Theta_2([\varphi_2],[\psi_2]) = \Theta_1([\varphi_1],[\psi_1]) + O((\delta + e^{-2\pi^2/\varepsilon})ABe^{C_0/\nu}\delta_1^{-C_0''}\delta_2^{-C_0}) + O(ABe^{C_0/\nu}\delta_1^{-C_0''}\delta_2^{1/C_0}).$$

Given the bound on δ_1, ν, for suitable δ_2 and C_1, we can make this

$$\Theta_2([\varphi_2],[\psi_2]) = \Theta_1([\varphi_1],[\psi_1]) + O((\delta^{1/C_1} + e^{-1/C_1\varepsilon})AB),$$

as required.

16.8. Corollary of 16.6. — *Let $[\varphi]$, $[\xi]$, $[\psi] \in (\mathcal{T}(Y))_{\geqslant\nu}$. Let $q_1(z)dz^2$, $q_2(z)dz^2$ be the quadratic differentials at $[\varphi]$, $[\xi]$ for $d([\varphi],[\xi])$ and $[d([\xi],[\psi])$. Let $p_1(z)dz^2$ be the stretch of $q_1(z)dz^2$ at $[\xi]$. Let the residues of p_1 and q_2 agree to within $O(\delta)$. Then*

$$d([\varphi],[\psi]) - (d([\varphi],[\xi]) + d([\xi],[\psi]) = O((\delta^{1/C_1}e^{C_1/\nu}).$$

Proof. — Let ψ_2 be such that $d([\xi],[\psi_2]) = d([\xi],[\psi])$ and $p_1(z)dz^2$ is the quadratic differential at $[\xi]$ for $d([\xi],[\psi_2])$. Then by 16.6,

$$d([\psi],[\psi_2]) = O((\delta^{1/C_1}e^{C_1/\nu}).$$

The result follows. □

16.9. An infinitesimal triangular equality. — 16.8 can be regarded as a reverse of the Triangular Lemma 8.9. It is possible to prove it by a method resembling that of 16.2. Here, we prove an infinitesimal version, in a slightly different context — outside $(\mathcal{T}(Y))_{\geqslant\nu}$. We shall use this in Chapter 18.

Lemma. — *The following holds for a suitable constant $C > 0$ given M. Let $\alpha \subset \overline{\mathbf{C}}$, Y_0, Y_1 be as in 16.2. For $1 \leqslant i \leqslant 3$, let*

$$[\varphi_i] \in \mathcal{T}(\partial\alpha,\varepsilon) \smallsetminus \cup\{\mathcal{T}(\gamma,\nu) : \gamma \cap \alpha = \varnothing\} \subset \mathcal{T}(Y_1).$$

Let $\delta \leqslant e^{-C_0/\nu}$. For $i<j$, let $q_{i,j}(z)dz^2$ be the quadratic differentials for $d([\varphi_i],[\varphi_j])$ at $[\varphi_i]$, with stretch $p_{i,j}(z)dz^2$ at $[\varphi_j]$. Normalise φ_i as in 16.2, with corresponding normalisations of $q_{i,j}(z)dz^2$, $p_{i,j}(z)dz^2$ $(i<j)$. Let the zeros of $q_{i,j}$ $(i<j)$ have the same properties as in 16.2. For $i<j$, let

$$d_\alpha([\varphi_i],[\varphi_j]) \leqslant d([\varphi_i],[\varphi_j]) - M^{-1} \leqslant M - M^{-1},$$

$$d([\varphi_1],[\varphi_2]) + d([\varphi_2],[\varphi_3]) - d([\varphi_1],[\varphi_3]) \leqslant \delta.$$

Let A and B be defined as in 16.2, but using $q_{1,2}$ instead of q_1. Let

$$\operatorname{Res}(q_{2,3} - p_{1,2}, \varphi_2(y)) = O(\delta A) \quad (y \in Y_0).$$

Then

$$\operatorname{Res}(q_{1,3}, \varphi(y)) = \operatorname{Res}(q_{1,2}, \varphi_1(y)) + O((\delta^{1/C} + e^{-1/C\varepsilon})AB),$$

and similarly for $p_{1,3}$, $p_{2,3}$.

Proof. — The method is very similar to that of 16.2. By the Triangular Lemma 8.9.

$$\int \theta(q_{1,2}, q_{1,3})^2 |q_{1,3}| = O(\delta).$$

Then (easy) Same Shape 9.5 implies that residues $q_{1,2}$ and $q_{1,3}$ are within $O(\delta^{1/C_0'})$ on $S(\overline{\mathbf{C}}\smallsetminus\alpha, [\varphi_1], \nu)$ (normalising so that points of $\varphi_1(A(\overline{\mathbf{C}}\smallsetminus\alpha))$ are bounded and bounded apart). Write $\frac{1}{2}\log K = d([\varphi_1],[\varphi_3])$, $\frac{1}{2}\log K_1 = d([\varphi_1],[\varphi_2])$, $\frac{1}{2}\log K_2 = d([\varphi_2],[\varphi_3])$. Then our hypotheses imply that for $\gamma \subset \alpha$,

$$\int_{\varphi_1(\gamma)} \sqrt{q_{1,2}} - \begin{pmatrix} 1/\sqrt{K_1K_2} & 0 \\ 0 & \sqrt{K_1K_2} \end{pmatrix} \int_{\varphi_3(\gamma)} \sqrt{p_{2,3}} \\ - \begin{pmatrix} 1/\sqrt{K_1} & 0 \\ 0 & \sqrt{K_1} \end{pmatrix} \Big(\int_{\varphi_2(\gamma)} \sqrt{p_{1,2}} - \sqrt{q_{2,3}} \Big) = O(\delta),$$

and hence

$$\int_{\varphi_1(\gamma)} \sqrt{q_{1,2}} - \begin{pmatrix} 1/\sqrt{K} & 0 \\ 0 & \sqrt{K} \end{pmatrix} \int_{\varphi_3(\gamma)} \sqrt{p_{2,3}} = O(\delta).$$

Define $Q_{i,j}$ from $q_{i,j}$ in the same way as Q_j is defined from q_j in 16.2. In Case 1 of 16.2, the corresponding coefficients $a_{i,j,y}$ differ from the residues by $O(e^{-1/M\varepsilon})$. In Case 2, the $a_{i,j,y}(\varphi_j(y))^{-1}$ differ from the residues by $O(e^{-2\pi^2/\varepsilon})$. Let

$$S_{i,j} = \{(z,w) : Q_{i,j}(z) = w^2\}.$$

Then, again in analogy, we obtain, for $\gamma \subset \alpha$,

$$\int_{\varphi_1(\gamma)} \sqrt{Q_{1,2}} - \begin{pmatrix} 1/\sqrt{K} & 0 \\ 0 & \sqrt{K} \end{pmatrix} \int_{\varphi_3(\gamma)} \sqrt{P_{2,3}} = O(\delta) + O(Ae^{-2\pi^2/\varepsilon}).$$

Then $Q_{1,3}$, $P_{1,3}$ are those perturbations of $Q_{1,2}$, $P_{2,3}$ which give equality in the above. Working as in 16.4-5, we obtain $Q_{1,3}$, $P_{1,3}$ with the required bounds on $Q_{1,3} - Q_{1,2}$, $P_{1,3} - P_{2,3}$ if the zeros of $Q_{1,2}$ are bounded apart and residues bounded from 0. We obtain the general case by arguing as in 16.6-7. ☐

16.10. The infimum for the mapping class map. — We now give two results which we shall need later and which concerns the map

$$G : [\varphi] \longmapsto d([\varphi], [\varphi \circ \psi]) : \mathcal{T}(Y) \longrightarrow \mathcal{T}(Y)$$

where $[\psi] \in \mathrm{MG}(\overline{\mathbf{C}}, Y)$ (the mapping class group, as in Chapter 1). The calculus of G was studied by Bers [**Bers**]. As we know, if the isotopy class of $[\psi]$ is pseudo-Anosov, then G has a non-zero minimum value, achieved uniquely on a geodesic in $\mathcal{T}(Y)$ invariant under $[\varphi] \mapsto [\varphi \circ \psi]$. If $[\psi]$ is reducible but has a pseudo-Anosov component, then the infimum of G is non-zero. In both cases, we call the infimum $\kappa([\psi])$. Now let $r(z)dz^2$ be the quadratic differential at $[\varphi]$ for $d([\varphi], [\varphi \circ \psi])$ with stretch $t(z)dz^2$ at $[\varphi \circ \psi]$. Then as we also know from 8.4,

$$DG = 2\pi \operatorname{Re}(\operatorname{Res}(r - t, \varphi(y)))$$

(in suitable coordinates). One can then ask about the size of $\|DG\|$ when $G - \kappa([\psi])$ is small. If G were real analytic, we would have $G - \kappa([\psi]) = o(\|DG\|)$, but it is unclear if this is the case. So it is unclear if, in the pseudo-Anosov case, one can find a path of finite length from a given point along which G decreases to the minimum value. Our first lemma says that we can at least bound the diameters of such paths if, for example, we start near a minimum value. We use the local coordinates on $\mathcal{T}(Y)$ introduced in 8.4.

16.11. Lemma. — *Let $Y = \{y_i : 1 \leqslant i \leqslant n\}$. Let $\psi : (\overline{\mathbf{C}}, Y) \to (\overline{\mathbf{C}}, Y)$ be pseudo-Anosov. The following holds for $\delta > 0$ sufficiently small. Let $[\varphi_0] \in \mathcal{T}(Y)$ be a point on the minimizing geodesic for $G([\varphi]) = d([\varphi], [\varphi \circ \psi])$, for $[\psi]$ pseudo-Anosov. Let $\varphi_0(y_i) \neq \infty$, $1 \leqslant i \leqslant n-3$. Let $r_0(z)dz^2$ be the quadratic differential for G at $[\varphi_0]$. Let*

$$\operatorname{Res}(r_0, \varphi_0(y_i)) = a_i, \quad 1 \leqslant i \leqslant n-3.$$

Then G has no singular points on

$$U = \left\{[\varphi_0 + \underline{h}] : |h(y_i)| < \delta,\ i \leqslant n-3,\ h(y_i) = 0,\ i > n-3,\ \operatorname{Re}\left(\textstyle\sum_{i=1}^{n-3} \overline{a}_i \varphi(y_i)\right) = 0\right\}.$$

Remark. — One can then find a vector field w on U and tangent to U such that $DG(w([\varphi]) < 0$ for all $[\varphi] \in U \smallsetminus \{[\varphi_0]\}$. One can then construct an open neighbourhood $U_1 \subset U$ of $[\varphi_0]$ such that the w-flow forward orbit of U_1 is contained in U, as follows. Take $\delta_1 > 0$ to be the minimum value of G on ∂U. Then choose $U_1 \subset U$ such that $G < \delta_1$ on U_1.

Proof. — Take any $[\varphi] \in U$. As above, let $r(z)dz^2$ be the quadratic differential at $[\varphi]$ for $G([\varphi])$ with stretch $t(z)dz^2$ at $[\varphi \circ \psi]$. Then given the form of G, we need to show that there is *no* $\lambda \in \mathbf{R}$ with

$$(1) \qquad (\operatorname{Res}(r - t, \varphi(y_i)) = \lambda(a_i) = \lambda \operatorname{Res}(r_0, \varphi_0(y)), \quad 1 \leqslant i \leqslant n-3.$$

Normalise so that $\varphi_0(y_n) = \varphi(y_n) = \infty$.

Using the relations

$$\sum_{i=1}^{n-1} \mathrm{Res}(r, \varphi(y_i)) = 0, \quad \sum_{i=1}^{n-1} \varphi(y_i)\,\mathrm{Res}(r, \varphi(y_i)) = 0$$

and similarly for r_0, φ_0, and using $\varphi(y_i) - \varphi_0(y_i) = O(\delta)$, (1) implies

$$\mathrm{Res}(r - t, \varphi(y_i))x = \lambda\,\mathrm{Res}(r_0, \varphi_0(y_i)) + O(\lambda\delta), \quad i > n - 3,$$

and hence (since G is C^2 by 12.1)

$$\mathrm{Res}(r - t, \varphi(y_i)) = \lambda\,\mathrm{Res}(r, \varphi(y_i)) + O(\lambda\delta) \quad \text{for all } i \tag{2}$$

If we did not know that G were C^2 we would get an error term $O(\lambda\delta_2)$, which would be good enough. Then (2) implies

$$\int \frac{(r - t)\overline{r}}{|r|} = \lambda(1 + O(\delta)) = \int |r - t|(1 + O(\delta)). \tag{3}$$

But

$$\int |r| = \int |t| = 1.$$

Assuming δ is small enough, $r - t$ is small enough to expand,

$$\int |t| = \int |r + (t - r)| = \int |r| + 2\,\mathrm{Re}\Big(\int \frac{(t - r)\overline{r}}{|r|}\Big) + O\Big(\int \frac{|t - r|^2}{|r|}\Big) + O\Big(\Big(\int |t - r|\Big)^2\Big)$$

and hence

$$\mathrm{Re}\Big(\int \frac{(t - r)\overline{r}}{|r|}\Big) = o\Big(\int |t - r|\Big),$$

which contradicts (3) above. So (1) does not hold for any $\lambda \in \mathbf{R}$, as required. □

16.12. Bounding $G - \kappa([\psi])$ by DG. — The following is the best estimate I can find at the moment. For simplicity, it is stated here in the pseudo-Anosov case. In the applications in Chapter 19 the reducible case will also be considered.

Lemma. — *Let $[\psi]$ be pseudo-Anosov. Let $k = \#(Y) - 4$ and let M be given. Let $\|DG\| \leqslant \delta$, $G([\varphi]) \leqslant M$ Then*

$$|G([\varphi]) - \kappa([\psi])| = O(\delta^{1/2+1/2k}).$$

Proof. — We have $[\varphi] \in (\mathcal{T}(Y))_{\geqslant \nu}$ for ν depending only on M, since $[\psi]$ is pseudo-Anosov. We have $\mathrm{Res}(r - t, \varphi(y)) = O(\delta)$ for all $y \in Y$, by assumption.

But $t(z)dz^2$ is also the stretch of $r(z)dz^2$ by λ, where $\frac{1}{2}\log\lambda = d([\varphi], [\varphi \circ \psi])$, that is, unstable foliation leaves of $r(z)dz^2$ are stretched by λ, and stable foliation leaves are contracted by λ^{-1}. Now for closed paths γ in $\overline{\mathbf{C}} \smallsetminus Y$ we need to compare *stable* and *unstable lengths* as defined in 14.8. We write $|\varphi(\gamma)|_-$ and $|\varphi(\gamma)|_+$ for the stable and unstable lengths with respect to the quadratic differential $r(z)dz^2$. It is probably worth emphasizing that the unstable and stable lengths are *not* the same as the real and imaginary parts of

$$\int_{\varphi(\gamma)} \sqrt{r},$$

(whose definition is, in any case, unclear) although $|\varphi(\gamma)|_+$ (for example) is given by putting $\varphi(\gamma)$ in good position (14.5) — as a union of arcs α_k — and taking

$$\sum_k \mathrm{Re}\Big(\int_{\alpha_k} \sqrt{r}\Big),$$

with positive sign for each term. Let $k = \max(2, \#(Y) - 4)$. Then any quadratic differential on $\overline{\mathbf{C}} \smallsetminus \varphi(Y)$ has zeros of multiplicity at most k. (Of course, there are no multiple zeros at all unless $\#(Y) \geqslant 6$.) Then we claim that

$$(1) \qquad |\varphi \circ \psi(\gamma)|_{0,+} = \lambda|\varphi(\gamma)|_{0,+} + O(\delta^{1/2+1/2k}).$$

This estimate is clear when r is not close to having multiple zeros, or zero residues at points of $\varphi(Y)$, when we obtain simply $O(\delta)$. So to obtain (1), we need to estimate local path integrals of the form

$$(2) \qquad \int \sqrt{r(z) + O(\delta)}dz, \ \int \sqrt{z^{-1}(r(z) + O(\delta))}dz, \quad \text{where } r(z) = \prod_{j=1}^{k}(z - z_j)$$

over paths where $|r(z)| = O(\delta)$ and $|r(z)| \geqslant R\delta$ for a suitable r. Changing coordinates if necessary, assume that $z_1 = 0$. Number the zeros z_j as in 16.6, and let β, m be as in 16.6, except that enlarge m if necessary so that

$$|z_m|^m\beta = O(\delta), \quad |z_{m+1}|^m\beta \geqslant R\delta$$

for a sufficiently large $R > 1$. Then we need to estimate the integrals over paths on which $|r(z)| = O(\delta)$ and paths on which $|r(z)| \geqslant R\delta$. This means estimating on paths on which $|z| = O((\delta/\beta)^{1/m})$ and on which $|z| \geqslant R_1(\delta/\beta)^{1/m}$. We have

$$\beta \geqslant |z_{m+1}|^{k-m}, \quad |z_{m+1}|^m\beta \geqslant R\delta.$$

This yields (since $R \geqslant 1$)

$$\beta \geqslant (R\delta/\beta)^{(k-m)/m}, \quad \beta \geqslant \delta^{(k-m)/k}, \quad (\delta/\beta) \leqslant \delta^{m/k}, \quad (\delta/\beta)^{1/m} \leqslant \delta^{1/k}.$$

Then the integrals of (2) over paths on which $r(z) = O(\delta)$ become $O(\delta^{1/2+1/k})$, $O(\delta^{1/2+1/2k})$. On paths on which $|r(z)| \geqslant R_1\delta$ for $R_1 \geqslant 2$, and thus $|z| \geqslant R_2(\delta/\beta)^{1/m}$ we have

$$\begin{aligned}\int \sqrt{r(z) + O(\delta)}dz &= \int \sqrt{r(z)}dz + \int O(\delta\beta^{-1/2}z^{-m/2})\\ &= \int \sqrt{r(z)}dz + O(\delta\beta^{-1/2}(\delta/\beta)^{1/m-1/2}),\end{aligned}$$

(with $(\delta/\beta)^{1/m-1/2}$ replaced by $\log(\delta/\beta) = O(\log \delta)$ if $m = 2$),

$$\int \sqrt{z^{-1}(r(z) + O(\delta))}dz = \int \sqrt{z^{-1}r(z)}dz + O(\delta\beta^{-1/2}(\delta/\beta)^{1/2m-1/2}).$$

In both cases the error term is $O(\delta^{1/2}(\delta/\beta)^{1/2m}) = O(\delta^{1/2+1/2k})$. This gives (1), as required. But (1) says that λ is within $O(\delta^{1/2+1/2k})$ of an eigenvalue (with positive-entry eigenvector) of a certain integer-valued matrix, which is one of the matrices defining the piecewise linear action of ψ on the projective space of measured foliations on $\overline{\mathbf{C}} \setminus Y$ (see [**F-L-P**]). This means that

$$\log \lambda = \kappa([\psi]) + O(\delta^{1/2+1/2k})$$

as claimed. □

PART III

PROOF OF THE TOPOGRAPHER'S VIEW

CHAPTER 17

DISTANCE AND THE PULLBACK MAP

17.1. Let B be a branched covering space of degree two type, $\mathcal{T} = \mathcal{T}(B)$ (6.2) and

$$F(x) = d(x, \tau(x)).$$

This chapter is devoted to some estimates which will be needed in the proof of the Level-κ Tool in Chapters 18-21. Some of these are estimates on F in subsets $\mathcal{T}(\Gamma, \varepsilon)$, where (f_0, Γ) is invariant in (17.4-6, 17.9). These are basically refinements of an easy lower bound (17.4) in terms of combinatorial data associated to (f_0, Γ). Some invariants of (f_0, Γ) depending on combinatorial data are defined in 17.2. These are related to some earlier definitions in Chapter 2. The other estimates in this chapter involve the quantity $m_\gamma(x)$ of 9.1, and its behaviour along a sequence $\{\tau^n(x)\}$. The results on this topic are given in 17.7-8. Various related definitions are given in 17.3.

17.2. Definition of $\kappa(\alpha)$, $\kappa_0(\Gamma)$. — Let (f_0, Γ) be invariant. Let α be any loop or gap of Γ of period p with periodic orbit $[\alpha]$. We define

$$\kappa(\alpha) = \lim_{\varepsilon \to 0} \inf\{\max_{\alpha' \in [\alpha]} d_{\alpha'}(y, \tau(y)) : y \in \mathcal{T}(\Gamma, \varepsilon)\}.$$

Let $\Gamma' \subset \Gamma$ be the maximal loop set such that (f_0, Γ') satisfies the Invariance and Levy Conditions. We suppose that $\Gamma' \neq \varnothing$. Let Ω be the fixed union of Γ'. Let Δ be any periodic gap of Γ', of period p, such that v_2 is not in its periodic orbit. If Δ is homeomorphic, let $[\psi_\Delta]$ be the gap map of Δ (see 2.13). If Δ is of degree two, let τ_Δ be the pullback map (6.6-7) on the associated critically finite branched map space $B(f_0, \Gamma, \Delta)$ (see 2.18). Write d_Δ for the Teichmüller metric on the associated Teichmüller space $\mathcal{T}(A(\Delta))$. (See 9.1.) Depending on whether Δ is homeomorphic or degree two,

$$\kappa(\Delta) = p^{-1} \inf\{d_\Delta(y, y \cdot [\psi_\Delta]) : y \in \mathcal{T}(A(\Delta))\},$$
$$\kappa(\Delta) = p^{-1} \inf\{d_\Delta(y, \tau_\Delta(y)) : y \in \mathcal{T}(A(\Delta))\}.$$

Then we define

$$\kappa_0(\Gamma) = \mathrm{Max}\{\kappa(\Delta) : \Delta \text{ is a gap of } \Gamma,\ \Delta \subset \Omega\}.$$

Note that $\kappa(\gamma) = 0$ for any loop $\gamma \subset \Omega$, $\gamma \in \Gamma$. This is reminiscent of our definition of $\kappa(f_0, \Gamma)$ in 7.2 — but not identical to it, except when (f_0, Γ) is minimal nonempty. For Γ, Γ' as above, we have $\kappa_0(\Gamma) = \kappa_0(\Gamma')$. If (f_0, Γ_i) satisfy the Invariance and Levy Conditions for $i = 1, 2$, and $[f_0, \Gamma_1] = [f_0, \Gamma_2]$ (see 3.8), then $\kappa_0(\Gamma_1) = \kappa_0(\Gamma_2)$. Indeed if α is a gap of Γ_1 which is a union of gaps for Γ_2, then

$$\kappa(\alpha) = \max\{\kappa(\beta) : \beta \subset \alpha,\ \beta \text{ a gap or loop of } \Gamma_2\}.$$

Now let $\Gamma' \subset \Gamma$ be maximal such that (f_0, Γ') satisfies the Invariance and Levy conditions.Let $\gamma \in \Gamma \setminus \Gamma'$ be periodic of period n. Let d be the degree of $f_0 \mid \gamma'$, where γ' is the (unique) component of $f_0^{-1}(\gamma)$ which is homotopic in $\overline{\mathbf{C}} \setminus Z$ to a loop in $[\gamma]$. Then

$$\kappa(\gamma) = \frac{1}{dn} \log d.$$

We define

$$\kappa_0'(\Gamma) = \max\{\kappa(\gamma) : \gamma \in \Gamma\}.$$

17.3. m_γ and discrete loop sets. — From now on, we work with $\mathcal{T}$ rather than $\mathcal{T}/G$. Let $x \in \mathcal{T}(Y)(\gamma, \varepsilon_0)$ for any simple nontrivial nonperipheral loop. Let π_γ be the projection to $\mathcal{T}(A(\gamma)) = \{z : \mathrm{Im}(z) > 0\}$ as in 9.1. We recall that

$$m_\gamma(x) = \log \mathrm{Im}\, \pi_\gamma(x).$$

We remark that there is a constant C such that if ε is the length of γ, then

$$\frac{2\pi^2}{\varepsilon} - C \leqslant m_\gamma(x) \leqslant \frac{2\pi^2}{\varepsilon} + C.$$

We also remark that, again for suitable $C > 0$, if $x = [\varphi]$, and the points of $\varphi(A(\gamma))$ are normalised so that two on one side of $\varphi(\gamma)$ are at (say) 1 and ∞, and a third is at 0, then the fourth is distance $\leqslant Ce^{-2\pi^2/\varepsilon}$ from 0 and distance $\geqslant (1/C)e^{-2\pi^2/\varepsilon}$ from 0. If $x \notin \mathcal{T}(Y)(\gamma, \varepsilon_0)$ we define $m_\gamma(x) = 0$. Note that this definition depends on the choice of $A(\gamma)$ (a very little) but not on the choice of α_2 (see 9.1).

Now we need to define a function $m_{1,\Gamma}$ for any invariant (f_0, Γ), $\Gamma \subset \overline{\mathbf{C}} \setminus Y$, with $\Gamma_2 = \Gamma_2(f_0, \Gamma) \neq \varnothing$ (see 2.5) and $x \in \mathcal{T}(\Gamma, \varepsilon_0)$. Let $\Omega = \Omega(f_0, \Gamma)$ be the maximal connected union of periodic homeomorphic gaps which containing the fixed set of Γ. (See 2.8.) We call Ω the *fixed union of* Γ. (We reserve P for an irreducible fixed gap.) Let $\gamma_0 \subset \partial\Omega$ be the loop separating Ω from v_2. Let γ_0' be the loop of Γ homotopic to γ_0 in $\overline{\mathbf{C}} \setminus Z$ but not in $\overline{\mathbf{C}} \setminus Y$, such that a component of $f_0^{-1}(\gamma_0')$ is homotopic in $\overline{\mathbf{C}} \setminus Z$ to the component of $f_0^{-1}(\gamma)$ in $\partial\Omega$ up to Z-preserving isotopy, if such a loop exists. If no such loop exists, we define $m_{\gamma_0'}(x) = 0$. Let $\Gamma' \subset \Gamma$ be the smallest invariant loop set containing $\Gamma_2(f_0,\Gamma)$, but not including γ_0', if γ_0' exists. Then we define

$$\begin{aligned}
m_{\mathrm{int},\Gamma}(x) &= \sum_{\gamma\subset\mathrm{int}(\Omega)} m_\gamma(x),\\
m_{\partial,\Gamma}(x) &= \frac{m_{\gamma_0}(x)}{m_{\gamma_0}(x)+m_{\gamma_0'}(x)}\sum_{\gamma\subset\partial\Omega} m_\gamma(x),\\
m_{1,\Gamma}(x) &= m_{\mathrm{int},\Gamma}(x)+m_{\partial,\Gamma}(x),\\
m_{2,\Gamma}(x) &= \sum_{\gamma\in\Gamma\smallsetminus\Gamma'} m_\gamma(x)+\frac{m_{\gamma_0'}(x)}{m_{\gamma_0}+m_{\gamma_0'}(x)}\sum_{\gamma\subset\partial\Omega} m_\gamma(x),\\
m_{\partial,0,\Gamma}(x) &= m_{\gamma_0}(x).
\end{aligned}$$

If α is a periodic loop or gap, we write $[\alpha]$ for the periodic orbit. If γ is a loop, we write

$$m_{[\gamma]}(x)=\sum_{\gamma'\in[\gamma]} m_{\gamma'}(x).$$

For $\gamma\subset\mathrm{int}(\Omega)$, let A be the set of loops of Γ in $\mathrm{int}(\Omega)$ separating γ from Δ_0'. Then we define

$$m_{\gamma,\partial,\Gamma}(x)=\sum_{\gamma'\in A} m_{\gamma'}(x)+m_{\partial,\Gamma}(x).$$

We define

$$m_{[\gamma],\partial,\Gamma}(x)=\min\{m_{\gamma',\partial,\Gamma}(x):\gamma'\in[\gamma]\}.$$

Let $x=[\varphi]$. We define $m_1([\varphi])$ for $m_{1,\Gamma}([\varphi])$, where Γ is the largest invariant loop set in which all loops of $\varphi(\gamma)$ have length $<\varepsilon_0$. We define $m_2([\varphi])$ similarly. Apart from this, we shall drop the suffix Γ where the context is clear.

Again assuming that $\Gamma_2(f_0,\Gamma)\neq\varnothing$, we say that (f_0,Γ) is *discrete* if $\Delta_0'(f_0,\Gamma)$ is equal to, or adjacent to Ω. It is possible, in the discrete case, to have $v_2\notin\Delta_0'$. In that case, the reduced map space $B(\Delta_0',f_0,\Gamma)$ (2.18) is critically finite, and may possibly have Euclidean orbifold (see 6.11). In this case, we shall say that (f_0,Γ) is *Euclidean.* Otherwise, we shall say that (f_0,Γ) is *non-Euclidean.*

We shall say that (f_0,Γ) is $(L_1,L_2,\varepsilon,\nu)$-*adapted* to x (or sometimes just (L_1,L_2)-adapted) if it is discrete and invariant, $x\in\mathcal{T}(\Gamma,\varepsilon)$, $L_1\varepsilon\leqslant\nu$, $L_2\nu\geqslant\varepsilon_0$, and the following hold. The loop set $\Gamma_2(f_0,\Gamma)\neq\varnothing$. Let α be a periodic degree two gap of Γ. Let $\zeta\subset\partial\alpha$ and $\gamma\subset\mathrm{int}(\alpha)$. Then $m_\gamma(x)\leqslant m_\zeta(x)/L_1$ and $x\notin\mathcal{T}(\gamma,\nu)$.

17.4. A Lemma showing certain thin parts cannot be re-entered

We continue with the notation of 17.2-3. Recall that $F(x)=d(x,\tau(x))$

Lemma. — *The following hold for C sufficiently large and C' sufficiently small and $\delta>0$ sufficiently small. Let $x\in\mathcal{T}(\Gamma,\delta)$ where (f_0,Γ) satisfies the Invariance and*

Levy conditions, with $\Gamma_2(f_0,\Gamma)\neq\varnothing$. *Let* ε' *be the lengths of the shortest loop for* x *in* $\partial\Omega$. *Let* μ *be the minimal nonempty node with* $\mu\leqslant[f_0,\Gamma]$. *Then*

$$F(x)\geqslant\kappa(\Omega)-Ce^{-2\pi^2/\varepsilon'}.$$

Proof. — Write $[\psi_\Omega]$ (as usual: 2.13) for the isotopy class of $f_0\mid\Omega$. Then

$$d_\Omega(\tau(x),x\cdot[\psi_\Omega])\leqslant Ce^{-2\pi^2/\varepsilon'}. \tag{1}$$

To get this, write $\pi_\Omega(x)=[\varphi]$, $\pi_\Omega(\tau(x))=[\varphi']$. We can normalise so that all points of $\varphi(A(\Omega))$ are bounded and at least two are bounded apart, and so that, for any y the distance between points $\varphi'(y)$ and $\varphi\circ\psi_\Omega(y))$ $(y\in A(\Omega))$ is $O(me^{-2\pi^2/\varepsilon'})$, where m is the distance of the next nearest point of $\varphi'(Y)$ to $\varphi'(y)$, which gives (1). Then

$$F(x)=d(x,\tau(x))\geqslant d_\Omega(x,\tau(x)\geqslant d_\Omega(x,x\cdot[\psi_\Omega])-Ce^{-2\pi^2/\varepsilon'}\geqslant\kappa(\mu)-C'e^{-2\pi^2/\varepsilon'}.$$

□

17.5. Lemma. — *Take the same hypotheses as 17.4. If* $\overline{P}\subset\operatorname{int}(\Omega)$ *is a connected union of gaps fixed up to isotopy by* f_0, *and* ε *is the length of the shortest loop in* ∂P *for* x, *then*

$$F(x)\geqslant\kappa(P)+C'e^{-2\pi^2/\varepsilon}.$$

Proof. — Suppose that $[\psi_\Omega]$ is reducible and $P\subset\Omega$ is fixed (not necessarily irreducible). Let

$$d(x,\tau(x))\leqslant\kappa(P)+C'e^{-2\pi^2/\varepsilon} \tag{2}$$

for a small $C'>0$. The idea is to obtain a contradiction by showing that for some C'' and some choice of $A(P)$

$$d_{A(P)}(x,\tau(x))\geqslant\kappa(P)+C''e^{-2\pi^2/\varepsilon}.$$

Write $x=[\varphi]$ and $\tau(x)=[\varphi']$. By (1) and (2) we have

$$d_\Omega([\varphi],[\varphi\circ\psi_\Omega])\leqslant\kappa(P)+C'e^{-2\pi^2/\varepsilon}+Ce^{-2\pi^2/\varepsilon'}.$$

If C' is small, this is only possible if the loops $\varphi(\partial P)$ have length $\leqslant\zeta'$ for a small ζ', because (as we shall see in 20.14, but this is well-known) the function

$$[\varphi]\longmapsto d_\Omega([\varphi],[\varphi\circ\psi_\Omega])$$

has no local minima. Given a constant $L_1>0$, we can choose ζ' so that there are ν with $L_1\zeta'\leqslant\nu$, and a gap union Q containing P such that all loops of $\varphi(\partial Q)\cup\varphi(\partial P)$ are homotopic to geodesics of length $\leqslant\zeta'$ but any other geodesic homotopic to $\varphi(\gamma)$ for some $\gamma\subset Q$ has length $\geqslant\nu$. Let n, nk be the numbers of components of ∂P, ∂Q respectively, and ψ_P, ψ_Q the gap maps. Normalise so that, for any choice of $A(P)$, the points of $\varphi(A(P))$ are bounded and bounded apart. Fix a choice of $y\in A(Q)$ such that $\varphi(y)$ is separated from $\varphi(P)$ by the shortest loop of $\varphi(\partial P)$ (of length ε). Then we can choose

$$A(Q)=\{f_0^i(y):0\leqslant i<nk\}.$$

Then in the notation of Chapter 8, we have

$$[\varphi']_P = [\varphi\circ\psi_Q + O(e^{-2\pi^2(1/\varepsilon+1/\zeta')}]_P = [\varphi\circ\psi_P + \underline{h}_0 + O(e^{-2\pi^2(1/\varepsilon+1/\zeta')})]_P.$$

Here $\underline{h}_0 = (h_0(y'))$ is a vector with $h_0(y') = 0$ for $y' \neq y$ and

$$h_0(y) = \varphi(f_0^n(y)) - \varphi(y),$$

so that $|h_0(y)| \geqslant C^{-1}e^{-2\pi^2(1/\varepsilon+1/\nu)}$. Similarly we can define $\underline{h}_j$ $(0 \leqslant j < k)$ using $f_0^{nj}(y)$ instead of y. Thus,

$$\sum_j h_j = O(e^{-2\pi^2(1/\varepsilon+1/\zeta')}).$$

Define

$$A_j(P) = \{f_0^{i+nj}(y) : 0 \leqslant i < n\}.$$

By the First Derivative Formula 8.4, we see that

$$\begin{aligned}
d_{A_j(P)}([\varphi],[\varphi']) &= d_{A_j(P)}([\varphi],[\varphi\circ\psi_P + \underline{h}_j + O(e^{-2\pi^2(1/\varepsilon+1/\zeta')})]) \\
&= d_{A_j(P)}([\varphi],[\varphi\circ\psi_P]) + 2\pi\,\mathrm{Re}(ch_j(y)) + O(e^{-2\pi^2(1/\varepsilon+1/\zeta')}) \\
&\geqslant \kappa(P) + 2\pi\,\mathrm{Re}(ch_j(y)) + O(e^{-2\pi^2(1/\varepsilon+1/\zeta')}),
\end{aligned}$$

where c is the residue of the quadratic differential for $d_{A_j(P)}([\varphi],[\varphi\circ\psi_P])$ at $[\varphi]$. It follows from the fact that ψ_P cyclically permutes the points of $A_j(P)$ that the residues at all points $\varphi(A_j(P))$ are bounded from 0. So c is bounded from 0. So either we can find a lower bound on $d_P([\varphi],[\varphi'])$ by suitable choice of j and h_j to make $\arg(ch_j)$ boundedly in the right half plane, or for all choices of j,

$$d_{A_j(P)}([\varphi],[\varphi']) = \kappa(\mu) + o(e^{-2\pi^2(1/\varepsilon+1/\nu)}).$$

This actually means that the points $\varphi(f_0^{nj}(y))$ are approximately collinear. Let q_1 denote the quadratic differential for $d([\varphi],[\varphi'])$ at $[\varphi]$. Choose j — and then choose $A(P)$ — so that q_1 has at least one pole distance $\geqslant C^{-1}e^{-2\pi^2(1/\varepsilon+1/\nu)}$ from $\varphi\circ\psi_Q^{jn}(y) \in A(P)$. Now we use π_P to denote projection from $\mathcal{T}(A(Q))$ to $\mathcal{T}(A(P))$. Let q_2 denote the best quadratic differential for $d_P([\varphi],[\varphi'])$. Let χ_1, χ_2 be the best quasiconformal maps. Let $\theta = \arg(q_1) - \arg(q_2)$. Then by 8.3,

$$d([\varphi],[\varphi']) \geqslant d_P([\varphi],[\varphi']) + C_1\int|\theta|^2 \geqslant \kappa(\mu) + C_2 m e^{-2\pi^2/\zeta'}.$$

This gives the required contradiction. □

17.6. Lemma. — *Take the same hypotheses as in 17.4, 17.5 and let C' be as in 17.5. Now let*

$$F(x) \leqslant \kappa(\mu) + C'e^{-2\pi^2/\varepsilon}.$$

Then Ω is irreducible. Further, if γ, γ' are loops in $\partial\Omega$ then

$$|m_\gamma(x) - m_{\gamma'}(x)| \leqslant C, \quad |m_\gamma(x) - m_\gamma(\tau(x))| \leqslant C.$$

In particular (f_0,Γ) is discrete, so that Δ_0' is adjacent to Ω.

Proof. — We now know that $\Omega = P$ is irreducible. So

$$d_P(x, \tau(x)) \geqslant \kappa(\mu) - Ce^{-2\pi^2/\varepsilon}.$$

Suppose for contradiction that there is $\gamma \subset \partial P$ such that

$$|m_\gamma(\tau(x)) - m_\gamma(x)| \geqslant \Delta, \quad m_\gamma(x) \leqslant 2\pi^2/\varepsilon + O(1). \tag{3}$$

It follows that there is an annulus A of modulus $\geqslant C_2\Delta$ (where C_2 depends only on $\kappa(\mu)$ and C) homotopic to $\varphi(\gamma)$ ($x = [\varphi]$) such that the quadratic differential $q(z)dz^2$ for $d(x, \tau(x))$ at x satisfies the Pole Zero Condition on A, that is, has two more poles than zeros in each component of $\overline{\mathbf{C}} \smallsetminus A$. It follows from 9.5 that, for C_3 bounded from 0,

$$F(x) \geqslant d_P(x, \tau(x)) + C_3 e^{C_2\Delta} e^{-2\pi^2/\varepsilon} \geqslant \kappa(\mu) + C' e^{-2\pi^2/\varepsilon}$$

if Δ is large enough given C'. So by our hypothesis, (3) does not hold. It follows that Γ is discrete. It also follows that, for suitable Δ given C',

$$\sum_{\gamma \subset \partial P} m_\gamma(\tau(x)) \geqslant \sum_{\gamma \subset \partial P} m_\gamma(x) - \Delta.$$

But if γ, $\gamma' \subset \partial P$ with $\gamma \subset f_0^{-1}(\gamma')$, then

$$m_\gamma(\tau(x)) \leqslant m_{\gamma'}(x) + O(1).$$

So we must also have

$$m_{\gamma'}(x) \leqslant m_\gamma(\tau(x)) + O(\Delta),$$

that is, for suitable C (given C' in this case),

$$|m_\gamma(x) - m_{\gamma'}(x)| \leqslant C.$$

□

17.7. Lemma. — *Let (f_0, Γ) be invariant with $\Gamma_2(f_0, \Gamma) \neq \varnothing$. Let Ω be as in 17.3. The sets $A(\gamma)$ ($\gamma \in \Gamma$) can be chosen so that the following holds. Let α, α' be any loops in $\overline{\Omega}$ with $\alpha' \subset f_0^{-1}(\alpha)$. Let $m_{\alpha',Z}(\tau(y)$ denote $\operatorname{Im} \pi_{\alpha',Z}(\tau(y)$ where α' is taken up to Z-preserving isotopy, $A(\alpha', Z) \subset Z$ has four points, two in each component of $\overline{\mathbf{C}} \smallsetminus \alpha'$ and $\pi_{\alpha',Z}$ is projection to $\mathcal{T}(A(\alpha', Z))$. Then if $y \in \mathcal{T}(\Gamma, \varepsilon_0)$,*

$$m_\alpha(y) - m_{\alpha',Z}(\tau(y)) = O(m_{[\alpha]}(y))^{-1}(e^{-m_{\alpha,\partial}(y)} + e^{-m_{\alpha',\partial}(y)}). \tag{1}$$

Hence,

$$\sum_{[\alpha] \subset \mathrm{Int}(\Omega)} |m_{[\alpha]}(y) - m_{[\alpha]}(\tau(y))| \leqslant C_1^{-1} e^{-m_\partial(y)}. \tag{2}$$

Proof. — Clearly, (2) follows from (1). We can choose the sets $A(\beta)$ so that, for any loops α, $\alpha' \subset \overline{\Omega}$ with $\alpha' \subset f_0^{-1}(\alpha')$, each point of $f_0(A(\alpha'))$ is in the same component of $\overline{\mathbf{C}} \smallsetminus \Omega$ as a point of $A(\alpha)$, and $f_0(A(\alpha') = A(\alpha)$ if α does not separate Δ_0' from the centre of P'. Now fix α, α' as above. Write $y = [\varphi]$. Let s be the holomorphic map used to define $\tau(y)$. Write

$$\tau(y) = [\varphi'] = [s^{-1} \circ \varphi \circ f_0].$$

Recall that $m_\alpha([\varphi]) = \mathrm{Im}(\pi_\alpha([\varphi]))$. Now

$$[\varphi']_{A(\alpha')} = [s_1^{-1} \circ \varphi \circ g]_{A(\alpha')} = [s_1^{-1} \circ \varphi]_{g(A(\alpha'))},$$

where s_1, g are homeomorphisms as follows. We take g to agree with f_0 on the component of $f_0^{-1}(P')$ homotopic to Ω, and to map $A(\alpha')$ to $f_0(A(\alpha'))$. We take s_1 to agree with s on

$$s^{-1}(\cup_{\beta\subset\overline{\Omega}} S(\beta, [\varphi], \varepsilon_0))$$

(see 9.3) and to be of bounded distortion in the complement. If α does not separate Δ_0' from the centre of Ω then $g(A(\alpha')) = A(\alpha)$ and

$$\int K(s_1^{-1})|q| = 1 + O((m_\alpha(y))^{-1}e^{-m_{\alpha.\partial}(y)}) \tag{3}$$

where $q(z)dz^2$ is the quadratic differential for $d_\alpha([\varphi], [s_1^{-1} \circ \varphi])$ at $[\varphi]$. We get this because s_1^{-1} is conformal except on a set of q-measure $O((m_\alpha(y))^{-1}e^{-m_{\alpha,\partial}(y)})$. So then the bound on $d_\alpha([\varphi], [s_1^{-1} \circ \varphi])$ follows from 8.2. If α does separate Δ_0' and the centre of Ω then $[s_1^{-1} \circ \varphi]_{A(\alpha)} = [s_1^{-1} \circ \chi \circ \varphi]_{A(\alpha)}$ where

$$\int K(s_1^{-1} \circ \chi)|q| = 1 + O((m_\alpha(y))^{-1}(e^{-m_{\alpha.\partial}(y)} + e^{-m_{\alpha'.\partial}(y)}).$$

Again, the required bound on $d_\alpha([\varphi], [s_1^{-1} \circ \chi \circ \varphi])$ follows from 8.2. □

17.8. We need to consider $m_\partial(x)$ for $x \in \mathcal{T}(\Gamma, \varepsilon_0)$ for a Euclidean discrete invariant (f_0, Γ). 17.7 helps for all but one of the loops that separate Δ_0' from Ω. The following Lemma implies that F-decreasing, F-between, or a property we might term m_∂-increasing, hold for x and (f_0, Γ) discrete Euclidean.

Lemma. — *Let (f_0, Γ) be discrete invariant Euclidean, $\Gamma_2(f_0, \Gamma) \neq \varnothing$. Let $x \in \mathcal{T}(\Gamma, \varepsilon_0) \cap \mathcal{T}(\partial\Delta_0', \varepsilon)$, $x \notin \mathcal{T}(\gamma, \nu)$ for $\gamma \subset \mathrm{Int}(\Delta_0')$, $L_1\varepsilon \leqslant \nu$. Let $F(x) \leqslant \kappa$. Let $\gamma_0 = \partial\Delta_0' \cap \partial\Omega$. There are $\varepsilon_0' > 0$ and an integer k_0, both depending only on κ, such that the following hold. Let $k_0 \leqslant k \leqslant 2k_0$. Then $\tau^i(x)) \notin \mathcal{T}(\gamma, \varepsilon_0')$ for $\gamma \subset \mathrm{Int}(\Delta_0')$ for $i \leqslant k$ and*

$$m_{[\gamma_0]}(\tau^k(x)) > m_{[\gamma_0]}(x) + \tfrac{1}{4}k \log 2.$$

Proof. — We start by considering $B(\Delta_0', f_0, \Gamma)$. Write $\pi_0(x) = x_1 = [\varphi_1]$. Write $B(\Delta_0', f_0, \Gamma) = B(Y_1, f_1)$ with associated pullback τ_1. Let s_1, $s_{1,n}$ be the holomorphic maps used to define $\tau_1(x_1)$, $\tau_1^n(x_1)$ from x_1. Normalise φ_1 so that the points of $\varphi_1(Y_1)$ corresponding to $A(\Delta_0') \cap A(\gamma_0)$ are at 0, 1, ∞, and similarly for $\tau_1^n(y_1)$, $n \geqslant 0$. We claim that, for n sufficiently large (depending only on κ),

$$|s_{1,n}'(0)| > 2^{n/4}, \tag{1}$$

We claim also that for $\varepsilon_0' > 0$ depending only on κ, $\tau_1^n(y) \in (\mathcal{T}(Y_1))_{\geqslant 2\varepsilon_0'}$ for all $n \geqslant 0$.

By our assumptions, f_1 is of a very particular type. We can regard $\overline{\mathbf{C}}$ as

$$((\mathbf{C}/(\mathbf{Z} + i\mathbf{Z}))/z \sim -z) = \{[z] : z \in \mathbf{C}\}.$$

Identifying $\mathbf{C}$ with $\mathbf{R}^2$, we can regard 2×2 real matrices as linear maps of $\mathbf{C}$. Then for $A \in GL(2, \mathbf{Z})$, $[z] \mapsto [Az]$ is well defined. Then for some such A of determinant 2, $f_1([z]) = [Az)] = [\widetilde{f}_1(z)]$ up to Thurston equivalence. We also have a bound on A in terms of M. In fact, we can take A of the form

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} A_1 = \sqrt{2} B$$

for $B \in \mathrm{SL}(2, \mathbf{Z})$. The choices of f_1 and A are restricted by the requirement that both critical points of f_1 be in the backward orbit of the same fixed point.

The Teichmüller space $\mathcal{T}(Y_1)$ then identifies with the Teichmüller space of the torus $\mathbf{C}/(\mathbf{Z}+i\mathbf{Z})$, and with the upper half-plane H. Then we can identify $x_1 \in \mathcal{T}(Y_1)$ with a bounded matrix $X \in \mathrm{SL}(2, \mathbf{R})$, with a torus $\mathbf{C}/X(\mathbf{Z}+i\mathbf{Z})$, and with element $X^{-1} \cdot i = \rho(X)$ of H, using the usual action of $GL_+(2, \mathbf{R})$ on H by Möbius transformations. (This is the same identification, up to scale, as in 9.1.) Write φ_X for the $\mathbf{R}$-linear map of $\mathbf{C}$ given by the matrix X. This descends to a map, also denoted φ_X,

$$\varphi_X : \mathbf{C}/(\mathbf{Z} + i\mathbf{Z}) \longrightarrow \mathbf{C}/X(\mathbf{Z} + i\mathbf{Z}).$$

We denote the isotopy class by $[\varphi_X]$. Then taking $\widetilde{s}_1$ to be the lift of s_1 to $\mathbf{C}/X_1(\mathbf{Z} + i\mathbf{Z})$, $\tau_1([\varphi_X]) = [\varphi_{X_1}]$ where $\widetilde{s}_1 \circ \varphi_{X_1} = \varphi_X \circ f_1 = \varphi_{XA}$. Similarly, $\tau_1^n([\varphi_X]) = [\varphi_{X_n}]$ where $\widetilde{s}_{1,n} \circ \varphi_{X_n} = \varphi_{XA^n}$. Write $B = \Sigma A_1 \in \mathrm{SL}(2, \mathbf{R})$. Since we are assuming that A is bounded, so is B, and for each n there are $C_n \in \mathrm{SL}(2, \mathbf{Z})$ and bounded D_n (independently of n) such that $B^n = D_n C_n$. The eigenvalues of B are distinct (because $A \in GL_+(2, \mathbf{Z})$ cannot have trace $2\sqrt{2}$ and either both on the unit circle or both real). If the eigenvalues of B are on the unit circle, then $B \neq \pm I$ has a unique fixed point in H. If the eigenvalues are real, one then uses the fact that quadratic numbers have bounded continued fraction expansions. This ensures that each point on the geodesic in H fixed by the Möbius action of B is a bounded distance from the $\mathrm{SL}(2, \mathbf{Z})$ orbit of, say, $i \in H$. (If we are lucky, B^2 will itself be in $\mathrm{SL}(2, \mathbf{Z})$, but this is not guaranteed.)

The bound on

$$d(x_1, \tau_1(x_1)) = \tfrac{1}{2} d_P(\rho(X), \rho(XA)),$$

where d_P is Poincaré distance, gives a bound on the torus $\mathbf{C}/X(\mathbf{Z} + i\mathbf{Z})$. The point $\rho(X)$ is a bounded distance from the geodesic or point fixed by the Möbius action of B, that is if the torus is of bounded type. So there is $\varepsilon_0' > 0$ depending only on M such that $x_1 \in (\mathcal{T}(Y_1))_{\geqslant 2\varepsilon_0'}$, and similarly for $\tau_1^n(x)$ for all $n \geqslant 0$. So the second claim is proved.

So then $\widetilde{s}_{1,n} \circ \varphi_{X_n} = \varphi_{2^{n/2} X D_n} \circ \varphi_{C_n}$. If we choose $\widetilde{s}_{1,n}$ so that the lift to $\mathbf{C}$ is given by $z \mapsto 2^{n/2} z$, then $X_n \,\mathrm{SL}(2, \mathbf{Z})$ is a bounded element of $\mathrm{SL}(2, \mathbf{R}) / \mathrm{SL}(2, \mathbf{Z})$. It follows that the holomorphic map $\sigma_n : \mathbf{C} \to \overline{\mathbf{C}}$ which identifies all points of

$$\mathbf{C} + X_n(\mathbf{Z} + i\mathbf{Z}) \pm z,$$

and maps 0 to 0 lies in a compact set of such maps, and

$$\sigma_n'(0) = 0, \quad \sigma_n''(0) = \beta_n$$

for β_n bounded and bounded from 0. Then noting that

$$s_{1,n} \circ \sigma_n(z) = \sigma_0(2^{n/2}z),$$

for n large enough depending only on M,

$$|s_{1,n}'(0)| > 2^{n/4}.$$

So (1) holds, as claimed.

Taking ε larger if necessary, we can assume for $n \leqslant k$ that

$$\tau^n(x) \in \mathcal{T}(\Gamma, \varepsilon_0) \cap \mathcal{T}(\partial\Delta_0', \varepsilon), \quad \tau^n(x) \notin \mathcal{T}(\gamma, \varepsilon_0'/2) \text{ for } \gamma \subset \mathrm{Int}(\Delta_0').$$

We shall write $x = [\varphi]$, $\tau(x) = [\varphi']$. Let s be the holomorphic map used to define $\tau(x)$, from x. Write $\gamma_j = \partial\Delta_j' \cap \partial\Omega$. We take the same choice of sets $A(\gamma_j)$ as in 17.7. Normalise $S(\Delta_j', x, \nu)$ so that the points of $\varphi(\gamma_j) \cap \varphi(A(\Delta_j'))$ are at 0, 1, ∞, with two points of $\varphi(A(\gamma_j)$ near 0. Normalise $S(\Delta_j', \tau^n(x), \nu)$ similarly. Then s^{-1} and s_1^{-1} are close, not just on $S(\Delta_0', x, \nu)$ but near 0 as well. In fact, they are C^1-close to within $e^{-1/C_1\varepsilon}$ on any bounded set. Recall (9.3) that if $[\varphi] \in \mathcal{T}(\gamma, \varepsilon_0)$, $m_\gamma([\varphi])$ is defined as $\mathrm{Im}(\pi_\gamma([\varphi]))$, where π_γ is projection to $\mathcal{T}(A(\gamma))$ and this is identified with the upper half-plane H. If the four points of $\varphi(A(\gamma))$ are at 0, 1, ∞ and a for a near 0, it follows that $m_\gamma([\varphi]) = -\log|a| + O(1)$. As explained in 9.1, the identification with H was chosen so that $m_\gamma([\varphi])$ was the modulus of $S(\gamma, [\varphi], \varepsilon_0)$ to first order. A calculation with elliptic integrals confirms this and shows that

$$m_\gamma([\varphi]) = -\log|a| - \log 2 + O(a\log(a)).$$

It follows that

$$m_{\gamma_{p-1}}(\tau(x)) = m_{\gamma_0}(x) + \log|(s_1^{-1})'(0)| + O(e^{-1/C_1\varepsilon}).$$

We also have estimates on $m_{\gamma_{j-1}}(\tau(x)) - m_{\gamma_j}(x)$ for $0 < j \leqslant p-1$ from 17.7. So similarly, for bounded n using $[n/p]$ to denote the largest integer $\leqslant n/p$ as usual,

$$m_{[\gamma_0]}(\tau^n(x)) = m_{[\gamma_0]}(x) + \log|s_{1,[n/p]+1}^{-1}(0)| + O(ne^{-1/C_1\varepsilon}).$$

So for a suitable k_0 depending only on κ, and C_1 sufficiently large given k_0, (1) gives the required result. □

17.9. Lemma. — *Take the same hypotheses as in 17.6. Then (f_0, Γ) is nonEuclidean.*

Proof. — Let p be the period of Δ_0' and write

$$Q = P \cup \Big(\bigcup_{i=0}^{p-1} \Delta_i'\Big).$$

By our definitions, $\kappa(\mu) = \kappa(P)$. Suppose that

$$F(x) \leqslant \kappa(\mu) + C'e^{-2\pi^2/\varepsilon}.$$

There are two cases to consider:

(i) $\kappa(\Delta_0') = \kappa(\mu)$,

(ii) $\kappa(\Delta_0') < \kappa(\mu)$.

(i) We consider the quadratic differentials $q_P(z)dz^2$, $q_{\Delta_0'}(z)dz^2$ for $d_P(x,\tau(x))$ and $d_{\Delta_0'}(x,\tau(x))$ at x, and compare these with the quadratic differential $q(z)dz^2$ for $d(x,\tau(x)$ at x. Let θ_P, $\theta_{\Delta_0'}$ be the angles between q and q_P, $q_{\Delta_0'}$. Consider the annulus $S(\partial\Delta_0, x, \varepsilon_0)$ (9.3), which has modulus $2\pi^2/\varepsilon + O(1)$. Let A be a subannulus of $S(\partial\Delta_0, x, \varepsilon_0)$ of modulus bounded from 0 and separated from each boundary by an annulus of modulus $\pi^2/\varepsilon + O(1)$. Then the measure of A with respect to both q_P, $q_{\Delta_0'}$ is $\geqslant C_1 e^{-\pi^2/\varepsilon}$ for a suitable $C_1 > 0$. One of θ_P, $\theta_{\Delta_0'}$ is bounded from 0 in A. So for $C_2 > 0$ independent of ε,

$$\int_A |\theta_P|^2 |q_P| \geqslant C_2 e^{-\pi^2/\varepsilon} \quad \text{or} \quad \int |\theta_{\Delta_0'}|^2 |q_{\Delta_0'}| \geqslant C_2 e^{-\pi^2/\varepsilon}.$$

So by 9.5, for C_3, $C_4 > 0$,

$$\begin{aligned} F(x) = d(x,\tau(x)) &\geqslant \max(d_P(x,\tau(x)), d_{\Delta_0'}(x,\tau(x)) + C_3 e^{-\pi^2/\varepsilon} \\ &\geqslant \kappa(\mu) - C_4 e^{-2\pi^2/\varepsilon} + C_3 e^{-\pi^2/\varepsilon} \geqslant \kappa(\mu) + (C_3/2) e^{-\pi^2/\varepsilon}. \end{aligned}$$

This yields the required contradiction.

(ii) First, we claim that, if C' is sufficiently small,

$$d_Q(\tau(x), \tau^2(x)) \leqslant \kappa(\mu) - 2C' e^{-2\pi^2/\varepsilon}.$$

To see this, we use an argument similar to 6.11. Let $q(z)dz^2$, $q_1(z)dz^2$ be the quadratic differentials for $d(x,\tau(x)$, $d_Q(\tau(x),\tau^2(x)$ at x, $\tau(x)$. Let s be the holomorphic branched covering used to define $\tau(x)$ from x. Let $\theta(z)$ be the angle between $s^*q(z)$ and $q_1(z)$. Then for $C_1 > 0$ bounded from 0, by 8.3 (as used in 6.11)

$$d_Q((\tau(x), \tau^2(x)) \leqslant d(x, \tau(x) - C_1 \int |\theta|^2 |q_1|.$$

But, because $\kappa(\Delta_0') < \kappa(P)$, $s^*q(z)dz^2$ has an approximate triple pole at $s^{-1}(S(\partial\Delta_0, x, \varepsilon_0))$ — where q_1 has at most a simple pole. So considering the integral over $S(\Delta_{p-1}', \tau(x), \varepsilon_1)$ for a sufficiently small ε_1, we obtain, for $C' > 0$ sufficiently small

$$d_Q(\tau(x), \tau^2(x)) \leqslant \kappa(\mu) - 2C' e^{-2\pi^2/\varepsilon}.$$

Now since we are assuming that (f_0, Γ) is Euclidean, $v_2 \notin Q$. Now we can choose $A(Q)$ so that $v_2 \notin A(Q)$ and $f_0(A(Q)) \subset A(Q)$. Let f_Q be the branched covering which has both critical values in $A(Q)$ and $f_Q \mid f_0^{-1}(Q) = f_0 \mid f_0^{-1}(Q)$. Write $\tau(x) = [\varphi_2]$. Write $v_{2,Q}$ for the critical value of f_Q in the same component of $\overline{\mathbf{C}} \setminus Q$ as v_2. Let τ_Q be the pullback corresponding to f_Q. Then $\varphi_2(v_2)$, $\varphi_2(v_{2,Q})$ lie in a disc of measure $O(e^{-2\pi^2/\varepsilon - 2\pi^2/\delta})$ with respect to the quadratic differential $q_{1,Q}(z)dz^2$ for $d_Q(\tau(x), \tau^2(x))$ at $\tau(x)$. Similarly of $\tau^2(x) = [\varphi_3]$ and $\tau_Q(\tau(x)) = [\varphi_4]$ then

$\varphi_3(y)$, $\varphi_4(y)$ lie in a disc of measure $O(e^{-2\pi^2/\varepsilon-2\pi^2/\delta})$ with respect to the stretch of $q_{1,Q}(z)dz^2$ at $\tau^2(x)$. Then by 9.8, assuming $\delta > 0$ is sufficiently small, and $x' = \tau(x)$

$$d_Q(x', \tau_Q(x')) \leqslant d_Q(\tau(x), \tau^2(x)) - C_1 e^{-2\pi^2/\varepsilon-2\pi^2/\delta} \leqslant \kappa(\mu) - C' e^{-2\pi^2/\varepsilon}$$

So now, replacing Y by $A(Q)$, τ by τ_Q, x by $\tau(x)$ assume that

$$d(x, \tau(x)) \leqslant \kappa(\mu) - C' e^{-2\pi^2/\varepsilon},$$

and we shall obtain a contradiction. Fix an integer k with $k > 20C + \log C'^{-1}$ for C as in 17.4, 17.6. By 17.8, if ε is sufficiently large given k,

$$m_1(\tau^k(x)) \geqslant m_1(x) + \tfrac{1}{4} k \log 2.$$

The lengths of any two loops in ∂P differ by at most C, by 17.6, so for any $\gamma \subset \partial P$ we have

$$m_\gamma(\tau^k(x)) \geqslant m_\gamma(x) - 2C + \tfrac{1}{4} \log k.$$

Then if ε_k is the shortest loop in ∂P for $\tau^k(x)$,

$$2\pi^2/\varepsilon_k \geqslant 2\pi^2/\varepsilon + C + \log C'^{-1}.$$

So (assuming $C > \log C$)

$$F(\tau^k(x)) \leqslant \kappa(\mu) - C' e^{-2\pi^2/\varepsilon} < \kappa(\mu) - C e^{-2\pi^2/\varepsilon_k},$$

contradicting 17.4.

17.10. Let (f_0, Γ) be as in 17.4. We now know that if

$$F(x) \leqslant \kappa(\mu) + C' e^{-2\pi^2/\varepsilon}$$

and C' is sufficiently small then $[f_0, \Gamma] = \mu$ for $[f_0, \Gamma]$ satisfying the Node Condition.

Lemma. — *Take the same hypotheses as in 17.4. Let $\overline{\alpha} \subset \mathrm{int}(\Omega)$ be a gap of Γ and*

$$F(x) \leqslant \kappa(\alpha) + e^{-m}.$$

Then for all $\gamma \subset \partial\alpha$, $m_\gamma(x) \leqslant m/C$.

Proof. — Let Q be the smallest connected subsurface containing the orbit of α. Then by 17.5, for any $\gamma \subset \partial Q$, $m_\gamma(x) \geqslant m - C$ if C is large enough. These γ include all but one loop in $\partial\alpha'$ for each $\alpha' \in [\alpha]$. Suppose that $m_\gamma(x) \leqslant m/C$ for the final boundary component γ of α. We can assume this holds for all $\gamma' \in [\gamma]$ since $d(x, \tau(x)) = F(x)$ is bounded. Let $q(z)dz^2$, $q_{\alpha'}(z)dz^2$ denote the quadratic differentials for $d(x, \tau(x))$, $d_{\alpha'}(x, \tau(x))$ at x and $\theta_{\alpha'}$ the angle between them. For at least one α' we must have

$$\int \theta_{\alpha'}^2 |q_{\alpha'}| \geqslant e^{-m/\sqrt{C}}.$$

Then by 9.5,

$$d(x, \tau(x)) \geqslant d_{\alpha'}(x, \tau(x)) + C^{-1} e^{-m/\sqrt{C}},$$

which gives a contradiction. □

CHAPTER 18

PUSHING THE PULLBACK MAP

18.1. In this chapter, we return to the proof of one of the main results of this work, namely the Topographer's View, which was stated in 5.10. More precisely, we return to the proof of one of the results to which the Topographer's View was reduced, namely, the Level κ Tool of 7.7. We show that it suffices to construct a modification $\tau' : \mathcal{T}/G \to \mathcal{T}/G$ of the pullback map τ. We produce the first versions of the properties of τ' in this chapter — in 18.3, and deduce the Level κ Tool in 18.4. We give a second version in 18.5, and show that it implies the first in 18.9. We give the third version (which is based on the second) in 18.11. In Chapter 19 we shall show that the third version implies the second, and produce a fourth version, in terms of a vector field, and show that the fourth version implies the third. The vector field will then be constructed in Chapters 20-21.

18.2. Introduction of τ'. — Throughout this chapter we fix $B = B(Y, f_0)$ of degree two type (see 1.9). As usual, the structure includes a set $Z \subset Y$ invariant under f_0, and we write $G = \pi_1(B, f_0)$. We write $\mathcal{T} = \mathcal{T}(B) = \mathcal{T}(Y)$ (see 6.2), and we use the pullback map $\tau : \mathcal{T} \to \mathcal{T}$ (see 6.6, 6.7). We write d for the usual Teichmüller distance on $\mathcal{T}(Y)$, we write d_Z for the semimetric defined by projection to $\mathcal{T}(Z)$, and as in 7.7, we define

$$F(x) = d(x, \tau(x)) = d_Z(x, \tau(x)).$$

Initially, this is defined for $x \in \mathcal{T}$, but since F is G-invariant, we can also regard it as a function on $\mathcal{T}/G$. We can also, of course, regard τ as a function on $\mathcal{T}/G$. We recall that the Level κ Tool is concerned with a continuous map $\alpha : \Delta \to \mathcal{T}/G$, where Δ is the unit interval, disc or circle, with $F(\alpha(x)) \leqslant \kappa$ for all $x \in \Delta$, and possibly some restriction on $\alpha \mid \partial$ for some $\partial \subset \partial\Delta$. Note that if Δ is the circle then α and $\tau \circ \alpha$ are freely homotopic. If $\alpha(\partial) \subset V$, then $\tau \circ \alpha \mid \partial = \alpha \mid \partial$. If Δ is a disc with $\partial = \partial\Delta$, α and $\tau\alpha$ are homotopic via a homotopy preserving $\partial\Delta$, since $\mathcal{T}/G$ is a $K(\pi, 1)$. Thus it would be natural, in trying to modify α, to consider the sequence $\tau^i\alpha$. In fact we need to define

$$\tau' : \mathcal{T} \longrightarrow \mathcal{T}$$

where τ' is a G-invariant modification of τ with certain additional properties. Thus, τ' can also be regarded as a function on $\mathcal{T}/G$.

18.3. First Version Properties of τ'

Numbers used in the properties. — The number ε_0 is chosen less than the Margulis constant. Numbers $\varepsilon_2 < \varepsilon_1 < \varepsilon_0$ and $\delta_1 > 0$ are fixed, and

$$\eta_1 = E_2^{-1} e^{-2\pi^2/\varepsilon_1}.$$

The number E_2 was introduced in 7.7, together with constants E_0, E_1. So far, the only restriction on it is that E_0 be sufficiently large, depending only on $\#(Y)$, and E_1, E_2 sufficiently large given E_0 and E_1. The number δ_1 is independent of ε_1 and ε_2. It actually depends on δ_0 of 7.8.

Basic Property of τ'. — Identifying τ' with its lift to $\mathcal{T}$,

$$\tau' = \sigma \circ \tau \circ \sigma_1,$$

where σ, σ_1 are the identity on $\widetilde{V}$. Thus, $\tau' = \tau$ on $\widetilde{V}$. Moreover,

$$F(\tau'(x)) \leqslant F(x),$$

with strict inequality whenever $\tau'(x) \neq \tau(x)$. From Property 3 of 6.6 it then follows that, for a k depending only on $\#(Y)$, $F(\tau'^k(x)) < F(x)$ whenever $F(x) > 0$.

Then we can choose N large enough that, if $\delta_1 \leqslant F(x) \leqslant \kappa$ and $\tau'^i x \in \mathcal{T}_{\geqslant \varepsilon_2}$ for all $i \leqslant N$ then

$$F(\tau'^N x) \leqslant \kappa - \eta_1.$$

We then have to consider what happens if $\tau'^i(x) \in \mathcal{T}_{<\varepsilon_2}$ for some $i \leqslant N$. This is the purpose of the following property. It says that entry into $\mathcal{T}_{<\varepsilon_2}$ occurs in the right way — which, as we shall see, essentially follows from the properties of τ which are stated and proved in Chapter 6. More importantly, it says that any subsequent exit from $\mathcal{T}_{<\varepsilon_1}$ occurs in the right way — through the "plugs" constructed in 7.7. The constant $L > 0$ will be suitably chosen. The subsets K_0, K_1 of $\mathcal{T}/G$ are as in the statement of the Level κ Tool in 7.7, but have not yet been defined: they will be described in 18.8 and defined precisely in 19.3. By abuse of notation, we write K_i also for the preimage of K_i in $\mathcal{T}$. The integer N_0 will be determined only by Y and f_0, as we shall see in 17.8. The integer k_1 depends only on Y and E_0.

Exit through Plug Property for τ'. — Let $x \in \mathcal{T}_{<\varepsilon_2}$ with $\delta_1 \leqslant F(x) \leqslant \kappa$, where either $x \in K_0$, or $x = \tau'^m(x')$ with $x' \in \mathcal{T}_{\geqslant \varepsilon_2}$ and $m \geqslant N_0$ is minimal with $\tau'^m(x') \in \mathcal{T}_{<\varepsilon_2}$. Then there is a minimal $\mu = [f_0, \Gamma]$ such that $x \in \widetilde{T}_{\geqslant \mu}(L\varepsilon_2)$. Let $\varepsilon \leqslant \varepsilon_1$ with $\varepsilon/\varepsilon_2$

sufficiently large. Let i be minimal with $\tau'^i(x) \notin \widetilde{T}_{\geqslant\mu}(\varepsilon)$. Then

$$\tau'^i(x) \in \widetilde{T}_{\geqslant\mu}(\varepsilon/(1-E_0\varepsilon)), \qquad \tau'^j(x) \in K_1(\mu,\varepsilon) \quad \text{for } i \leqslant j \leqslant i+k_1,$$
$$F(\tau'^{i+k_1}(x)) \leqslant \kappa(\mu) - E_0 e^{-2\pi^2/\varepsilon}, \qquad \tau'^{i+k_1}(x) \notin \widetilde{T}_{\geqslant\mu}(\varepsilon/(1-E_0\varepsilon)).$$

It then follows from the Basic Property that $F(\tau'^j(x)) \leqslant \kappa(\mu) - E_0 e^{-2\pi^2/\varepsilon}$ for all $j \geqslant i+k_1$. The Exit through Plug Property then implies that for some ε'_j and i_1, $\tau'^j(x) \in K_1(\mu,\varepsilon'_j)$ for $i \leqslant j \leqslant i_1$, with $\varepsilon'_{i_1} = \varepsilon_1$: for example, for i, ε as above and some j, $i < j \leqslant i+k_1$, the number ε'_j satisfies $\varepsilon'_j \leqslant \varepsilon/(1-E_0\varepsilon)$.

18.4. Proof of the Level κ Tool given 18.3. — Recall that the Level κ Tool of 7.7 concerns $\alpha : \Delta \to \mathcal{T}/G$ with $F(\alpha(x)) \leqslant \kappa$ for all $x \in \Delta$ and $\alpha(\partial) \subset T''_\kappa(\varepsilon,\varepsilon')$ for $\partial \subset \partial\Delta$. We recall (from 7.7) that

$$\begin{aligned} T''_\kappa(\varepsilon) = \bigcup\{K_1(\mu,\varepsilon) : \mu \text{ minimal nonempty}, \kappa(\mu) \leqslant \kappa\} \\ \cup\, T'_\kappa(\varepsilon) \cup \{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon}\}, \\ T''_\kappa(\varepsilon,\varepsilon') = \bigcup\{K_0(\mu,\varepsilon'') : \mu \text{ minimal nonempty}, \kappa(\mu) \leqslant \kappa,\ \varepsilon' \leqslant \varepsilon'' \leqslant \varepsilon\} \\ \cup\, T'_\kappa(\varepsilon') \cup \{x : F(x) \leqslant \kappa - E_0 e^{-2\pi^2/\varepsilon}\}. \end{aligned}$$

These were originally defined as subsets of $\mathcal{T}/G$. We identify them with their preimages in $\mathcal{T}$. The Exit through Plug Property then implies that, if ε is sufficiently small and ε/ε' sufficiently large,

$$\tau'(T_\kappa(\varepsilon,\varepsilon')) \subset T''_\kappa(\varepsilon),$$

and if $\varepsilon_2 \leqslant \varepsilon'$ and $\tau'^i(x) \notin \mathcal{T}_{<\varepsilon_2}$ for $i \leqslant N_0$ (N_0 as in 17.3), but $\tau'^n(x) \in \mathcal{T}_{<\varepsilon_2}$ for a least n, then $\tau'^m(x) \in T''_\kappa(\varepsilon)$ for all $m \geqslant n$. Now choose ε_2 sufficiently small that $\tau'^i(x) \in \mathcal{T}_{\geqslant\varepsilon_2}$ for $i \leqslant N_0$ and $x \in \alpha(\Delta)$. Let N be as in 18.3. Inductively, we can define α_i by $\alpha_0 = \alpha$, and α_{i+1} is an extension of $\tau'\alpha_i$ so that $\alpha_i = \alpha$ on ∂, and

$$\operatorname{Im}(\alpha_{i+1} \smallsetminus \tau'\alpha_i) \subset T''_\kappa(\varepsilon).$$

Then we take $\alpha' = \alpha_N$. If ε is sufficiently small that

$$D_0 e^{-2\pi^2/\varepsilon} \leqslant D_2^{-1} e^{-2\pi^2/\varepsilon_1},$$

then $\alpha'(\Delta) \subset T''_\kappa(\varepsilon)$, as required.

18.5. Second Version of Properties of τ'. — We shall now describe the *Second Version* of the properties of $\tau' = \sigma \circ \tau \circ \sigma_1$, $\sigma = \sigma_2 \circ \sigma_1$. We shall see that they imply the First Version Properties of τ', in 18.6-10. Let the numbers $\varepsilon_0 > \varepsilon_1 > \varepsilon_2 > 0$ be as in the First Version Properties. The Second Version Properties are for x with $F(x) \leqslant \kappa$ and depend on $\delta_1 > 0$, and constants C_1, L_1, L_2, L'_1, L'_2. These constants are yet to be chosen. The constant C_1 will depend only on κ and $\#(Y)$. The constant L_2 will depend on L_1, and L'_2 on L'_1. We shall have $L_1 \leqslant L'_1$, $L_2 \leqslant L'_2$. We use the definitions of 17.2-3.

Basic Properties. — As before, we have

$$F(\tau'(x)) \leqslant F(x), \quad F(\sigma_1(x)) \leqslant F(x). \tag{1}$$

We have $\sigma(x) = \sigma_1(x) = x$, except when $x \in \mathcal{T}(\Gamma, \varepsilon_0) \cap \mathcal{T}(\partial\Delta_0, \varepsilon)$ for some $\varepsilon \leqslant \varepsilon_0$ and (f_0, Γ) which is (L_1', L_2')-adapted to x, and $m_1(x) \geqslant \varepsilon_1^{-1/2}$.

If $\gamma \in \Gamma \smallsetminus \Gamma_2(f_0, \Gamma)$ (2.5) for all such Γ, then

$$m_\gamma(\sigma_j(x)) \leqslant m_\gamma(x) + C_1. \tag{2}$$

(3) For all loops $\gamma \subset \mathrm{int}(\Omega)$ for the fixed union Ω of some such Γ, and all $\gamma \subset \overline{\Omega}$ if $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) \geqslant \kappa_0(\Gamma)$, one of (3a), (3b) holds:

Loop. — *F-between.* $\gamma \subset \partial\beta$ for some gap $\overline{\beta} \subset \mathrm{int}\Omega_1$ for the fixed union Ω_1 of some such $\Gamma_1 \supset \Gamma$, and

$$F(\sigma_j(x)) \leqslant \kappa(\beta) + e^{-m_{[\gamma]}(\sigma_j(x))/C_1}. \tag{3a}$$

$m_{[\gamma]}$ *nondecreasing*

$$m_{[\gamma]}(\sigma_j(x)) \geqslant m_{[\gamma]}(x). \tag{3b}$$

Let (f_0, Γ) be (L_1, L_2)-adapted to x and $m_1(x) \geqslant \varepsilon_1^{-3/4}$. Let $F(x) \geqslant \delta_1$. In addition, one of the following holds.

Nondiscrete or Euclidean. — There exists (f_0, Γ_1) which is (L_1, L_2)-adapted to $\tau'(x)$ with $\kappa_0(\Gamma_1) \geqslant \kappa_0(\Gamma)$ which is either nondiscrete, or discrete Euclidean. If (f_0, Γ_1) is discrete Euclidean then for any $\alpha \subset \Delta_j'$ or $\alpha \subset \partial\Omega$, $d(\pi_\alpha \circ \sigma_1(x), \pi_\alpha(x)) \leqslant C_1^{-1}$ and $d(\pi_\alpha \circ \tau'(x), \pi_\alpha \circ \tau \circ \sigma_1(x)) \leqslant C_1^{-1}$.

F-between

$$m_{\mathrm{int}}(\tau'(x)) \geqslant m_1(x)/C_1^2 \tag{4}$$

$$\kappa_0(\Gamma) + e^{-C_1^2 m_1(\tau'(x))} \leqslant F(\tau'(x)) \leqslant \kappa_0(\Gamma) + e^{-m_1(\tau'(x))/C_1^{20}}.$$

m_1-control. — Either $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) > \kappa_0(\Gamma)$, or (4) holds. Let $\gamma \subset \mathrm{int}(\Omega)$. Then either (3a) holds; or $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) \leqslant \kappa_0(\Gamma)$ and

$$m_{[\gamma]}(\tau'(x)) + m_{[\gamma],\partial}(\tau'(x)) \leqslant m_{\mathrm{int}}(\tau'(x))/C_1^2 m_{\partial,0}(\tau'(x));$$

or

$m_{[\gamma]}$-*increasing*

$$m_{[\gamma]}(\tau'(x)) \geqslant m_{[\gamma]}(\tau \circ \sigma_1(x)) + C_1^2 e^{-m_{[\gamma],\partial}(\tau\circ\sigma_1(x))}.$$

In addition, if $[\gamma] = [\gamma_0]$ and $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) > \kappa_0(\Gamma)$, then

$$m_{[\gamma]}(\tau'(x)) \geqslant m_{[\gamma]}(\tau \circ \sigma_1(x)) + m_{\gamma_0'}(x) + 1.$$

Minimal. — The loop set (f_0, Γ) is minimal nonempty, $\kappa_0'(\Gamma) < \kappa_0(\Gamma)$ and

$$F(\tau'(x)) \leqslant \kappa_0(\Gamma) + e^{-m_1(\tau'(x))/C_1}.$$

If in addition $F(x) \leqslant \kappa_0(\Gamma) + e^{-C_1 m_1(x)}$, then

$$F(\tau'(x)) \leqslant F(x) - e^{-m_{\partial,0}(x) - C_1^5 L_2/\varepsilon_0}.$$

Preliminary definition of K_i

$$\{\sigma(x) : x \in \widetilde{T}_\mu(\varepsilon_1),\ \mu \text{ minimal},\ m_2(x) \leqslant E_i\} \subset K_i.$$

We shall give the full definition in 18.12.

18.6. Lemma. — *Let N_0 be sufficiently large given $\#(Y)$, κ. Let the Second Version Basic Property hold. Let $\varepsilon_2 > 0$ be sufficiently small. Let $x' \in \mathcal{T}$, $F(x') \leqslant \kappa$. Let $\tau'^i(x') \in \mathcal{T}_{<\varepsilon_2}$ for a least $i = i_0 \geqslant 0$. Let $i_0 \geqslant N_0$. Write $x = \tau'^{i_0}(x')$. Then for a constant C_3 depending only on C_1, and some minimal nonempty μ, $m_2(x) \leqslant C_3 m_1(x)$, $x \in \widetilde{T}_{\geqslant\mu}(C_3\varepsilon_2)$.*

Proof. — This is very similar to 6.13 (proof of property 5 of τ). Note that $\tau' = \tau$ except on a union of sets $\mathcal{T}(\Gamma, \varepsilon_0)$ with $\Gamma_2(f_0, \Gamma) \neq \varnothing$, where, indeed, $\Gamma_2(f_0, \Gamma)$ includes all loops of length $\leqslant \varepsilon_0/L_2'$. The Second Version Basic Property gives $m_\gamma(\tau'(x)) \leqslant m_\gamma(\tau(x)) + C_1$ when γ contributes to either $m_2(\tau(x)$ or $m_2(\tau'(x))$. We also have $m_1(\tau'(x)) \geqslant m_1(\tau(x))/C_1^5$. So, exactly as in 6.13, the Second Version basic property implies that, for an integer N_0 and $\lambda_1 \in (0,1)$ depending only on (Y, f_0); C_4 depending only on (Y, f_0) and on C_1, ε_0; $\varepsilon \leqslant \varepsilon_0$ sufficiently small and $F(y) \leqslant \kappa$, $\tau'^{N_0}(y) \in \mathcal{T}_{<\varepsilon}$,

$$m_2(\tau'^{N_0}(y)) \leqslant \lambda_1 m_2(y) + C_4 \sum_{i=0}^{N_0} m_1(\tau'^i(y)). \tag{3}$$

The contribution from m_1 occurs only if there are loops $\gamma_0 \subset \partial\Omega$ and $\gamma_0' \in \Gamma$ isotopic to γ_0 in $\overline{\mathbf{C}} \smallsetminus Z$ with v_2 separating γ_0 and γ_0'. In fact, given $\lambda_1 \in (0,1)$, we can choose N_0 and C_4 so that (3) holds. We also have, if γ is in the backward orbit of some Levy cycle then some component of $f_0^{-N_0}(\gamma)$ is in the fixed set, and hence

$$m_1(\tau'^{N_0}(y)) \geqslant m_\gamma(y)/C_4.$$

Then if λ_1 is small enough (as we can assume), and C_3 is suitably chosen given C_1, if $\tau'^i(y) \notin \mathcal{T}_{<\varepsilon}$ for $0 \leqslant i < N_0$, and $y' = \tau'^{N_0}(y)$, then $m_2(y') \leqslant C_3 m_1(y')$, and $y' \in \mathcal{T}(\Gamma, C_3\varepsilon)$ for (f_0, Γ) as above. So then there is a minimal nonempty μ with $\mu \leqslant [f_0, \Gamma_2(f_0, \Gamma)]$. (See Chapters 2-3, in particular 2.5, 2.16.). By definition (see 7.2) this means that $y' \in \widetilde{T}_{\geqslant\mu}(C_3\varepsilon)$. We apply this with $y = \tau'^{i_0 - N_0}(x')$, $y' = \tau'^{i_0}(x') = x$. □

18.7. Indiscrete increases m_1. — Now we show that m_1-increasing holds automatically if (f_0, Γ) is not discrete.

Lemma. — *Suppose that L_1 is suitably chosen (depending only on $\#(Y)$). Let $x \in \mathcal{T}(\Gamma, \varepsilon)$ where (f_0, Γ) is not discrete, $L_1\varepsilon \leqslant \varepsilon_0$. Then*

$$m_{1,\Gamma}(\tau(x)) > m_{1,\Gamma}(x) + \pi^2/\varepsilon,$$
$$m_{\partial,\Gamma}(\tau(x)) \geqslant m_{\partial,\Gamma}(x) + \pi^2/\varepsilon,$$
$$m_{\partial,\Gamma}(\tau'(x)) \geqslant m_{\partial,\Gamma}(x) + \pi^2/\varepsilon.$$

Proof. — Let γ_0' (if it exists) be the loop separating v_1 and v_2 which is homotopic in $\overline{\mathbf{C}} \smallsetminus Z$ to a loop in $\partial\Omega$ and with $x \in \mathcal{T}(\gamma_0', \varepsilon_0)$. Because $m_\gamma(x)$ is the modulus of $S(\gamma, x, \varepsilon_0)$ (see 9.3) to within $O(1/\varepsilon_0)$, there is a matrix $(a(\gamma, \delta))$ such that for all $\gamma \in \Gamma$,

$$\sum_\delta a(\gamma,\delta)m_\delta(x) + O(1/\varepsilon_0) \leqslant m_\gamma(\tau(x)) \leqslant \sum_\delta a(\gamma,\delta)m_\delta(x) + O(1/\varepsilon_0).$$

Here, all entries of $a(\gamma, \delta)$ are nonnegative integers independent of x except when $\gamma = \gamma_0 = \partial\Delta_0(f_0, \Gamma)$ or $\gamma = \gamma_0'$. If $\gamma \subset \overline{\Omega}$, then $a(\gamma, \delta) = 1$ for precisely one $\delta \subset \overline{\Omega}$. Since Γ is not discrete, $a(\gamma, \delta) \geqslant 1$ for at least one $\gamma \subset \overline{\Omega}$ and $\delta \in \Gamma$, $\delta \not\subset \overline{\Omega}$. So then, if L_1 is large enough.

$$m_{\partial,\Gamma}(\tau(x)) \geqslant m_{\partial,\Gamma}(x) + a(\gamma,\delta)m_\gamma(x) - O(1/\varepsilon_0) \geqslant m_{\partial,\Gamma}(x) + \frac{3\pi^2}{2\varepsilon},$$

and similarly for $m_{1,\Gamma}$. The Second Version Basic Properties imply that the $m_{\partial,\Gamma}$ inequality holds for τ' replacing τ, with π^2/ε replacing $3\pi^2/2\varepsilon$. □

18.8. Lemma. — *Let (f_0, Γ) be discrete Euclidean. Let $x \in \mathcal{T}(\Gamma, \varepsilon_0) \cap \mathcal{T}(\partial\Delta_0', \varepsilon)$, $x \notin \mathcal{T}(\gamma, \nu)$ for $\gamma \subset \mathrm{Int}(\Delta_0')$, $L_1\varepsilon \leqslant \nu$. Let $F(x) \leqslant \kappa$. Let $\gamma_0 = \partial\Delta_0' \cap \partial\Omega$. Let k_0 be as in 17.8. Let*

$$d(\pi_\alpha \circ \tau'^{n+1}(x)), \pi_\alpha \circ \tau \circ \sigma_1 \circ \tau^n(x)) \leqslant C_1^{-1}, \quad d(\pi_\alpha \circ \sigma_1 \circ \tau'^n(x), \pi_\alpha \circ \tau'^n(x)) \leqslant C_1^{-1}$$

for all $0 \leqslant n \leqslant 2k_0$ and $\alpha = \Delta_j'$ or $\alpha \subset \partial\Omega$. Then for $k_0 \leqslant k \leqslant 2k_0$, if C_1 is sufficiently large given k_0,

$$m_{[\gamma_0]}(\tau'^k(x)) > m_{[\gamma_0]}(x).$$

Proof. — By induction on n, assuming C_1 is large enough, for $n \leqslant 2k_0$,

$$d(\pi_\alpha \circ \tau^n(x), \pi_\alpha \circ \tau'^n(x)) \leqslant nC_1^{-3/4} \leqslant C_1^{-1/2}.$$

Then the result follows from that for τ in 17.8. □

18.9. Lemma. — *Assume that the Second Version properties hold. Let $x \in \widetilde{T}_{\geqslant\mu}(C_3\varepsilon)$ with $m_2(x) \leqslant C_3 m_1(x)$. The following holds for suitable constants C_7 depending only on C_1, C_3 and $\#(Y)$ and C_2 depending only on C_7, C_1, C_3, and $\#(Y)$. Let j_1,*

$0 \leqslant j_1 \leqslant +\infty$, *be defined by:* $j = j_1$ *is the first index with* $C_7 m_1(\tau'^j(x)) \leqslant m_1(x)$. *Then for* $0 \leqslant j \leqslant j_1$,

$$m_2(\tau'^j(x)) \leqslant C_2 m_1(\tau'^j(x)),$$

and if $y = \tau'^{j_1+1}(x)$,

$$F(y) \leqslant \kappa(\mu) - e^{-m_{\partial.0}(y) - L_2 C_1^5/\varepsilon_0}.$$

Proof. — The same method as 18.6 gives, for a suitable $\lambda_2 \in (0,1)$, if C_4 is large enough,

$$(1) \qquad m_2(\tau'^n(x)) \leqslant \lambda_2^n m_2(x) + \sum_{i=0}^{n} C_4 \lambda_2^{n-i} m_1(\tau'^i(x)).$$

This gives a bound on $m_2(\tau'^n(x))/m_1(\tau'^n(x))$ if $n \leqslant j_1$ and the quantity

$$\sum_{i=0}^{n} \frac{\lambda_2^{n-i} m_1(\tau'^i(x))}{m_1(\tau'^n(x))}$$

is bounded. We consider this for $n \leqslant j_0$, where $j = j_0$ (possibly with $j_0 = +\infty$) is the first index such that Minimal occurs for $y = \tau'^j(x)$. We obviously obtain a bound

$$(2) \qquad m_2(\tau'^n(x)) \leqslant C_2 m_1(\tau'^n(x)), \quad n \leqslant j_0$$

if C_2 is sufficiently large given C_4, C_5 and λ_2, and if, as we claim, for $0 \leqslant i \leqslant n \leqslant j_0$ and a constant C_5,

$$(3) \qquad C_5 m_1(\tau'^n(x)) \geqslant m_1(\tau'^i(x)), \quad 0 \leqslant i \leqslant n \leqslant j_0.$$

We shall show that this is true, for C_5 depending only on C_1 and $\#(Y)$. By the Properties of 18.5, for each $0 \leqslant i \leqslant n$ there exists a loop set (f_0, Γ_i) which is (L_1, L_2)-adapted to $\tau'^i(x)$ with fixed union Ω_i.

Now take $0 \leqslant j \leqslant j_0 - k_0$, where k_0 is as in 17.8. For $1 \leqslant \ell < q$, let $\Gamma_{j+\ell}$ be nondiscrete or discrete Euclidean. If Γ_j is discrete Euclidean then either, by m_1-control,

$$m_\partial(\tau'^{j+1}(x)) \geqslant m_\partial(\tau'^j(x)) + C_1^2 e^{-m_{\partial.0}(\tau'^j(x))}$$

or we can take $q \geqslant k_0$ for k_0 as in 17.8 with Second Version Euclidean holding for $j+\ell$, $0 \leqslant \ell < q$. Suppose we have the latter. Then the fixed union $\Omega_{j+\ell}$ is increasing for $0 \leqslant \ell < q$. Since $\Omega_{j+\ell}$ can only strictly increase boundedly finitely often for $0 \leqslant \ell < q$, we can assume that it is constant. By 17.8, 18.7,

$$(4) \qquad m_\partial(\tau'^{j+q}(x)) \geqslant m_\partial(\tau'^j(x)) + \tfrac{1}{2} q \log 2.$$

We claim that

$$(5) \qquad m_1(\tau'^{j+q}(x)) \geqslant C_1^{-3} m_1(\tau'^j(x)).$$

Replacing j by some $j+\ell$ if necessary, we can assume that $m_1(\tau'^{j+\ell}(x)) \leqslant m_1(\tau'^j(x))$ for $1 \leqslant \ell \leqslant q$. Then (5) is immediate unless $m_{\mathrm{int}}(\tau'^j(x)) \geqslant C_1^2 m_\partial(\tau'^j(x))$. So now assume this. Take any $\gamma \in \Gamma_j$, $\gamma \subset \mathrm{Int}(\Omega_j)$. We need to obtain a lower bound on $m_{[\gamma]}(\tau'^{j+q}(x))/m_{[\gamma]}(\tau'^j(x)$. We can assume without loss of generality that (3a) of 8.5

does not hold for γ and any $j+\ell$, $0\leqslant \ell\leqslant q$. For if it does hold for $j+\ell$, $j+\ell'$ with $j+\ell<j+\ell'$ then we have, using 17.10 for the lefthand inequality, for $j+\ell\leqslant i\leqslant j+\ell'$

$$\kappa(\beta)+e^{-C_1 m_{[\gamma]}(\tau'^i(x))}\leqslant F(\tau'^i(x))\leqslant F(\tau'^{j+\ell}(x)\leqslant \kappa(\beta)+e^{-m_{[\gamma]}(\tau'^{j+\ell}(x))/C_1}.$$

We can remove all such intervals $\{i : j+\ell\leqslant i\leqslant j+\ell'\}$ from consideration. If what remains is bounded there is nothing to prove. If an interval have length $\geqslant k_0$ then we can assume it is the original interval $\{i : i\leqslant i\leqslant j+q\}$. So now, given γ, assume (3b) of 8.5 holds for all $j+\ell$, $0\leqslant\ell\leqslant q$. Then by 17.7-8, 18.7,

$$\begin{aligned} m_{[\gamma]}(\tau'^{j+q}(x)) &\geqslant m_{[\gamma]}(\tau'^{j+1}(x))-\sum_{\ell=1}^{q-1} C_1 e^{-m_{[\gamma].\partial}(\tau'^{j+\ell}(x))} \\ &\geqslant m_{[\gamma]}(\tau'^{j}(x))-\sum_{\ell=0}^{q-1} C_1 e^{-m_{[\gamma].\partial}(\tau'^{j+\ell}(x))}\geqslant m_{[\gamma]}(\tau'^{j}(x))-C_1^{3/2} \end{aligned}\tag{6}$$

So we obtain (5). Note that if we use (4), and (6) for $[\gamma']$ between γ and $\partial\Omega$, to obtain a better estimate on $m_{[\gamma],\partial}(\tau'^{j+\ell}(x)$ then we can improve use (6) to obtain

$$m_{[\gamma]}(\tau'^{j+q}(x)\geqslant m_{[\gamma]}(\tau'^{j+1}(x))-C_1^{3/2}e^{-m_{[\gamma].\partial}(\tau'^{j}(x))}.\tag{7}$$

This gives (3) in some cases (with $i=j$ and $n=j+q$). Now we work for (3) in general.

We can assume that F-between does not hold for $i\leqslant j\leqslant n$, because if F-between holds for $\tau'^j(x)$, $\tau'^\ell(x)$, $j<\ell$,

$$\kappa_0(\Gamma)+e^{-C_1^5 m_1(\tau'^\ell(x))}\leqslant F(\tau'^\ell(x))\leqslant F(\tau'^j(x))\leqslant \kappa_0(\Gamma)+e^{-m_1(\tau'^j(x))/C_1^{23}},$$

and $m_1(\tau'^\ell(x))\geqslant m_1(\tau'^j(x))/C_1^{28}$.

We can assume there is a first index i_1, $i\leqslant i_1\leqslant n$ such that (f_0,Γ_{i_1}) is discrete nonEuclidean. Then we can assume that $i_1=i$ by (4). Then m_1-control or nondiscrete-Euclidean occurs for all $i\leqslant j\leqslant n$. We can assume that

$$m_{\mathrm{int},\Gamma_i}(\tau'^i(x))\geqslant m_{1,\Gamma_i}(\tau'^i(x)/C_1.\tag{8}$$

Otherwise take the first $i=i_2\geqslant i_1$ for which this happens (if at all). By Second Version m_1-control we have, for $i_1\leqslant j<i_2$,

$$m_{\partial,\Gamma_{j+1}}(\tau'^{j+1}(x))\geqslant m_{\partial,\Gamma_j}(\tau'^j(x)).$$

So now assume (8). We can also assume that $\Omega(\Gamma_i)\subset\Omega(\Gamma_j)$ for $i\leqslant j\leqslant n$. Otherwise we can find boundedly finitely many indices i_k with $i=i_2$, and $\Omega(\Gamma_{i_k})\subset\Omega(\Gamma_j)$ for $i_k\leqslant j<i_{k+1}$, but $\Omega(\Gamma_{i_{k+1}})$ strictly smaller than $\Omega(\Gamma_{i_k})$. But then by m_1-control, $m_1(\tau'^{i_{k+1}}(x))\geqslant m_1(\tau'^{i_k}(x))/C_1$. Now for $i\leqslant j\leqslant n$, let $\Gamma_{i,j}$ be the set of loops γ of Γ_i such that $\gamma\subset\mathrm{int}(\Omega(\Gamma_i))$ such that either nondiscrete/discrete Euclidean or Loop-F-between or $m_{[\gamma]}$-increasing of 8.5 occurs for $\tau'^k(x)$, $i\leqslant k\leqslant j$. By the definition, $\Gamma_{i,j}$ is decreasing in j. If $\gamma\in\Gamma_{i.n}$ then

$$m_{[\gamma]}(\tau'^n(x))\geqslant m_{[\gamma]}(\tau'^i(x))/C_1^4.$$

We see this as follows. As before, we can bound $m_{[\gamma]}(\tau'^{j}(x))/m_{[\gamma]}(\tau'^{\ell}(x))$ if $i \leqslant j \leqslant \ell \leqslant n$ (3a) holds for j and some j', $\ell \leqslant j' \leqslant n$. So we can assume this never happens for γ. By (7), there is little decrease in $m_{[\gamma]}$ over a string of nondiscrete/discrete Euclidean, and $m_{[\gamma]}$ has a compensating increase at each preceding occurrence of $m_{[\gamma]}$-increasing. So it remains to show that for each $i \leqslant j \leqslant n$, $\Gamma_{i,j} \neq \varnothing$ and contains at least one loop γ with $m_{[\gamma]}(\tau'^{j}(x)) \geqslant m_1(\tau'^{i}(x))/C_1^N$ for some $N \leqslant 4\#(Y)$. We do this by induction. It is true for $j = i$ with $N = 1$. We only need to show that each time $\Gamma_{i,j+1}$ is strictly smaller than $\Gamma_{i,j}$, it contains a loop of length $\geqslant \min(m_1(\tau'^{i}(x)), m_1(\tau'^{j}(x))/C_1^2$. We can only have $\gamma \in \Gamma_{i,j} \setminus \Gamma_{i,j+1}$ if $m_{[\gamma]}(\tau'^{j+1}(x)) + m_{[\gamma],\partial}(\tau'^{j+1}(x))$ is small compared to more centrally based loops. In this case, we get the required lower bound on $m_{[\gamma']}(\tau'^{j+1}(x))$ for at least one γ' more central than γ. So the proof of (3) is completed.

By the definition of j_0, we also have $F(y) \leqslant \kappa(\mu) + e^{-m_1(y)/C_1}$, for $y = \tau'^{j_0}(x)$. If C_2 is sufficiently large given a constant C_7, and $j = j_1$ is the first index $\geqslant j_1$ with $C_7 m_1(\tau'^{j}(x)) \leqslant m_1(y)$, then we have (2) for $n = j_1$ also, because (3) holds for $n = j_1$ and a suitable constant C_5. We also have, for $y' = \tau'^{j_1}(x)$,

$$F(y') \leqslant \kappa(\mu) + e^{-C_7 m_1(y')/C_1}.$$

By 17.7, assuming C_7 is large enough, the fixed union for y' is irreducible, (f_0, Γ_{j_1}) is discrete, and non-Euclidean. Then Minimal of 18.5 gives, for $y'' = \tau'^{j_1+1}(x)$,

$$F(y'') \leqslant \kappa(\mu) - e^{-m_{\partial,0}(y'') - L_2 C_1^5/\varepsilon_0}. \qquad \square$$

18.10. Proof of the Exit Through Plug Property. — The following completes the proof of the Exit Through Plug Property of 18.3, assuming that τ' satisfies the Second Version Properties. It follows from the lemma that for suitable k_1, E_0, E_1 and E_2, and if ε_2, ε_1 are sufficiently small given ε_1, ε_0, and $\varepsilon_1 \geqslant \varepsilon \geqslant C_{13}\varepsilon_2/\log\varepsilon_2$, the First Version Properties hold.

Lemma. — *We continue with the same notation and hypotheses as in 18.9. Let $j_2 \geqslant 0$ be minimal with $m_1(\tau'^{j_2}(x)) \leqslant m_1(x)/(\log m_1(x))^2$. Let j_3 be minimal with $\tau'^{j}(x) \in \widetilde{T}_{\geqslant\mu}(\varepsilon_1)$ for $0 \leqslant j \leqslant j_3$ and $\tau'^{j_3+1}(x) \notin \widetilde{T}_{\geqslant\mu}(\varepsilon_1)$. Then, enlarging C_2 if necessary (but still depending only on C_1, C_3, $\#(Y)$)*

$$m_2(\tau'^{j}(x)) \leqslant C_2 m_1(\tau'^{j}(x)) \quad \text{for } 0 \leqslant j \leqslant j_3,$$

and for a suitable constant E_0' depending only on $\#(Y)$, if ε_2 is sufficiently small,

$$m_2(\tau'^{j}(x)) \leqslant E_0', \quad j_2 \leqslant j \leqslant j_3.$$

Proof. — By 17.7, for $j_1 < j \leqslant j_3$, the fixed union P is irreducible and Γ_j is discrete. Let ε_j be the length of the longest loop in ∂P for $\tau'^{j}(x)$. By induction on j, and the Minimal property,

$$F(\tau'^{j+1}(x)) \leqslant F(\tau'^{j}(x)) - e^{-C_1 L_2/\varepsilon_0 - 2\pi^2/\varepsilon_j},$$

and $B(\Delta_0', f_0', \Gamma)$ has nonEuclidean orbifold. So by 17.4, it follows that $\{1/\varepsilon_j\}$ is essentially decreasing, that is, for an integer k_2 depending only on C_2 and ε_0, and suitable C_{10},

$$1/\varepsilon_{j'} \leqslant (1/\varepsilon_j) - 1 \text{ for } p \geqslant j' \geqslant j + k_2, \quad 1/\varepsilon_{j'} \leqslant (1/\varepsilon_j) + C_{10} \text{ for } j \leqslant j' \leqslant j + k_2.$$

Also by 17.6, for all γ, $\gamma' \subset \partial P$, $m_\gamma(x)$, $m_{\gamma'}(x)$, $m_\gamma(\tau(x))$ are the same to within a bounded constant. Hence, if γ_0', exists, $m_{\gamma_0'}(x)$ must be bounded. It then follows that for suitable $\lambda_2 \in (0,1)$ and $j_1 \leqslant j \leqslant j_3$, if $\varepsilon_0/\varepsilon$ is large enough, and for a suitable constant E_0', depending only on $\#(Y)$,

$$m_2(\tau'^{j+1}(x)) \leqslant \mathrm{Max}(E_0', \lambda_2 m_2(\tau'^{j}(x)), \; m_1(\tau'^{j+1}(x))) \geqslant m_1(\tau'^{j}(x)) - C_{11}/\varepsilon_0.$$

It follows that if $j_2 \leqslant j \leqslant j_3$ then $m_2(\tau'^{j}(x)) \leqslant E_0'$. □

18.11. More definitions. — Let (f_0, Γ) be invariant and let $x \in \mathcal{T}(\Gamma, \varepsilon)$. Let π_α denote the projection of $\mathcal{T}$, first, to $\mathcal{T}(Y)$ and then to $\mathcal{T}(A(\alpha))$. Then for any gap or loop α of Γ with one exception, we define

$$F_\alpha(x) = d_\alpha(\pi_\alpha(x), \pi_\alpha(\tau(x))),$$

and if α is periodic with orbit $[\alpha]$ (see 17.3)

$$F_{[\alpha]}(x) = \sum_{\alpha' \in [\alpha]} F_{\alpha'}(x).$$

The exception is the periodic loop γ (if it exists, in which case there is at most one) such that there is $\gamma' \in \Gamma$ such that γ, γ' are homotopic in $\overline{\mathbf{C}} \setminus Z$ but not in $\overline{\mathbf{C}} \setminus Y$ and a component of $f_0^{-1}(\gamma')$ is homotopic to the component of $f_0^{-1}(\gamma)$ in $[\gamma]$, up to Z-preserving isotopy. In that case we take $A(\gamma, Z) \subset Z$ and $\pi_{\gamma,Z}$ to be projection to $\mathcal{T}(A(\gamma, Z)$ and

$$F_\gamma(x) = d_{\gamma,Z}(\pi_{\gamma,Z}(x), \pi_{\gamma,Z}(\tau(x)))$$

and $F_{[\gamma]}(x)$ as above.

For α periodic or nonperiodic, write $[[\alpha]]$ for the components of $\cup_{i\geqslant 0} f_0^{-i}\alpha$. This set might contain loops if α is a gap. Then we define

$$F_{[[\alpha]]}(x) = \sum_{\beta \in [[\alpha]]} F_\beta(x).$$

We write $a(\alpha, x)$ for the q-area of $S(\alpha, x, \varepsilon_0)$, where $q(z)dz^2$ is the quadratic differential at x for $d(x, \tau(x))$. (See definition in general in 9.4.) If α is periodic with orbit α_i $(1 \leqslant i \leqslant k)$, we write

$$a([\alpha], x) = \sum_{i=1}^{k} a(\alpha_i, x).$$

For a loop γ, let $\kappa(\gamma)$, $\kappa_0'(\Gamma)$ be as in 17.2. Let (f_0, Γ) be discrete and $\kappa_0'(\Gamma) > 0$. If $B(\Delta_0', f_0, \Gamma)$ (2.18) is of periodic type (1.9) then the unique loop cycle $[\gamma]$ with

$\kappa(\gamma) = \kappa_0'(\Gamma)$ is in $\partial[\Delta_0']$. If $B(\Delta_0, f_0, \Gamma)$ is of eventually fixed type (1.9) then there is one loop of $\Gamma_2(f_0, \Gamma)$ (2.5) separating such a γ from $[\Delta_0']$.

Now let (f_0, Γ) be discrete. Then Δ_0' is adjacent to the fixed union Ω, and periodic. Then we define a set Q as follows. Note that, by its definition in 17.2, $\kappa(\Delta_0') > 0$ only if (f_0, Γ) is Euclidean. If $\kappa_0'(\Gamma) < \max(\kappa_0(\Gamma), \kappa(\Delta_0'))$ then take

$$Q = \overline{\Omega} \cup [\Delta_0'].$$

In this case we also write $\beta_0 = \Delta_0'$. If $\kappa_0'(\Gamma) \geqslant \max(\kappa_0(\Gamma), \kappa(\Delta_0')$ then choose $\gamma_1$with $\kappa(\gamma_1) = \kappa_0'(\Gamma)$ and γ_1 separating v_1 from Δ_0' and let β_0 be the gap adjacent to γ_1 separated from Δ_0' by γ_1. Write β_i for the gaps with $\beta_i \subset f_0^{-1}(\beta_{i+1})$ up to trivial and peripheral loops, $0 \leqslant i \leqslant n-2$, and $\beta_{n-1} \subset f_0^{-1}(\beta_0)$. Then let Q be the union of Ω, the β_j, and all gaps and loops separating Ω from some β_j, $0 \leqslant j \leqslant n-1$.

We define quantities $\omega_j(x,\Gamma)$ $(1 \leqslant j \leqslant 3)$, $\omega(x,\Gamma)$ if (f_0,Γ) is $(L_1, L_2, \varepsilon, \nu)$-good for x. We define this only if (f_0, Γ) is discrete and if $\max(\kappa_0(\Gamma), \kappa(\Delta_0'), \kappa_0'(\Gamma) > 0$ and either $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) \geqslant \kappa_0(\Gamma)$ or $m_{\gamma_0'}(x) \geqslant C_1$ or the orbit $[\alpha] \subset \Omega$ adjacent to $\partial\Omega$ satisfies

$$\kappa(\alpha) = \max(\kappa_0(\Gamma), \kappa(\Delta_0'), \kappa_0'(\Gamma)). \tag{1}$$

Define

$$\omega_1(x,\Gamma) = \frac{\omega_{1,1}(x,\Gamma)}{\omega_{1,2}(x,\Gamma)}$$

where

$$\omega_{1,1}(x,\Gamma) = C_1^{-1} e^{-m_{\partial,0}(x)} \text{ or } e^{-m_1(x)/C_1^5} \text{ or } C_1^{-1}$$

in the following cases. We use the first possibility if (1) holds and there is no $\overline{\beta} \subset \operatorname{int}(\Omega)$ such that (1) holds with β replacing α. We use the second possibility if (1) holds and there is such a β. We use the third possibility if (1) does not hold, $\beta_0 = \Delta_0'$ and we define

$$\omega_1(x,\Gamma) = \frac{e^{-m_{\partial,0}(x)}}{1 + m_{\gamma_0'}(x)}.$$

We define

$$\omega_{1,2}(x,\Gamma) = 1 + \sum m_\gamma(x),$$

where there are at most three loops in the sum. If $\beta_0 = \Delta_0'$ then the only term in the sum is $m_{\gamma_0'}(x)$. If $\beta_0 \neq \Delta_0'$ then we take the (at most 3) loops which are equal to γ_0', γ_1, or separate γ_0' and γ_1. If $\kappa(\alpha) < \kappa_0(\Gamma)$ (in which case $\max(\kappa_0'(\Gamma), \kappa(\Delta_0') \geqslant \kappa_0(\Gamma)$ or $m_{\gamma_0'}(x) \geqslant C_1$), then we replace the numerator in the above expression for $\omega_1(x,\Gamma)$ by $e^{-C_1^7/\nu}$.

Then we define

$$\omega_2(x,\Gamma) = e^{-m_{\mathrm{int}}(x)/C_1^4}\omega_1(x,\Gamma) \text{ or } e^{-C_1^3/\nu}\omega_1(x,\Gamma),$$

$$\omega_3(x,\Gamma) = e^{-m_{\mathrm{int}}(x)/C_1^{7/2}}\omega_1(x,\Gamma) \text{ or } e^{-C_1^{7/2}/\nu}\omega_1(x,\Gamma),$$

$$\omega(x,\Gamma) = e^{-m_{\mathrm{int}}(x)/C_1^3}\omega_1(x,\Gamma) \text{ or } e^{-C_1^4/\nu}\omega_1(x,\Gamma),$$

depending on whether or not (1) holds.

If (f_0, Γ) is minimal, then of course $m_{\text{int}}(x) = 0$.As usual, we take $m_{\gamma_0'}(x) = 0$ if γ_0' (17.3) does not exist. In many cases, of course, $m_{\gamma_0'}(x)$ will be bounded. In current applications, $m_{\gamma_0'}(x)$ and $m_{\gamma_1}(x)$ will be bounded in terms of $m_1(x)$. Therefore, the above expressions could be slightly simplified. But in Chapter 29 we will need to consider cases where $m_2(x)$ is not bounded in terms of $m_1(x)$. Therefore, we leave the expressions as they are.

Now we describe the *Triangle Conditions* on x, for x such that $\omega(x, \Gamma)$ is defined. Now let $q(z)dz^2$ be the quadratic differential at $x = [\varphi]$, for $d(x, \tau(x))$, with stretch $p(z)dz^2$ at $\tau(x)$. Let s be the holomorphic branched covering used to define $\tau(x)$. Assume $[\varphi]$ is normalised so that the components of $\partial\beta_j$ are bounded and bounded apart. Then the *Triangle Conditions* are as follows.

For each component D of $\overline{\mathbf{C}} \setminus \beta_j$ disjoint from Ω,

(1)
$$\sum_{y \in D \cap Y} |\operatorname{Res}(q - s_* p, \varphi(y))| \leqslant e^{-C_1/\nu} a(\beta_j, q), \quad 1 \leqslant j \leqslant n-1.$$

For any gap or loop $\beta \subset \overline{\mathbf{C}} \setminus Q$ or loop $\beta \subset \partial Q$,

(2)
$$a(\beta, q) \leqslant e^{-C_1/\nu} a(\beta_j, q), \quad 0 \leqslant j \leqslant n-1.$$

There is a $e^{-C_1/\nu}$-dominant area surface U' for $q(z)dz^2$ containing β_{n-1} such that

(3)
$$a(\beta_{n-1}, q)/a(U', q) \geqslant \omega_1(x, \Gamma).$$

18.12. The Third Version Properties of σ_1. — We now need to reduce the Second Version Properties to more manageable ones. Recall that $\sigma = \sigma_2 \circ \sigma_1$. The homeomorphisms σ_1, σ_2 are the identity except on the set (described in 17.8) where σ is not the identity. The homeomorphism σ_2 is actually the identity except on a much smaller set. We shall construct σ_2 explicitly in 21.13, spelling out its properties in 19.3.

Third Version Properties. — Write $x' = \tau(x)$ and $x'' = \sigma_1(x')$. The Second Version Basic Properties are as before.

In addition, whenever (f_0, Γ) is $(L_1, L_2, \varepsilon, \nu)$-adapted to x, and $F(x) \geqslant \delta_1$, either Second Version nondiscrete/Euclidean holds, or the following Third version F-between or F-small-decrease holds.

F-between
$$\kappa_0(\Gamma) + e^{-C_1^{3/2} m_1(\sigma_1(x))} \leqslant F(\sigma_1(x)) \leqslant \kappa_0(\Gamma) + e^{-m_1(\sigma_1(x))/C_1^{39/2}}.$$

F-small decrease. — If (f_0, Γ) is not minimal, then
$$m_{\text{int}}(\sigma_1(x)) \geqslant m_1(\sigma_1(x))/C_1.$$

If (f_0, Γ) is minimal the Triangle Conditions hold for $\sigma_1(x)$. In addition, in all cases, if the triangle conditions do not hold for both x and $\sigma_1(x)$ then
$$F(\sigma_1(x)) \leqslant F(x) - \omega_3(\sigma_1(x), \Gamma).$$

Definition of K_i. — The set K_i consists of $x \in \cup\{\widetilde{T}_\mu(\varepsilon_0) : \mu \text{ minimal }\}$ such that

$$m_2(x) \leqslant E_i', \quad |m_\gamma(x) - m_{\gamma'}(x)| \leqslant E_i' \text{ for } \gamma,\, \gamma' \subset \partial P,$$

where P is the fixed set of $\mu = (f_0, \Gamma)$, and the Triangle conditions hold for a constant C_i' replacing C_1. In particular,

$$\{\sigma_1(x) : x \in \widetilde{T}_\mu(\varepsilon_1),\, \mu \text{ minimal},\, m_2(x) \leqslant E_i'\} \subset K_i$$

for a suitable constant $E_i' > 0$. In chapter 19, we shall see that the Third Version Properties do imply the Second Version Properties, for a suitable construction of σ_2.

CHAPTER 19

PUSHING AND THE GOOD VECTOR FIELD

19.1. We need to show that the Third Version Properties of σ_1 in 18.12, together with a suitable construction of σ_2, imply the Second Version Properties of τ'. To do this, we show in 19.2 that the Triangle Conditions often force F-small-decrease.We then state the properties of σ_2 in 19.3. In 19.4 we see that these are sufficient to make the Third Version Properties of σ_1 imply the Second Version properties of τ. We shall complete the construction of σ_2 only in 21.13. In the rest of this chapter, we reduce the construction of τ' further, to the construction of a vector field (of which σ_1 is the time one map) satisfying the *Fourth Version Properties*. The following two chapters will then be devoted to the construction of this vector field.

19.2. Lemma. — *Given κ, there is a constant C_1 such that the following holds, whenever L_1 is sufficiently large. Let (f_0, Γ) be discrete non-Euclidean and $(L_1, L_2, \varepsilon, \nu)$-adapted to $x = [\varphi]$. Let $F(x) \leqslant \kappa$. Let $q(z)dz^2$ be the quadratic differential for $d(x, \tau(x))$ at x with stretch $p(z)dz^2$ at $\tau(x) = [\varphi']$. Let $q'(z)dz^2$ be the quadratic differential for $d(\tau(x), \tau^2(x))$ at $\tau(x)$ with stretch $p'(z)dz^2$ at $\tau^2(x)$. Let (1) to (3) of the Triangle conditions of 18.11 be satisfied for each of x, $\tau(x) = x'$ replacing x. Let U' be a $e^{-C_1/\nu}$-dominant area piece for $q'(z)dz^2$ containing $S(\beta_{n-1}, \tau(x), \varepsilon_0)$ and homotopic to $\varphi'(R')$ for $R' \subset Q$, Q as in 18.11. Then*

$$F(\tau(x)) \leqslant F(x) - \frac{a(\beta_{n-1}, \tau(x), q'))e^{-C_1/\nu}}{a(U', \tau(x), q')}. \tag{1}$$

Proof. — Let $S(\beta_0, \nu)$ be as in 9.3. Now we need the concept of an *m-approximate pole* or zero in $S(\beta_0, \nu)$. Let $w_0 \in \overline{\mathbf{C}}$. An *$m$-good annulus at w_0 for q* is an annulus $A \subset S(\beta_0, \nu)$ of modulus $\geqslant m$ round w_0 such that A contains no zeros of q nor points of $\varphi(Z)$, and $a(A, q) \geqslant m$. Let the disc containing w_0 and bounded by A have p poles and n zeros. Then w_0 as an *m-approximate pole of order $p - n$* — or, equivalently a zero of order $n - p$. The sum of orders of poles is always 4, for any quadratic differential.

Let s be the holomorphic branched covering used to define $\tau(x)$ from x. Depending on κ, we can find a constant K such that an m-good annulus for any one of q, p, s^*q gives one for one of the others. Since $[\varphi] \in (\mathcal{T}(Y_1))_{\geqslant \nu/2}$, if C_1 is sufficiently large, then there is some $m \geqslant e^{-C_1/2\nu}$ such that any m-good annulus at a point can be extended to a Km-good one, for all the quadratic differentials we are considering. Our aim is to show that, if (1) does not hold, then $q(z)dz^2$ on $S(\beta_0, x, \nu)$ and $s^*q(z)dz^2$ on $S(\beta_{n-1}, \tau(x), \nu)$ have the same number of approximate poles. This is impossible, since s is of degree 2 on $S(\beta_{n-1}, \tau(x), \nu)$. Therefore, (1) will hold.

By the Triangle properties of 18.11, as weakened slightly above, $q(z)dz^2$ and $s_*p(z)$ are close on $S(\beta_i, x, \nu)$ for each $1 \leqslant i \leqslant n-1$, as are $q'(z)dz^2$ and $s_*p'(z)dz^2$ on $S(\beta_i, \tau(), \nu)$, $1 \leqslant i \leqslant n-1$. In fact, the Triangle Condition gives more than this. There are two terms in s_*p (and similarly for s_*p'), and one of them (corresponding to $f_0^{-1}(\beta_i) \setminus \beta_{i-1}$) is negligible, by the Triangle Condition (2). Therefore, $q(z)dz^2$ has the same number of approximate poles on $S(\beta_i, x, \nu)$ as $p(z)dz^2$ on $S(\beta_{i-1}, \tau(x), \nu)$, and therefore the same number as $q(z)dz^2$ on $S(\beta_{i-1}, x, \nu)$, $1 \leqslant i \leqslant n-1$. Hence $q(z)dz^2$ has the same number of approximate poles on all $S(\beta_i, x, \nu)$, $0 \leqslant i \leqslant n-1$. A similar result holds for $q'(z)dz^2$ on all $S(\beta_i, \tau(x), \nu)$.

The next step is to show that $q(z)dz^2$ has the same number of approximate poles on $S(\beta_i, x, \nu)$ as $q'(z)dz^2$ on $S(\beta_j, \tau(x), \nu)$. If we can do this for one i and j, then by the above, we can do it for all. We use the fact that Δ_0' — and hence certainly β_0 — has period > 1. This is the hardest point of the proof, in that we have to use the Hard Same Shape Theorem. Let U_i' be of $e^{-2C_1/\nu}$-dominant area for $q'(z)dz^2$, and homotopic to $\varphi'(R_i')$, where $\beta_i \subset R_i'$. (Thus, $U' = U_{n-1}'$.) By the Triangle Condition (2) of 18.11, we can choose $R_i' \subset Q$. Of course, we have $R_i' = R_j'$ if $\beta_j \subset R_i'$. Write $R' = R_{n-1}'$, $U' = U_{n-1}'$. Choose $R_i \subset Q$ ($1 \leqslant i \leqslant n$, $R_0 = R_n$) with $R_i' \subset f_0^{-1}(R_{i+1})$. We can choose $A(R_i)$ with $v_2 \in A(R_0)$, $A(R_i') \subset f_0^{-1}(A(R_{i+1})$. Write $R_i'' = f_0^{-1}(R_{i+1})$. By simple properties of pullback,

$$d_{A(R_i')}(\tau(x), \tau^2(x)) \leqslant d_{A(R_{i+1})}(x, \tau(x)) \leqslant F(x).$$

Let $r_i(z)dz^2$, $r_i'(z)dz^2$ be the quadratic differentials for $d_{A(R_i)}(x, \tau(x))$ at x, $d_{A(R_i')}(\tau(x), \tau^2(x))$ at $\tau(x)$. The quadratic differential for $d_{A(R_i'')}(\tau(x), \tau^2(x))$ at $\tau(x)$ is $s^*r_{i+1}(z)dz^2$. By Easy Same Shape 9.5, $r_i'(z)dz^2$ has the same number of approximate poles in $S(\beta_i, \tau(x), \nu)$ as $q'(z)dz^2$. Trivially, for $0 \leqslant i \leqslant n-2$, $s^*r_{i+1}(z)dz^2$ has the same number of approximate poles in $S(\beta_i, \tau(x), \nu)$ as $r_{i+1}(z)dz^2$ in $S(\beta_i, x, \nu)$. By Hard Same Shape 16.2, $s^*r_{i+1}(z)dz^2$ has the same number of approximate poles in $S(\beta_i, \tau(x), \nu)$ as $r_i'(z)dz^2$, that is, as $q'(z)dz^2$. Also by Hard Same Shape, $r_i(z)dz^2$ has the same number of approximate poles in $S(\beta_i, x, \nu)$ as $q(z)dz^2$. For this, we use the Triangle Condition (2) of 18.11 for x, which implies that, although U_i itself may not be of dominant area for $q(z)dz^2$, there is at least a dominant area piece for $q(z)dz^2$ containing $S(\beta_j, x, \nu)$ and not intersecting the component of $\overline{\mathbf{C}} \setminus S(\beta_j, x, \nu)$ disjoint from $\varphi(Q)$. So, in summary, $q'(z)dz^2$ and $r_0'(z)dz^2$ in $S(\beta_0, \tau(x), \nu))$ and

$q(z)dz^2$, $r_1(z)dz^2$ in $S(\beta_1,x,\nu)$ have the same number of approximate simple poles, and hence the same is true for all $S(\beta_i,x,\nu)$, $r_i(z)dz^2$, $q(z)dz^2$ and $S(\beta_j,\tau(x),\nu)$, $r'_j(z)dz^2$, $q'(z)dz^2$.

Now let $[\varphi']=[\chi\circ\varphi]$ where χ minimizes distortion up to isotopy constant on $\varphi(A(R_0))$. The direction of maximum distortion of $s'^{-1}\circ\chi\circ s$ is the direction of the unstable foliation of $s^*r_0(z)dz^2$,. So applying 8.3 with $\chi_1=s'^{-1}\circ\chi\circ s$, if $\theta(s^*r_0,r'_{n-1})$ denotes the angle between the unstable foliations, for a suitable constant L_3,

$$d_{R'_{n-1}}(\tau(x),\tau^2(x))\leqslant d_{R_0}(x,\tau(x))-L_3\int_{S(\beta_{n-1},\tau(x),\nu)}\theta(s^*r_0,r'_{n-1})^2|r'_{n-1}|.$$

So if (1) does not hold we have

$$\begin{aligned}L_3\int_{S(\beta_{n-1},\tau(x),\nu)}\theta(s^*r_0,r'_{n-1})^2|r'_{n-1}| &\leqslant \frac{a(\beta_{n-1},\tau(x),q'))e^{-C_1/\nu}}{a(U',\tau(x),q')}\\ &\leqslant C_2e^{-C_1/\nu}a(\beta_{n-1},\tau(x),r'_{n-1})\end{aligned}$$

with the last inequality given by Same Shape 9.5, for a suitable C_2 independent of C_1, if C_1 is sufficiently large. If C_1 is sufficiently large, this implies that r'_{n-1} and $s^*r_0(z)dz^2$ have the same number of approximate simple poles in $S(\beta_{n-1},\tau(x),\nu)$, and hence so do $q'(z)dz^2$, and $s^*q(z)dz^2$. Then, as we have seen, $q(z)dz^2$ in $S(\beta_0,x,\nu)$, and $s^*q(z)dz^2$ in $S(\beta_{n-1},\tau(x),\nu)$, have the same numbers of approximate poles, which is impossible. □

19.3. The construction of σ_2. — This lemma gives the conditions that σ_2 will satisfy, and ensures that m_1-increasing and all the basic properties hold for $\tau'(x)$ if Triangle or F-small-decrease holds for each of $\sigma_1(x)$, $\sigma_1\circ\tau\circ\sigma_1(x)$. Take β_j $(0\leqslant j\leqslant n-1)$ and $\omega(x)=\omega(x,\Gamma)$ as in 18.11.

Lemma. — *Let (f_0,Γ) be discrete invariant, $(L_1,L_2,\varepsilon,\nu)$-adapted to x, and let Ω be reducible. Write $x''=\sigma_1(x')=\sigma_1\circ\tau\circ\sigma_1(x)$. A map σ_2 can be constructed such that $\sigma_2(x'',x)=x''$ except on a set where (f_0,Γ) is not minimal, $\omega(x'',\Gamma)$ is defined (18.11),*

$$F(x'')-F(x)\leqslant-\omega(x'',\Gamma)/2.$$

All the Third Version properties Basic properties hold for $\sigma_2(x'')$ replacing $\sigma_1(x)$, x'' replacing x except that the basic property $F(\sigma_1(x))\leqslant F(x)$ is replaced by:

$$F(\sigma_2(x'',x))<F(x'')+\omega(x,\Gamma)/2.$$

In addition either Second Version nondiscrete-Euclidean holds or Third Version F-between holds with $\sigma_2(x'',x)$ replacing $\sigma_1(x)$ or the following holds if $F(x'')\leqslant F(x)-\omega(x'',\Gamma)$. Take any loop $\gamma\in\Gamma$, $\gamma\subset\operatorname{int}(\Omega)$, with $m_\gamma(x'')+m_{[\gamma],\partial}(x'')\geqslant m_{\rm int}(x'')/C_1^2$. Then

$$m_\gamma(\sigma_2(x'',x))\geqslant m_\gamma(x'')+2C_1^2e^{-m_{\gamma,\partial}(x'')}.$$

If $\kappa_0'(\Gamma) > \kappa_0(\Gamma)$, $\gamma \subset \partial\Omega$ and γ_0' exists, then

$$m_{[\gamma]}(\sigma_2(x'', x)) \geqslant m_{[\gamma]}(x'') + C_1 m_{\gamma_0'}(x'') + 1.$$

This will be proved in 21.13.

19.4. Corollary. — *Assuming that σ_2 can be constructed as in 19.3, the Third Version Properties of σ_1 in 18.12 imply the Second Version properties of τ' in 18.5.*

Proof. — The basic properties are the same. The map σ_2 has been chosen so that the Basic Properties will hold. So we need to show that if $F(x) \geqslant \delta_1$ then one of Second Version F-decrease Nondiscrete-or-Euclidean, F-between, m_1-increasing, or minimal, holds, and we need to ensure that $F(\tau'(x)) < F(x)$ when σ_2 is not the identity. If Third Version F-between holds for $\sigma_1(x)$ or $\sigma_1 \circ \tau \circ \sigma_1(x)$, then Second Version F-between holds for $\tau'(x)$. So suppose that $\sigma_1(x)$ and $\sigma_1 \circ \tau \circ \sigma_1(x)$ satisfy Third Version F-small-decrease. Then by the formulation of F-small decrease, and the definition of $\omega(x, \Gamma)$, we have

$$(1) \qquad \begin{aligned} F(\sigma_1 \circ \tau \circ \sigma_1(x)) &\leqslant F(x) - \min(\omega_3(\sigma_1(x), \Gamma), \omega_3(\sigma_1 \circ \tau \circ \sigma_1(x), \Gamma)) \\ &\leqslant F(x) - \omega(\sigma_1 \circ \tau \circ \sigma_1(x), \Gamma) \end{aligned}$$

unless both $\sigma_1(x)$ and $\tau \circ \sigma_1(x)$ satisfy the Triangle Conditions of 18.11. But then by 19.2 and the definition of the Triangle Conditions, we can obtain (1) with $F(x)$, $F(\sigma_1 \circ \tau \circ \sigma_1(x))$ replaced by $F(\sigma_1(x))$, $F(\tau \circ \sigma_1(x))$. So then by 9.3, the Second Version Properties will indeed hold for $\tau'(x)$.

Finally, we need to check for Second Version minimal. The Third Version Triangle Properties (and the definition of $\omega(\tau'(x))$ give the required negative upper bound on $F(\tau'(x)) - F(x)$, since, as noted in 18.10, $m_{\gamma_0'}(x)$ is bounded. ☐

19.5. The Product Structure of $\mathcal{T}(\Gamma, \varepsilon)$ and a Product Norm. — We now describe a decomposition of the tangent space over $\mathcal{T}(\Gamma, \varepsilon)$. Let Σ denote the set of loops and gaps of Γ. Recall (from 9.1) that $\mathcal{T}(\Gamma, \nu)$ is naturally isomorphic to a subspace of $\prod_{\alpha \in \Sigma} \mathcal{T}(A(\alpha))$, and that the tangent space of $\mathcal{T}(A(\alpha))$ is naturally isomorphic to $\mathbf{C}^{A'(\alpha)}$, where $A'(\alpha)$ is $A(\alpha)$ minus three points x_α, y_α, z_α, and, if α is a loop, then y_α and z_α are in the same component of $\overline{\mathbf{C}} \setminus \alpha$.

Then we identify $[\varphi] \in \mathcal{T}(\Gamma, \varepsilon)$ with

$$([\varphi_\alpha]) \in \prod_{\alpha \in \Sigma} \mathcal{T}(A(\alpha)).$$

We normalize so that $\varphi_\alpha(x_\alpha) = 0$, $\varphi_\alpha(y_\alpha) = 1$, $\varphi_\alpha(z_\alpha) = \infty$.

We define a norm $\|\cdot\|$ on the tangent space of $\mathcal{T}(A(\alpha))$ for all α. If α is a gap, we have (from the above) a local identification of $\mathcal{T}(A(\alpha))$ with $(\mathbf{C} \setminus \{0,1\})^{A'(\alpha)}$, and thus of the tangent space at each point with $\mathbf{C}^{A'(\alpha)}$. If α is a gap, we simply use the standard Euclidean norm on $\mathbf{C}^{A'(\alpha)}$. If α is a loop then we identify $\mathcal{T}(A(\alpha)$ with the upper half plane as described in 9.1, and use the *Euclidean* (not Poincaré) norm on the tangent space. Thus an $O(1)$ change in modulus can be achieved by a path of length $O(1)$.

19.6. Preliminaries to the Fourth Version Properties in terms of the Vector Field w. — It is now time to express properties of σ_1 entirely in terms of the vector field w. The map σ_1 will be the time 1 map of w. In order to describe w, we need the notion of a *gap* independently of a loop set. A *gap* at x is simply a subsurface of $\overline{\mathbf{C}} \setminus Y$ such that $x \in \mathcal{T}(\partial\alpha, \varepsilon_0)$. We shall be interested only in *invariant* gaps, that is, $\partial\alpha \subset \Gamma$ for some invariant (f_0, Γ). Thus, α is a gap of Γ, and probably of other invariant loop sets too. We can then talk of α being periodic, nonperiodic, equal to Δ_0', in the fixed union (for some loop set) and so on. Similarly, we can talk of *invariant* loops, and talk of a loop being in the fixed union (for some invariant loop set) and so on.

Let (f_0, Γ) be discrete and invariant and $x \in \mathcal{T}(\Gamma, \varepsilon)$ for some $\varepsilon \leqslant \varepsilon_0$. Let Σ be the set of loops and gaps of Γ. Let Γ' be the largest subset of Γ such that (f_0, Γ') satisfies the Invariance and Levy Conditions. We shall use the notation of 17.5-6, 18.5.

This is reminiscent of the definitions of 17.5. If $\gamma \in \Gamma$, $\gamma \subset \text{int}(\Omega)$ is a loop adjacent to α with $\kappa(\alpha)$ maximal, then we define H_γ to be a function $g(m_{[\gamma]}(x))$ where g is C^1 with $g(t) = -1/\log t$ if if $t \leqslant m_1(x)/C_1$, $g(t) = -1/\log t_0$ is constant if $t \geqslant t_0 = 2m_1(x)/C_1$, and g is an increasing function for $t < t_0$.

The set of *decreasable functions* for (f_0, Γ) is:

$$F, \quad F_{[[\alpha]]}\ (\alpha \in \Sigma), \quad \frac{-1}{\log(2 + m_{[\gamma]}(x))}\ (\gamma \in \Gamma \setminus \Gamma'), \quad H_\gamma\ (\gamma \subset \partial\alpha,\ \overline{\alpha} \subset \text{int}(\Omega)).$$

Here, we include γ_0' in $\Gamma \setminus \Gamma'$ if it exists (see 17.3). Note that every decreasable function is bounded below and above. The number of decreasable functions at x, even if (f_0, Γ) is allowed to vary, is bounded in terms of $\#(Y)$.

The functions $n_i(\alpha, .)$. — If α is a gap or loop in the full orbit of $\beta \cup \partial\beta$ for some $\beta \subset \Omega(f_0, \Gamma)$ for some invariant (f_0, Γ) where Ω is the fixed union, and $\kappa(\beta) = \max(\kappa_0(\Gamma), \kappa_0'(\Gamma), \kappa(\Delta_0'))$, define

$$n_1(\alpha, x) = e^{-m_{\partial\alpha}(x)/C_1^3} = n_2(x, \alpha), \quad n_3(\alpha, x) = e^{-m_{\partial\alpha}(x)/C_1^6}.$$

Let $\alpha = \beta_j$ for some $0 \leqslant j \leqslant n-1$ for β_j as in 18.11, and let Q, A_j be as in 18.11. Let ν be the length of the shortest loop in $\cup_j \text{int}(\beta_j)$. Let $q(z)dz^2$ be the quadratic differential for $d(x, \tau(x))$ at x, with stretch $p(z)dz^2$ at $\tau(x)$. Let s be the holomorphic branched covering used to define $\tau(x)$. Then

$$n_1(\alpha, x) = e^{-C_1^7/\nu} \max\{|\textstyle\sum_{y \in D} \text{Res}(q - s_* p, \varphi(y))| : \\ D \text{ is a component of } \overline{\mathbf{C}} \setminus \cup_j \beta_j \text{ disjoint from } Q\}.$$

We define $n_2(\alpha, x)$ similarly but with C_1^7 replaced by C_1^8.

For all other gaps, define $n_1(\alpha, x) = C_1^{-2}$ and $n_2(\alpha, x) = 1$. For all loops γ, define

$$n_1(\gamma, x) = m_\gamma(x)^{-1}(\log(m_\gamma(x)))^{-2}, \quad n_2(\gamma, x) = 1$$

19.7. Fourth Version Properties in terms of the Vector Field w

The vector field w is the sum of vector fields w_Γ where (f_0, Γ) is discrete and $\Gamma_2(f_0, \Gamma) \neq \varnothing$. We have $w_\Gamma(x) = 0$ except when (f_0, Γ) is (L_1'', L_2'')-adapted to x. We take Γ' to be the largest invariant subset of Γ satisfying the Levy Condition.

Write $w_\Gamma = w_\Gamma(x)$. Let Σ be the set of loops and gaps of Γ for any Γ as above. Let Γ' be the loop set generated by $\Gamma_2(f_0, \Gamma)$.

Basic Properties. — The most important Basic Bound is

(1) $$DF(w_\Gamma) \leqslant 0.$$

(2) $$Dm_\gamma(w_\Gamma) \leqslant C_1^{-1} \quad \text{if } \gamma \notin \Gamma'.$$

(3) Let γ have length $L_1^{-1/2}\varepsilon_0$ and $\gamma \subset \operatorname{int}(\Omega(f_0, \Gamma))$, or $\gamma \subset \overline{\Omega(f_0, \Gamma)}$ if $\max(\kappa_0'(\Gamma), \kappa(\Delta_0')) \geqslant \kappa_0(\Gamma)$. (This may include some loops not in Γ.) Then one of the following holds.

F-between $\kappa(\beta)$. — For some gap β with $\beta \subset \Omega'$ for Ω' the fixed union of some loop set, and $\gamma \subset \operatorname{int}(\beta)$,

(3*a*) $$F(x) \leqslant \kappa(\beta) + e^{-m_{[\gamma]}(x)/C_1}.$$

$m_{[\gamma]}$ non-decreasing

(3b) $$Dm_{[\gamma]}(w_\Gamma) \geqslant 0.$$

(4) $$\sum_{\alpha\in\Sigma} n_2(\alpha, x)a(\alpha, x)\|D\pi_\alpha(w_\Gamma)\| \leqslant -C_1 DF(w_\Gamma) \leqslant C_1^2 \sum_{\alpha\in\Sigma} a(\alpha, x)\|D\pi_\alpha(w_\Gamma)\|.$$

The lefthand inequality can be strengthened if β_j (as in 18.11) is in a $e^{-C_1/\nu}$-dominant area piece U' homotopic to $\varphi(R)$ ($x = [\varphi]$) with $a(\beta_j, x) \geqslant \omega_1(x, \Gamma)a(U', x)$. In this case

(5) $$\sum_{\alpha\subset R} \|D\pi_\alpha(w_\Gamma)\| a(\alpha, x)/a(U', x) n_2(\alpha, x) \leqslant -C_1 DF(w_\Gamma).$$

If H is a decreasable function,

(6*a*) $$DH(w_\Gamma) \leqslant C_1^{-1}.$$

If $H = F_{[\alpha]}$

(6*b*) $$DH(w_\Gamma) \leqslant e^{-m_{\partial[\alpha]}(x)/C_1^2}.$$

In addition the following hold when (f_0, Γ) is (L_1', L_2')-adapted to x.

Decreasable Conditions. — Let $F(x) \geqslant \delta_1$. Let one of the following *Decreasable Conditions* (7)-(13) hold. Then (15) holds. If none of (7)-(13) holds but (14) holds then (16) holds, for $\omega_2(x,\Gamma)$ as in 18.11.

(7) For some nonperiodic loop β outside Q, or loop β in Ω, $F_\beta(x) \geqslant \delta_0$.

(8) For some gap β outside Q or periodic loop outside Q (hence in a non -Levy cycle) $F_\beta(x) \geqslant F(x) - \delta_0$.

(9) For some periodic gap β with $\kappa(\beta) < \max(\kappa_0(\Gamma), \kappa_0'(\Gamma), \kappa(\Delta_0'))$, $F_\beta(x) \geqslant F(x) - \delta_0$.

(10) For some (possibly reducible) periodic gap $\alpha \subset \Omega$ with

$$\kappa(\alpha) = \max(\kappa_0(\Gamma), \kappa_0'(\Gamma), \kappa(\Delta_0')), \quad F_\alpha(x) \geqslant \kappa(\alpha) + n_3(\alpha, x).$$

(11) For some periodic gap orbit $[\alpha] \subset Q$, and some α, $\alpha' \in [\alpha]$, $a(\alpha', x)/a(\alpha, x) \notin [1 - C_1^{-1}, 1 + C_1^{-1}]$.

(12) The loop γ_0' (17.3) is such that a component of $f_0^{-1}(\gamma_0')$ is homotopic in $\overline{\mathbf{C}} \setminus Z$ to the component of $f_0^{-1}(\gamma_0)$ in Ω and

$$F_{[\gamma_0]}(x) \geqslant \frac{1}{p} \log \frac{m_{\gamma_0}(x) + m_{\gamma_0'}(x)}{m_{\gamma_0}(x) + \frac{1}{2} m_{\gamma_0'}(x)} + \delta_0$$

where γ_0 is of period p.

(13) We have $a(\Omega, x) \geqslant C_1^{-1}$, and there is a gap β with $\overline{\beta} \subset \operatorname{int}(\Omega)$ with $\kappa(\beta) = \max(\kappa_0(\Gamma), \kappa_0'(\Gamma), \kappa(\Delta_0'))$ but no such gap in Ω adjacent to $\partial\Omega$, and $m_{\partial[\beta]}(x) \leqslant m_1(x)/C_1$ for all such β.

(14) There is a subsurface U of $e^{-C_1^3/\nu}$ dominant area containing β_{n-1}, with $a(\beta_{n-1}, x) \geqslant \omega_1(x,\Gamma) a(U, x)$ (as in 18.11) and $a(U, x) \geqslant C_1^{-1}$, but the Triangle conditions (1) and (2) of 18.11 are not satisfied, with $e^{-C_1/\nu}$ replaced by $e^{-C_1^3/\nu}$ (a stronger condition).

Decreasable consequences

$$DH(w_\Gamma) \leqslant -C_1. \tag{15}$$

$$DF(w_\Gamma) \leqslant -\omega_2(x, \Gamma). \tag{16}$$

19.8 Lemma. — *For suitable choice of C_1, the time 1 map σ_1 of w satisfies the Third Version Properties.*

Proof. — Write σ_t for the flow of w. We have $\sigma_1(x) = x$ on the set where $w = 0$. So the set where $\sigma_1(x) \neq x$ is as required by the Basic Properties. We need to show that if the Fourth Version properties hold for $w(x)$ and $x = \tau(x_1)$, then the Third Version Properties hold for x_1 and $\sigma_1 \circ \tau(x_1) = \sigma_1(x)$. Consistent numbering (1) to (3) is chosen in 19.10 with the properties of 18.5. The connections are clear in each case.

Now we need to check the other properties of 18.12 (which refer back to the properties of 18.5). So let (f_0, Γ) be invariant discrete non-Euclidean and (L_1, L_2)-adapted to x. Let Γ' be the largest subset satisfying the Invariance and Levy Conditions. Then the Basic Bounds imply that for $\gamma \in \Gamma'$, $0 \leqslant t \leqslant 1$, $m_\gamma(\sigma_t(x)) \geqslant C_1^{-2} m_\gamma(x)$.

$\sigma_t(x) \in \mathcal{T}(\partial\Delta_0'(f_0,\Gamma'), C_1^3/L_1)$ for all $t \in [0,1]$. Then assuming that $m_1(x)$ is sufficiently large, if (f_0,Γ) is (L_1,L_2)-adapted to x and L_2 is sufficiently large, there is a loop set $\Gamma_1 \subset \Gamma$ which is (L_1,L_2)-adapted to $\sigma_t(x)$ for all $t \in [0,1]$, and all loops of $\Gamma \setminus \Gamma_1$ have length $\leqslant L_2\varepsilon_0$ at $\sigma_t(x)$ for all $t \in [0,1]$. It suffices to work with Γ_1.

Assuming C_1 is large enough, there is $t < 1$ such that no decreasable function decreases for $\sigma_t(x)$. Then we look at dominant area pieces (9.4) for the quadratic differential $q_t(z)dz^2$ for $d(\sigma_t(x), \tau\circ\sigma_t(x))$. If none of the decreasable conditions hold, then there can be no dominant area pieces outside Q. Hence $a(Q,q_t)$ is bounded from 0. Since none of the decreasable conditions hold, for $\beta \subset \Omega$, $a(\beta,x)$ can only be bounded from 0 if β is a gap with $\kappa(\beta) = \max(\kappa_0(\Gamma), \kappa(\Delta_0'), \kappa_0'(\Gamma))$ maximal. If there is such an β but no $\alpha \subset \Omega$ adjacent to $\partial\Omega$ with $\kappa(\alpha) = \kappa(\beta)$ then since neither (13) nor (10) holds, we have

$$F(x) \leqslant \kappa(\beta) + e^{-m_1(x)/C_1^2}$$

and hence by 17.5 we have F-between. We can also assume there is no Ω_1 with $\overline{\Omega}_1 \subset \Omega$, $a(\Omega_1,x) \geqslant C_1^{-1}$ and $m_{\partial\Omega_1}(x) \geqslant m_1(x)/C_1^5$. For if Ω_1 exists then by (11) it is a component of $f_0^{-1}\Omega_1$ up to isotopy, by (9) $\kappa(\Omega_1)$ is maximal and by (10) and 17.5 we again get F-between. So area must be concentrated in $Q \setminus \Omega$ or near $\partial\Omega$. In fact U must exist as in (14), and the strengthened Triangle Conditions of 18.11 must hold for $\sigma_t(x)$, that is, with $e^{-C_1^7/\nu}$ replacing $e^{-C_1/\nu}$. Then we want to show that

$$F(\sigma_1(x)) \leqslant F(x) - \omega_3(x,\Gamma).$$

Write $U = U(t)$. Then $U(t)$ is homotopic to $\varphi_t(R)$ for some $R \subset \Omega$. Then $\sigma_s(x)$ will satisfy (1)-(3) of the Triangle Conditions of 18.11 if there is $U(s)$ homotopic to $\varphi_s(R)$ which is $e^{-C_1^2/\nu}$-dominant for $d(\sigma_s(x), \tau\circ\sigma_s(x))$ at $\sigma_s(x)$ and all loops in $\varphi_s(\beta_j)$ have length $\geqslant \nu/C_1$, and (1) and (2) of the Triangle Condition hold for $\sigma_s(x)$ with $e^{-C_1^2/\nu}$ replacing $e^{-C_1/\nu}$. Then we need to show that if these conditions do not hold for all $s \in [0,t]$ then for some $s < t$,

$$F(\sigma_t(x)) \leqslant F(\sigma_s(x)) - \omega_3(x,\Gamma) \tag{1}$$

because $F(x) \geqslant F(\sigma_s(x))$ and $F(\sigma_1(x)) \leqslant F(\sigma_t(x))$. Now

$$F(\sigma_t(x)) \leqslant F_{U(t)}(\sigma_t(x)) + e^{-C_1^2/\nu} a(\beta_{n-1},q_t)/a(U(t),q_t) \tag{2}$$

by (easy) Same Shape 9.5. Of course $F(\sigma_s(x)) \geqslant F_{U(t)}(\sigma_s(x))$. By the bound on $a(\beta_{n-1},q_t)/a(U(t),q_t)$ in (3) of the Triangle Conditions in 18.11, we shall certainly obtain the result unless

$$F(\sigma_s(x)) \leqslant F_{U(t)}(\sigma_s(x)) + 2e^{-C_1^2/\nu} a(\beta_{n-1},q_t)/a(U(t),q_t). \tag{3}$$

Now by 9.8, the quadratic differentials $q_{U(t)}$ for $d_{U(t)}(\sigma_t(x), \tau\circ\sigma_t(x))$ and $q_{U(t),s}$ $d_{U(t)}(\sigma_s(x), \tau\circ\sigma_s(x))$ are the same shape so long as all quantities

$$a(\beta, q_{U(t)}) d_\beta(\sigma_s(x), \sigma_t(x)) = o(a(\beta_{n-1}, q_{U(t)}) \tag{4}$$

that is as long as this is true, $a(\beta, q_{U(t)})/a(\beta, q_{U(t),s})$ are boundedly proportional to 1. In particular, $a(\beta_{n-1}, q_{U(t),s})$ is boundedly proportional to $a(\beta_{n-1}, q_{U(t)})$. So now as long as this is true we can assume that $U(t) = U(s)$ is of dominant area for q_s. So take the smallest possible $s \leqslant t$ so that (4) holds, that is, for small δ,

$$\int_s^t \|D\pi_\beta(w(\sigma_u(x))\|a(\beta, q_{U(t)})du = \delta a(\beta_{n-1}, q_{U(t)}).$$

We have w is a sum of vector field such as w_{Γ_1}, for each of which the Basic bound (5) of 19.7 holds, but these loop sets differ by loops of length $\geqslant \varepsilon_0/L_2$, which are of very small modulus compared with $m_1(\sigma_t(x)) \geqslant \varepsilon_1^{-1/2}$ (assuming ε_1 is sufficiently small given ε_0). So we have

$$-DF(w(\sigma_u(x)) \geqslant \sum_{\beta \subset R} a(\beta, q_{U(t)})e^{-m_{\partial\beta}(x)/C_1^3}.$$

Then

$$\int_s^t -DF(w(\sigma_u(x))e^{m_{\text{int}}(x)/C_1^4}du \geqslant \delta a(\beta_{n-1}, q_{U(t)}).$$

with the term $e^{m_{\text{int}}(x)/C_1^4}$ replaced by $e^{L_2 C_1^6/\varepsilon_0}$ if $\kappa_0(\Gamma) < \max(\kappa(\Delta_0'), \kappa_0'(\Gamma))$. Since $a(\beta_{n-1}, q_{U(t)}) \geqslant \omega_1(\sigma_u(x), \Gamma)$, this gives (1) as required. □

CHAPTER 20

CONSTRUCTION OF THE GOOD VECTOR FIELD: PART 1

20.1. The Good Vector Field Theorem. — *Let (f_0, Γ) be invariant discrete non-Euclidean with $\Gamma_2(f_0, \Gamma) \neq \varnothing$. Then there exists a vector field w satisfying the properties of 19.7.*

We devote this chapter and the following one to proving this proposition.

20.2. Idea of the Proof and some notation. — Let (f_0, Γ) be fixed, as above. As usual, we let Σ denote the set of loops and gaps α of Γ. We use the natural product structure of $\mathcal{T}(\Gamma, \varepsilon_0)$ as described in 9.1, which involves choosing sets $A(\alpha) \subset Y$ for each $\alpha \in \Sigma$. If α is a loop then $A(\alpha)$ contains exactly four points. If α is a gap then $A(\alpha)$ contains at least 3 points. We fix a decomposition

$$A(\alpha) = \{x_\alpha, y_\alpha, z_\alpha\} \cup A'(\alpha).$$

If α is a loop then y_α, z_α are in the same component of $\overline{\mathbf{C}} \setminus \alpha$. We use the norm on the tangent space over $\mathcal{T}(\Gamma, \varepsilon_0)$ described in 19.5. Now we describe, for each $\alpha \in \Sigma$, a subspace $V(\alpha)$ of the tangent space of $\mathcal{T}(Y)$, which can also be regarded as a space of vector fields over $\mathcal{T}(\Gamma, \varepsilon_0)$. Given α, normalise all $[\varphi] \in \mathcal{T}(Y)$ so that $\varphi(x_\alpha) = 0$, $\varphi(y_\alpha) = 1$, $\varphi(z_\alpha) = \infty$. First, let α be a gap. Then let $V(\alpha)$ denote the set of all $\theta = (\theta(y)) \in \mathbf{C}^Y$ such that $\theta(y) = \theta([\varphi])(y)$ is constant for y in each component of $\overline{\mathbf{C}} \setminus \alpha$, and $\theta(y) = 0$ for y in the same component of $\overline{\mathbf{C}} \setminus \alpha$ as a point of $A(\alpha) \setminus A'(\alpha)$. Now let α be a loop. Then we take $\theta = (\theta(y))$ where $\theta(y) = 0$ if y is in the same component of $\overline{\mathbf{C}} \setminus \alpha$ as z_α (and y_α) or if $y = x_\alpha$, and $\theta(y) = \lambda \varphi_\alpha(y)$ if $y \neq x_\alpha$ (that is, $\varphi_\alpha(y) \neq 0$) but y is in the same component of $\overline{\mathbf{C}} \setminus \alpha$ as x_α, for some fixed λ.

We also define $V([\alpha])$ to be the direct sum of $V(\alpha')$ ($\alpha' \in [\alpha]$) if α is periodic, and $V([[\alpha]])$ to be the direct sum of $V(\alpha')$ for $\alpha' \in [[\alpha]]$.

Then we have the following. Recall (from 9.1 onwards) that if β is a loop,

$$m_\beta([\varphi]) = \mathrm{Im}(\pi_\beta([\varphi])).$$

20.3. Lemma. — *Given $\eta > 0$ there is a constant M depending on ε such that the following holds. Let $\theta \in V(\alpha)$. Then:*

$$\|D\pi_\beta(\theta([\varphi]))\| \leqslant M\|D\pi_\alpha(\theta([\varphi]))\|e^{-\Delta},$$

where

$$\Delta = \sum_\gamma m_\gamma([\varphi]),$$

where the sum is over loops γ separating α and β if α is a gap, and includes α if α is a loop but does not include β if β is a loop.

Proof. — Fix $[\varphi]$. First suppose that α is a gap. We can assume without loss of generality that $\#(A(\alpha)) \geqslant 4$. Let $[\psi_t \circ \varphi]$ be the image of φ under the time t flow of θ for small t, where we normalise so that ψ_t fixes $0, 1, \infty$. Thus, for y in the same component of $\overline{\mathbf{C}} \smallsetminus \alpha$ as a point w of $A'(\alpha)$, there is a constant c_w such that

$$\psi_t(\varphi(y)) = \varphi(y) + c_w t.$$

We can assume without loss of generality that $c_w \neq 0$ for a single $w \in A'(\alpha)$. If β is a gap, ψ_t translates at most one point of $A(\beta)$, or all but at most one. If β is a loop, ψ_t translates none, two or four of the points of $A(\beta)$. Let ρ_t be the Möbius transformation such that

$$\rho_t \circ \psi_t \circ \varphi(x_\beta) = 0, \quad \rho_t \circ \psi_t \circ \varphi(y_\beta) = 1, \quad \rho_t \circ \psi_t \circ \varphi(z_\beta) = \infty.$$

Then $\rho_t \circ \psi_t$ moves points of $\varphi_\beta(A'(\beta))$ by Euclidean distance at most $tM|c_w|e^{-\Delta}$. If β is a loop, then the Euclidean distance moved by the single point $\rho_t \circ \psi_t(A'(\beta))$ is at most $tM'|c_w|e^{-\Delta-\Delta'}$, where $\Delta' = m_\beta([\varphi])$. This means that, regarding $\pi_\beta([\psi_t \circ \varphi])$ as an element of the upper half-plane in the usual way (9.1) $\pi_\beta([\psi_t \circ \varphi]) = \pi_\beta([\varphi]) + O(tM'e^{-\Delta})$. These give the required bounds on $\|\pi_\beta(\theta)\|$ is α is a gap.

Now suppose that α is a loop. Again, normalise so that $\varphi = \varphi_\alpha$ on x_α, y_α, z_α, in particular, $\varphi(x_\alpha) = 0$. Then ψ_t is of the form

$$\psi_t(\varphi(y)) = e^{ct}\varphi(y)$$

for y in the same component of $\overline{\mathbf{C}} \smallsetminus \alpha$ as x_α, $\psi_t(\varphi(y) = \varphi(y)$ otherwise. Then choose ρ_t as above. Then, once again, $\rho_t \circ \psi_t$ moves points of $\varphi_\beta(A'(\beta))$ by a distance of at most $M|c|te^{-\Delta}$ or $M|c|te^{-\Delta-\Delta'}$, depending on whether β is a gap or a loop, and we obtain the same estimates as above. □

20.4. The Formula for $DF(\theta)$ if $\theta \in V(\alpha)$. — We use the following notation in the next few sections. Fix $x = [\varphi] \in \mathcal{T}$, $\tau(x) = [\psi]$ and let $q(z)dz^2$ be the quadratic differential for $F(x) = d(x, \tau(x))$ at x, with stretch $p(z)dz^2$ at $\tau(x)$. Let s be the holomorphic branched covering used to define $\tau(x)$. We recall (8.11) that

$$(DF)_x(\underline{h}) = 2\pi \sum_{y \in Y} \mathrm{Re}(\mathrm{Res}(q - s_* p, \varphi(y))h(y)).$$

This, of course, underlies our calculations in this chapter. Given α, let ρ_α be the Möbius transformation such that

$$\rho_\alpha(\varphi(x_\alpha)) = 0, \quad \rho_\alpha(\varphi(y_\alpha)) = 1, \quad \rho_\alpha(\varphi(z_\alpha)) = \infty.$$

Write

$$q^\alpha = (\rho_\alpha)_* q.$$

We define $(s_* p)^\alpha$ similarly. As in Chapter 9, we write $q_\alpha(z)dz^2$ for the quadratic differential for $d_\alpha(\pi_\alpha([\varphi]), \pi_\alpha([\psi]))$ at $\pi_\alpha([\varphi])$.

Lemma

(1) *Let α be a gap. For each $y \in A'(\alpha)$ let $\gamma(y)$ denote the component of $\partial\alpha$ separating y from α. Then for $\theta \in V(\alpha)$,*

$$DF(\theta([\varphi])) = \operatorname{Im}\Big(\sum_{y\in A'(\alpha)} \int_{\varphi(\gamma(y))} \theta(y)(q^\alpha(z) - (s_*p)^\alpha(z))dz\Big).$$

(2) *Let α be a loop. Then if $\theta \in V(\alpha)$ and λ is the scalar used in the definition of θ,*

$$DF(\theta) = \operatorname{Im}\Big(\lambda \int_{\varphi(\alpha)} z(q^\alpha(z) - (s_*p)^\alpha(z))dz\Big).$$

Proof

(1) Let α be a gap. Define $D(y)$ to be the set of $y' \in Y$ in the same component of $\overline{\mathbf{C}} \setminus \alpha$ as y. Then for $\theta \in V(\alpha)$ the Derivative Formula 8.11 gives

$$\begin{aligned} DF(\theta) &= 2\pi \operatorname{Re}\Big(\sum_{y\in A'(\alpha)} \sum_{y'\in D(y)} \theta(y) \operatorname{Res}(q^\alpha - (s_*p)^\alpha, \varphi(y'))\Big) \\ &= \operatorname{Re}\Big(\sum_{y\in A'(\alpha)} \theta(y) \int_{\varphi(\gamma(y))} -i(q^\alpha(z) - (s_*p)^\alpha(z))dz\Big), \end{aligned}$$

which, by definition, we write as

$$2\pi \operatorname{Re}\Big(\sum_{y\in A'(\alpha)} \theta(y)(c(\alpha, q, y) - c(\alpha, s_*p, y))\Big) = 2\pi \operatorname{Re}\Big(\sum_{y\in A'(\alpha)} \theta(y)c(\alpha, y)\Big).$$

(2) Now let α be a loop. Let D denote the set of $y \in Y$ in the same component of $\overline{\mathbf{C}} \setminus \alpha$ as ∞. Then the First Derivative Formula gives

$$\begin{aligned} DF(\theta) &= 2\pi \operatorname{Re}\Big(\lambda \sum_{y\in D} \varphi(y)(\operatorname{Res}(q^\alpha - (s_*p)^\alpha, \varphi(y))\Big) \\ &= \operatorname{Re}\Big(\lambda \int_{\varphi(\alpha)} -iz(q^\alpha(z) - (s_*p)^\alpha(z))dz\Big), \end{aligned}$$

which, by definition, we write as

$$2\pi \operatorname{Re}(\lambda(c(\alpha, q) - c(\alpha, s_*p))) = 2\pi \operatorname{Re}(\lambda c(\alpha)). \qquad \square$$

In future, we shall write $c(\alpha, q)$ for the vector $(c(\alpha, q, y))$, for any quadratic differential $q(z)dz^2$.

20.5. A Formula for $DF_\zeta(\theta)$ if $\theta \in V(\alpha)$. — Let $\zeta \in \Sigma$, $x \in \mathcal{T}(\Gamma, \varepsilon_0)$. Let s be the holomorphic branched covering used to define $\tau(x)$. We recall that

$$F_\zeta(x) = d_\zeta(x \cdot \tau(x)) = d(\pi_\zeta(x), \pi_\zeta \circ \tau(x)).$$

Let ρ_α be as in 20.4. If $\zeta \subset f_0^{-1}(\alpha)$ then we define $s_{\alpha,\zeta} = \rho_\alpha \circ s \circ \rho_\zeta^{-1}$. We have the following.

Lemma. — *Let $\theta \in V(\alpha)$. Let $x = [\varphi] \in \mathcal{T}(\Gamma, \varepsilon_0) \cap \mathcal{T}(\partial\alpha, \varepsilon) \cap \mathcal{T}(\partial\zeta, \varepsilon)$, $\varphi_\alpha = \rho_\alpha \circ \varphi$, $F(x) \leqslant \kappa$. Then depending on whether of not $\alpha \subset f_0^{-1}(\alpha)$,*

$$(1) \qquad DF_\alpha(\theta) = 2\pi \operatorname{Re}\Big(\sum_{y \in A'(\alpha)} \theta(y) \operatorname{Res}(q_\alpha - (s_{\alpha,\alpha})_* p_\alpha, \varphi_\alpha(y))\Big) + O(e^{-2\pi^2/\varepsilon}),$$

$$(2) \qquad DF_\alpha(\theta) = 2\pi \operatorname{Re}\Big(\sum_{y \in A'(\alpha)} \theta(y) \operatorname{Res}(q_\alpha, \varphi_\alpha(y))\Big) + O(e^{-2\pi^2/\varepsilon}).$$

If $\zeta \subset f_0^{-1}(\alpha)$, $\alpha \neq \zeta$, then

$$(3) \qquad DF_\zeta(\theta) = -2\pi \operatorname{Re}\Big(\sum_{y \in A'(\alpha)} \theta(y) \operatorname{Res}((s_{\alpha,\zeta})_* p_\zeta, \varphi_\alpha(y))\Big) + O(e^{-2\pi^2/\varepsilon}).$$

If $\zeta \neq \alpha$, $\zeta \not\subset f_0^{-1}(\alpha)$,

$$(4) \qquad DF_\zeta(\theta) = O(e^{-2\pi^2/\varepsilon}).$$

Proof. — The formulae given are all invariant under affine change of coordinates, that is, if we replace $A'(\alpha)$ by a different choice $A''(\alpha) \subset A(\alpha) \smallsetminus \{z_\alpha\}$ and replace φ_α by $\sigma \circ \varphi_\alpha$ for $\sigma(z) = \lambda z + \mu$ for $\sigma \circ \varphi_\alpha(A(\alpha) \smallsetminus A''(\alpha)) = \{0, 1, \infty\}$ then the formulae above hold for q_α, $A'(\alpha)$ if and only if they hold for $\sigma_* q_\alpha$, $A''(\alpha)$ (with the appropriate changes to $s_{\alpha,\alpha}$ and so on.) So we may assume that $\theta(y) = 0$ for $y = v_1, v_2$, the critical values of f_0. Now let $\alpha \subset f_0^{-1}(\beta)$. Then

$$(5) \qquad D(\pi_\alpha \circ s^{-1})(\theta)(y) = (s_{\beta,\alpha}^{-1})'(\varphi_\beta(f_0(y))) D\pi_\beta(\theta(f_0(y))).$$

If $\beta \neq \alpha$ then we have $D\pi_\beta(\theta(y')) = O(e^{-2\pi^2/\varepsilon})$ by 20.4. If $\beta = \alpha$, $f_0(y) \in D(y')$ for $y' \in A(\alpha)$ and $D(y')$ as in 20.4, then

$$(s_{\beta,\alpha}^{-1})'(\varphi_\beta(f_0(y))) D\pi_\beta(\theta(f_0(y))) = (s_{\alpha,\alpha}^{-1})'(\varphi_\alpha(y'))\theta(y') + O(e^{-2\pi^2/\varepsilon}).$$

Then by the First Derivative Formula

$$DF_\alpha(\theta) = 2\pi \sum_{y \in A(\alpha) \smallsetminus \{z_\alpha\}} \operatorname{Re}(\operatorname{Res}(q_\alpha, \varphi_\alpha(y))\theta(y) - \operatorname{Res}(p_\alpha, \varphi_\alpha(y)) D(\pi_\alpha \circ s^{-1})(\theta)(y)).$$

Then (1) and (2) follow, depending on whether $\alpha = \beta$ or $\alpha \neq \beta$. The calculation for $D(\pi_\zeta \circ s^{-1})(\theta)$ when $\zeta \subset f_0^{-1}(\alpha)$ is exactly like (5) above, with ζ replacing α and α

replacing β. The calculation of $D(\pi_\zeta \circ s^{-1})(\theta)$ when $\zeta \subset f_0^{-1}(\beta)$ for $\beta \neq \alpha$ is exactly like (5) above with ζ replacing α.

20.6. The Orbit Dominant Area Condition. — We now state the *Orbit Dominant Area Condition.* This is a condition on gaps and loops. As the name suggests, it has a resemblance to the Dominant Area Condition of 9.4, but, roughly speaking, dominance has to extend over the backward orbit. For technical reasons, our orbit dominant area condition for loops is actually weaker, in some respects, than our original dominant area condition for loops.

We fix a discrete (f_0, Γ) with $x \in \mathcal{T}(\Gamma, \varepsilon_0)$. We use $F_{[\alpha]}$, $F_{[[\alpha]]}$ of 18.11, m_α (from 9.1 onwards) and $a(\alpha, q)$ as in 9.4. Let $\Gamma' \subset \Gamma$ be the largest subset such that (f_0, Γ') satisfies the Invariance and Levy Conditions. Let Ω be the fixed union.

We say that α satisfies the *Orbit Dominant Area Condition for* $q(z)dz^2$ and a constant D_1 if the following hold. First, let α be a gap or a union of gaps and loops containing at least one gap. Then the conditions are simply that

$$a(\beta) \leqslant D_1 a(\gamma)$$

whenever β is a nonperiodic gap or loop in $\cup_{n>0} f_0^{-n}(\alpha)$, or $\beta \subset \cup_{n\geqslant 0} f_0^{-n}(\partial\alpha)$, and $\gamma \subset \mathrm{int}(\alpha)$ is any bounded loop. Now let α let any loop in Γ which is not a periodic loop outside Ω, and and suppose that α is not in the periodic orbit of $\partial\Delta_0$. Then the conditions are that

$$F_\alpha(x) \geqslant \frac{1}{D_1 m_\alpha(x)}, \tag{1}$$

$$a(\beta) \leqslant D_1 a(\alpha) \tag{2}$$

whenever β is nonperiodic in $\cup_{n>0} f_0^{-n}\alpha$ or $\beta \in [\alpha]$ (for α periodic) and

$$F_\beta(x) \leqslant \frac{1}{\sqrt{D_1} m_\beta(x)}.$$

Note that these are automatically satisfied if, for a suitable $D_2 > 0$,

$$a(\beta) \leqslant e^{D_2 m_\alpha(x)} a(\alpha)$$

whenever β is a nonperiodic loop in $\cup_{n>0} f_0^{-n}\alpha$ or a gap adjacent to $\cup_{n\geqslant 0} f_0^{-n}\alpha$.

Now let α be in the periodic orbit of $\partial\Delta_0 = \gamma_0$. Let the period be n, with orbit $\{\gamma_i : 0 \leqslant i \leqslant n-1\}$ and a component of $f_0^{-1}(\gamma_0)$ homotopic in $\overline{\mathbf{C}} \smallsetminus Z$ to γ_{n-1}. Let $d(x, \tau(x)) = \frac{1}{2}\log K$. The conditions are as above unless a loop γ_0' exists such that $x \in \mathcal{T}(\gamma_0', \varepsilon_0)$, γ_0' is isotopic to γ_0 in $\overline{\mathbf{C}} \smallsetminus Z$ but not in $\overline{\mathbf{C}} \smallsetminus Y$. In the above, we then take d_{γ_0} to be a semimetric on $\mathcal{T}(Z)$, that is, we are only interested in the homotopy class of γ_0 in $\overline{\mathbf{C}} \smallsetminus Z$. So $d_{\gamma_0} = d_{\gamma_0'}$. The conditions for Orbit Dominant Area are exactly as above unless a component of $f_0^{-1}(\gamma_0')$ is homotopic to γ_{n-1}. Then the conditions are that, for $\Delta = D_1^{-1} + \frac{1}{2} m_{\gamma_0'}(x)$,

$$F_{\gamma_{n-1}}(x) \geqslant \log \frac{m_{\gamma_{n-1}}(\tau(x)) + (\sqrt{K}-1)\Delta}{m_{\gamma_{n-1}}(\tau(x))}$$

and (2) above, with the same conditions as before.

Now let $\alpha \in \Gamma$ be a periodic loop outside Ω. Then the conditions are much stronger. The loop α is (D_1, D_1')-orbit dominant if there is an annulus homotopic to $\varphi(\alpha)$ $(x = [\varphi])$ of modulus $\geqslant (1 - D_1')m_\alpha(x)$, satisfying the Pole Zero Condition (9.4), $a(\beta) \leqslant D_1^{-1/2} a(\alpha)$ for all $\beta \in [\alpha]$, and $a(\beta) \leqslant D_1 a(\alpha)$ whenever $\beta \in [\alpha] \setminus [[\alpha]]$ or the maximal annulus homotopic to $\varphi(\beta)$ satisfying the Pole Zero Condition has modulus $\leqslant (1 - D_1'/D_1) m_\beta(x)$.

20.7. In what follows, $n_1(\alpha, x)$ is as in 19.6. We use the notation of 20.6. Let $m_{\gamma,\alpha}(x)$ be the sum of $m_\zeta(x)$ with ζ separating α from $[\gamma]$ $(\zeta \in \Gamma)$ and including α if α is a loop. Note that

$$e^{-m_{\gamma,\alpha}(x)}(n_1(\alpha, x))^{-1} \leqslant (n_1(\gamma, x)))^{-1}.$$

In fact, if $\gamma \not\subset \partial\alpha$, $\gamma \neq \alpha$,

$$e^{-m_{\gamma,\alpha}(x)}(n_1(\alpha, x))^{-1} \leqslant n_1(\gamma, x) = o(1).$$

Proposition. — *Let $D_1 > 0$, $D_1' > 0$ be given. Let $\varepsilon > 0$ be sufficiently small given these and $x \in \mathcal{T}(\Gamma, \varepsilon)$. Then there exists a gap or loop α of D_1-orbit dominant area for x, and of (D_1, D_1')-orbit dominant area if α is a periodic loop outside the fixed union Ω. Moreover, for any loop $\gamma \subset \mathrm{Int}(\Omega \cup f_0^{-1}(\Omega))$, with $\gamma \subset f_0^{-1}(\gamma')$,*

(1) $$a(\gamma) = o(\min(e^{m_{\gamma,\alpha}}, e^{m_{\gamma',\alpha}}) a(\alpha) n_1(\alpha, x)).$$

If the only possible such α are periodic gaps with $\kappa(\alpha)$ maximal, then α is of D_1''-orbit dominant area for $D_1'' = e^{-m_{\partial\alpha}(x)/C_1}$. If in addition there is no irreducible gap $\beta \subset \Omega$ adjacent to $\partial\Omega$ with $\kappa(\beta)$ maximal then we can find α reducible with $D_1'' = e^{-m_{\alpha,\partial}(x)/C_1}$. If the only possibility for such α is in a U containing some β_j with U of $e^{-C_1/\nu}$-dominant area and satisfying Triangle Condition (3) of 18.11, then for at least one such U we have $a(U, x) \geqslant 1/\#(Y)$.

Now let α be of D_1-orbit dominant area, and of (D_1, D_1')-orbit dominant area if α is a periodic loop outside Ω. Suppose that $[\alpha]$ does not intersect some U with $\beta_{n-1} \subset U$ for some U satisfying Triangle Condition (3) of 18.11. If α is periodic irreducible with $\kappa(\alpha)$ maximal, suppose that $F(x) > \kappa(\alpha) + D_1^{1/20\#(Y)}$. Then there exists $\theta \in V[[\alpha]] \oplus \bigoplus_{\gamma \in \Gamma} V(\gamma)$ and a decreasable function H such that $DH(\theta) \leqslant -C_1^{-1}\|\theta\|$, and θ satisfies all the basic bounds (1)-(6) of 19.7. If α is contained in U as above, and Triangle Conditions (1)-(2) of 18.11 are not satisfied, and this is the only possibility for α, then there exists $\theta \in V([[\beta_0]]) \oplus \oplus_{\gamma \in \Gamma} V(\gamma)$ satisfying all the Basic Bounds and $DF(\theta) \leqslant -e^{-C_1^5/\nu} C_1^{-1} \|\theta\| a(\beta_{n-1})$.

20.8. Proof of the Good Vector Field Proposition. — To find w_Γ with the properties of 19.7, we then take the sum of $\theta \in V[[\alpha]] \oplus \oplus_{\gamma\in\Gamma} V(\gamma)$ corresponding to α of D_1-dominant area or (D_1, D_1')-dominant area, and of $\theta \in V([[\beta_0]]) \oplus \oplus_{\gamma\in\Gamma} V(\gamma)$ if U exists as in 20.7 with $a(U,x) \geqslant C_1^{-1}$ but only including θ corresponding to for α as above if $a(U,x) \geqslant C_1^{-1}$. The existence of α as in the first paragraph is proved in 21.12, and therefore the sum will always be nonempty. The existence of θ is proved in 21.2-11.

The main part of the proof of the Orbit Dominant Area Proposition involves comparing DF and $DF_{[[\alpha]]}$, which we proceed to do, in a number of lemmas. From now on in this chapter, let $x = [\varphi] \in \mathcal{T}(\Gamma, \varepsilon)$ and let $q(z)dz^2$ be the quadratic differential at x for $d(x, \tau(x)) = F(x)$, with stretch $p(z)dz^2$ at $\tau(x)$.

20.9. Lemma. — *Let α be a gap or loop. Then*

$$c(\alpha, q) = O(a(\alpha, q)), \quad c(\alpha, s_*p) = O(a(f_0^{-1}\alpha, p)).$$

Proof. — The second estimate follows from the first, since $a(\alpha, s_*p) = O(a(f_0^{-1}\alpha, p))$ by change of variable. For gaps, the result follows immediately from the definition of $c(q,\alpha)$, because $a(q,\alpha) = a(q^\alpha, \alpha)$ (change of coordinates does not change area), and since $S(\alpha, [\varphi^\alpha], \varepsilon_0)$ is normalised to have diameter bounded and bounded from 0, the area is proportional to the maximum of q^α on $S(\alpha, [\varphi^\alpha], \varepsilon_0)$. Now let α be a loop. Normalise so that $S(\alpha, \varepsilon_0)$ contains, and is homotopic to, $\{z : 1 \leqslant |z| \leqslant R\}$ for a large R. Normalise also so that $\{z : 1 \leqslant |z| \leqslant 2\}$ has area boundedly proportional to $a(\alpha, q)$. This is possible, by the definition of $a(\alpha, q)$. (See 9.4.). Then for a constant $C \leqslant C_0$, some C_0 depending only on $\#(Y)$, we can write

$$q(z) = A\frac{\prod_{i=1}^k (z - c_i)\prod_{i=1}^\ell (1 - \lambda_i z)}{\prod_{i=1}^m (z - b_i)\prod_{i=1}^n (1 - \mu_i z)}$$

for bounded A, $|c_i| \leqslant C$, $|b_i| \leqslant \frac{1}{2}$ and $|\lambda_i| \leqslant C^{-4\#(Y)}$, $|\mu_i| \leqslant R^{-1}$. Of course, k, ℓ, m, n are all $\leqslant \#(Y)$. Then $q(z) = Az^{k-m}(1 + o(1))$ for $C^{2\#(Y)} \leqslant |z| \leqslant 2C^{2\#(Y)}$. So we can reduce area by moving to ann annulus around $\{z : |z| = C^{2\#(Y)}\}$ if $m - k \geqslant 3$. So $m - k \leqslant 2$. Then $a(q, \alpha)$ is boundedly proportional to A. To compute $c(\alpha, q)$ we need the coefficient of z^{-2} in the Laurent expansion of q near $\{z : |z| = 1\}$. Since the c_i are all bounded, this is $O(A)$, as required. □

20.10. Lemma. — *Let α be a gap, or connected union containing at least one gap, satisfying the Orbit Dominant Area Conditions for D_1. Assume without loss of generality that $D_1 \geqslant e^{-2\pi^2/\varepsilon}$. Then, using the notation of 20.4,*

$$\|c(\alpha, q) - a(\alpha, q)c(\alpha, q_\alpha)\| = O(\sqrt{D_1}a(\alpha)), \tag{1}$$

and, if α is periodic with periodic preimage β,

$$\|c(\alpha, s_*p) - a(\beta, q)c(\alpha, (s_{\alpha,\beta})_*p_\beta)\| = O(\sqrt{D_1}a(\alpha)). \tag{2}$$

Proof. — Recall that $q(z)dz^2$ is a quadratic differential at $x \in \mathcal{T}(\partial\alpha, \varepsilon_0)$. As before, we write $x = [\varphi]$, normalising with respect to α as in 20.2. Formula (1) comes directly from the last part of the Same Shape Lemma 9.5, and the integral interpretations of the vectors $c(\alpha, q)$ in 20.4. For (2), we need to estimate $(s^{-1})'(z)/(s_{\alpha,\beta}^{-1})'(z)$ for z on a contour for one of the integrals of 20.4. Fix such a contour. We can normalise so that the contour has diameter bounded and bounded from 0. Because $x \in \mathcal{T}(\partial\alpha, \varepsilon)$, in this normalisation we can assume that each critical value of s is distance $O(e^{-2\pi^2/\varepsilon})$ from a critical value of $s_{\alpha,\beta}$. It then follows that, on the contours,

$$\left|\frac{(s^{-1})'(z)}{(s_{\alpha,\beta}^{-1})'(z)} - 1\right| = O(e^{-2\pi^2/\varepsilon}).$$

We remark that $a(\beta, p) = a(\beta, q) + O(\sqrt{D_1}a(\alpha))$ by the Orbit Dominant Area Condition. The result follows. □

20.11. Lemma

Let $\log\sqrt{K} = d(x, \tau(x))$. *Let* α *be a loop with* $F_\alpha(x) \geqslant 1/(D_1 m_\alpha(x))$. *Then there is a good boundary annulus (9.4) A homotopic to* $\varphi(\alpha)$ *of modulus* $\Delta \geqslant (D_1(\sqrt{K}-1))^{-1}$ *such that in each component of* $\overline{\mathbf{C}} \setminus A$,

$$\#(\mathrm{poles}(q)) - \#(\mathrm{zeros}(q)) = 2,$$

and

$$2\pi i \int_{\varphi(\alpha)} zq^\alpha(z)dz = (1 + O(e^{-\Delta/2}))\Big(\int_{\varphi(\alpha)} \sqrt{q^\alpha}dz\Big)^2.$$

Proof. — Suppose for contradiction that Δ is the largest possible modulus of such an annulus, and that $\Delta \leqslant (D_1(\sqrt{K}-1))^{-1}$. Let $[\varphi_\alpha]$, $[\psi_\alpha]$ be the projections of x, $\tau(x)$ to $\mathcal{T}(A(\alpha))$. As in 14.3, let $|\varphi_\alpha(\gamma)|'$ denote length of the geodesic homotopic to $\varphi_\alpha(\gamma)$ in $\overline{\mathbf{C}} \setminus \varphi_\alpha(A(\alpha))$ with respect to the *Euclidean* metric arising from the identification of $\overline{\mathbf{C}} \setminus \varphi_\alpha(A(\alpha))$ with $(\mathbf{C}/\Gamma)/(z \sim -z)$ for $\Gamma \leqslant \mathbf{C}$ a lattice such that $\mathbf{C}/\Gamma$ has area 2. Define $|\psi_\alpha(\gamma)|'$ similarly. Then, because $\mathcal{T}(A(\alpha))$ identifies with the Teichmüller space of the torus,

$$d_\alpha([\varphi], [\theta]) = \sup\{|\log|\psi_\alpha(\gamma)|' - \log|\varphi_\alpha(\gamma)|'| : \gamma \text{ nontrivial nonperipheral}\}.$$

So

$$\begin{aligned} F_\alpha(x) &\leqslant \tfrac{1}{2}\log\frac{m_\alpha(x) - \Delta + \sqrt{K}\Delta + O(1)}{m_\alpha(x) - \Delta + (1/\sqrt{K})\Delta + O(1)} \\ &\leqslant \frac{\frac{1}{2}\Delta(\sqrt{K} + \sqrt{K^{-1}} - 2) + O(1)}{m_\alpha(x)} \leqslant \frac{(\sqrt{K}-1)\Delta}{m_\alpha(x)}. \end{aligned}$$

Our normalization is chosen so that, for some $r \leqslant 1$,

$$q^\alpha(z) = \frac{\mu}{z^2}\prod_{j=1}^n(1 - \zeta_j z)\prod_{j=1}^n\Big(1 - \frac{\xi_j}{z}\Big)^{-1},$$

where $|\zeta_j| \geqslant r$ for all j and $\xi_j = O(re^{-\Delta})$. Then $\varphi(\alpha)$ can be taken as the circle $\{z :| z |= re^{-\Delta/2}\}$. Then on this circle we have

$$q^\alpha(z) = \frac{\mu}{z^2}(1 + O(e^{-\Delta/2})).$$

So

$$\int_{\varphi(\alpha)} zq^\alpha(z)dz = 2\pi i\mu(1 + O(e^{-\Delta/2})), \quad \int_{\varphi(\alpha)} \sqrt{q^\alpha}dz = 2\pi i\sqrt{\mu}(1 + O(e^{-\Delta/2})),$$

as required. □

20.12. Reduced Map Spaces again. — Let α be a periodic gap of period n. In 20.10, we have seen that if α has orbit dominant area, then the vectors $c(\alpha, q)$ and $a(\alpha)c(\alpha, q_\alpha)$ and $c(\alpha, s_*p)$, $a(\alpha)c(\alpha,(s_{\alpha,\beta})_*p_\beta)$ are approximately equal. We need to work further with these.

Write $\alpha = \alpha_0$. Inductively, let α_{i+1} be the gap with $\alpha_i \subset f_0^{-1}(\alpha_{i+1})$. Thus, $\alpha_n = \alpha_0$. If v_2 is in this cycle of gaps, it is convenient to choose α so that $v_2 \in \alpha$. We have previously (since 2.15) called this gap E_2. Corresponding to this cycle of gaps, f_0 induces a space B_α of branched coverings of a disjoint union of n copies of $\overline{\mathbf{C}}$ with marked points, with the spheres cyclically permuted. (This is a slight generalization to a union of marked spheres of the definition of Chapter 1.) We can fix a branched covering $f_\alpha \in B_\alpha$ which leaves $A(\alpha)$ invariant. Recall that $\mathcal{T}(Y)$ has a projection to $\mathcal{T}(A(\alpha))$. Then we define

$$\mathcal{T}(B_\alpha) = \prod_{j=0}^{n-1} \mathcal{T}(A(\alpha_j)).$$

Then we can project $\mathcal{T}$ to $\mathcal{T}(B_\alpha)$ in the natural way. Of course, if α is a homeomorphic gap, then $\mathcal{T}(B_\alpha)$ is isomorphic to $\prod_j \mathcal{T}(A(\alpha_j))$. Now we can define a pullback

$$\tau_\alpha : \mathcal{T}(B_\alpha) \longrightarrow \mathcal{T}(B_\alpha)$$

as in Chapter 6. If α is homeomorphic then

$$\tau_\alpha(([\varphi_j])) = ([\varphi_j \circ \chi_\alpha]),$$

where χ_α is the isotopy class of homeomorphism of n disjoint copies of $\overline{\mathbf{C}}$ induced by f_0. If α is nonhomeomorphic then put $s_\alpha = s_{\alpha_{j+1},\alpha_j}$ on the j'th sphere. Then by abuse of notation

$$\tau_\alpha(([\varphi_j])) = ([s_\alpha^{-1} \circ \varphi_j \circ f_\alpha]).$$

20.13. Lemma. — *Let α be a periodic gap satisfying the Orbit Dominant Area Condition with D_1. Assume without loss of generality that $e^{-2\pi^2/\varepsilon} = o(D_1)$. Let α_j $(0 \leqslant j \leqslant n-1)$ denote the periodic orbit with $\alpha_0 = \alpha_n$. As usual, write $x = [\varphi]$. Write $[\varphi_j]$, q_j, p_j, s_j for $[\varphi_{\alpha_j}]$, q_{α_j}, p_{α_j}, $s_{\alpha_{j+1},\alpha_j}$. Let $r_j(z)dz^2$ and $t_j(z)dz^2$ denote* *the quadratic differentials at* $[\varphi_j]$, $\pi_{\alpha_j}(\tau_\alpha(([\varphi_j])))$ *for* $d(([\varphi_j]),\tau_\alpha(([\varphi_j])))$. *Then for every periodic* α_j *in the orbit of* α,

$$\|c(\alpha_j,r_j)-c(\alpha_j,q_j)\|=O(\sqrt{D_1}),$$
$$\|c(\alpha_j,(s_\alpha)_*t_{j-1})-c(\alpha_j,(s_j)_*p_{j-1})\|=O(\sqrt{D_1}).$$

Consequently,

$$\|(c(\alpha_j,q)-c(\alpha_j,s_*p))-(a(\alpha_j)c(\alpha_j,r_j)-a(\alpha_{j-1})c(\alpha_j,(s_\alpha)_*t_{j-1}))\|=O(\sqrt{D_1}a([\alpha])).$$

Proof. — By 20.10, the third inequality follows from the first two. Also, as in 20.10, for all j we have

$$d_{\alpha_j}(\pi_{\alpha_j}(\tau_\alpha(([\varphi_k])),\pi_{\alpha_j}(\tau(([\varphi_k]))))=O(\exp(-2\pi^2/\varepsilon)).$$

Then by 9.8 we have

$$\|c(\alpha_j,r_j)-c(\alpha_j,q_j)\|=O(\exp(-\pi^2/\varepsilon)),$$
$$\|c(\alpha_j,s_{\alpha *}t_{j-1})-c(\alpha_j,(s_j)_*p_{j-1})\|=O(\exp(-\pi^2/\varepsilon)).$$

The result follows. □

20.14. Critical Points of $G_{[\alpha]}$. — We define $G_{[\alpha]}$ on $\mathcal{T}(B_{[\alpha]})$ by

$$G_{[\alpha]}(x)=\frac{1}{n}\sum_{j=0}^{n-1}d_{\alpha_j}([\varphi_j],\pi_{\alpha_j}(\tau_\alpha(([\varphi_k])))).$$

Proposition

(1) *If* α *is homeomorphic then* $G_{[\alpha]}$ *has no critical points unless* χ_α *is irreducible. In this case, the critical points are all minima, on which* $G_{[\alpha]}$ *takes the value* $\kappa(\alpha)$, *and comprise a closed geodesic in the quotient of* $\mathcal{T}(B_\alpha)$ *by the modular group. If* $[\chi_\alpha]$ *is reducible, then* $G_{[\alpha]}$ *takes bounded values only in*

$$\mathcal{T}(B_\alpha)_{\geqslant\delta}\cup\bigcup_{\gamma\in\Gamma}\mathcal{T}(\gamma,\delta)$$

where Γ *is a unique maximal* χ_α*-invariant set of disjoint simple loops.*

In both cases, if $\kappa(\alpha)>0$, *there is a constant* C *and an integer* k *depending only on* $\#(Y)$ *such that if* $\|(DG_{[\alpha]})_x\|\leqslant\delta$ *for small* δ *then*

$$|G_{[\alpha]}(x)-\kappa(\alpha)|\leqslant C\delta^{1/2+1/2k}.$$

Now let α *be nonhomeomorphic.*

(2) *Let* f_α *be critically finite equivalent to a rational map or of polynomial type. Then* $G_{[\alpha]}$ *has no critical points on the set where* $G_{[\alpha]}\neq 0$. *Given* $\delta>0$ *there are constants* M_1 *and* M_2 *such that* $G_{[\alpha]}$ *is bounded above by* M_1 *only on*

$$\mathcal{T}_{\geqslant\delta}\cup\bigcup_{\Gamma}\mathcal{T}(\Gamma,M_2\delta)$$

where the union is over loop sets Γ *such that* (f_0, Γ) *satisfies the Invariance Condition but has no subset satisfying the Levy Condition.*

(3) *If* f_α *is critically finite degree two irreducible not equivalent to a rational map, then critical points are all minima on which* $G_{[\alpha]}$ *takes the value* $\kappa(\alpha)$*, and comprise a closed geodesic in the quotient of* $\mathcal{T}(B_\alpha)$ *by the modular group. There is a constant* C *such that if* $\|(DG_{[\alpha]})_x\| \leqslant \delta$ *for small* δ *then* $|G_{[\alpha]}(x) - \kappa(\alpha)| \leqslant C\delta^2$.

(4) *Let* $\alpha = \Delta_0'$. *and let* $([\varphi_j]) \in \prod \mathcal{T}(A(\Delta_j'))$. *Then* $([\varphi_j])$ *is a critical point for* $G_{[\alpha]}$ *restricted to the set where* $[\varphi_0]$ *is constant if and only if all* $[\varphi_j]$ *and* $[\tau_\alpha([\varphi_0])]$ *lie on the same geodesic with* $[\varphi_j]$ *between* $[\varphi_{j-1}]$ *and* $[\varphi_{j+1}]$ *if* $j < n$*, and with* $[\varphi_{n-1}]$ *between* $[\varphi_{n-2}]$ *and* $\tau_\alpha([\varphi_0])$. *There is an integer* k *depending only on* $\#(Y)$ *such that if* $\|(DG_{[\alpha]})_x\| \leqslant \delta$ *for small* δ *then*

$$|F_{[\Delta_0']}(x) - \kappa_1(\Gamma, x)| \leqslant C\delta^{1/k}.$$

Proof. — At a critical point we have

$$r_j = s_{\alpha *} t_{j-1}$$

for all j.

(1) Let α be homeomorphic. We deduce that all r_j, t_j lie on the same geodesic, and that r_j is a stretch of itself. This is impossible if $[\chi_\alpha]$ preserves some simple disjoint loop set, that is, if $[\chi_\alpha]$ is reducible. Therefore, $[\chi_\alpha]$ is irreducible. Then the stable and unstable measured foliations of r_j are preserved by the pseudo-Anosov $[\chi_\alpha]$, up to multiple, and hence are unique up to multiple [**F-L-P**]. Hence the geodesic determined by the r_j is unique, and there is a unique closed geodesic of minima.

Now suppose that, for small δ,

$$\|(DG_{[\alpha]})_x\| \leqslant \delta.$$

We may as well assume that $s_\alpha(z, j) = (z, j+1)$ (since s_α permutes spheres and is Möbius in each coordinate). Then the residues of r_j and t_{j-1} agree within $O(\delta)$. Write

$$G_{[\alpha]}(x) = d(x, x \cdot [\chi_\alpha]) = \sum_j d([\varphi_j], [\varphi_j \circ \chi_j]) = \sum_j \log \lambda_j.$$

Here, $t_j(z)dz^2$ is the stretch of $r_j(z)dz^2$ by λ_j, that is, unstable foliation leaves of $r_j(z)dz^2$ are stretched by λ_j, and stable foliation leaves are contracted by λ_j^{-1}. Now, following the method of 16.12, for closed paths γ in $\overline{\mathbf{C}} \setminus Y$ we need to compare *stable* and *unstable lengths* as defined in 14.8. We write $|\varphi_j(\gamma)|_{j,-}$ and $|\varphi_j(\gamma)|_{j,+}$ for the stable and unstable lengths with respect to the quadratic differential $r_j(z)dz^2$. Let $k = \max(2, \#(Y) - 4)$. Then any quadratic differential on $\overline{\mathbf{C}} \setminus \varphi(Y)$ has zeros of multiplicity at most k. (Of course, there are no multiple zeros at all unless $\#(Y) \geqslant 6$.) Then we claim that

$$(1) \qquad |\varphi_j(\gamma)|_{j,+} = \lambda_{j-1}|\varphi_{j-1}(\gamma)|_{j-1,+} + O(\delta^{1/2+1/2k}), \quad 1 \leqslant j \leqslant n-1,$$

$$(2) \qquad |\varphi_0 \circ \chi(\gamma)|_{0,+} = \lambda|\varphi_0(\gamma)|_{0,+} + O(\delta^{1/2+1/2k}),$$

where $\chi = \chi_0 \circ \cdots \circ \chi_{n-1}$ and $\log \lambda = \sum_j \log \lambda_j$. Similar equations hold with $|\cdot|_{j,+}$ replaced by $|\cdot|_{j,-}$ and λ_j replaced by λ_j^{-1}. This estimate is clear when r_j is not close to having multiple zeros, or zero residues at points of $\varphi_j(Y_j)$, when we obtain simply $O(\delta)$. An application of Rouché's Theorem

$$\frac{1}{2\pi i}\int_\alpha \frac{r_j'}{r_j} = \#(\text{zeros inside } \Gamma)$$

gives that r_j and t_{j-1} have the same number of zeros inside any contour Γ on which $|r_j| \geqslant M\delta$ for M sufficiently large. In particular, there is a pairing of the zeros of r_j and t_{j-1} (up to multiplicity) such that any pair is $O(\delta^{1/k})$ apart. So to obtain (1), we need to estimate local path integrals of the form

$$(3) \qquad \int \sqrt{r(z) + O(\delta)}dz, \quad \int \sqrt{z^{-1}(r(z) + O(\delta))}dz, \quad r(z) = \prod_{j=1}^{k}(z - z_j)$$

over paths where $|r(z)| = O(\delta)$ and $|r(z)| \geqslant M\delta$. This is done exactly like the estimate of (2) of 16.12. This gives (1), as required. Then (2) follows. But (2) says that λ is within $O(\delta^{1/2+1/2k})$ of an eigenvalue (with positive-entry eigenvector) of a certain integer-valued matrix, which is one of the matrices defining the piecewise linear action of χ on the projective space of measured foliations on $\overline{\mathbf{C}} \setminus Y_0$ (see[**F-L-P**]). This means that

$$\log \lambda = \kappa(\alpha) + O(\delta^{1/2+1/2k})$$

as claimed. This works equally well, whether or not $[\chi_\alpha]$ is irreducible.

(2)-(4). First suppose that $DG_{[\alpha]} = 0$ but $G_{[\alpha]} \neq 0$. We have

$$r_j = s_{\alpha*}t_{j-1} = s_{\alpha*}s_\alpha^* r_j$$

for all j. Now if s_α has degree > 1 on the $j-1$'th sphere (that is, degree two in all the cases we are considering) then the terms in the sum

$$s_{\alpha*}t_{j-1}(z) = (s_1^{-1})'(z))^2 t_{j-1}(s_1^{-1}(z)) + ((s_2^{-1})'(z))^2 t_{j-1}(s_2^{-1}(z))$$

must have the same argument for all z. It follows that

$$t_{j-1} = s_\alpha^* r_j$$

for all j. This is impossible in case 2. In case 3, v_1, v_2 are strictly preperiodic, and we have a degree two branched covering of the union of spheres by tori. The map s_α lifts to a nonbranched degree two covering. Then in the cover the condition $r_j = s_{\alpha*}t_{j-1}$ implies that the r_j all lie on a unique geodesic in the Teichmüller space of the torus, that is, in the upper half-plane. The bound on $G_{[\alpha]} - \kappa(\alpha)$ relative to $\|DG_{[\alpha]}\|$ in this case holds because the second derivative has co-rank one (which is not true in all cases of 1). In case 4, the condition also implies that the r_j (or $[\varphi_j]$) lie on a geodesic, in the order specified.

In case 2, if $DG_{[\alpha]}$ is close to 0, we must be in $\mathcal{T}(B_\alpha)(\Gamma, M_2\delta)$ for a maximal invariant (f_α, Γ), and Γ contains no Levy cycles by 2.7.

(4) Let n be the period of Δ_0'. Write $[\varphi_n] = \tau([\varphi_0])$. We consider the bound on $F_{[\Delta_0']}(x) - \kappa_1(\Gamma, x)$. This essentially uses an inverse of the Triangular Lemma 8.9. Extend the geodesic joining $[\varphi_j]$ and $[\varphi_{j+1}]$ a distance $\sum_{\ell=j}^{n-1} d([\varphi_\ell], [\varphi_{\ell+1}])$ to a point $[\psi_j]$. Thus, $[\varphi_n] = [\psi_n]$. By the hypothesis, the quadratic differentials $r_j(z)(z)dz^2$, $t_{j-1}(z)dz^2$ at $[\varphi_j]$ for $d([\varphi_j], [\psi_j])$ and $d([\varphi_j], [\psi_{j-1}]$ are within $O(\delta)$. Note that

$$\kappa_1(\Gamma, x) = d([\varphi_0], [\psi_n]), \quad F_{[\Delta_0']}(x) = \sum_{j=0}^{n-1} d([\varphi_j], [\varphi_{j+1}]) = d([\varphi_0], [\psi_0].$$

So we need to show that

$$d([\psi_0], [\psi_n]) = O(\delta^{1/k}).$$

So, for fixed j, we need to show that

$$d([\psi_j], [\psi_{j-1}]) = O(\delta^{1/k}).$$

This follows from 16.8. □

CHAPTER 21

CONSTRUCTION OF THE GOOD VECTOR FIELD: PART 2

21.1. The construction of a good vector field has now been reduced to Proposition 20.7. There are a number of different cases to consider, which we proceed to do in this chapter. Throughout this chapter, we fix a discrete invariant (f_0, Γ) with $\Gamma_2(f_0, \Gamma) \neq \varnothing$, fixed union Ω and maximal $\Gamma' \subset \Gamma$ satisfying the Invariant and Levy conditions. We also fix $x \in \mathcal{T}(\Gamma, \varepsilon_0)$.

In the next few lemmas, to simplify notation, we write $\alpha = [\alpha]$ if α is non-periodic.

21.2. Lemma. — *Let α be a a gap or loop. Let $x \in \mathcal{T}(\gamma, \varepsilon)$ for all $\gamma \subset \partial\beta$, $\beta \in [[\alpha]]$. Let α be of D_1-orbit dominant area. Let $\theta_1 \in V[\alpha]$ with*

$$D_1 C_1^2 \|\theta_1\| \leqslant 1, \quad DF(\theta_1) \leqslant -2C_1 a(\alpha), \quad DF_{[\alpha]}(\theta_1) \leqslant -2C_1.$$

Then there exists $\theta_2 \in \oplus_{\beta \in [[\alpha]] \setminus [\alpha]} V(\beta)$ such that $\theta = \theta_1 + \theta_2$ satisfies

$$\|\theta\| \leqslant C_1 \|\theta_1\|, \quad DF(\theta) \leqslant -C_1 a(\alpha), \quad DF_{[\alpha]}(\theta) \leqslant -C_1,$$

$$DF_{[[\beta]]}(\theta) \leqslant C_1 e^{-2\pi^2/\varepsilon} \|\theta_1\| \quad \text{for all } \beta \in [[\alpha]] \setminus [\alpha].$$

Proof. — Take any $\theta \in V[[\alpha]] = \oplus_{\beta \in [[\alpha]]} V(\beta)$ and write $\theta = \sum_{\beta \in [[\alpha]]} \theta_\beta$ with $\theta_\beta \in V(\beta)$. Then by 20.5, if $\beta \in [[\alpha]] \setminus [\alpha]$, then

$$DF_\beta(\theta) = DF_\beta(\theta_\beta + \theta_{f_0(\beta)}) + O(e^{-2\pi^2/\varepsilon} \|\theta\|),$$

and

$$DF_{[[\alpha]]}(\theta) = \sum_{\beta \in [[\alpha]]} DF_\beta(\theta) = DF_{[\alpha]}(\theta_1) + \sum_{\beta \in [[\alpha]] \setminus [\alpha]} (DF_\beta(\theta_\beta + \theta_{f_0(\beta)})) + O(e^{-2\pi^2/\varepsilon} \|\theta\|),$$

and for $\beta \in [[\alpha]] \setminus [\alpha]$,

$$DF_{[[\beta]]}(\theta) = \sum_{\beta' \in [[\beta]]} DF_{\beta'}(\theta_{\beta'} + \theta_{f_0(\beta')}) + O(e^{-2\pi^2/\varepsilon} \|\theta\|).$$

By 20.10,

$$DF(\theta) = DF(\theta_1) + \sum_{\beta \in [[\alpha]] \setminus [\alpha]} O(a(\beta) \|\theta\|) = DF(\theta_1) + O(D_1 \|\theta\| a(\alpha)).$$

So having chosen $\theta_1 = \sum_{\alpha' \in [\alpha]} \theta_{\alpha'}$, we can define $\theta_2 = \sum_\beta \theta_\beta$, the sum being over $\beta \in [[\alpha]] \setminus [\alpha]$, by $\theta_{\beta'} = -\theta_{f_0(\beta')}$ for successive preimages. Then, assuming that C_1 is large enough $\|\theta\| \leqslant C_1 \|\theta_1\|$ and we have the required estimates. □

21.3. For any gap or loop α, we let $n_1(\alpha, x)$ be as in 19.6. Let $m_{\gamma,\alpha}(x)$ be the sum of $m_\zeta(x)$ with ζ separating α from $[\gamma]$ ($\zeta \in \Gamma$) and including α if α is a loop. Note that

$$e^{-m_{\gamma,\alpha}(x)}(n_1(\alpha, x))^{-1} \leqslant (n_1(\gamma, x)))^{-1}.$$

In fact, if $\gamma \not\subset \partial\alpha$, $\gamma \neq \alpha$,

$$e^{-m_{\gamma,\alpha}(x)}(n_1(\alpha, x))^{-1} \leqslant n_1(\gamma, x) = o(1).$$

Lemma. — *Let α be a gap or loop. Suppose that for any loop $\gamma \subset \mathrm{Int}(\Omega \cup f_0^{-1}(\Omega))$, with $\gamma \subset f_0^{-1}(\gamma')$,*

$$a(\gamma) = o(\min(e^{m_{\gamma,\alpha}}, e^{m_{\gamma',\alpha}})a(\alpha)n_1(\alpha, x)).$$

Let $\theta_1 \in V[[\alpha]]$ with

$$\|\theta_1\| \leqslant n_1(\alpha, x)^{-1}, \quad DF(\theta_1) \leqslant -a(\alpha)\|\theta_1\|/C_1.$$

Then there exists $\theta_2 = \sum_{\gamma\in\Gamma'} \theta_\gamma$, $\theta_\gamma \in V(\gamma)$, such that if $\theta = \theta_1 + \theta_2$,

$$\|\theta_\gamma\| \leqslant e^{-m_{\gamma,\alpha}(x)} C_1\|\theta_1\| \leqslant (n_1(\gamma, x))^{-1},$$
$$DF(\theta) \leqslant -a(\alpha)\|\theta_1\|/2C_1, \quad Dm_{[\gamma]}(\theta) \geqslant 0 \text{ for all loops } \gamma \subset \mathrm{Int}(\Omega).$$

Proof. — Take any θ decomposing as in the statement above, with bounds on $\|\theta_\gamma\|$ as above. By 20.5, we have $\|\pi_\gamma(\theta - \theta_\gamma)\| = O(e^{-m_{\gamma,\alpha}}\|\theta_1\|)$ for all loops γ. So $Dm_\gamma(\theta - \theta_\gamma) = O(e^{-m_{\gamma,\alpha}}\|\theta_1\|)$. So we can indeed choose $\theta_\gamma \in V(\gamma)$ with $\|\theta_\gamma\| = O(e^{-m_{\gamma,\alpha}}\|\theta_1\|)$ for $\gamma \in \Gamma'$, $\gamma \subset \mathrm{Int}(\Omega)$, and letting θ_2 be the sum of these, we can ensure that

$$Dm_{[\gamma]}(\theta_1 + \theta_2) \geqslant 0 \quad \text{for all } \gamma \in \Gamma,\ \gamma \subset \mathrm{Int}(\Omega).$$

If we consider the formula for $DF(\theta_1 + \theta_2)$, we see that we have

$$DF(\theta_1 + \theta_2) = DF(\theta_1) + 2\pi \,\mathrm{Re}\Big(\sum\Big\{\int_{\varphi(\gamma)} c_\gamma z q(z)dz : \gamma \subset \mathrm{Int}(\Omega)\Big\}\Big)$$
$$+ 2\pi \,\mathrm{Re}\Big(\sum\Big\{\int_{\psi(\gamma')} c_{\gamma'} z p(z)dz : \gamma' \in f_0^{-1}(\gamma) \setminus [\gamma],\ \gamma \subset \mathrm{Int}(\Omega)\Big\}\Big),$$

where

$$c_\gamma = O(\|\theta_\gamma\|) = O(e^{-m_{\gamma,\alpha}}\|\theta_1\|) = O(e^{-m_{\gamma,\alpha}} n_1(\alpha, x)^{-1}),$$

and if $\gamma' \in f_0^{-1}(\gamma)$ then

$$c_{\gamma'} = O(\|\theta_\gamma\|) = O(e^{-m_{\gamma,\alpha}}\|\theta_1\|) = O(e^{-m_{\gamma,\alpha}} n_1(\alpha, x)).$$

So the additional terms are

$$O(a(\gamma)e^{-m_{\gamma,\alpha}}\|\theta_1\|, \quad O(a(\gamma')e^{-m_{\gamma,\alpha}}\|\theta_1\|,$$

which are both $o(a(\alpha)\|\theta_1\|/C_1)$, as required. □

21.4. Lemma. — *Let α be nonperiodic. Then there exists $\theta \in V(\alpha)$ with*

$$\|\theta\| \leqslant C_1^2, \quad DF(\theta) \leqslant -a(\alpha)C_1, \quad DF_\alpha(\theta) \leqslant -C_1.$$

Proof. — Let α be nonperiodic. Take any $\theta \in V(\alpha)$. So by 20.5, 20.10, and the definition of D_1-orbit dominant area,

$$DF(\theta) = a(\alpha)DF_\alpha(\theta) + O(\sqrt{D_1}a(\alpha)).$$

We have seen in 20.5 that $DF_\alpha \mid V(\alpha)$ is (in suitable coordinates) the vector $2\pi c(\alpha, q_\alpha) = 2\pi(\mathrm{Res}(q_\alpha, \varphi_\alpha(y))$, which is $\geqslant Ca(\alpha)$ for a constant $C > 0$. So we can choose θ as required. □

21.5. Lemma. — *Let α be a periodic gap. Let $x \in \mathcal{T}(\gamma, \varepsilon)$ for $\gamma \subset \partial\alpha'$, $\alpha' \in [\alpha]$. Let $D_1 \geqslant e^{-2\pi^2/\varepsilon}$. Let $k = \#(Y)$. Let α be of D_1-orbit dominant area. Let*

$$G_{[\alpha]}(x) \geqslant \kappa(\alpha) + D_1^{1/20k}.$$

Then there exists $\theta \in V[\alpha]$ with

$$\|\theta\| \leqslant D_1^{-1/4}, \quad DF(\theta) \leqslant -C_1 a(\alpha), \quad DF_{[\alpha]}(\theta) \leqslant -C_1.$$

Proof. — Write $[\alpha] = \{\alpha_j : 0 \leqslant j < n\}$ with $\alpha = \alpha_0$. Write $a(\alpha_j, q) = a(\alpha_j)$. We have seen in 20.5 and 20.13 that $DF_{[\alpha]} \mid V[\alpha]$ and $DF \mid V[\alpha]$ are given by the vectors

$$\underline{x}_1 = (c(\alpha_j, r_j - (s_\alpha)_* t_{j-1})) + O(e^{-2\pi^2/\varepsilon}),$$
$$\underline{x}_2 = (a(\alpha_j)c(\alpha_j, r_j) - a(\alpha_{j-1})c(\alpha_j, (s_\alpha)_* t_{j-1})) + O(\sqrt{D_1}a([\alpha]).$$

If it is impossible to find θ then either $\underline{x}_1 = O(D_1^{1/8})$ or $\underline{x}_2 = O(D_1^{1/8}a([\alpha])$ or the angle between $\underline{x}_1$ and $-\underline{x}_2$ is $O(C_1 D_1^{1/8})$. Either of the last two gives, for some $\mu > 0$,

$$(1) \quad (c(\alpha_j, r_j) - c(\alpha_j, (s_\alpha)_* t_{j-1})) = -\mu(a(\alpha_j)c(\alpha_j, r_j) - a(\alpha_{j-1})c(\alpha_j, (s_\alpha)_* t_{j-1})) \\ + O(C_1 D_1^{1/8}(1 + \mu a([\alpha])).$$

But this implies that

$$(1 + \mu a(\alpha_j))r_j = (1 + \mu a(\alpha_{j-1})(s_\alpha)_* t_{j-1} + \sum_{y \in A(\alpha_j)} \frac{O(C_1 D_1^{1/8}(1 + \mu a([\alpha]))}{z - \varphi_j(y)}$$

which gives, by integrating modulus, for all j (replacing $j-1$ by $n-1$ if $j = 0$)

$$1 + \mu a(\alpha_j) \geqslant 1 + \mu a(\alpha_{j-1}) + O(C_1 D_1^{1/8}(1 + \mu a([\alpha]))),$$

and thus

$$(2) \qquad \mu a(\alpha_j) = \mu a(\alpha) + O(C_1 D_1^{1/8}(1 + \mu a([\alpha]))) \quad \text{for all } j.$$

Then from (1) and (2) we obtain either

$$(3) \qquad c(\alpha_j, r_j - (s_\alpha)_* t_{j-1}) = O(C_1 D_1^{1/16})$$

or $\mu a([\alpha]) \geqslant D_1^{1/16}/C_1$, in which case

$$a(\alpha_j) = a(\alpha)(1 + O(D_1^{1/16})),$$

which, when substituted into (1) again gives (3). By 20.14, this is only possible if $G_{[\alpha]} \leqslant \kappa(\alpha) + D_1^{1/20k}$.

21.6. Lemma. — *Let α be a periodic gap. Let $x \in \mathcal{T}(\gamma, \varepsilon)$ for all $\gamma \subset \partial\alpha'$, $\alpha' \in [\alpha]$. Write $k = \#(Y)$. Let $C_1^{-8k} \geqslant D_1 \geqslant e^{-2\pi^2/\varepsilon}$, and let D_1 and ε be sufficiently small. Let α be of D_1-orbit dominant area. Let $a(\alpha'')/a(\alpha') - 1 \geqslant D_1^{1/5}a[\alpha]$ for some α', $\alpha'' \in [\alpha]$, or let $\kappa(\alpha)$ be nonmaximal. Then there exists $\theta \in V[\alpha]$ with*

$$\|\theta\| \leqslant D_1^{-1/4}C_1^2, \quad DF(\theta) \leqslant -a([\alpha])C_1.$$

Proof. — If θ does not exist then, using the formula of 20.13 as in 21.5, for all j

$$a(\alpha_j)c(\alpha_j, r_j) - a(\alpha_{j-1})c(\alpha_j, (s_\alpha)_* t_{j-1}) = O(D_1^{1/4}a([\alpha])).$$

Integrating as in 21.5 we obtain

$$a(\alpha_{j-1}) \geqslant a(\alpha_j) + O(D_1^{1/4})$$

and hence $a(\alpha_j) = a(\alpha)(1 + O(D_1^{1/4}))$ for all j. In particular, all α_j are of $(2D_1)$-orbit dominant area. So then, for all j

$$c(\alpha_j, r_j) - c(\alpha_{j-1}, (s_\alpha)_* t_{j-1}) = O(D_1^{1/4}).$$

So then by 20.14 (again as in 21.5)

$$G_{[\alpha]}(x) = \kappa(\alpha) + O(D_1^{1/4k}).$$

So for all j

$$F_{\alpha_j}(x) = \kappa(\alpha) + O(D_1^{1/4}) + O(e^{-2\pi^2/\varepsilon}).$$

But α_j is of $2D_1$-dominant area. So by 9.5

$$F(x) \leqslant F_{\alpha_j}(x) + CD_1 \leqslant \kappa(\alpha) + O(D_1^{1/4}) + O(e^{-2\pi^2/\varepsilon}).$$

So $\kappa(\alpha)$ must be maximal. □

21.7. Lemma. — *Let α be a loop in Ω of D_1-orbit dominant area. Write*

$$m = \min\{m_{\alpha'}(x) : \alpha' \in [\alpha]\}.$$

Then there exists $\theta \in V[\alpha]$ with

$$DF(\theta) \leqslant -a(\alpha)\|\theta\|/C_1, \quad Dm_{[\alpha]}(\theta) \geqslant (m_{[\alpha]}(x))^{-1}\|\theta\|/C_1.$$

If in addition $d_\alpha(x, \tau(x)) \geqslant \frac{1}{2}d(x, \tau(x))$, and $a(\alpha') = o(a(\alpha))$ whenever $d_{\alpha'}(x, \tau(x)) = o(d_\alpha(x, \tau(x))$ then

$$Dm_{\alpha'}(\theta) \leqslant C_1 e^{-m}\|\theta\| \quad \text{for all } m_{\alpha'} \text{ maximal,}$$
$$DF_{[\alpha]}(\theta) \leqslant -(m_{[\alpha]}(x))^{-1}\|\theta\|/C_1.$$

Proof. — Let $\alpha = \alpha_0$ be a periodic loop in the central part, of period n and with orbit $\{\alpha_j : 0 \leqslant j \leqslant n-1\}$ with α_j homotopic to a component of $f_0^{-1}(\alpha_{j+1})$: write $\alpha_{-1} = \alpha_{n-1}$. For the moment, we assume that $[\alpha] \neq [\partial\Delta_0]$. Write

$$\int_{\varphi(\alpha_j)} \sqrt{q} = c_j + id_j.$$

Take $\frac{1}{2}\log K = F(x) = d(x,\tau(x))$ and $\frac{1}{2}\log K_j = d_{\alpha_j}(x,\tau(x))$. Then by 20.11 and the Orbit Dominant Area Condition we have

$$\begin{aligned}
2\pi c(\alpha_j, q) &= -(2\pi)^{-1}(c_j + id_j)^2 + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)) \\
&= -(2\pi)^{-1}(c_j^2 - d_j^2 + 2ic_jd_j) + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)), \\
2\pi c(\alpha_j, s_*p) &= 2\pi c(\alpha_{j-1}, p) + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)) \\
&= -(2\pi)^{-1}(\sqrt{K}c_{j-1} + id_{j-1}/\sqrt{K})^2 + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)) \\
&= -(2\pi)^{-1}(Kc_{j-1}^2 - d_{j-1}^2/K + 2ic_{j-1}d_{j-1} + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)).
\end{aligned}$$

Similarly let $q_{\alpha_j}(z)dz^2$ be the quadratic differential for $d_{\alpha_j}(x,\tau(x))$ at x and write $\frac{1}{2}\log K_j = d_{\alpha_j}(x,\tau(x))$ and

$$\int_{\varphi(\alpha_j)} \sqrt{q_{\alpha_j}} = c_j' + id_j'.$$

Then

$$2\pi c(\alpha_j, q_{\alpha_j}) = -(2\pi)^{-1}(c_j' + id_j')^2 + O(e^{-m/2(\sqrt{K_j}-1)}m^{-1}),$$

and similarly for $s_*p_{\alpha_j}$. We are required to find λ_j with

$$\text{(1)} \qquad \sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j) \leqslant -C_1^{-1} \sum_{j=0}^{n-1} |\lambda_j|,$$

$$\text{(2)} \qquad (2\pi)^{-1} \sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j(Kc_{j-1}^2 - d_{j-1}^2/K + 2ic_{j-1}d_{j-1} - (c_j^2 - d_j^2 + 2ic_jd_j))) + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)) \leqslant -C_1^{-1}a(\alpha)\sum_{j=0}^{n-1}|\lambda_j|,$$

$$\text{(3)} \qquad (2\pi)^{-1} \sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j(K_jc_{j-1}'^2 - d_{j-1}'^2/K_j + 2ic_{j-1}'d_{j-1}' - (c_j'^2 - d_j'^2 + 2ic_j'd_j'))) + O((e^{-m/2(\sqrt{K_j}-1)}m^{-1}) \leqslant -C_1^{-1}m^{-1}\sum_{j=0}^{n-1}|\lambda_j| \max_j(K_j - 1),$$

$$\mathrm{Re}(\lambda_j) \leqslant 0 \text{ if } m_{\alpha_j}(x) \text{ is maximal.}$$

We can clearly achieve (1) and (2) with $\lambda_j = -1$ for all j. More care is needed to obtain (1)-(4) simultaneously in the case when $d_\alpha(x,\tau(x)) \geqslant \frac{1}{2}d(x,\tau(x))$. The required lower bound on $Dm_{\alpha_j}(\theta)$ will then follow from (4) by 20.4.

Fix a path γ_j which is nontrivial nonperipheral in $\overline{\mathbf{C}} \setminus A(\alpha_j)$ and intersects α_j exactly twice. We can choose the γ_j so that $f_0^{-1}(\gamma_{j+1})$ is homotopic to γ_j in $\overline{\mathbf{C}} \setminus A(\alpha_j)$, $f_0^{-1}(\gamma_0)$ is nontrivial nonperipheral in $\overline{\mathbf{C}} \setminus A(\alpha_{n-1})$ and has bounded number of intersections with γ_{n-1}. Write $x = [\varphi]$ and $\tau(x) = [\psi]$. Normalize so that $\varphi(A(\alpha_j))$, $\psi(A(\alpha_j))$ contain 0, 1, ∞ and one other point close to 0. Let Δ_j be the modulus of the largest annulus satisfying the Pole-Zero Condition (9.4) and homotopic to $\varphi(\alpha_j)$. By the assumption $d_\alpha(x, \tau(x)) \geqslant \frac{1}{2} d(x, \tau(x))$, Δ_0 is boundedly proportional to $m_\alpha(x)$, and hence to m. We claim that the (signed) difference between the number of crossings of the positive real axis by $\varphi(\gamma_j)$ and $\psi(\gamma_j)$ is boundedly proportional to $c_j d_j \Delta_j (c_j^2 + d_j^2)^{-1}$. We see this as follows. Put $\varphi(\gamma_j)$, $\psi(\gamma_j)$ in good position (14.5) with respect to the quadratic differentials $q(z)dz^2$, $p(z)dz^2$. Then we need to calculate the difference between the number of times that $\varphi(\gamma_j)$, $\psi(\gamma_j)$ cross the real axis. It is probably easier to calculate with γ_j replaced by γ_j' where $\varphi(\gamma_j')$ is approximately perpendicular to $\varphi(\alpha_j)$, that is, for some $\lambda > 0$

$$\int_{\varphi(\gamma_j')} \sqrt{q(z)}dz = \lambda(2c_j d_j + i(d_j^2 - c_j^2)) + O(c_j^2 + d_j^2).$$

Then

$$\int_{\psi(\gamma_j')} \sqrt{q(z)}dz = \lambda(2c_j d_j + i(d_j^2/K - Kc_j^2)) + O(c_j^2 + d_j^2).$$

Then λ is boundedly proportional to Δ_j. Assuming for the moment that $|d_j| \leqslant |c_j|$, the difference between real axis crossings is boundedly proportional to $(d_j/c_j)\Delta_j$, and hence to $c_j d_j \Delta_j (c_j^2 + d_j^2)^{-1}$. We get the same estimate (with the same sign) if $|c_j| \leqslant |d_j|$.

So for numbers C_j, $C_j' > 0$, bounded and bounded from 0

$$\mathrm{Re}(\pi_{\alpha_j}([\varphi])) - \mathrm{Re}(\pi_{\alpha_j}([\psi])) = C_j c_j d_j \Delta_j (c_j^2 + d_j^2)^{-1} = C_j' c_j' d_j' m^2 + o(1)$$

But the bound on $f_0^{-1}(\gamma_j) \cap \gamma_{j+1}$ yields (with $j+1$ replaced by 0 if $j = n-1$)

$$\mathrm{Re}(\pi_{\alpha_j}([\varphi])) - \mathrm{Re}(\pi_{\alpha_{j+1}}([\psi])) = O(1).$$

Therefore

$$\sum_{j=0}^{n-1} C_j c_j d_j \Delta_j (c_j^2 + d_j^2)^{-1} = \sum_{j=0}^{n-1} C_j' c_j' d_j' m^2 + o(1) = O(1).$$

But $c_j'^2 + d_j'^2$ is boundedly proportional to m^{-1}. So if $|c_j' d_j' m^2| \geqslant \Delta$ for a sufficiently large Δ, this is also true for some other $k \neq j$, where $c_j' d_j'$ and $c_k' d_k'$ have opposite signs. Also by our assumptions, for a suitable $C > 0$, $|c_j' d_j' m^2| \geqslant \Delta$ if and only if $|c_j d_j| \geqslant C\Delta m^{-1}(c_0^2 + d_0^2)$, whenever $\Delta/m \geqslant \frac{1}{2} d(x, \tau(x))$. Arguing as in 21.5, either we can find λ_j purely imaginary such that (5) and (6) hold, or there is $\mu > 0$ such that

(7) holds:

$$\sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j(ic_jd_j - ic_{j-1}d_{j-1})) \leqslant -\delta \sum_{j-0}^{n-1} |\lambda_j| \sum_{j=0}^{n-1} |c_jd_j| \tag{5}$$

$$\sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j(ic'_jd'_j - ic'_{j-1}d'_{j-1})) \leqslant -\delta \sum_{j-0}^{n-1} |\lambda_j| \sum_{j=0}^{n-1} |c'_jd'_j| \tag{6}$$

$$c'_jd'_j + \mu c_jd_j = c'_{j-1}d'_{j-1} + \mu c_{j-1}d_{j-1} + O(\delta \max_k |c'_kd'_k - c'_{k-1}d'_{k-1}|). \tag{7}$$

But $c'_jd'_j$ and c_jd_j have the same sign whenever $|c'_jd'_j| \geqslant \Delta m^{-2}$. So (7) does not hold — and (5) and (6) do hold — for small δ, unless $|c'_jd'_j| \leqslant \Delta m^{-2}$ for all j.

So we can solve (1) to (4) simultaneously unless $|c'_kd'_k| \leqslant D_1m^{-1}$ for all k. So now suppose this holds. Then for each k exactly one of c'^2_k, $d'^2_k \leqslant D_1m^{-1}$ and the other is boundedly proportional to m^{-1}. We can find j such that $c'^2_j \leqslant D_1m^{-1}$ and $d'^2_{j-1} \leqslant D_1m^{-1}$, (with $j-1$ replaced by $n-1$ if $j=0$) because

$$\mathrm{Im}(\pi_{\alpha_j}(\tau(x))) = \mathrm{Im}(\pi_{\alpha_{j+1}}(x)) + O(1).$$

(We can interpret this suitably if $\ell > n-1$.) Then for a large Δ',

$$m_{\alpha_j}(x) < m_{\alpha_{j-1}}(x) - \Delta' = m_{\alpha_j}(\tau(x)) - \Delta' + O(1).$$

We must have similar inequalities for c_j, d_j, c_{j-1}, d_{j-1}: that is, $c_j^2 \leqslant CD_1d_j^2$ and $d_{j-1}^2 \leqslant CD_1c_{j-1}^2$. Then we can take $\lambda_\ell = 0$ for $\ell \neq j$ and λ_j real and negative to solve (1) (3) and (4) simultaneously, and also (2) unless $c_j^2 + d_j^2 = o(a(\alpha))$. If $c_j^2 + d_j^2 = o(c_0^2 + d_0^2)$ for some $k \leqslant j \leqslant \ell$ then we can add on smaller nonzero λ_ℓ, $\mathrm{Re}(\lambda_\ell) \leqslant 0$ for some ℓ to solve (1) to (4) simultaneously.

Now we consider the case $[\alpha] = [\partial\Delta_0]$. Write $\partial\Delta_0 = \alpha_0$. This is done exactly as above, but more care is needed to get the estimate

$$\begin{aligned} 2\pi c(\alpha_0, s_*p) &= 2\pi c(\alpha_{n-1}, p) + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)) \\ &= -(2\pi)^{-1}(c_{n-1} + id_{n-1})^2 + O(e^{-1/(2(\sqrt{K}-1)D_1)}a(\alpha)). \end{aligned}$$

This is proved, essentially via 20.11, but using the method of the first part of 20.11 with $\tau(x)$ replacing x: if $\tau(x) = [\psi]$, there is a good boundary annulus homotopic to $\psi(\gamma_{n-1})$ of modulus $\geqslant \frac{1}{2}m_{\gamma'_0}([\varphi]) + D_1^{-1}$ satisfying the Pole-Zero Condition. Then an annulus satisfying the Pole-Zero Condition and of modulus $\geqslant D_1^{-1} - O(1)$ maps under s to the homotopy class of γ_0. On this annulus s is of degree 1. So we have the same estimates as before.

21.8. Lemma. — *Let $\alpha \subset \Omega(f_0, \Gamma)$ be reducible and of D_1-orbit dominant area with $\kappa(\alpha)$ maximal. Let $\beta \subset \alpha$ adjacent to the outer boundary have $\kappa(\beta)$ nonmaximal. Let $\zeta \subset \alpha$ be separated from β by a single loop γ, and let $\kappa(\zeta)$ maximal. Then there exists $\theta \in V[\alpha]$ with*

$$\|D\pi_\beta(\theta)\| \leqslant n_1(\beta, x)^{-1} \quad \text{for all gaps and loops } \beta \subset \alpha$$

(for some decomposition of α into gaps and loops: see 19.6 for $n_1(\beta,x)$) and

$$DF(\theta) \leqslant 0, \quad Dm_{[\gamma]}(\theta) \geqslant m_{[\gamma]}(x)(\log m_{[\gamma]}(x))^2. \tag{1}$$

Proof. — As usual, let $q(z)dz^2$ denote the quadratic differential for $d(x,\tau(x))$ at x, with stretch $p(z)dz^2$ at $\tau(x)$. In each of the following circumstances, we can find θ as above, given a sufficiently small D_1 (depending only on C_1).

1. Some $\gamma' \in [\gamma]$ is of D_1-orbit dominant area: in this case $\theta \in V[\gamma]$.
2. Some $\beta' \in [\beta]$ has D_1-dominant area and $a([\gamma]) \leqslant e^{-m_\gamma(x)/C_1} a([\beta])$: in this case $\theta = \theta_1 + \theta_2$ with $\theta_1 \in V[\beta]$ and $\theta_2 \in V[\gamma]$.
3. For some ζ', $\zeta'' \in [\zeta]$ $|a(\zeta')/a(\zeta'') - 1| \geqslant D_1^{1/5}$. In this case $\theta = \theta_1 + \theta_2$ with $\theta_1 \in V[\zeta]$ and $\theta_2 \in V[\gamma]$.
4. $|F(x) - \kappa(\zeta)| \geqslant e^{-m_\gamma(x)/C_1}$.

Case 1 follows from 21.7. By 21.6 we can find θ_1 with $\|\theta_1\| \leqslant D_1^{-1/4} C_1^2$ in cases 2 and 3 with $DF(\theta_1) \leqslant -C_1 a([\beta])$ or $DF(\theta_2) \leqslant -C_1 a([\zeta])$. Then assuming case 1 does not hold we can add on θ_2, assuming only that $m_\gamma(x)$ is sufficiently large given D_1. If case 4 holds and none of the previous cases hold, then either ζ itself has $e^{-m_\gamma(x)/C_1}$-orbit dominant area or some other gap or loop $\eta \subset \alpha$ satisfies the condition of 21.5 or 21.6, with $a([\gamma]) \leqslant e^{-m_\gamma(x)/C_1} a([\eta])$.

Reduce α if necessary, so that $\beta \subset \alpha$ is adjacent to the outer boundary. Let $[\alpha] = \{\alpha_j : 0 \leqslant j \leqslant m-1\}$ with $\alpha_j \subset f_0^{-1}(\alpha_{j+1})$ and $\alpha_0 = \alpha$. Write $[\varphi_j] = \pi_{\alpha_j}(x)$. Let $\chi_j = f_0 \mid \alpha_j$ and write, as usual,

$$G_{[\alpha]}(x) = \sum_{j=0}^{m-1} d_{\alpha_j}([\varphi_j], [\varphi_{j+1} \circ \chi_j]).$$

Let

$$[\gamma] = \{\gamma_j : 0 \leqslant j \leqslant n-1\} \quad \text{with } \gamma_j \subset f_0^{-1}(\gamma_{j+1}) \text{ and } \gamma = \gamma_0,$$
$$[\beta] = \{\beta_j : 0 \leqslant j \leqslant n-1\} \quad \text{with } \beta_j \subset f_0^{-1}(\beta_{j+1}) \text{ and } \beta_0 = \beta.$$

Thus $\gamma_j \subset \partial\beta_j$ separates β_j from $[\zeta]$. Choose $A(\beta_k)$ with $y_k \in A(\beta_k)$ separated from β_k by γ_k. Choose the y_k so that $y_k = y_\ell$ if β_k, $\beta_\ell \subset \alpha_j$ for some j. Normalize φ_j so that $\varphi_j(y_k) = \infty$. Otherwise normalise as in 20.2. Let $r_j(z)dz^2$ denote the quadratic differential at $[\varphi_j]$ for $d_{\alpha_j}([\varphi_j], [\varphi_{j+1} \circ \chi_j])$, with stretch $t_j(z)dz^2$ at $[\varphi_{j+1} \circ \chi_j]$.

Let $\theta \in V[\gamma] = \oplus_{\gamma' \in [\gamma]} V(\gamma')$ be defined by $\lambda = e$ for each $\gamma' \in [\gamma]$ (20.2). Let σ_u be the integral flow of this θ. Write $\sigma_u(x) = ([\varphi_{u,j}])$ in analogy with the above. Normalize in the same way as for $[\varphi_j]$. Let $r_{u,j}(z)dz^2$ be the quadratic differential for $d_{\alpha_j}([\varphi_{u,j}], [\varphi_{u,j+1} \circ \chi_j])$ at $[\varphi_{u,j}]$, with stretch $t_{u,j}(z)dz^2$ at $[\varphi_{u,j+1} \circ \chi_j]$. By Hard Same Shape 16.2 (where a different scaling was used),

$$\mathrm{Res}(r_{u,j}, \varphi_{u,j}(y)) = \lambda_u(\mathrm{Res}(r_j, \varphi_j(y)) + o(a(\beta_j, r_j))) \tag{4}$$

and similarly for $t_{u,j}(z)dz^2$.With similar notation to the above, let

$$G_{[\alpha]}(u) = \sum_{j=0}^{m-1} d_{\alpha_j}([\varphi_{u,j}], [\varphi_{u,j} \circ \chi_j]).$$

Let

$$H(u) = F(\sigma_u(x)).$$

Then to complete the proof, it suffices to show that

$$H'(0) \leqslant 0.$$

The usual derivative formula, as in 20.4, gives

$$G'_{[\alpha]}(u) = \sum_{k=0}^{m-1} \int_{\varphi_{u,k}(\gamma_k)} z(r_{u,k}(z) - t_{u,k-1}(z))dz$$

which, by 16.2 (where a different normalisation was used) gives

$$G'_{[\alpha]}(u) = e^{-u} \sum_{k=0}^{m-1} \int_{\varphi_k(\gamma_k)} z(r_k(z) - t_{k-1}(z)) + o(e^{-u}a([\beta], r)) = o(a([\beta], r_u).$$

Also by 16.2

$$a([\beta], r_u) = e^{-u}a([\beta], r)(1 + o(1)).$$

By 20.4 and 20.13, however,

$$H'(0) = \sum_{k=0}^{m-1} \int_{\varphi_k(\gamma_k)} z(r_k(z) - t_{k-1}(z)) + O(\sqrt{D_1}a([\beta]).$$

Now using 9.5 when β has dominant area and 17.5 otherwise, for a suitable constant $C_0 > 0$,

$$\kappa(\alpha) + C_0^{-1}e^{-u}a([\beta], r) \leqslant G_{[\alpha]}(u) \leqslant \kappa(\alpha) + C_0e^{-u}a([\beta], r)$$

So for a constant $C'_0 > 0$ we must have

$$\sum_{k=0}^{m-1} \int_{\varphi_k(\gamma_k)} z(r_k(z) - t_{k-1}(z)) \leqslant -C'_0a([\beta]).$$

So then $H'(0) \leqslant 0$ as required. □

21.9. Lemma. — *Let $\alpha \in \Gamma \setminus \Gamma'$ be periodic of (D_1, D'_1)-orbit dominant area. Let*

$$m = \min\{m_{\alpha'}(x) : \alpha' \in [\alpha]\}.$$

Suppose that $|F(x) - \kappa(\alpha)| \geqslant D'_1/D_1$. Then there exists $\theta \in V[\alpha]$ with

$$DF(\theta) \leqslant -D'_1C_1^{-1}a(\alpha)\|\theta\|, \quad Dm_{[\alpha]}(\theta) \leqslant C_1e^{-m}\|\theta\|,$$
$$DF_{[\alpha]}(\theta) \leqslant 0 \text{ or } F_{[\alpha]}(x) \leqslant \kappa(\alpha) + D'_1/D_1,$$

and one of

$$DF_{[\alpha]}(\theta) \leqslant -D'_1C_1^{-1}m_\alpha(x)^{-1}\|\theta\|, \quad Dm_{[\alpha]}(\theta) \leqslant -D'_1C_1^{-1}\|\theta\|.$$

Proof. — Use the notation of 21.7. Write n_j for the degree of $f_0 \mid \alpha_{j-1}$. Then $c(\alpha_j, s_*p) = n_j^{-1}c(\alpha_{j-1}, p) + O(e^{-m/2}a(\alpha)) = n_j^{-1}(c_{j-1} + id_{j-1})^2 + O(e^{-m/2}m^{-1}))$.
The reason for the first equality is that

$$\sum_{k=1}^{n_j-1}((d/dz)(e^{2\pi ik/n_j}z^{1/n_j}))^2 z/(e^{2\pi ik/n_j}z^{1/n_j})^{-2} = n_j^{-1}z^{-2}.$$

Similarly to 21.7, we need λ_j solving (1) and (2), (3) whenever $F_{[\alpha]}(x) \geqslant \kappa(\alpha)+D_1'/D_1$, and one of (4), (5). Of course, (4) is stronger than (3) and (5) is stronger than (1).

$$\sum_j \mathrm{Re}(\lambda_j) \geqslant 0 \tag{1}$$

$$\begin{aligned}\sum_{j=0}^{n-1}(\mathrm{Re}\,\lambda_j(n_j^{-1}(Kc_{j-1}^2 - d_{j-1}^2/K + 2ic_{j-1}d_{j-1}) - (c_j^2 - d_j^2 + 2ic_jd_j)))\\ + O(e^{-m/2}m^{-1}) \leqslant -D_1'C_1^{-1}a(\alpha)\sum_{j=0}^{n-1}|\lambda_j|.\end{aligned} \tag{2}$$

$$\begin{aligned}\sum_{j=0}^{n-1}(\mathrm{Re}\,\lambda_j(n_j^{-1}(K_{j-1}c_{j-1}'^2 - d_{j-1}'^2/K_{j-1} + 2ic_{j-1}'d_{j-1}')\\ - (c_j'^2 - d_j'^2 + 2ic_j'd_j'))) \leqslant 0\end{aligned} \tag{3}$$

$$\begin{aligned}\sum_{j=0}^{n-1}(\mathrm{Re}\,\lambda_j(n_j^{-1}(K_{j-1}c_{j-1}'^2 - d_{j-1}'^2/K_{j-1} + 2ic_{j-1}'d_{j-1}')\\ - (c_j'^2 - d_j'^2 + 2ic_j'd_j'))) \leqslant -D_1'C_1^{-1}m^{-1}\sum_{j=0}^{n-1}|\lambda_j|.\end{aligned} \tag{4}$$

$$\sum_j \mathrm{Re}(\lambda_j) \geqslant D_1'C_1^{-1}\sum_{j=0}^{n-1}|\lambda_j| \tag{5}$$

Here $n-1$ replaces $j-1$ if $j=0$. Using the method of 21.7, we can solve (1) to (4) simultaneously unless $c_j'd_j' = O(D_1'm^{-1})$ for all j, and $c_jd_j = O(D_1'a(\alpha))$ for all j. So now suppose this holds. Arguing similarly to 21.7, we can solve (1) to (4) simultaneously unless $d_j'^2 = O(D_1'm^{-1})$ for all j, in which case $c_j'^2$ is boundedly proportional to m^{-1}. (In contrast to 21.7, if $c_j'^2 = O(D_1'm^{-1})$ we would need to take λ_j real and positive.) So now assume this. Then

$$K_{j-1}m_{\alpha_{j-1}}(\tau(x)) = n_jm_{\alpha_j}(x)(1 + O(D_1')),$$

that is $F_{[\alpha]}(x) = \kappa(\alpha)+O(D_1')$. So (3) is satisfied. We only need to satisfy (2) and (5) simultaneously. We can clearly do this if $c_j^2 = o(c_k^2)$ for some j, k. So now suppose all c_j^2 are boundedly proportional to $a(\alpha)$, Again arguing as in 21.7, we can solve (1)-(3) and (5) simultaneously unless for some $\mu \geqslant 0$

$$n_j^{-1}Kc_{j-1}^2 - c_j^2 = -\mu + O(D_1'a(\alpha)).$$

Taking the product yields

$$\prod_j n_j K^{-n} = 1 + O(D_1'),$$

that is,

$$F(x) = \kappa(\alpha) + O(D_1'),$$

as required. □

21.10. Lemma. — *Let $\alpha \in \Gamma'$, $\alpha \not\subset \Omega$, be a periodic loop of (D_1, D_1')-orbit dominant area, $D_1' < D_1$. Let $m_\gamma(x) \geqslant \sqrt{D_1} m_{[\alpha]}(x)$ for some $\gamma \notin [\alpha]$ with $f_0^{-1}(\gamma) \cap [\alpha] \neq \varnothing$. Then there exists $\theta \in V[\alpha]$ with*

$$DF(\theta) \leqslant -D_1' a(\alpha) \|\theta\| / C_1,$$

$$DF_{[\alpha]}(\theta) \leqslant -D_1' \|\theta\| m_\alpha(x)^{-1} \quad \text{or} \quad F_{[\alpha]}(x) \leqslant F(x) - D_1'/D_1.$$

Proof. — Again, this is very similar to 21.7, 21.9. We use the same notation as in 21.7, 21.9. As in 21.9, we can find θ unless either all $c_k'^2 \leqslant D_1' m^{-1}$ or all $d_k'^2 \leqslant D_1' m^{-1}$. Similar inequalities hold for the c_k, d_k. Suppose all $c_k'^2 \leqslant D_1' m^{-1}$. Arguing as in 21.5, and using the expressions for $DF_{[\alpha]}(\theta)$, $DF(\theta)$ of 21.7, 21.9 we can find θ with the required bounds on $DF_{[\alpha]}(\theta)$, $DF(\theta)$ unless for some $\mu \geqslant 0$

$$-d_{j-1}'^2 / n_j K_{j-1} + d_j'^2 = -\mu(-d_{j-1}^2 / n_j K + d_j^2) + O(D_1' m^{-1}).$$

This cannot happen because $\prod_j (n_j K_j)^{-1}$ is boundedly < 1 and similarly for $\prod_j (n_j K)^{-1}$. So we only need to consider the case $d_j'^2 = O(D_1' m^{-1})$ for all j. Then as in 21.9 we obtain

$$F_{[\alpha]}(x) = \sum \log m_{\alpha_j}(x) - \log m_{\alpha_j}(\tau(x)) \leqslant \sum_j \log n_j - C\sqrt{D_1}$$

for C bounded from 0 because for some j

$$m_{\alpha_j}(\tau(x)) \geqslant m_{\alpha_{j-1}}(x) + m_\gamma(x) - O(1).$$

Thus

$$\sum_j \log(K_j / n_j) \leqslant -C\sqrt{D_1}.$$

Then we can find θ with the required bounds unless for some $\mu \geqslant 0$

$$K_{j-1} c_{j-1}'^2 / n_j - c_j'^2 = -\mu(K c_{j-1}^2 / n_j - c_j^2) + O(D_1' m^{-1}). \tag{1}$$

We have $K = K_j + O(D_1')$ whenever $a(\alpha_j) \geqslant a(\alpha)/D_1$ by the definition of (D_1, D_1')-orbit dominant area (and 9.5 which then gives $F(x) = F_{\alpha_j}(x) + O(D_1')$). So (1) gives

$$K_{j-1}(c_{j-1}'^2 + \mu c_{j-1}^2)/n_j = c_j'^2 + \mu c_j^2 + O((D_1 + D_1')m^{-1}).$$

Then taking the product yields a contradiction as above.

21.11. Lemma. — *Let Γ be $(L_1, L_2, \varepsilon, \nu)$-adapted to $x = [\varphi]$. Let U, β_{n-1} satisfy Triangle Condition (3) of 18.11, but suppose that the triangle Conditions (1) and (2) of 18.11 do not both hold with $e^{-C_1^2/\nu}$ replacing $e^{-C_1/\nu}$. Let U be homotopic to $\varphi(R)$ Suppose that any dominant area α is periodic with $[\alpha] \cap R \neq \varnothing$. Then there is $\theta \in \oplus_{\beta \subset R} V[\beta]$ with*

$$DF(\theta) \leqslant -C_1^{-1} e^{-C_1^2/\nu} a(\beta_{n-1}).$$

In addition, if there is another loop set Γ_1 which is (L_1, L_2)-adapted to x, possibly with loops in Int(R), *then we can ensure that θ satisfies the basic bounds.*

Proof. — By our assumption there is no dominant area α in $[[R]] \smallsetminus [R]$. So Triangle Condition (2) of 18.11 holds. So we consider a dominant $\alpha \subset [R]$. Then $a(\beta_{n-1}) = O(a(\alpha))$ by the properties of U. We can use 21.4-10 if α is a loop or $\alpha \subset \Omega$ is a gap with $\kappa(\alpha)$ nonmaximal or $\kappa(\alpha)$ maximal with $F(x) \geqslant \kappa(\alpha) + n_3(\alpha, x)$ or $a(\alpha)/a(\alpha')$ bounded from 1 (some $\alpha' \in [\alpha]$). If none of these hold then in fact $a(\alpha)/a(\alpha') \in [C_1^{-1}, C_1]$ for all $\alpha \subset [R]$, $\alpha' \subset [\alpha]$. Then if we take $\theta \in \bigoplus_j V(\beta_j)$ we have

$$DF(\theta) = \sum_{j=0}^{n-1} 2\pi \operatorname{Re}(c(\beta_j, q - s_* p)\theta_j) + o(e^{-C_1^2/\nu} a([\beta_0])).$$

If there is an orbit $[\gamma]$ of Γ_1 in Int$(\cup_j \beta_j)$ satisfying the orbit dominant condition then we can include in θ a component in $V[\gamma]$ and can ensure that $Dm_{[\gamma]}(\theta) \geqslant 0$ or that the F-between condition holds for γ, that is the Basic bounds of 19.7 hold. ☐

21.12. Proof of 20.7. — We need to produce the dominant area gap or loop of 20.7. First, let $Q = Q(f_0, \Gamma)$ be as in 18.11, and let

$$a = \max\{a(\alpha, x), a(\gamma, x) m_\gamma(x) : \alpha,\ \gamma \text{ are gaps and loops outside } Q\}.$$

If the maximum occurs at a loop γ outside $\overline{Q}$, then we can find $\gamma' \in [[\gamma]]$ such that $a(\gamma'') \leqslant D_1 a(\gamma')$ for all $\gamma'' \in [[\gamma']] \smallsetminus \{\gamma'\}$, $m_{\gamma'}(x) a(\gamma') \geqslant K^{-r} D_1^r$ for some $r \leqslant \#(Y)$ and $K = \frac{1}{2}\log F(x)$. Then γ' is of D_1-dominant area, and of (D_1, D_1') dominant area if ε is sufficiently small given D_1'. The same is true if $m_\gamma(x) a(\gamma) \geqslant D_1^{3\#(Y)} a$ for some loop γ outside $\overline{Q}$. If this does not happen, and $a(\partial Q) \leqslant D_1^{2+\#(Y)} a$, then there is α of D_1-dominant area outside Q with $a(\alpha) \geqslant D_1^{\#(Y)} a$, and $a(\partial Q) \leqslant D_1^2 a(\alpha)$.

So now we assume that $a(\partial Q) \geqslant D_1^{2+\#(Y)} a$. But $m_\gamma(x) a(\gamma) \leqslant D_1^{3\#(Y)} a$ for all γ in the backward orbit of ∂Q. We can assume that no loop of ∂Q is of D_1-dominant area. Then $m_{\partial Q}(x) a(\partial Q, x)$ is small, and a is small. Now let

$$a' = \max\{a(\alpha, x), a(\gamma, x) m_\gamma(x) : \alpha,\ \gamma \text{ are gaps and loops in } Q\}.$$

Then $a' \geqslant 1/2$ is bounded from 0, because a is small but the sum a and a' is bounded below by approximately 1 (the area of $\overline{\mathbf{C}}$). If $a(\gamma) \geqslant a' e^{-D_1 m_\gamma(x)}$ for some loop $\gamma \subset \overline{\Omega}$ then we can find $\gamma' \in [[\gamma]]$ of D_1-orbit dominant area, arguing as before, assuming ε is sufficiently small given D_1. So now assume that $a(\gamma) \leqslant a' e^{-D_1 m_\gamma(x)}$

for all $\gamma \subset \partial\Omega$. If we only have $a(\alpha) \geqslant 1/4\#(Y)$ for α with $m_{\alpha,\partial}(x) \leqslant m_{\rm int}(x)/C_1^6$, then either we can find some gap $\alpha \subset \Omega$ of D_1-dominant area with $a(\alpha)/a(\alpha') \notin [1 - C_1^{-1}, 1 + C_1^{-1}]$ for some $\alpha' \in [\alpha]$, or we can find U of $e^{-C_1/\nu}$-dominant area containing some β_{n-1} (as in 18.11) with $a(U) \geqslant a'/4\#(Y)$. If not, then a' is achieved at a gap $\alpha \subset \Omega$, with $m_{\alpha,\partial}(x) \geqslant m_{\rm int}(x)/C_1^6$ and $a(\partial\alpha) \leqslant e^{-m_{\partial\alpha}(x)/3\#(Y)}$ and $a(\beta) \leqslant e^{-m_{\partial\alpha}(x)/4\#(Y)}a(\alpha)$ for all $\beta \in [[\alpha]] \smallsetminus [\alpha]$. Adding in extra gaps β to α if necessary, but all with $m_{\beta,\partial}(x) \geqslant m_{\rm int}(x)/C_1^5$, we can assume α is of D_1'' dominant area for $D_1'' = e^{-m_{\partial\alpha}(x)/C_1}$. If there is no D_1-orbit dominant area periodic β with $\kappa(\beta)$ nonmaximal, then $a(\beta) \leqslant e^{-\frac{1}{2}m_{\alpha,\beta}(x)}a(\alpha)$ for α with $\kappa(\alpha)$ maximal and we can find an orbit dominant α with $\kappa(\alpha)$ maximal, as required by (1) of 20.7.

21.13. Proof of 19.3: construction of σ_2. — If there is no $\beta \subset \Omega$ with $\kappa(\beta) = \max(\kappa_0(\Gamma), \kappa_0'(\Gamma), \kappa(\Delta_0'))$, let Γ_1 be the set of all loops in $\overline{\Omega}$. . Otherwise, let Γ_1 be the set of all loops $\gamma \subset {\rm Int}(\Omega)$ with $m_\gamma(x'') + m_{\gamma,\partial}(x'') \geqslant m_1(x'')/C_1^2$. Let $q(z)dz^2$ be the quadratic differential for $F(x'') = d(x'', \tau(x''))$ at x'' with stretch $p(z)dz^2$ at $\tau(x'')$. We obtain $\sigma_2(x'')$ by *modifying* the vector field w to a vector field $w_2 = w_{2,x,x''}$ *which depends continuously on x and x''.* We then take $\sigma_2(x'')$ to be the time one map of $w_{2,x,x''}$. We define $w_2(z)$ by adding in an extra term to $w(z)$. Let θ_γ be the vector field of 20.2 with $1 = e$. For $\gamma \subset \partial\Omega$ let $g(t)$ be a positive function which is $t(\log(t))^2$ for $t \leqslant C_1 m_{\gamma_0'}(x))$, and 0 for $t \geqslant 2C_1 m_{\gamma_0'}(x))$. Define $w_2(z)$ to be the sum of $w(z)$ and a multiple of

$$\sum_{\gamma\in\Gamma_1\smallsetminus\partial\Omega} C_1^3 e^{-m_{[\gamma],\partial}(z)}\theta_\gamma(z) + \sum_{\gamma\in\Gamma_1\cap\partial\Omega} C_1 L_\gamma(m_\gamma(z)))\theta_\gamma(z),$$

where the multiple of this term is bounded from 0 only when F-between does not hold. Let $\sigma_{2,t}$ be the flow of w_2. Then

$$DF(w_2(z)) \leqslant \sum_{\gamma\in\Gamma_1\smallsetminus\partial\Omega} C_1^4 a(\gamma)e^{-m_{[\gamma],\partial}(x)} + \sum_{\gamma\in\Gamma_1\cap\partial\Omega} C_1^2 m_{\gamma_0'}(z)a(\gamma) + DF(w(z)),$$

or if F-between holds then $DF(w_2) \leqslant DF(w)$. We claim that this implies

$$DF(w_2(z)) \leqslant \omega(x,\Gamma)/2.$$

This will clearly be true so long as $a(\gamma) \leqslant e^{-m_{\rm int}(x)/C_1^3}$ for $\gamma \in \Gamma_1 \smallsetminus \partial\Omega$ and $a(\gamma) \leqslant e^{-m_\gamma(z)/C_1}$ for $\gamma \in \Gamma_1 \cap \partial\Omega$. If one of these does not hold then there is a larger negative term in $DF(w(z))$. □

CHAPTER 22

PROOF OF DESCENDING POINTS: STRATEGY

22.1. The next three chapters are devoted to proving the Descending Points Theorem of 7.7. The main thrust of the statement of 7.7 is contractibility of components of a set $K_i(\mu, \varepsilon)$ within a set $K_{i+1}(\mu, \varepsilon)$, $i = 0, 1$.

The basic strategy is to study the function $F(x) = d(x, \tau(x))$ for $x \in K_i(\mu, \varepsilon)$. We shall find that this function essentially depends only on $\kappa = \kappa(\mu)$ and on a projection of x. So, roughly speaking, we are able to study a new function $\Phi = \Phi_1$ of (x, κ) which becomes F (essentially) when κ is restricted to a discrete set of values, and a supplementary function Φ_2, which has to be added on for κ bounded from 0, to ensure compactness. In this chapter, we construct the functions Φ_1 and Φ_2, and establish some of their properties.

22.2. Projection to the Domain of Φ. — Let (f_0, Γ) be an invariant loop set with $[f_0, \Gamma]$ minimal, with fixed set P. As in 2.13, let $[\psi_P]$ be the isotopy class of $f_0 \mid P$, identifying the components of ∂P with points. As in 2.10, 2.15, let Δ_0, Δ_0' to be the components of $\overline{\mathbf{C}} \setminus P$, $\overline{\mathbf{C}} \setminus (\cup\Gamma)$ containing v_1, v_2 and Δ_i' $(0 \leqslant i \leqslant n-1)$ the gaps in the orbit of Δ_0'. In 9.1, we defined sets $A(\alpha) \subset Y$ for gaps and loops of Y, and identified $\mathcal{T}(\Gamma, \varepsilon_0)$ with a subset of a product of spaces $\mathcal{T}(A(\alpha))$. The projection of $\mathcal{T}(\Gamma, \varepsilon_0)$ to $\mathcal{T}(A(\alpha))$ was denoted by π_α. In the next three chapters, we shall need to make use of some somewhat different projections.

Let

$$Q = \cup_{i=0}^{n-1} \Delta_i' \cup P.$$

The usual convention on the sets $A(\alpha)$ ensures that each set $A(\Delta_i') \cap A(P)$ consists of two points, and $A(\partial \Delta_i') \subset A(\Delta_i') \cup A(P)$. It is therefore natural to choose $A(Q)$ to be the union of $A(P)$ and all $A(\Delta_i')$, $0 \leqslant i \leqslant n-1$. Now we define, for some $\mathcal{K} \subset \mathcal{T}(\Gamma, \varepsilon_0)$, a projection

$$\rho_P : \mathcal{K} \longrightarrow \mathcal{T}(A(P))$$

which is slightly differently from the projection π_P of 9.1, although it is the same to within $O(e^{-2\pi^2/\varepsilon})$ on $\mathcal{T}(\Gamma,\varepsilon)$, where defined. Without specifying $\mathcal{K}$, it is such that we can normalise so that for all $[\varphi]$ in the domain, $\varphi(Y)$ is bounded, points of $\varphi(A(P))$ are bounded apart. Then $\rho_P([\varphi]) = [\varphi'] \in A(P)$, where

$$\varphi'(\Delta_i \cap A(P)) = (1/\#(A(Q)\cap\Delta_i)) \sum_{y\in A(Q)\cap\Delta_i} \varphi(y),$$
$$\varphi'(y) = \varphi(y) \quad \text{for any other } y \in A(P).$$

(There is at most one point of $A(P)$ which is not in any Δ_i.) This definition is independent of the choice of local coordinates. This type of coordinate is used implicitly in Chapters 20 and 21, from 20.2 onwards.

We also define a projection ρ_i to $\mathcal{T}(A(\Delta_i'))$, which, again, is very slightly different from the projection $\pi_i = \pi_{\Delta_i'}$. We define $\rho_i([\varphi]) = [\varphi_i']$ to be the point close to $\pi_i([\varphi]) = [\varphi_i]$, with $\varphi_i'(y) = \varphi_i(y)$ for $y \in \Delta_i \cap A'(\Delta_i')$, and for the single point $y \in A(\Delta_i') \smallsetminus \Delta_i$, $\varphi_i'(y) = \varphi'(y)$, for φ' as above. Similarly, we can define projections ρ_i' to $\mathcal{T}(A(\Delta_i' \cup P))$ which are close to $\pi_i' = \pi_{\Delta_i'\cup P}$, using φ' on $A(P) \smallsetminus \Delta_i$. Then it is clear that $\prod_i \rho_i'$ is a homeomorphism on $\mathcal{T}(A(Q))$.

We also define a projection ρ_∂ to $\{z : \mathrm{Im}(z) > 0\}$ which, up to scale, is close, but not identical, to $\pi_{\partial\Delta_0}$. As above, write $\rho_P([\varphi]) = [\varphi']$. Let $q_P(z)dz^2$ be the quadratic differential for $d([\varphi'], [\varphi']\cdot[\psi_P])$ at $\rho_P([\varphi_P])$. Normalise $[\varphi]$ so that

$$\mathrm{Res}(q_P, \varphi'(A(P)\cap\Delta_0)) = 1.$$

Then

$$e^{i\rho_\partial([\varphi])} = \varphi(v_2) - \varphi(v_1).$$

Here, we are assuming, as we may do, that v_1, $v_2 \in A(\Delta_0')$. This, of course, only defines $\rho_\partial([\varphi])$ up to addition of an element of $2\pi\mathbf{Z}$, but if we make a choice at one point, we can extend continuously. It would be possible to define projections $\rho_{\partial\Delta_i}$ similarly for all i, if we fixed 2 points of $A(\Delta_i')$.

Let

$$\rho = \rho_0 \times \rho_\partial : \mathcal{K} \longrightarrow \mathcal{T}(A(\Delta_0')) \times \{z : \mathrm{Im}(z) > 0\}.$$

The space $\mathcal{T}(A(\Delta_0'))$ then identifies with the Teichmüller space $\mathcal{T}(B(f_0,\Gamma,\Delta_0'))$ of the branched map space $B(f_0,\Gamma,\Delta_0)$, which is defined in 2.18. Recall from 3.7 that $G(f_0,\Gamma)$ is the subgroup of G which leaves $\mathcal{T}(\Gamma,\varepsilon_0)$ invariant. Then the $G(f_0,\Gamma)$-action on $\mathcal{T}(\Gamma,\varepsilon_0)$ descends through ρ. Because (f,Γ) is minimal, the $G(f_0,\Gamma)$-action descends to an action of $G(B(f_0,\Gamma,\Delta_0'))\times\mathbf{Z}$, where the action of $\mathbf{Z}$ on $\mathcal{T}(A(\partial\Delta_0)) \cong H$ is given by $z \mapsto z + 2\pi n$ $(n \in \mathbf{Z})$. To simplify notation in the following, we define

$$\mathcal{T}_0 = \mathcal{T}(A(\Delta_0')) \times \{z : \mathrm{Im}(z) \geqslant 4\pi^2/\varepsilon_0\},$$

remembering that it identifies with most of $\rho(\mathcal{T}(\Gamma,\varepsilon_0))$. Then $\mathcal{T}_0 \times \mathbf{R}_+$ is the domain of the function Φ than we want to define.

22.3. The basic structure of Φ. — The basic form of $\Phi = \Phi_1$ is as follows. Choose a finite set

$$Y' = Z' \amalg \{v_2\} = A(\Delta_0') \cup A' \subset \overline{\mathbf{C}} \setminus \partial\Delta_0$$

such that $A' \cap A(\Delta_0')$ consists of two points $\{v_1, w_1\}$, with $w_1 \notin \Delta_0$. Then we shall define maps

$$x' : \mathcal{T}_0 \longrightarrow \mathcal{T}(Y'), \quad x'' : \mathcal{T}_0 \times \{t : t > 0\} \longrightarrow \mathcal{T}(Z'),$$

such that for $x' = x'([\varphi], z)$, $x'' = x''([\varphi], z, u)$,

$$\Phi_1([\varphi], z, u) + u = d_{Z'}(x', x'').$$

Moreover, this will be such that

$$\text{if } ([\varphi], z) = \rho(x),\ x \in K_i(\mu, \varepsilon) \text{ and } u = d(\rho_P(x), \rho_P(x) \cdot [\psi_P^n]),$$
$$F(x) - d(\rho_P(x), \rho_P(x) \cdot [\psi_P]) = \Phi_1(\rho(x), u) + o(e^{-2\pi^2/\varepsilon}).$$

So the next task is to define x' and x''. The rough idea is to "paste" $([\varphi], z)$ and its "pullback" into $\mathcal{T}(Y')$ and $\mathcal{T}(Z')$.

22.4. Construction of x', x''. — We use the notation of 22.2-3. Fix a half-geodesic $\{[\xi_t] : t \geqslant 0\}$ in $\mathcal{T}(A')$ with $d([\xi_0], [\xi_t]) = t$. Let $q_t(z)dz^2$ be the corresponding quadratic differential at y_t. We can normalise so that

$$\xi_t(v_1) = 0, \quad \xi_t(w_1) = \infty, \quad \operatorname{Res}(q_t, 0) = 1.$$

Definition of x'. — Now let $[\varphi] \in \mathcal{T}(A(\Delta_0'))$ and $\operatorname{Im}(z) \geqslant m$. Use the normalisation above with $t = 0$, and take $\varphi(w_1) = \infty$. Then $[\varphi'] = x' = x'([\varphi], z, u) = x'([\varphi], z)$ is chosen so that

$$\sum_{a \in \Delta_0 \cap A'(\Delta_0')} \varphi'(a) = 0, \quad \varphi'(v_2) - \varphi'(v_1) = e^{iz},$$
$$\varphi'(a) = \xi_0(a),\ a \in A' \setminus \Delta_0, \quad \rho_0([\varphi']) = [\varphi],$$

where ρ_0 is the projection to $\mathcal{T}(A(\Delta_0'))$ (defined similarly to 22.2). This defines $x'([\varphi], z)$ modulo composition on the right with a Dehn twist round $\partial\Delta_0$. To define $x'([\varphi], z)$ completely, we simply make a choice for some particular $([\varphi_0], z_0)$, and there is then a unique continuous extension so that $x'([\varphi], z)$ is defined and continuous in $([\varphi], z)$. This definition is such that $x'([\varphi], z + 2\pi n)$ is the composition of $x'([\varphi], z)$ with an integral Dehn twist round $\partial\Delta_0$, for all $n \in \mathbf{Z}$. We shall sometimes regard $x'([\varphi], z) = [\varphi']$ as a homeomorphism up to isotopy, that is, the precise scaling of φ' given above will sometimes be important. Note also that, if $P = \overline{\mathbf{C}} \setminus \Delta_0$ and ρ_P, ρ_∂ are defined similarly to 22.2,

$$\rho_\partial([\varphi']) = z, \quad \rho_P([\varphi']) = [\xi_0].$$

Our definition of $x''([\varphi], z, u) = [\varphi'']$ will be in terms of x' (and u). If we change our definition of x' by composition with a Dehn twist round $\partial\Delta_0$, then the definition of x'' will change by composition with the same Dehn twist.

First Preliminary to defining x'': restrictions of homeomorphisms to discs. — Suppose that $x' = [\varphi'] = [\psi']$, and consider φ', ψ' restricted to each of the discs Δ_0, $\overline{\mathbf{C}} \setminus \Delta_0$. We call these restrictions φ'_1, ψ'_1 (for Δ_0) and φ'_2, ψ'_2. Then there is a homeomorphism σ which is the identity on a neighbourhood of $\varphi'(Y')$ such that $\sigma \circ \varphi'_1 = \psi'_1$ on $\partial\Delta_0$ and such that $\sigma \circ \varphi'_1$, ψ'_1 are isotopic via an isotopy which is constant on $\partial\Delta_0$, and similarly for $\sigma \circ \varphi'_2$, ψ'_2. We fix such φ'_1, φ'_2. We can also assume (without changing $[\varphi'] = [\psi']$) that $\varphi'_1(\partial\Delta_0) = \varphi'_2(\partial\Delta_0)$ is a union of one stable segment and one unstable segment from the foliations for $q_0(z)dz^2$, close, but not too close, to $\varphi'(A(\Delta'_0) \cap \Delta_0)$. In order to describe $x'' = [\varphi'']$ completely, it suffices to describe it up to isotopy restricted to each of the discs Δ_0, $\overline{\mathbf{C}} \setminus \Delta_0$ modulo isotopies which are constant on $\partial\Delta_0$.

Second preliminary: the branched covering g_0. — First, we note that the reduced branched covering g_0 for Δ'_0 is defined up to isotopy constant on $\partial\Delta_0$. To do this we need to choose the isotopy class of $f_0^p \mid P$ suitably, where p is the period of Δ'_0, and assuming without loss of generality that $f_0^p(P) = P$. Let ψ be obtained from a pseudo-Anosov on $\overline{\mathbf{C}} \setminus A(P)$ by blowing up the points of $A(P)$ to discs, such that the blowups of singular leaves ending at points of $A(P)$ have endpoints on the components of ∂P. The point is that there is a unique way to choose $f_0^p \mid P$ so that ∂P is fixed pointwise by f_0^p *modulo an isotopy constant on ∂P*, so that f_0^p is isotopic via an isotopy which is constant on ∂P to ψ.

Third preliminary: the holomorphic branched covering s. — Now let s be a holomorphic branched covering with critical values at $\varphi'(v_1)$, $\varphi'(v_2)$ of the form

$$s(z) = \frac{(z - z_0)^2}{z + a} + a + 2z_0$$

for z_0, a. The critical points are z_0 and $-z_0 - 2a$. The critical values are $a + 2z_0$ and $-3a - 2z_0$. Then for z bounded from 0,

$$s(z) = (z - z_0)(1 + (a + z_0)(z - z_0)^{-1})^{-1} + a + 2z_0 = z + O(e^{-2m})$$

Then for z bounded from 0, the branch of s^{-1} fixing ∞ also has an expansion

$$s^{-1}(z) = z + O(e^{-2m}).$$

Then $s^{-1} \circ \varphi'_1 \circ g_0$ maps $\partial\Delta_0$ approximately to the union of a stable and unstable segment of the foliations for $q_u(z)dz^2$ (because of the normalisation).

Definition of x''. — Now we are ready to define $x'' = [\varphi''] \in \mathcal{T}(\partial\Delta_0, 2\varepsilon) \subset \mathcal{T}(Z')$. In fact, the definition will describe $[\varphi'']$ as an element of $\mathcal{T}(Z'')$ where

$$Z'' = g_0^{-1}(Y' \cap \Delta_0) \cup (A' \setminus \Delta_0).$$

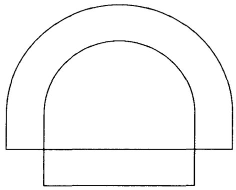
We are trying to abstract the pullback construction, and φ'' will be a lift of $\varphi' \circ g_0$ under s. Precisely we choose the isotopy class of $\varphi'' \mid \Delta_0$ close to that of $s^{-1} \circ \varphi_1' \circ g_0$, mapping $\partial \Delta_0$ exactly to the union of stable and unstable segments of $q_u(z)dz^2$, and so that

$$s \circ \varphi''(a) = \varphi' \circ g_0(a), \quad a \in A(\Delta_0') \cap \Delta_0,\ a \neq v_2.$$

Now we define $\varphi'' \mid \overline{\overline{\mathbf{C}} \setminus \Delta_0}$ up to isotopy. We take

$$\varphi''(a) = \xi_u(a), \quad a \in A' \setminus \Delta_0.$$

To complete the description, let χ_u be the map minimizing distortion given by $q_0(z)dz^2$, $q_u(z)dz^2$. Then $\chi_u(\varphi_1'(\partial\Delta_0)) \neq \varphi''(\partial\Delta_0)$, but is again a union of stable and unstable segment for $q_u(z)dz^2$, as shown:



Natural isotopy.

We then compose χ_u on the left by a map which sends the second stable and unstable segments to the first, pushing along unstable, stable leaves respectively, and this composition completes the description of $\varphi'' \mid \overline{\overline{\mathbf{C}} \setminus \Delta_0}$, as required.

22.5. Scalar Multiplication and Basic Properties of x', x'', Φ

The Lower Bound Property. — An obvious property of Φ — proved like 17.4 — is that, for a constant $C > 0$

$$\Phi([\varphi], z) \geqslant -Ce^{-\operatorname{Im}(z)}.$$

Equivariance. — The maps x', x'' are homomorphisms with respect to the natural $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-actions on $\mathcal{T}_0$, $\mathcal{T}(Y')$, $\mathcal{T}(Z')$, and Φ is invariant with respect to the natural action on $\mathcal{T}_0$.

Scalar multiplication. — Let $\varnothing \neq \Gamma \subset \overline{\mathbf{C}} \setminus Y$ be a set of simple disjoint loops such that all loops of Γ are in the boundary of a single component P of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$, and let Δ_i, $0 \leqslant i \leqslant p-1$, $p \geqslant 1$, be the other (disc) components of $\overline{\mathbf{C}} \setminus (\cup\Gamma)$. We are thinking of $Y = A(Q)$ for Q as in 22.2, or $Y = Y'$ as in 22.3-4. Fix sets $A(P)$, $A(\Delta_i)$. Normalise so that points of $\varphi(A(P))$ are bounded and the diameter of $\varphi(A(P))$ is bounded from 0. Let $[\varphi_0] \in \mathcal{T}(A(P))$ be close, or equal, to $\rho_P([\varphi])$. Define

$$\varphi_0(y) = \varphi_0(y') \quad \text{if } y,\ y' \in \Delta_i.$$

We can choose a representative φ_0 of $[\varphi_0]$ so that $(\varphi_0(y))$ and $(\varphi(y))$, as elements of $\overline{\mathbf{C}}^Y$, are close. Then the element

$$[\varphi_0] + \lambda([\varphi] - [\varphi_0]) = [\varphi_0 + \lambda(\varphi - \varphi_0)] = [\varphi_\lambda] \in \mathcal{T}(Y)$$

makes sense, where $[\varphi_\lambda]$ is defined by

$$\varphi_\lambda(y) = \varphi_0(y) + \lambda(\varphi(y) - \varphi_0(y)) \ (y \in Y),$$

and the isotopy class is the natural one, that is, $[\varphi_1] = [\varphi]$ and $\lambda' \mapsto \varphi_{\lambda'}$ $(\lambda' \in [\lambda, 1])$ is an isotopy from φ_λ to φ_1. This is reminiscent of the use of local coordinates to define isotopy classes $[\varphi + \underline{h}]$ in 8.4 and later. Such a coordinate is also used implicitly in Chapters 20 and 21, from 20.2 onwards.

Scalar multiplication Properties. — Then for all all $m' \geqslant 0$, the definitions of x' and x'' are such that

$$x'([\varphi], z + im') = [\xi_0] + e^{-m'}(x'([\varphi], z) - [\xi_0]),$$
$$x''([\varphi], z + im', u) = [\xi_u] + e^{-m'}(x''([\varphi], z, u) - [\xi_u]).$$

If $[\varphi_1]$, $[\varphi_2] \in \mathcal{T}(\Gamma, \varepsilon_0)$ with

$$\pi_Q([\varphi_2]) = \rho_P([\varphi_1]) + e^{-m'}(\pi_Q([\varphi_1] - \rho_P([\varphi_1])),$$

then

$$x'(\rho([\varphi_2])) = [\xi_0] + e^{-m'}(x'(\rho([\varphi_1]) - [\xi_0]),$$

and if $u = d(\rho_P([\varphi_1]), \rho_P([\varphi_1]) \cdot [\psi_P^n]))$,

$$x''(\rho([\varphi_2]), u) = [\xi_u] + e^{-m'}(x''(\rho([\varphi_1])) - [\xi_u]).$$

Our next task is to transfer these scalar multiplication properties to Φ, and also to relate Φ and F.

22.6. Formula for $D\Phi_1$.

— This is an analogue of the Pullback Derivative Formula of 8.11. In some respects this is an easier result, because we are only considering a restricted class of pullbacks. But it is different from 8.11, and it is important to get the details right.

Lemma. — *Let $q(z)dz^2$ and $p(z)dz^2$ be the quadratic differentials for $d(x', x'')$ at x', x'' for $x' = [\varphi'] = x'([\varphi], z, u)$, $x'' = [\varphi''] = x''([\varphi], z, u)$. Let $\mathrm{Im}(z) = m$. Then for $\underline{h}$ with $h(v_j) = w_0 + (e^{i\nu} - 1)\varphi'(v_j)$, $\nu \in \mathbf{R}$ and*

$$\sum_{a \in \Delta_0 \cap Y'} h(a) = 0,$$

we have

$$D\Phi_1([\varphi],z)+\underline{h},u)=\Phi_1([\varphi],z,u)+\sum_{a\in Y'\cap\Delta_0}2\pi\operatorname{Re}(h(a)\operatorname{Res}(q-s_*p,\varphi'(a)))$$
$$+\int_{\varphi'(\partial\Delta_0)}(\nu z-iw_0)(s_*p-p)+O(\|\underline{h}\|^2).$$

Remark. — The integral term is relatively small, because if we choose the contour of integration bounded from 0 then by the estimates on s in 22.4, $s_*p-p=O(e^{-2m})$ there.

Proof. — The functions x' and x'' are clearly holomorphic in $([\varphi],z)$, from their definitions in 22.4. Since the Teichmüller distance function is C^2 by Theorem 12.1, so is Φ_1. Let s be the holomorphic branched covering used to define $x''([\varphi],z,u)=[\varphi'']$ from $x'([\varphi],z)=[\varphi']$. Let g_0 be the branched covering fixing Δ_0 also used to define x'', with critical points c_1, c_2 and critical values v_1, v_2. Let s_1 be the holomorphic branched covering used to define $x''([\varphi+\underline{h}],m,u)$ from $x'([\varphi+\underline{h}],m,u)$ with $\underline{h}=o(e^{-m})$. Then s_1 is of the form

$$s_1(z)=e^{i\nu}s(e^{-i\nu}(w-w_0))+w_0.$$

Then $x'([\varphi+\underline{h}],z,u)=[\varphi'+\underline{h}]$, $x''([\xi+\underline{h}],z,u)=[\varphi''+\underline{h}']$, where

$$h'(c_1)=w_0+(i\nu+O(\nu^2))\varphi''(c_1),$$
$$h'(a)=w_0+i\nu\varphi''(a)-i\nu(s^{-1})'(\varphi'(g_0(a)))\varphi'(g_0(a))$$
$$+(s^{-1})'(\varphi'(g_0(a))(h(g_0(a))-w_0-i\nu\varphi'(a))+O(\|\underline{h}\|^2)$$

for $a\in\Delta_0\cap Y'$, $a\neq c_1$, v_2. Now by 8.4,

$$\Phi_1([\varphi+h],z,u)=\Phi_1([\varphi],z,u)$$
$$+\sum_{a\in Y'\cap\Delta_0}2\pi\operatorname{Re}(h(a)\operatorname{Res}(q,\varphi'(a))-h'(a)\operatorname{Res}(p,\varphi''(a)))+O(\|\underline{h}\|^2).$$

At this stage there is no contribution from $a=v_2$. We now reinterpret the terms in the second sum using the formulae for $h'(a)$:

$$\sum_{a\in\Delta_0\cap Y'}i\nu\varphi''(a)\operatorname{Res}(p,\varphi''(a))=\frac{1}{2\pi i}\int_{\varphi''(\partial\Delta_0)}i\nu z(p(z)-s_*p(z))dz$$
$$+\sum_{a\in\Delta_0\cap Y'}i\nu\varphi'(a)\operatorname{Res}(s_*p,\varphi'(a)),$$

$$w_0\sum_{a\in\Delta_0\cap Y'}\operatorname{Res}(p,\varphi''(a))-w_0\sum_{\substack{a\in Y'\cap\Delta_0\\ a\neq c_1,v_2}}(s^{-1})'(\varphi'(g_0(a))\operatorname{Res}(p,\varphi''(a))$$
$$=\sum_{j=1}^{2}w_0\operatorname{Res}(s_*p,v_j)+\frac{w_0}{2\pi i}\int_{\varphi'(\partial\Delta_0)}(p-s_*p)(z)dz,$$

which gives the result. □

22.7. Some Corollaries of the Hard Same Shape Theorem. — (Easy) Same Shape 9.5 and the Hard Same Shape results 16.2 and 16.9 yield closeness of residues of a number of quadratic differentials, as we now explain. These will be needed to relate F and Φ. Fix $x = [\varphi_1] \in K_i$, $x \in \mathcal{T}_{\geqslant \nu}$, $\nu \geqslant \varepsilon_0/L_2$, with corresponding

$$\tau(x) = [\varphi_2], \quad u = d(\rho_P(x), \rho_P(x) \cdot [\psi_P^n]), \quad x'(\pi(x), u) = [\varphi'], \quad x''(\pi(x), u) = [\varphi''].$$

Let $q(z)dz^2$, $p(z)dz^2$ be as in 22.6. Let $q_Q(z)dz^2$ be the quadratic differential for $d(\pi_Q(x), \pi_Q \circ \tau(x))$ at $\pi_Q(x)$, with stretch $p_Q(z)dz^2$ at $\pi_Q \circ \tau(x)$. Let π_j denote projection of $\mathcal{T}(Y)$ to $\mathcal{T}(A(P \cup \Delta_j'))$. Let $r_j(z)dz^2$ denote the quadratic differential for $d(\pi_j(x), \pi_j \circ \tau(x))$ at $\pi_j(x)$, with stretch $t_j(z)dz^2$ at $\pi_j \circ \tau(x)$. We assume that $A(\Delta_j' \cup P) \subset f_0^{-1}(A(\Delta_{j+1}' \cup P))$ for $0 \leqslant j \leqslant p-2$. This gives a natural projection from $\mathcal{T}(A(\Delta_{j+1}' \cup P))$ to $\mathcal{T}(A(\Delta_j' \cup P))$. So we can consider elements of $\mathcal{T}(A(\Delta_j' \cup P))$ as elements of $\mathcal{T}(A(\Delta_0' \cup P)$, for all $0 \leqslant j \leqslant p-1$. Using this identification, let $r(z)dz^2$ denote the quadratic differential at $\pi_0(x)$ for $d(\pi_0(x), \pi_{p-1} \circ \tau(x))$.

By Hard Same Shape 16.2, for all $y \in A(\Delta_j') \cap \Delta_j$,

$$\mathrm{Res}(r_j, \varphi_1(y) = \mathrm{Res}(q_Q, \varphi_1(y)) + O(e^{-2\pi^2/\varepsilon C_1}),$$
$$\mathrm{Res}(t_j, \varphi_2(y)) = \mathrm{Res}(p_Q, \varphi_2(y)) + O(e^{-2\pi^2/\varepsilon C_1}).$$

By the Triangle Conditions 18.11 (since $x \in K_i$) and Easy Same Shape relating $q_Q(z)dz^2$, $p_Q(z)dz^2$ to the quadratic differentials for $d(x, \tau(x))$,

$$\mathrm{Res}(q_Q - s_* p_Q, \varphi_1(y)) = O(e^{-C_1/\nu}), \quad y \in A(\Delta_j') \cap \Delta_j,\ 1 \leqslant j \leqslant p-1$$

(There is no term $a(\Delta_j')$, as there was in 18.11, because we have changed the normalisation.) This gives

$$\mathrm{Res}(r_j - s_* t_{j-1}, \varphi_1(y)) = o(1), \quad y \in A(\Delta_j') \cap \Delta_j,\ 1 \leqslant j \leqslant p-1.$$

Using the identification of $\mathcal{T}(A(\Delta_j' \cup P)$ with $\mathcal{T}(A(\Delta_0' \cup P))$, and $s(z) = z(1 + o(1))$ on $S(\Delta_j', \nu)$ this gives

$$\mathrm{Res}(r_j - t_{j-1}, \varphi_1(y)) = o(1), \quad y \in A(\Delta_0') \cap \Delta_0,\ 1 \leqslant j \leqslant p-1.$$

Then 16.9 gives that

$$\mathrm{Res}(r_0, \varphi(y)) = \mathrm{Res}(r, \varphi(y)) + o(1), \quad y \in A(\Delta_0') \cap \Delta_0,$$
$$\mathrm{Res}(t_{p-1}, \xi(y)) = \mathrm{Res}(t, \xi(y)) + O(e^{-2\pi^2/\varepsilon C_1}), \quad y \in A(\Delta_0') \cap \Delta_0.$$

But by Hard Same Shape 16.2, for all $y \in A(\Delta_0') \cap \Delta_0$,

$$\mathrm{Res}(r, \varphi_1(y)) = \mathrm{Res}(q, \varphi'(y)) + O(e^{-2\pi^2/\varepsilon C_1}),$$
$$\mathrm{Res}(t, \varphi_2((y)) = \mathrm{Res}(p, \varphi''(y)) + O(e^{-2\pi^2/\varepsilon C_1}).$$

The important deduction from this, for us, is that

$$\mathrm{Res}(q_Q, \varphi_1(y)) = \mathrm{Res}(q, \varphi'(y)) + o(1),$$
$$\mathrm{Res}(s_* p_Q, \varphi_1(y)) = \mathrm{Res}(s_* p, \varphi'(y)) + o(1).$$

Note that the same holomorphic branched covering s is indeed used to define $\tau(x)$ and x''.

22.8. Formulae for F and Φ. — We use the notation of 22.7. We also write $A'(\Delta_i') = A(\Delta_i') \cap \Delta_i$.

Lemma. — *Let $m = 2\pi^2/\varepsilon = \mathrm{Im}(z)$. Let x satisfy all the conditions for being in $K_i(\mu,\varepsilon)$, apart from the bound on $F(x)$. Let $x' = x'([\varphi],z)$, $x'' = x''([\varphi],z,u)$. Then, for a suitable $C>0$,*

$$(1)\quad F(x) - d(\rho_P(x), \rho_P(x)\cdot[\psi_P]) = \sum_{a\in A'(\Delta_0')} 2\pi\,\mathrm{Re}(\varphi_1(a)\,\mathrm{Res}(q_Q - s_*p_Q, \varphi_1(a))) + O(e^{-m(1+1/C)}),$$

$$(2)\quad d(x',x'') - u = \Phi_1([\varphi],z,u) = \sum_{a\in A'(\Delta_0')} 2\pi\,\mathrm{Re}(\varphi'(a)\,\mathrm{Res}(q - s_*p, \varphi'(a))) + o(e^{-m}).$$

If in addition $([\varphi],z) = \rho(x)$ and $u = d([\varphi_P],[\varphi_P\circ\psi_P^p])$, then the two sums are equal to within $o(e^{-m})$.

Proof. — First we consider (1). Write

$$F_Q(x) = d_Q(x,\tau(x)).$$

By (easy) Same Shape 9.5, for $x\in K_i(\mu,\varepsilon)$, for a suitable $C>0$,

$$F(x) = F_Q(x) + O(e^{-m(1+1/C)}).$$

Let s_λ be the holomorphic branched covering used to define $\tau([\varphi_\lambda])$, and write $s = s_1$. For $0<\lambda\leqslant 1$, by the Derivative Formula 8.11,

$$(d/d\lambda)(F([\varphi_\lambda])) = \sum_{i=0}^{p-1}\sum_{a\in A'(\Delta_i')} 2\pi\,\mathrm{Re}((\varphi(a)-\varphi_0(a))\,\mathrm{Res}(q_{Q,\lambda} - (s_\lambda)_*p_{Q,\lambda}, \varphi_\lambda(a))).$$

Here, if $b = \varphi_0(a)$ for $a\in A'(\Delta_{p-1}')$,

$$s_\lambda(z) = \lambda s(z - b/\lambda).$$

So $s_\lambda'(\varphi_\lambda(a)) = s'(\varphi(a))$ for $a\in A'(\Delta_{p-1}')$. For $i\neq p-1$, we have seen in 22.4 that

$$s_\lambda'(\varphi_\lambda(a)) = 1 + O(e^{-2m}),\quad a\in A'(\Delta_i').$$

By Hard Same Shape 16.2, for a suitable $C>0$,

$$\mathrm{Res}(q_{Q,\lambda},\varphi_\lambda(a)) = \mathrm{Res}(q_Q,\varphi(a)) + O(e^{-m(1+1/C)}),\ a\in\cup_i A'(\Delta_i'),$$

and similarly for $p_{Q,\lambda}$. Each of these results is actually obtained by 2 (or 3) applications of 16.2: one applies 16.2 to show that $q_{Q,\lambda}$ is close on $S(\Delta_i',[\varphi_\lambda],\varepsilon_0)$ to the quadratic differential for $d(\pi_i'([\varphi_\lambda]), \pi_i' \circ \tau([\varphi_\lambda]))$, and then considers the latter for varying λ. So we have

$$(d/d\lambda)(F([\varphi_\lambda]) = \sum_{i=0}^{p-1} \sum_{a \in A'(\Delta_i')} 2\pi \operatorname{Re}((\varphi(a) - \varphi_0(a)) \operatorname{Res}(q_Q - s_* p_Q, \varphi(a)) + O(e^{-m(1+1/C)}).$$

Also, for $i \neq 0$, $a \in \Delta_i$, we have seen that $\operatorname{Res}(q_Q, \varphi(a))$ and $\operatorname{Res}(s_* p_Q, \varphi(a))$ are close to $\operatorname{Res}(r_i, \varphi(a))$ and $\operatorname{Res}(t_{i-1}, \varphi(a))$ respectively and hence by the Triangle Conditions

$$\operatorname{Res}(q_Q - s_* p_Q, \varphi(a)) = O(e^{-m/C}).$$

Moreover, $\varphi_0(a) = 0$ for $a \in A'(\Delta_0')$. So altogether, integrating up, we have (1).

(2) is proved exactly similarly, using 22.6, but is somewhat simpler because there is only one disc Δ_0 to consider. The equality to within $o(e^{-m})$ of (1) and (2) in the stated circumstances follows from the closeness of quadratic differentials established in 22.7. □

22.9. Scalar multiplication of Φ. — Similar techniques to those in 22.8 establish the following.

Lemma. — *The following holds for a suitable constant C. If $\operatorname{Im}(z) = m$, $m' \geqslant 0$ and D_1 denotes derivative with respect to the first two coordinates,*

$$D_1\Phi([\varphi], z + m, u) = D_1\Phi([\varphi], z, u) + O(e^{-m/C}),$$
$$\Phi([\varphi], z + m', u) = e^{-m'}(\Phi([\varphi], z, u) + O(e^{-m(1+1/C)})).$$

Proof. — Let $q(z)dz^2$, $p(z)dz^2$ be as above and let $q'(z)dz^2$ $p'(z)dz^2$ be the quadratic differentials for $\Phi_1([\varphi], z + m')$ at $x'([\varphi], z + m')$ and $x''([\varphi], z + m' + u)$. Then the residues of q and q', and of p and p', are within $O(e^{-m/C})$ by 16.2. Then we apply (2) of 22.8. □

22.10. How to compactify: the function Φ_2. — One problem with the functions F, $\Phi = \Phi_1$, is that for u, $\kappa(\mu)$ bounded from 0, $\eta > 0$, quotients by the natural $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action of the sets

$$\{x \in \mathcal{T}(\Gamma, \varepsilon) : F(x) \leqslant \kappa(\mu) - \eta e^{-2\pi^2/\varepsilon}\},$$
$$\{([\varphi], z, u) : m_1 \leqslant \operatorname{Im} z \leqslant m_2 : \Phi([\varphi], z, u) \leqslant -\eta e^{-\operatorname{Im}(z)}\}$$

need not be compact, although, as we shall see, compactness does hold if $u > 0$ is sufficiently small. It is therefore necessary to introduce a function

$$\Phi_2 : \mathcal{T}_0 \longrightarrow \{t : t \geqslant 0\},$$

which we now do, for M suitably large yet to be chosen. Let $([\varphi], z) \in \mathcal{T}_0$. We take $\Gamma_1 = \Gamma_1([\varphi])$ to be the set of (disjoint) loops in $\Delta_0 \setminus Y'$ such that $[\varphi] \in \mathcal{T}(\gamma, \varepsilon_0)$ if and only if $\gamma \in \Gamma_1$. We include a loop γ_0' in $\operatorname{Int}(\Delta_0)$ which is homotopic to $\partial\Delta_0$ in

$\overline{\mathbf{C}} \setminus Z'$ but not in $\overline{\mathbf{C}} \setminus Y'$, if such a γ_0' exists with $[\varphi] \in \mathcal{T}(\gamma_0', \varepsilon_0)$. We do not include $\partial\Delta_0$ itself. Let $m_\gamma([\varphi])$ be defined as usual (9.1 onwards) using a set $A(\gamma)$ which is invariant under the $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action. This can be done, for example, by choosing the points of $A(\gamma)$ to be periodic under g_0. Let

$$\Phi_{2,\gamma}([\varphi], z) = e^{-m}\xi_1(m_\gamma([\varphi])),$$

where

$$\xi_1 : \mathbf{R} \longrightarrow [0, \infty),\ \xi_1 = 0 \text{ on } (-\infty, M],\ \xi_1(t) = t - M \text{ for } t \in [M+1, \infty),\ M > 1/\varepsilon_0,$$

and ξ_1 is C^1 with derivative bounded independently of M. Let

$$\Phi_2([\varphi], z, u) = \sum_{\gamma \in \Gamma_1} \Phi_{2,\gamma}([\varphi], z).$$

Then Φ_2 is continuous.

22.11. Properties of Φ_2. — The following are immediate from the definition.

Scalar Multiplication

$$\Phi_2([\varphi], z + im') = e^{-m'}\Phi_2([\varphi], z).$$

Invariance. — Φ_2 is invariant with respect to the $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action.

Compactness. — For any $C' > 0$, $M' > 0$ the quotient by the $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action on

$$\{([\varphi], z) : 0 \leqslant \operatorname{Im}(z) \leqslant M,\ \Phi_2(z) \leqslant C' e^{-\operatorname{Im}(z)}\}$$

is compact.

CHAPTER 23

PROOF OF DESCENDING POINTS: REDUCTIONS

23.1. In this chapter we reduce the proof of Descending Points. Two reductions are given in 23.2 and 23.4, the second being closely related to the first. The reductions are basically to results about sets

$$\{([\varphi], z) : (\Phi_1 + \delta\Phi_2)([\varphi], z, u) \leqslant 0\}$$

for varying u. The No Boundary Critical Points result of 23.5 then says that the sets for different u (and δ, with certain restrictions) are diffeomorphic, as is explained in 23.6. We then state some results about the structure of the set for u near 0 in 23.8. All these results will be proved in the next chapter, and complete the proof of Descending Points. The chapter ends with a construction of a pullback map for the function $\Phi = \Phi_1$, which is then used to give estimates on Φ in the thin part of $\mathcal{T}_0$. The estimates are similar to estimates on F obtained in Chapter 17, and proved in a similar way.

We use the notation of Chapter 22. In particular, given $([\varphi], z, u)$ we let $x' = x'([\varphi], z)$ and $x'' = x''([\varphi], z, u)$. We let $q(z)dz^2$ be the quadratic differential for $d(x', x'')$ at x', with stretch $p(z)dz^2$ at x''.

23.2. First Reduction in Descending Points. — We shall see that the proof of Descending Points reduces to the following, provided the constants used to define the sets K_i are suitably chosen.

First Reduction in Descending Points. — *The following holds for a suitable constant C_0. Let $u_0 \geqslant u > 0$, $\eta > 0$ be given. Let $\delta > 0$ be sufficiently small given u_0, or $\delta \geqslant 0$ sufficiently small if u_0 is sufficiently small. Then for $\eta' > 0$ sufficiently small given u_0 and δ, and all sufficiently large $m > 0$, any component of the set*

$$\{([\varphi], z) : \mathrm{Im}(z) \geqslant m,\ (\Phi_1 + \delta\Phi_2)([\varphi], z, u) \leqslant -\eta e^{-m}\}$$

is contractible within

$$\{([\varphi], z) : \mathrm{Im}(z) \geqslant m,\ (\Phi_1 + \delta\Phi_2)([\varphi], z, u) \leqslant -\eta' e^{-m}\}.$$

The stabilizer of any component in $G(B(f_0, \Gamma, \Delta_0'))$ is trivial. There is a natural one-to-one correspondence between components and lifts to $\mathcal{T}(B(f_0, \Gamma, \Delta_0')) = \mathcal{T}(A(\Delta_0'))$ of polynomials in $B(f_0, \Gamma, \Delta_0')$.

23.3. Proof of Sufficiency of the First Reduction. — We need to contract $K_i(\mu, \varepsilon) \subset \mathcal{T}$ G-equivariantly within $K_{i+1}(\mu, \varepsilon)$. So we shall construct

$$h : K_i(\mu, \varepsilon) \times [0, 1] \longrightarrow K_{i+1}(\mu, \varepsilon)$$

such that $h(x, 0) = x$, $h(K_{i+1}(\mu, \varepsilon) \times \{1\})$ is a topological line invariant under $x' \mapsto x' \cdot [\psi_P]$. We construct h by constructing $\rho_i'(h(x, t)) = x_{i,t}'$ $(0 \leqslant i \leqslant n-1)$ and $\pi_\alpha(h(x, t))$ for all gaps and loops $\alpha \not\subset Q$. As a preliminary to constructing the $\rho_i'(h(x, t))$ we construct $\rho_P(h(x, t)) = x_{P,t}$.

We choose $x_{P,t}$ so that:

(1) $(x \cdot [\psi_P])_{P,t} = x_{P,t} \cdot [\psi_P]$;

(2) $t \mapsto d(x_{P,t}, x_{P,t} \cdot [\psi_P])$ is decreasing in t for each $x \in K_i(\mu, \varepsilon)$;

(3) for $t \geqslant \frac{1}{2}$, $x_{P,t}$ is constant in t, and $\{x_{P,t} : x \in K_i(\mu, \varepsilon)\}$ is the unique geodesic on which $y \mapsto d(y, y \cdot [\psi])$ and $y \mapsto d(y, y \cdot [\psi^p])$ take the minimum values $\kappa(\mu)$ and $p\kappa(\mu)$ respectively.

By 16.11 we can take the paths $\{x_{P,t} : t \in [0, 1]\}$ of small diameter independent of x and ε. (The best estimate I can achieve is $O(e^{-2\pi^2/C\varepsilon})$ for some $C > 0$.) For $\alpha \not\subset Q$ we shall take $\pi_\alpha(h(x, t))$ such that

$$d(\pi_\alpha(h(x, t)), \pi_\alpha \circ \tau(h(x, t))) \leqslant F(h(x, t)) - E_i^{-1}$$

and $\pi_\alpha(h(K_i(\mu, \varepsilon)) \times \{1\})$ is a point. This, of course, depends on the definition of $x_{Q,t}$, but can clearly be done.

As a further preliminary to defining the $x_{i,t}'$, we define $\rho(h(x, t)) = x_{0,t}$. We shall write $\rho_\partial(h(x, t)) = x_{\partial,t}$. By the definition of K_i, 22.2 and 22.8, we have, for $x \in K_i(\mu, \varepsilon)$ and $u = d(\rho_P(x), \rho_P(x \cdot [\psi_P^n]))$,

$$\operatorname{Im}(\rho_\partial(x)) \geqslant \frac{2\pi^2}{\varepsilon} - 4\pi^2 E_i, \quad \Phi_1(\rho(x), u) \leqslant -(2E_i)^{-1} e^{-2\pi^2/\varepsilon},$$

and by the definition of Φ_2, for a suitable $\delta_i > 0$ given C_i', D_i,

$$\delta_i \Phi_2(\pi(x)) \leqslant (4E_i)^{-1} e^{-2\pi^2/\varepsilon}.$$

So then

$$(\Phi_1 + \delta_i \Phi_2)(\pi(x), u) \leqslant -(4E_i)^{-1} e^{-2\pi^2/\varepsilon}.$$

Then by 23.2, for suitable choice of E_{i+1}, we can choose $x_{0,t}$ so that $x_{0,t} = x$ for $t \leqslant \frac{1}{2}$, $\rho(\{x_{0,1} : x \in K_i(\mu, \varepsilon)\})$ is a point, and for all $x \in K_i(\mu, \varepsilon)$, $t \geqslant \frac{1}{2}$,

$$\operatorname{Im}(x_{\partial,t}) \geqslant \frac{2\pi^2}{\varepsilon} - 4\pi^2 E_i, \quad (\Phi_1 + \delta_i \Phi_2)(x_{0,t}, p\kappa(\mu)) \leqslant -2E_{i+1}^{-1} e^{-2\pi^2/\varepsilon}.$$

Then we construct $x_{0,t}'$ from $x_{0,t}$ and $x_{P,t}$ in the same way as $x'([\varphi], z)$ is constructed from $([\varphi], z)$ and $[\xi_0]$, but using $A(P \cup \Delta_0')$ instead of Y'. Construct $x_{0,t}''$ in the same way as x'', but using $x_{0,t}$ and $d(x_{P,t}, x_{P,t}\cdot[\psi_P^n])$. As noted before, $x''_{0,t}$ can be regarded as an element of $\mathcal{T}(Z'')$ where Z'' is somewhat larger than $A(\Delta'_0\cup P)$. In fact, using f_0^{-i}, $\mathcal{T}(A(\Delta'_i\cup P))$ can be regarded as a factor of $\mathcal{T}(Z'')$. For $t\leqslant\frac{1}{2}$, $x''_{0,t}$ is very close to $\rho'_{p-1}(\tau(x))$. For $t\leqslant\frac{1}{2}$ we take $x_{i,t}$ very close to $\rho'_i(x)$. For $t\geqslant\frac{1}{2}$ we take $x'_{i,t}$ close to the geodesic joining $x_{0,t}$ and $x''_{0,t}$ and equally spaced. Then $h(x,t)$ has been completely determined and $\pi'_{p-1}(\tau(x))$ is close to $x''_{0,t}$, $\pi'_i(h(x,t))$ is close to $x'_{i,t}$. So the π' triangle conditions of 18.11 for K_{i+1} hold for $h(x,t)$ for all t. Write

$$u_t=d(x_{P,t},x_{P,t}\cdot[\psi_P^n]).$$

Let $q_t(z)dz^2$ be the quadratic differential for $d(x'_t,x''_t)$ at x'_t with stretch $p_t(z)dz^2$ at x''_t. By 9.5,

$$F(h(x,t))=d(\pi_Q(h(x,t)),\pi_Q\circ\tau(h(x,t)))+o(e^{-m})$$

and by 22.8

$$\begin{aligned}F(h(x,t))-d(x_{P,t},x_{P,t}\cdot[\psi_P]) &= \sum_{a\in A'(\Delta_0)}\operatorname{Res}(q_t,s_*p_t,\varphi_t(a))+o(e^{-2\pi^2/\varepsilon})\\ &= d(x'_t,x''_t)-u_t+o(e^{-2\pi^2/\varepsilon}).\end{aligned}$$

So for $t\leqslant\frac{1}{2}$,

$$F(h(x,t))\leqslant F(x)+o(e^{-2\pi^2/\varepsilon})\leqslant\kappa(\mu)-E_{i+1}^{-1}e^{-2\pi^2/\varepsilon}.$$

For $t\geqslant\frac{1}{2}$,

$$\begin{aligned}F(h(x,t))-\kappa(\mu) &\leqslant \Phi_1(x_{1,t},p\kappa(\mu))+o(e^{-2\pi^2/\varepsilon})\\ &\leqslant(\Phi_1+\delta_i\Phi_2)(x_{1,t},p\kappa(\mu))+o(e^{-2\pi^2/\varepsilon})\leqslant -E_{i+1}^{-1}e^{-2\pi^2/\varepsilon}.\end{aligned}$$

Then we also have, for C as in 22.5,

$$\delta_i\Phi_2(x_{1,t})\leqslant Ce^{-2\pi^2/\varepsilon},$$

which gives the required bound on $m_2(h(x,t))$ for $t\geqslant\frac{1}{2}$, if E'_{i+1} is large enough given δ_i. We automatically have this bound for $t\leqslant\frac{1}{2}$. So $h(x,t)\in K_{i+1}(\mu,\varepsilon)$ for all $x\in K_i(\mu,\varepsilon)$, $t\in[0,1]$. □

23.4. The Second Reduction. — We can then reduce Descending Points a little further.

Second Reduction in Descending Points. *Let $0<u\leqslant u_0$. Let $\delta>0$ be sufficiently small given u_0, or $\delta\geqslant 0$ sufficiently small if u_0 is sufficiently small. Let η be sufficiently small given u_0 and δ. Let m be sufficiently large given u. Then any component of the set*

$$\{([\varphi],z):\operatorname{Im}(z)=m,\ (\Phi_1+\delta\Phi_2)([\varphi],z,u)\leqslant-\eta e^{-m}\}$$

is contractible within itself to a point. There is a natural one-to-one correspondence as in the First Reduction.

To obtain the First Reduction, since $\Phi_2 \geqslant 0$ recall from 22.5 that

$$(\Phi_1 + \delta\Phi_2)([\varphi], z, u) \geqslant \Phi_1([\varphi], z, u) \geqslant Ce^{-\operatorname{Im}(z)}.$$

But if $-Ce^{-\operatorname{Im}(z)} \leqslant -\eta e^{-m}$ then $\operatorname{Im}(z) \leqslant m + \log(C/\eta)$. So

$$\begin{aligned}\{([\varphi], z) : \operatorname{Im}(z) \geqslant m,\ (\Phi_1 + \delta\Phi_2)([\varphi], z, u) \leqslant -\eta e^{-m}\}\\ \subset \{z : m \leqslant \operatorname{Im}(z) \leqslant m + \log(C/\eta)\}\\ \subset \cup_{m \leqslant m' \leqslant m+\log(C/\eta)}\{([\varphi], z) : \operatorname{Im}(z) = m',\ (\Phi_1 + \delta\Phi_2)([\varphi], z, u) \leqslant -\eta e^{-m'}\}.\end{aligned}$$

23.5. We now exploit the variable u. The following will be proved in the next chapter.

No Critical Points in Boundary. — *Let u_0 be given. There are $\delta_0 > 0$, $C_1 > 0$ such that the following hold. Let $0 < u \leqslant u_0$, and $0 < \delta \leqslant \delta_0$, or $0 \leqslant \delta \leqslant \delta_0$ if u_0 is sufficiently small. Let m be sufficiently large given u_0. Let D_t denote the derivative tangent to the set $\{(y, z) : \operatorname{Im}(z) = m\}$. Let $|(\Phi_1 + \delta\Phi_2)(y, z, u)| \leqslant \eta e^{-m}$ for η sufficiently small given u and δ. Then*

$$\|D_t(\Phi_1 + \delta\Phi_2)\| \geqslant \delta^4 C_1^{-1},$$

and if u is sufficiently small,

$$\|D_t(\Phi_1 + \delta\Phi_2)\| \geqslant C_1^{-1}.$$

Then standard differential topology yields the following.

23.6. Corollary. — *Fix $u_0 > 0$, $u_1 > 0$. Let δ_0, be as in 23.5 given u_0. Fix $\delta_1 > 0$ with $\delta_0 > \delta_1 > 0$, or $\delta_1 = 0$ if u_0 is sufficiently small. Then for all $u_1 \leqslant u \leqslant u_0$, $\delta_1 \leqslant \delta \leqslant \delta_0$, $0 \leqslant \eta \leqslant \eta_0/2$, if m is sufficiently large, the sets*

$$\{(y, z) : \operatorname{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u) \leqslant -\eta e^{-m}\}$$

are diffeomorphic.

Proof of the Corollary. — We shall show that the sets are diffeomorphic for varying u: the proof for varying δ or η is similar. It suffices to prove that, for u' sufficiently close to u, for all fixed sufficiently large m the sets

$$\{(y, z) : \operatorname{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u') \leqslant -\eta e^{-m}\}$$

are all diffeomorphic. By 22.11, the quotient of any such set by the action of $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$ is compact. (This is why Φ_2 has to be introduced.)

Fix u, and $\eta < \eta' \leqslant \eta_0$. If u' is sufficiently close to u given η and η', then

$$\begin{aligned}\{(y, z, u') : \operatorname{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u'e) \leqslant -\eta' e^{-m}\}\\ \subset \{(y, z, u) : \operatorname{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u) < -\eta e^{-m}\}.\end{aligned}$$

By the Implicit Function Theorem, if u' is sufficiently close to u, we can construct a C^1 function $\Psi(y, z, \eta'')$, on an open set containing

$$\{(y, z, \eta'') : \eta \leqslant \eta'' \leqslant \eta',\ \mathrm{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u') = -\eta'' e^{-m}\}$$

which is a diffeomorphism C^1-close to the identity between the two sets

$$\{(y, z) : \mathrm{Im}(z) = m,\ \Phi(y, z, u') = -\eta'' e^{-m}\},$$
$$\{(y, z) : \mathrm{Im}(z) = m,\ \Phi(y, z, u) = -\eta'' e^{-m}\}.$$

Now let t be a C^1 function which is 0 on $(-\infty, -\eta']$, 1 on $[-\eta, \infty)$ and strictly increasing on $(-\eta', -\eta)$. We can choose t with a bound $O((\eta' - \eta)^{-1})$ on the first derivative. Then if $u' - u$ is sufficiently small, the function

$$(y, z) \longmapsto t(\Phi(y, z, u'))\Psi(y, z, \eta'') + (1 - t((\Phi_1 + \delta\Phi_2)(y, z, u')))y$$

is the required diffeomorphism between the sets

$$\{(y, z) : \mathrm{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u') \leqslant -\eta e^{-m}\},$$
$$\{(y, z) : \mathrm{Im}(z) = m,\ (\Phi_1 + \delta\Phi_2)(y, z, u) \leqslant -\eta e^{-m}\}.$$

Note that it is the identity on the set

$$\{(y, z) : (\Phi_1 + \delta\Phi_2)(y, z, u') \leqslant -\eta' e^{-m}\}. \qquad \square$$

23.7. The functions α and β. — It follows from 23.5 and 23.6 that the Second Reduction of Descending Points only needs to be proved for sufficiently small $u > 0$ and $\delta = 0$ (and m sufficiently large given such u).

Let g_0 be the branched covering, and s the holomorphic branched covering, used to define $[\varphi'']$ from $[\varphi']$ in 22.4. Recall that $[\varphi]$ and $\tau([\varphi]) = [s^{-1} \circ \varphi \circ g_0]$ are the projections of $[\varphi']$, $[\varphi'']$ to $\mathcal{T}(A(\Delta_0))$. Choose constants $\alpha = \alpha([\varphi], z)$ and $\beta = \beta([\varphi], z)$ so that $t \mapsto s(\alpha^{-1}(t - \beta))$ fixes $\varphi'(v_1)$ and $\varphi'(c_1)$. Take $\alpha = 1$ if $\#(Z' \cap \Delta_0) = 1$ (which is true if $\varphi'(v_1) = \varphi'(c_1)$). Then α and β are uniquely determined. Furthermore, $\beta([\varphi], \lambda z) = \lambda\beta([\varphi], z)$

23.8. Small Parameter Value. — *Let $C > 0$ be a sufficiently large constant. Let $u > 0$ be sufficiently small and m sufficiently small given u.*

(1) *Any component W' of*

$$\{([\varphi], z) : \mathrm{Im}(z) = m,\ \Phi([\varphi], z, u) \leqslant 0\}$$

is contained in

$$\{[\varphi], z) : d([\varphi], \tau([\varphi])) \leqslant u,\ |\alpha([\varphi], z) - 1| \leqslant Cu|\log u|,\ \mathrm{Re}(\beta([\varphi], z)) \geqslant -e^{-m(1+1/C)}\},$$

and is thus disjoint from its iterates under the $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action.

(2) *For $\delta > 0$ sufficiently small independent of u, any singular point in a component W' as above is contained in*

$$\{([\varphi], z) : d([\varphi], \tau([\varphi])) \leqslant u(1-\delta),\ |\alpha([\varphi], z) - 1| \leqslant Cu,\ \operatorname{Re}\beta([\varphi], z) \geqslant C^{-1}\delta e^{-m}\} = W(u, \delta, C, m).$$

(3) *If m is sufficiently large given (u, δ, C),*

$$W(u, \delta, C, m) \subset \{([\varphi], z) : \Phi([\varphi], z, u) \leqslant -C^{-1}\delta e^{-m},\ \operatorname{Im}(z) = m\}.$$

(4) *The set $W(u, \delta, C, m)$ is contractible within $W(u, \delta C^{-1}, C^2, m)$ to*

$$\{([\varphi], z) : d([\varphi], \tau([\varphi])) = 0,\ \alpha([\varphi], z) = 1,\ \beta = \beta([\varphi], z) \in \mathbf{R},\ \beta \geqslant C^{-1}e^{-m}\}.$$

This will be proved in Chapter 24.

23.9. Corollary. — *The orbits under the $G(B(f_0, \Gamma, \Delta_0')) \times \mathbf{Z}$-action of components of $\{([\varphi], z) : \Phi([\varphi], z) \leqslant 0\}$ are in 1-1 correspondence with polynomials in $B(f_0, \Gamma, \Delta_0') = B(A(\Delta_0'), g_0)$, and are contractible to points.*

Proof. — The components of the set to which $W(u, \delta, C, m)$ contracts are points. If $[\varphi]$ is such a point and $s_\beta(\zeta) = s(\zeta - \beta)$ is the rational map such that

$$[\varphi] = [s_\beta^{-1} \circ \varphi \circ g_0],$$

then s_β has a parabolic fixed point at ∞, and a finite critical orbit with the same dynamics as v_1 under g_0. Such a map is on the boundary of the hyperbolic component of a unique polynomial in $B(A(\Delta_0'), g_0) = B(f_0, \Gamma, \Delta_0')$ — whose Thurston equivalence class can be immediately computed from s_β, by joining the infinite-orbit critical point of s_β to its image under s_β, by an arc in the parabolic basin at ∞. Then we can take a flow Ψ_t with vector field w transverse to $\partial W'$ and to $\partial W(u, \delta, C, m)$ such that $D\Phi_{([\varphi],z)}(w) \leqslant -C_1 e^{-m}\|w\|\|D\Phi_{([\varphi],z)}\|$ for a constant $C_1 > 0$. (To get this for C and C_1 independent of m we can use the scaling properties of Φ.) Then $\Psi_t(W') \subset W'$ for all $t \geqslant 0$. Let W'' be a component of $W(u, \delta, C, m)$. For small $t > 0$, we have either $\Psi_t(\partial W'') \subset \operatorname{int}(W'')$ or $\Psi_t(\partial W'' \cap W'' = \varnothing$, that is, either $\Psi_t(W'') \subset W''$ for at least one (and hence all) $t > 0$ or $\Psi_t(\partial W'') \cap W'' = \varnothing$ for all $t > 0$. We must have the former, because all singular points of Φ in W' are in W'', and all positive Ψ_t flow orbits in W' must pass arbitrarily close to singular points, otherwise Φ decreases by arbitrarily much. So there is T such that $\Psi_t(W') \subset W''$ for all $t \geqslant T$. Then by 3 of 23.8, W'' is contractible within W' to a point in W''. So W' is also contractible within W' to this point. □

To obtain the Second Reduction from this Corollary, simply apply 23.6.

23.10. A pullback for Φ. — We recall that the key property of $F(x) = d(x, \tau(x))$ is $F(\tau(x)) \leqslant F(x)$. We want something of this nature to be true for $\Phi = \Phi_1$. Now $\Phi_1([\varphi], z, u) = d(x', x'')$ where $x' = x'([\varphi], z) \in \mathcal{T}(Y')$ and $x'' = x''([\varphi], z, u) \in \mathcal{T}(Z')$. As in 6.7, we can extend x'' uniquely to an element of $\mathcal{T}(Y')$ by the condition

$$d_{Y'}(x', x'') = d_{Z'}(x', x'').$$

Let the branched covering g_0 of 22.4 be defined to fix $\overline{\mathbf{C}} \smallsetminus \Delta_0$ pointwise, and hence fix the 3 points of $A' \smallsetminus \Delta_0$. This definition is only unique up to a Dehn twist, but the Dehn twists of g_0 are conjugate to g_0, as can easily be checked. We have a pullback $\tau_1 : \mathcal{T}(Y') \to \mathcal{T}(Z')$ defined using g_0. See 6.7 for the general definitions of pullbacks, but note that we have taken the range as $\mathcal{T}(Z')$, which is simpler than the (currently irrelevant) definition with range $\mathcal{T}(Y')$. As usual with pullbacks, we have

$$d_{Z'}(\tau_1(x'), \tau_1(x'')) \leqslant d_{Y'}(x', x''). \tag{1}$$

Recall that a pullback (which we can think of as τ_1) was used in 22.4 to define x'' from x'. Define

$$\tau([\varphi], z, u) = (\rho_0 \times \rho_\partial)(x''([\varphi], z, u).$$

We have the following.

23.11. Lemma. — *If* $\mathrm{Im}(z) = m$,

$$\Phi_1(\tau([\varphi], z, u), u) \leqslant \Phi_1([\varphi], z, u) + O(e^{-2m}).$$

Proof. — The key is (1) above together with the formula for $d(x', x'')$ of 22.8. Write $x' = [\varphi']$, $x'' = [\varphi'']$ as usual, and

$$x_1' = [\varphi_1'] = \tau_1(x'),\ x_1'' = [\varphi_1''] = \tau_1(x''),$$
$$x_2' = [\varphi_2'] = x'(\tau([\varphi], z, u)),\ x_2'' = [\varphi_2''] = x''(\tau([\varphi], z, u), u).$$

Let s, s_1, $s_{1,1}$, s_2 be the holomorphic branched coverings used to define x'' from x', x_1' from x', x_1'' from x'' and x_2'' from x_2'. (See 22.4 for the first and fourth, and 6.7, if necessary, for the second and third.) Thus,

$$(\rho_0 \times \rho_\partial)([s^{-1} \circ \varphi' \circ g_0]) = (\rho_0 \times \rho_\partial)([\varphi'']),$$
$$[s_1^{-1} \circ \varphi' \circ g_0] = [\varphi_1'], \quad [s_{1,1}^{-1} \circ \varphi'' \circ g_0] = [\varphi_1''], \quad [s_2^{-1} \circ \varphi_2' \circ g_0] = [\varphi_2''].$$

Then s, s_1 have the same critical values. So we can take $s_1 = s$. Then

$$s_1^{-1} \circ \varphi' \circ g_0(a) = \varphi'(a) + O(e^{-2m})$$

for $a \in A' \smallsetminus \Delta_0$. We can choose $s_{1,1}$ so that the corresponding relation holds for φ'' and $s_{1,1}$. From the definition of $[\varphi_2']$, for some $h = O(e^{-m}) \in \mathbf{C}$, and the fact that $s_1 = s$,

$$\varphi_2'(a) + h = \varphi''(a) = \varphi_1'(a)\ (a \in \Delta_0 \cap Z'), \quad \varphi_2'(v_2) + h = \varphi''(v_2). \tag{2}$$

The h occurs because $(\varphi_2'(a))$ is defined to be $(\varphi''(a)+h)$ with

$$\sum_{a\in\Delta_0\cap Y'}\varphi_2'(a)=0.$$

In particular (2) holds for the critical values v_1, v_2 of g_0. But $\varphi''(v_1)$, $\varphi''(v_2)$ are the critical values of $s_{1,1}$ and $\varphi_2'(v_1)$, $\varphi_2'(v_2)$ are the critical values of s_2. Then if we choose both $s_{1,1}$ and s_2 to be within $O(e^{-2m})$ of the identity away from 0 (which we are bound to do for s_2, at least, by the definitions in 22.4) we have

$$s_2(z)=s_{1,1}(z+h)-h.$$

So then

$$\varphi_2''(a)+h=\varphi_1''(a),\ a\in Z'\cap\Delta_0. \tag{3}$$

We also have

$$\begin{aligned}\varphi_2'(a)&=\varphi'(a)=\varphi_1'(a)+O(e^{-2m}),\quad a\in A'\setminus\Delta_0,\\ \varphi_2''(a)&=\varphi''(a)=\varphi_1''(a)+O(e^{-2m}),\quad a\in A'\setminus\Delta_0,\end{aligned}$$

and from the definitions and restrictions on Dehn twists, $[\varphi_1']$ and $[\varphi_2']$ are close, as are $[\varphi_1'']$ and $[\varphi_2'']$. Let $x_{2,t}',\,x_{2,t}''\in\mathcal{T}(Z')$ be defined by (2), (3) with h replaced by th, and by being continuous. Let $q_{2,t}(z)dz^2$ be the quadratic differential for $d(x_{2,t}',x_{2,t}'')$ at $x_{2,t}'$ with stretch $p_{2,t}(z)dz^2$ at $x_{2,t}''$. Write $q_2=q_{2,0}$ and $p_2=p_{2,0}$. By Hard Same Shape 16.2, the residues of $q_{2,t}$, $p_{2,t}$ are close to those of q_2, p_2 at all points. Then we can apply the Derivative Formula 8.4 to get

$$\begin{aligned}d_{Z'}(x_1',x_1'')&=d_{Z'}(x_2',x_2'')+\int_0^1(d/dt)(d_{Z'}(x_{2,t}',x_{2,t}''))dt\\ &=d_{Z'}(x_2',x_2'')+o(e^{-m})+\sum_{a\in\Delta_0\cap Z'}2\pi\operatorname{Re}(h\operatorname{Res}(q_2-(s_2)_*p_2,\varphi_2'(a)))\\ &=d_{Z'}(x_2',x_2'')+o(e^{-m})\end{aligned}$$

because the sums of residues for both q_2 and $(s_2)_*p_2$ are $1+o(1)$.

23.12. Invariant loops sets in Δ_0. — Now we need to consider, for small $\varepsilon_0'>0$,

$$\{([\varphi],z):\Phi([\varphi],z,u)\leqslant 0\}\cap\bigcup\{\mathcal{T}(\gamma,\varepsilon_0'):\gamma\subset\operatorname{int}(\Delta_0)\}.$$

Given $L_1>0$ we can find L_2 such that the following holds. Given $u>0$, take any $([\varphi],z)$ with $(\Phi_1+\delta\Phi_2)([\varphi],z,u)\leqslant 0$. Let Γ_1 be the possibly empty loop set of 22.10. Let g_0 be the branched covering of 22.4. Then we can find $\varepsilon_0'\geqslant\varepsilon_0/L_2$, and $\Gamma_2\subset\Gamma_1$ which is invariant under g_0, such that $x'\in\mathcal{T}(\gamma,\varepsilon_0')$ for all $\gamma\in\Gamma_2$, and such that if $\gamma\notin\Gamma_2$ then $x'\notin\mathcal{T}(\gamma,L_1\varepsilon_0')$. Let $u\leqslant u_0$. Then assuming L_1 is sufficiently large given u_0, Γ_2 contains no Levy cycles. This follows from an analogue of 17.5, and is proved in much the same way as 17.5, replacing $F(x)=d(x,\tau(x))$ by $\Phi(\pi(x),u)=\Phi([\varphi],z,u)=d(x',x'')$, as we now outline. Let s, $[\xi_0]$, $[\xi_u]$ $q_0(z)dz^2$ and $q_u(z)dz^2$ be as in 22.4. Suppose that Γ_2 contains Levy cycles. Then in the coordinate of 22.4, $\varphi(v_1)$, $\varphi(v_2)$ are within $o(e^{-m})$ of each other. Then the formula in 22.4 shows that the branch of s^{-1} fixing ∞ is close to the identity except within $o(e^{-m})$ of $\varphi(v_1)$, $\varphi(v_2)$. Hence $\varphi''(a)$ is within $o(e^{-m})$ of $\varphi'(g_0(a))$ for $a \in Y' \cap \Delta_0$. In analogy to 17.5, this shows that if Γ_2 contains Levy cycles then $\Phi([\varphi], z, u)$ is bounded below by $d([\xi_0 + \underline{h} + \underline{h}'], [\xi_u + \underline{h}')$ where $\underline{h}' = O(\|\underline{h}\|)$, $Ce^{-m} \leqslant \|\underline{h}\| \leqslant \delta$ for δ small and C bounded from 0 and $\underline{h}$ varies over vectors in directions which sum to 0. To do this, we need to know that

$$d([\xi_0 + \underline{h}'], [\xi_u + \underline{h}']) = u + O(\|\underline{h}'\|^2).$$

This is true, by the Derivative formula 8.4, because the local coordinates are chosen so that $\mathrm{Res}(q_0, 0) = \mathrm{Res}(q_u, 0)$.

We also claim that the gap Δ_0' of Γ_2 adjacent to $\partial\Delta_0$ does not have Euclidean branched map space. This is proved using analogues of 17.8 and 17.9 and the pullback $\tau([\varphi], z, u)$ of 23.10. Since u is a continuous variable we have to consider the cases $\kappa(\Delta_0') \geqslant u - \delta$ for small δ and $\kappa(\Delta_0') \leqslant u - \delta$.

Let orbit dominant area be as in 20.6. We can use this concept for gaps of Γ_2, using $q(z)dz^2$. We shall write $a(\alpha)$ for $a(\alpha, q)$.

Let $\mathrm{Im}(z) = m$. By the same argument as in 17.6, for a constant $C > 0$, $a(\partial\Delta_0) \leqslant Ce^{-m}/u$. Because of our normalisation, it is possible that there is a large modulus annulus homotopic in $\overline{\mathbf{C}} \setminus \varphi'(Z)$ to $\varphi'(\partial\Delta_0)$ and satisfying the Pole Zero Condition 9.4 — but if it does exist then its diameter is $O(e^{-m}/u)$.

CHAPTER 24

PROOF OF DESCENDING POINTS: CRITICAL POINTS

24.1. In this chapter, we prove the No Boundary Critical Points and Small Parameter Value results of 23.5, 23.8. This then completes the proof of Descending Points, as explained in Chapter 23. We continue to use the notation $x' = [\varphi'] = x'([\varphi], z)$, $x'' = [\varphi''] = x''([\varphi], z, u)$ and $q(t)dt^2$ for the quadratic differential for $d(x'', x'')$ at x', with stretch $p(t)dt^2$ at x''.

24.2. Lemma. — *There is a constant $C > 0$ such that if $e^{-m/3C} \leqslant u \leqslant u_0$, and D_t denotes tangential derivative as in 23.5, then the following holds. Let $\mathrm{Im}(z) = m$, $\|D_t(\Phi_1 + \delta\Phi_2)([\varphi], z)\| \leqslant \zeta$ and let $(\Phi_1 + \delta\Phi_2)([\varphi], z) = -\zeta' e^{-m}$. Then*

$$\|D(\Phi_1 + \delta\Phi_2)([\varphi], z)\| = O(\zeta + \zeta' + e^{-m/C})$$

with respect to natural local coordinates.

Now let $q(t)dt^2$ be the quadratic differential for $d(x' \cdot x'')$ at x' ($x' = x'([\varphi], z)$, $x'' = x''([\varphi], z, u)$) replace the condition $(\Phi_1 + \delta\Phi_2)([\varphi], z) = -\zeta' e^{-m}$ by the existence of an annulus of modulus $\geqslant -2\log\zeta'$ satisfying the Pole-Zero Condition 9.4 for $q(t)dt^2$ which intersects $\{t : |t| \geqslant \zeta'^{-2} e^{-m}\}$. Then

$$\|D(\Phi_1 + \delta\Phi_2)([\varphi], z)\| = O(u + \zeta + \zeta' + e^{-m/C}).$$

Proof. — We use the local coordinates $(\varphi'(a))$ $(a \in Y' \cap \Delta_0)$, where $[\varphi'] = x'([\varphi], z)$. Then the tangent space of $\{([\psi], w) : \mathrm{Im}(w) = m\}$ at $([\varphi], z)$ identifies with the real codimension 3 vector space W given by

$$\Big\{(\xi(a)) : \textstyle\sum_{a \in Y' \cap \Delta_0} \xi(a) = 0,\ \mathrm{Re}(\varphi'(v_1) - \varphi'(v_2))\overline{\xi(v_1) - \xi(v_2)} = 0\Big\}.$$

We are assuming that $D\|(\Phi_1 + \delta\Phi_2)(\underline{w})\| \leqslant \zeta\|\underline{w}\|$ for all $\underline{w} \in W$. But we claim that there is a boundedly complementary 3-dimensional subspace on which $D(\Phi_1 + \delta\Phi_2)$ is $O(e^{-m/C})$ for suitable C. The following (real) vectors are in the kernel of $D\Phi_2$:

$$\begin{aligned} \underline{x} &= (x(a)) \quad \text{with } x(a) = 1\ (a \in Y' \cap \Delta_0), \\ \underline{y} &= (y(a)) \quad \text{with } y(a) = i\ (a \in Y' \cap \Delta_0), \\ \underline{\varphi'} &= (\varphi'(a)). \end{aligned}$$

Let $\underline{v}$ be any of these vectors, which can also be regarded as a vector field. To see that $\underline{v} \in \mathrm{Ker}(\Phi_2)$, we need to consider m_γ for $\gamma \in \Gamma_1$, Γ_1 as in 22.10. Let φ'_t be an orbit of the vector field $\underline{v}$, with $\varphi'_0 = \varphi'$. Then one point of $\varphi'_t(A(\gamma))$ is fixed at ∞, for $\underline{v} = \underline{x}$ the others are all translated to t, for $\underline{v} = \underline{y}$ they are translated by ti, and for

$\underline{v} = \underline{\varphi}'$ they are scaled by e^t. So in each case, $m_\gamma([\varphi'_t])$ is constant and $Dm_\gamma(\underline{v}) = 0$. So all such $\underline{v}$ are in $\mathrm{Ker}(\Phi_2)$.

We claim that

$$\sum_{a\in Y'\cap\Delta_0} \mathrm{Res}(q, \varphi'(a)) = 1 + O(u^{-1/2}e^{-m/2}) = 1 + O(e^{-m/C}) \tag{1}$$

for suitable C, and similarly for s_*p. From (1) and the formula for $D\Phi_1$ in 22.6,

$$D\Phi_1(\underline{x}) = O(e^{-m/C}) = O(e^{-m/C}\|\underline{x}\|), \quad D\Phi_1(\underline{y}) = O(e^{-m/2}) = O(e^{-m/C}\|\underline{y}\|).$$

We prove (1) as follows. Let q_0, q_u be as in 22.4. Then recall that q_0, q_u were both normalised to have residue 1 at 0, and

$$\int |q_0| = \int |q_u| = 1.$$

By the Same Shape Lemma 9.5, if θ' denotes the angle between q_0 and q, and θ'' denotes the angle between q_u and p,

$$u\int \theta'^2|q_0| = O(e^{-m}), \quad u\int \theta''^2|q_u| = O(e^{-m}), \tag{2}$$

and hence

$$\sum_{a\in Y'\cap\Delta_0} \mathrm{Res}(q, \varphi'(a)) - \mathrm{Res}(q_0, \varphi_0(A'\cap\Delta_0) = O(u^{-1/2}e^{-m/2}),$$

and similarly for p, q_u, and (1) for q and s_*p follows, since by 22.4 $s^{-1}(z) - z = O(e^{-2m})$ away from 0. From (2) we also have

$$\mathrm{Res}(q, \varphi'(a)) = O(u^{-1}e^{-m}) = O(e^{-2m/C}) \tag{3}$$

for suitable C, and similarly for p, s_*p.

Now we need to obtain estimates on $D\Phi_1(\underline{\varphi})$. First, suppose that

$$(\Phi_1 + \delta\Phi_2)([\varphi], z) = -\zeta' e^{-m}.$$

We have, by 22.6 and 22.8,

$$-\zeta' e^{-m} = \Phi_1([\varphi], z, u) = D\Phi_1(\underline{\varphi}') + o(e^{-m}).$$

In fact, from 16.2 and (3), the $o(e^{-m})$ term is $O(e^{-m(1+1/C)})$ for suitable $C > 0$. This gives

$$D\Phi_1(\underline{\varphi}') = O((\zeta + e^{-m/C})e^{-m}) = O((\zeta' + e^{-m/C})\|\underline{\varphi}'\|),$$

because $|\varphi'(v_1) - \varphi'(v_2)| = e^{-m}$.

Now suppose that there is an annulus of modulus $\geqslant -2\log\zeta'$, intersecting $\{t : |t| \geqslant \zeta'^{-2}e^{-m}\}$, and satisfying the Pole-Zero Condition for q. Choose $r > 0$ such that the annulus contains $\{t : C_1\zeta' r \leqslant |t| \leqslant (C_1\zeta')^{-1}r\}$ for a $C_1 > 0$. Then take $\gamma = \{t : |t| = r\}$. Then

$$\int_\gamma tq(t)dt = (1 + O(u))\int_\gamma tp(t)dt = (1 + O(u) + O(\zeta'))\int_\gamma ts_*p(t)dt,$$

and hence

$$\sum_{a\in Y'\cap\Delta_0}\varphi'(a)\,\mathrm{Res}(q-s_*p,\varphi'(a))=(O(u)+O(\zeta'))e^{-m},$$

that is,

$$\|D\Phi_1([\varphi],z)\underline{\varphi}'\|=(O(u)+O(\zeta'))\|\underline{\varphi}'\|.$$

With respect to the standard inner product on $\mathbf{C}^n$, $n=\#(A(\Delta_0)\cap\Delta_0)$, we have

$$\|\underline{x}-W\|=\sqrt{n}=\|\underline{y}-W\|,\quad \|\underline{\varphi}'-W\|=|\varphi'(v_1)-\varphi'(v_2)|\geqslant C_0\|\underline{\varphi}'\|,$$

for a suitable constant $C_0>0$. This last follows from 23.12: for $u\leqslant u_0$ there is $\varepsilon_0'>0$ such that the following holds. If $\Phi_1([\varphi],z,u)\leqslant 0$, then by 23.12, there cannot be a loop γ separating v_1 and v_2 from some other points of $Y'\cap\Delta_0$ with $[\varphi']\in\mathcal{T}(\gamma,\varepsilon_0')$. So then the required bound on $\|D(\Phi_1+\delta\Phi_2)(\underline{v})\|$ for all vectors $\underline{v}$ follows from the bound on $\|D_t(\Phi_1+\delta\Phi_2)\|$. □

24.3. Lemma. — *Let $e^{-3m/C}\leqslant u\leqslant u_0$. Let L_1 be sufficiently large and D_1 sufficiently small given u_0, and $\delta_0>0$ be sufficiently small given L_1. Let $\delta\leqslant\delta_0$. Let $([\varphi],z)$ satisfy either of the hypotheses of 24.2, with $u=O(\zeta')$ under the second hypothesis, with ζ, $\zeta'=o(\delta^4)$ if $\delta>0$, and ζ, $\zeta'\geqslant 0$ sufficiently small otherwise. Let Γ_2 be the (possibly empty) invariant loop set of 23.3 with the stated properties relative to L_1. Then:*

(1) *for all $\gamma\in\Gamma_2$, $m_\gamma(x')=O(\delta^{-1/2}\log\delta)$;*

(2) *for a suitable $C_2>0$, there are no D_1-orbit dominant area gaps α of Γ_2 in* $\mathrm{Int}(\Delta_0)$ *with $a(\alpha)\geqslant C_2(\zeta+\zeta'+e^{-m/C})e^{-m}$.*

Proof

1. Let α be a loop or gap of Γ_2. Let $V(\alpha)$ be as in 20.2, and similarly for $V[\alpha]$ if α is periodic. As in 20.2, use the norm of 19.8 on $V(\alpha)$, $V[\alpha]$. Then by 24.2, our assumption on $D(\Phi_1+\delta\Phi_2)$ translates to

$$|D(\Phi_1+\delta\Phi_2)_{([\varphi],z)}(\theta)|\leqslant C_0\|\theta\|e^{-m}(\zeta+\zeta'+e^{-m/C})\tag{3}$$

for a suitable constant C_0 and any $\theta\in V(\alpha)$, any loop or gap α.

Let $\delta>0$. Suppose that α is a periodic loop (not necessarily dominant). Then for a constant $C_3>0$, either $m_\alpha(x')\leqslant M+1$ or

$$C_3^{-1}e^{-m}\leqslant\max_{\theta\in V[\alpha]}|D\Phi_2(\theta)|/\|\theta\|\leqslant C_3e^{-m}.$$

But for $\theta\in V[\alpha]$

$$D\Phi_1(\theta)=O(a(\alpha)\|\theta\|)=O(e^{-m}m_\alpha(x')^{-1}\|\theta\|),$$

so we obtain $m_\alpha(x')=O(\delta^{-1})$. We then have a similar bound on $m_{\alpha'}(x')$ for all $\alpha'\in[\alpha]$. In fact we can do better than this, as we shall see. In any event, if $a(\alpha)\leqslant m_\alpha(x')^{-2}$ we have $m_\alpha(x')=O(\delta^{-1/2})$.

Now suppose that α is a D_1-orbit dominant nonperiodic loop or gap. Let

$$m'(x) = m_\alpha(x) \text{ or } m'(x') = \min\{m_\gamma(x') : \gamma \subset \partial\alpha\},$$

depending on whether α is a loop or a gap. Then by 20.4

$$(D\Phi_2)_{([\varphi],z)}(\theta) = O(e^{-m}e^{-m'(x')}\|\theta\|).$$

Suppose also that

$$a(\alpha) \geqslant \max(\delta^{1/2}e^{-m-m'(x')}, C_2(\zeta + \zeta' + e^{-m/C})e^{-m}).$$

Then by the analogue of 21.4 for $d(x', x'')$, for a suitable constant C_1, we can choose $\theta \in V[\alpha]$ with

$$D\Phi_1(\theta) \leqslant -a(\alpha)\|\theta\|/C_1,$$
$$D(\Phi_1 + \delta\Phi_2)(\theta) \leqslant -a(\alpha)\|\theta\|/2C_1 < -C_0(\zeta + \zeta' + e^{-m/C})e^{-m}\|\theta\|,$$

if C_2 is large enough given C_0 and C_1, which contradicts (3).

Now let $D_1' > 0$ be sufficiently small depending only on $\#(Y')$, and let $\delta > 0$ be sufficiently small given D_1'. Suppose that α is a (D_1, D_1')-orbit dominant periodic loop. Note that if δ is sufficiently small given D_1', then by the above estimate on nonperiodic loops, this includes all loops α with $a(\alpha) \geqslant m_\alpha(x)^{-2}e^{-m}$ and $m_\alpha(x) \geqslant \delta^{-1/2}$, because we already know that $m_\alpha(x) = O(\delta^{-1})$. Let $[\alpha] = \{\alpha_j : 0 \leqslant j \leqslant n-1\}$ and let $\theta_j \in V(\alpha_j)$ as in 20.2 with $\lambda = 1$. Then $Dm_{\alpha'}(i\theta_j) = O(e^{-m'})$, for $m' = \min_j m_{\alpha_j}(x')$. This comes from the asymptotic formula $m_\gamma([\psi]) = -\log|a| + O(1) + O(a\log a)$ mentioned in 17.8, if $\psi(A(\gamma)) = \{0, 1, \infty, a\}$. So $D\Phi_2(i\theta_j) = O(e^{-m'})$. So we have $D\Phi_1(i\theta_j) = O((e^{-m'-m})$ for all j, assuming that m is large enough for $\varepsilon_m = O(e^{-m/C}) = O(e^{-m'}) = O(e^{-1/\delta})$. As in 21.7, 21.9-10, let

$$c_j + id_j = \int_{\varphi'(\alpha_j)} \sqrt{q}\, dz.$$

(This coordinate is actually invariant under scaling.) Then as in 21.7 and 21.9-10 we can find θ with both $D\Phi_1(\theta) < 0$, $D\Phi_2(\theta) \leqslant 0$ unless $c_jd_j = O(a(\alpha)m_\alpha(x)^{-1})$ for all j, and hence unless $d_j^2 = O(a(\alpha)m_\alpha(x)^{-1}))$ for all j, and $F(x) = \kappa(\alpha) + O(D_1')$. So $\kappa(\alpha)$ must be maximal and $\alpha \subset \partial\Delta_0'$. Now let $a(\alpha) \geqslant m_\alpha(x)^{-2}$, $m_\alpha(x) \geqslant \delta^{-1/2}\log\delta$. Then

$$a(\alpha_j) = O(m_{\alpha_j}(x)^{-1}) = O(\delta^{1/2}(-\log\delta))$$

for all j. Then we have α is (D_1, D_1') dominant with $D_1' = O(\log m_\alpha(x)(m_\alpha(x))^{-1/2})$. In analogy with the formula for $DF(\theta)$ in 21.9 we have, for $\theta = \sum_j \lambda_j\theta_j$,

$$D\Phi_1(\theta) = \sum_{j=0}^{n-1} \mathrm{Re}(\lambda_j(n_j^{-1}Kc_{j-1}^2 - c_j^2)) + O(m_\alpha(x)^{-1}a(\alpha)\|\theta\|).$$

So we can solve with $D\Phi_2(\theta) = 0$, $D\Phi_1(\theta) \leqslant -Ca(\alpha)m_\alpha(x))^{-1} \leqslant -C\delta^4e^{-m}$ for $C > 0$ bounded from 0, unless $n_j^{-1}Kc_{j-1}^2 - c_j^2 = O(m_\alpha(x)^{-1}a(\alpha))$ for all j, which gives $D\Phi_1(\theta) = O(a(\alpha)m_\alpha(x)^{-1}\|\theta\|)$ for all θ. So for $D(\Phi_1 + \delta\Phi_2)(\theta) = O(\delta^5e^{-m}\|\theta\|)$

for all θ, $a(\alpha)(m_\alpha(x))^{-1}$ must be $\geqslant C_1\delta$ for $C_1 > 0$ bounded from 0. So $m_\alpha(x) \leqslant \delta^{1/2}$. So altogether we have $m_\alpha(x) = O(\delta^{1/2}\log\delta)$.

2. Now suppose that α is a periodic D_1-orbit dominant gap with $a(\alpha, q) \geqslant C_2(\zeta + \zeta' + e^{-m/C})e^{-m}$. By the analogue of 21.6 for $d(x, x')$, there exists $\theta \in V[\alpha]$ with $D\Phi_1(\theta) \leqslant -e^{-C_1/\varepsilon_0'}a(\alpha)\|\theta\|$ which again contradicts (1), assuming C_2 is large enough given C_1 and ε_0'. □

24.4. Corollary. — *Take the same hypotheses as in 24.3. For a constant $C_1 > 0$ depending only on u_0:*

$$\int |s_*p| \leqslant 1 - C_1^{-1}a(\partial\Delta_0, q)$$

Remark. — $e^{-m} = O(a(\partial\Delta_0, q))$.

Proof. — We consider the pointwise expression for s_*p

$$s_*p(z) = ((s_1^{-1})'(z))^2 p(s_1^{-1}(z)) + ((s_2^{-1})'(z))^2 p(s_2^{-1}(z)),$$

where s_1^{-1} and s_2^{-1} are local branches of s^{-1}. By 2 of 24.3, we can take a bounded modulus annulus A homotopic to $\varphi''(\partial\Delta_0)$ on which $C_3e^{-m} \leqslant |z| \leqslant C_2e^m$ for C_2 bounded and C_3 bounded from 0, and such that on A,

$$p(z) = z^{-j}(1 + o(1))$$

for $j = 1$ or 2. We have a similar expression for $((s_1^{-1})'(z))^2 p(s_1^{-1}(z))$ for suitable choice of branch s_1, using the expression for s from 22.4. Computing with s as in 22.4, we obtain, for $z \in A$

$$s_2^{-1}(z) = -a + \frac{(z_0 + a)^2}{z - a} + O(e^{-3m}(z - a)^{-2}).$$

Because the critical points are bounded apart, we have $|z_0 + a| \geqslant C_3e^{-m}$ for some $C_3 > 0$ bounded from 0. Then on a positive measure proportion of A, if C_2 is large enough,

$$p(s_2^{-1}(z)) = \Big(\frac{z - a}{-a(z - a) + (z_0 + a)^2}\Big)^j (1 + O(C_2^{-1})),$$

and hence

$$\begin{aligned}&((s_2^{-1})'(z))^2 p(s_2^{-1}(z))\\ &= \Big(\frac{z - a}{-a(z - a) + (z_0 + a)^2}\Big)^j (z_0 + a)^4 (z - a)^{-4} (1 + O(C_2^{-1}) + O(e^{-m}/(z - a))).\end{aligned}$$

Thus the two terms in s_*p are of the same order of magnitude on a positive proportion of A (to within $O(C_2^3)$) but their arguments differ by a bounded amount. Then integrating and using the change of variable formula we obtain, for a constant C_1' depending only on C_2, C_3,

$$\int_A |s_*p| \leqslant \int_{s^{-1}(A)} |p|(1 - C_1'^{-1})$$

So then

$$\int |s_* p| \leqslant \int |p| - C_1^{-1} a(\partial\Delta_0, q) = 1 - C_1^{-1} a(\partial\Delta_0, q),$$

as required. □

24.5. Proof of No Critical Points. — The following lemma completes the proof of the No Critical Points result of 23.5, because it gives a contradiction to 24.4 if δ is sufficiently small, or if ζ, ζ' are sufficiently small in the case $\delta = 0$.

Lemma. — *Take the same hypotheses as in 24.3. Then, for a constant $C_1 > 0$, in the notation of 9.4 and*

$$\int |s_* p| = \int |q| + O((\delta^{1/2}\log\delta + \zeta^{1/2} + \zeta'^{1/2} + e^{-m/2C})e^{-m}.$$

The term $\delta^{1/2}\log\delta$ is omitted if $\delta = 0$.

Proof. — Let q_0, q_u be as in 22.4. We need to examine the relationship between q_0 and q, and between q_u and p. Recall that $\#(A') = 4$, so that q_0 and q_u have a particularly simple form. We are assuming that q_0 and q_u have residue 1 at the point $0 = \varphi_0(v_1) = \varphi_u(v_1)$, remembering that $\{v_1\} = Y' \cap \Delta_0$. Then for some $t \geqslant 1$ and $b_j = O(e^{-m})$ $(1 \leqslant j \leqslant t)$, ε_1, $\varepsilon_2 = O(e^{-m/C})$ (using 24.2),

$$q(z) = zq_0(z)(1+\varepsilon_1)r(z), \quad p(z) = zq_u(z)(1+\varepsilon_2)r'(z),$$

$$r(z) = \sum_{j=1}^{t} \frac{c_j}{(z-b_j)}, \quad r'(z) = \sum_{j=1}^{t} \frac{c'_j}{z-b'_j}, \quad \sum_{j=1}^{t} c_j = \sum_{j=1}^{t} c'_j = 1, \quad s(b'_j) = b_j.$$

Write

$$c = \sum_{j=1}^{t} c_j b_j.$$

Then we claim that

$$\sum_{j=1}^{t} c'_j b'_j = c + O((\zeta + \zeta' + e^{-m/C})e^{-m}).$$

The reason is as follows. We use the trick employed in 22.6, in replacing a residue sum by an integral, and also use 24.2:

$$\begin{aligned}\sum_{j=1}^{t} c'_j b'_j &= \sum_{a\in\Delta_0\cap Y'} \varphi''(a)\,\mathrm{Res}(p, \varphi''(a)) + O(e^{-m(1+1/C)})\\ &= \frac{1}{2\pi i}\int_{\varphi''(\partial\Delta_0)} zp(z)dz + O(e^{-m(1+1/C)})\\ &= \frac{1}{2\pi i}\int_{s\circ\varphi''(\partial\Delta_0)} zs_*p(z)dz + O(e^{-m(1+1/C)})\end{aligned}$$

$$= \sum_{a\in Y'\cap\Delta_0} \varphi'(a)\operatorname{Res}(s_*p, \varphi'(a)) + O(e^{-m(1+1/C)})$$
$$= \sum_{a\in Y'\cap\Delta_0} \varphi'(a)\operatorname{Res}(q, \varphi'(a)) + O(e^{-m}(\zeta + \zeta' + e^{-m/C}))$$
$$= \sum_{j=1}^{t} c_j b_j + O(e^{-m}(\zeta + \zeta' + e^{-m/C})) = c + O(e^{-m}(\zeta + \zeta' + e^{-m/C})).$$

For a suitable $\delta' = O(\zeta^{1/2} + \zeta'^{1/2} + e^{-m/2C})$ the following holds. The residues of q and s_*p at a pole of either in $\varphi'(\Delta_0)$ are within $O(\delta'^2)$. Let B be the ball of radius e^{-m}/δ' round 0. Then

$$\int_B |s_*p| = \int_B |q| + O(\delta' e^{-m}).$$

For a bounded constant C_u, we also have $q_u(z) = q_0(z) + C_u + O(\delta'^{-1}e^{-m})$ on B, and so, if $\delta' \geqslant e^{-m/3}$,

$$\int_B |q_u| = \int_B |q_0| + O(\delta'^{-2}e^{-2m}) = \int_B |q_0| + O(\delta' e^{-m}).$$

Now, on $\overline{\mathbf{C}} \setminus B$, and bounded from the points $\varphi''(A' \setminus \Delta_0)$,

$$s_*p = (1+\varepsilon_2)q_u(1 + (c + O(e^{-m}\delta'))/z).$$

Also, from the expression for s of 22.4, the poles of s_*p outside $\varphi''(\Delta_0)$ differ by $O(e^{-2m})$ from those of p, and residues at these points differ by $O(e^{-2m})$. So using our expression for p, we have

$$\int_{\overline{\mathbf{C}}\setminus B} |s_*p| = \int_{\overline{\mathbf{C}}\setminus B} |q_u| + \operatorname{Re}(\varepsilon_2) + \int |q_u| \operatorname{Re}(c/z) + O(e^{-m}\delta').$$

Similarly,

$$\int_{\overline{\mathbf{C}}\setminus B} |q| = \int_{\overline{\mathbf{C}}\setminus B} |q_0| + \operatorname{Re}(\varepsilon_1) + \int |q_0| \operatorname{Re}(c/z) + O(e^{-m}\delta').$$

So to complete the proof of the lemma, it suffices to show that

$$\operatorname{Re}(\varepsilon_1 - \varepsilon_2) + \int (|q_0| - |q_u|)\operatorname{Re}(c/z) = O(\delta^{1/2}\log\delta e^{-m}) + O(\delta' e^{-m}).$$

To do this, we use a simple version of some ideas employed in Chapters 10-13, and regard $\sqrt{q(z)}dz$, $\sqrt{p(z)}dz$ as elements of $H^1(\text{torus})$. To do this, take any two simple isotopically distinct nontrivial nonperipheral loops γ_1, γ_2 in $\overline{\mathbf{C}} \setminus (\Delta_0 \cup A')$, which thus generate $H_1(T)$ for a torus T. Then by integrating along $\varphi'(\gamma_j)$, $\sqrt{q(z)}dz$, can be regarded as an element of $H^1(T)$, and similarly for $\sqrt{p(z)}dz$, integrating along the $\varphi''(\gamma_j)$. Moreover, expanding out, we see that, in $H^1(T)$, we have, for z bounded and bounded from 0,

$$\sqrt{q(z)}dz = (1 + (\varepsilon_1/2))\sqrt{q_0(z)}dz + c(\sqrt{q_0(z)}/2z)dz + O(\delta' e^{-m}),$$
$$\sqrt{p(z)}dz = (1 + (\varepsilon_2/2))\sqrt{q_u(z)}dz + c(\sqrt{q_u(z)}/2z)dz + O(\delta' e^{-m}).$$

Now as computed in 11.5, in $H^1(T)$,

$$(\sqrt{q_0}/z)dz = -4\pi\overline{\sqrt{q_0(z)}\,dz} + \Big(\int \frac{|q_0|}{z}\,dz\Big)\sqrt{q_0(z)},$$

in H^1, and similarly for q_u. Identify complex numbers with matrices in the usual way, namely

$$a+ib \longmapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Write $u' = d(x', x'') = u + \Phi_1([\varphi], z, u)$.Then we also have

$$\sqrt{q_u(z)}\,dz = \begin{pmatrix} e^{u/2} & 0 \\ 0 & e^{-u/2} \end{pmatrix} \sqrt{q_0(z)}\,dz,$$

$$\sqrt{p(z)}\,dz = \begin{pmatrix} e^{u'/2} & 0 \\ 0 & e^{-u'/2} \end{pmatrix} \sqrt{q(z)}\,dz = (1+O(u'-u)) \begin{pmatrix} e^{u/2} & 0 \\ 0 & e^{-u/2} \end{pmatrix} \sqrt{q(z)}\,dz.$$

By 24.3, at our singular point we have

$$\Phi_2([\varphi], z) = O(e^{-m}\delta^{-1/2}\log\delta),$$

and hence, since $\Phi_1 + \delta\Phi_2([\varphi], z, u) = -\zeta' e^{-m}$,

$$u - u' = O((\delta^{1/2}\log\delta + \zeta')e^{-m}),$$

with $\delta^{1/2}\log\delta$ omitted if $\delta = 0$. Combining all these equations, we obtain

$$\begin{aligned}
&\varepsilon_1 + c\int \frac{|q_0|}{z} - 4\pi c \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&\quad = \begin{pmatrix} e^{-u/2} & 0 \\ 0 & e^{u/2} \end{pmatrix} \Big(\varepsilon_2 + c\int \frac{|q_u|}{z} - 4\pi c \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\Big) \begin{pmatrix} e^{u/2} & 0 \\ 0 & e^{-u/2} \end{pmatrix} \\
&\qquad + O(\delta^{1/2}\log\delta e^{-m}) + O(\delta' e^{-m}).
\end{aligned}$$

An effect of conjugating a matrix A by a diagonal matrix is to leave the diagonal entries of A unchanged. So considering the top left entry in the equation above, we obtain

$$\mathrm{Re}\Big(\varepsilon_1 + c\int \frac{|q_0|}{z} - 4\pi c\Big) = \mathrm{Re}\Big(\varepsilon_2 + c\int |q_u|/z - 4\pi c\Big) + O(\delta^{1/2}\log\delta e^{-m}) + O(\delta' e^{-m}),$$

as required. □

24.6. Proof of 1 of 23.8. — Suppose that $d(x', x'') = O(u)$. This will be true, in particular, if $\Phi([\varphi], z, u) \leqslant 0$, because

$$\Phi([\varphi], z, u) = d(x', x'') - u.$$

Let g_0 be the branched covering, and s the holomorphic branched covering, used to define $[\varphi'']$ from $[\varphi']$ in 22.4. Recall that $[\varphi]$ and $\tau([\varphi]) = [s^{-1}\circ\varphi\circ g_0]$ are the projections of $[\varphi']$, $[\varphi'']$ to $\mathcal{T}(A(\Delta_0))$. So $d([\varphi], \tau([\varphi])) = O(u)$. As in 23.7, let

$\alpha = \alpha([\varphi])$ and $\beta = \beta([\varphi], z)$ be such that $t \mapsto s(\alpha^{-1}(t-\beta))$ fixes $\varphi'(v_1)$ and $\varphi'(c_1)$. So (using 8.7) the set $B = \varphi'(Z' \cap \Delta_0)$ is in an $O(ue^{-m})$-neighbourhood of $\alpha s^{-1}(B) + \beta$.

Now s moves the points of $\varphi'(A(\Delta_0) \cap Y')$ by at most $O(e^{-m})$. So $\beta = O(e^{-m})$, and, as we have seen in 23.7, $\beta([\varphi], z + c) = e^{ic}\beta([\varphi], z)$. Note also that $\beta \neq 0$ if u is small, because s itself is close to a quadratic rational map for which both critical points are attracted to ∞ (because of the conditions imposed on the expansion of s near ∞ in 22.4). In fact for a constant $C > 0$, $|\beta| \geqslant Ce^{-m}$

The following lemma, together with the subsequent corollary, prove 1 of 23.8.

Lemma. — *Let* $\Phi([\varphi], z, u) \leqslant 0$. *Then*

$$\alpha([\varphi], z) = 1 + O(u \log u), \quad \mathrm{Re}(\beta([\varphi], z)) = O(e^{-2m}/u).$$

Proof. — We deduce this from $d([\varphi'], [\varphi'']) \leqslant u$, as follows. Let $\mathrm{Im}(z) = m$. Recall that $[\varphi']$, $[\varphi''] \in \mathcal{T}(Y')$, and there is a distinguished set $A' \subset Y'$ with $\#(A') = 4$. Recall also from 22.2, 22.4 that there is a projection ρ_P of $\mathcal{T}(Y')$ to $\mathcal{T}(A')$ such that if $\rho_P([\varphi'] = [\varphi_1']$ and $\rho_P([\varphi''] = [\varphi_1'']$ then the points $\varphi_1'(a)$ are within $O(e^{-m})$ of the points $\varphi'(a)$ for $a \in A'$, and similarly for φ'', φ_1'', and

$$u = d(\rho_P([\varphi']), \rho_P([\varphi''])).$$

So

$$d([\varphi'], [\varphi'']) \leqslant u \leqslant d([\varphi']_{A'}, [\varphi'']_{A'}) + Ce^{-m},$$

where C is bounded. Let θ denote the angle between the quadratic differential $q(z)dz^2$ for $d([\varphi'], [\varphi''])$ at $[\varphi']$ and $q_{A'}(z)dz^2$ for $d([\varphi']_{A'}, [\varphi'']_{A'})$ at $[\varphi']_{A'}$. Then by 9.5, we have

$$\int |\theta|^2 |q_{A'}| \leqslant C_1 e^{-m}/u.$$

The following holds for constants $C_j > 0$ bounded from 0, $2 \leqslant j \leqslant 5$. The annuli of maximal modulus homotopic to $\varphi'(\partial\Delta_0)$, $\varphi''(\partial\Delta_0)$ in $\overline{\mathbf{C}} \setminus \varphi'(Y')$, $\overline{\mathbf{C}} \setminus \varphi''(Y'')$, differ by at least a boundedly nonzero multiple of $|\alpha - 1|$. So there must be an annulus homotopic to $\varphi'(\partial\Delta_0)$ and satisfying the Pole-Zero Condition (9.4) for q of modulus $\geqslant C_2|\alpha - 1|/u$. So θ is bounded from 0 on this annulus. The inner boundary of this annulus has Euclidean diameter $r \geqslant C_3 e^{-m}$, and whose outer boundary has Euclidean diameter $\geqslant e^{C_3|\alpha-1|/u} r$. The integral of $|\theta|^2$ over this annulus is thus $\geqslant C_4 e^{C_3|\alpha-1|/u} e^{-m}$. We deduce that $|\alpha - 1| \leqslant -C_5 u \log u$, which $\to 0$ as $u \to 0$.

Now remembering that $v_1 \in A'$, let $[\xi_0]$, $[\xi_u]$ be as in 8.4. Let ξ_0', ξ_u' be close to ξ_0, ξ_u respectively with $\xi_0'(v_1) = \xi_u'(v_1) = \varphi'(v_1)$. Then by the First Derivative Formula 8.4

$$d([\xi_0'], [\xi_u']) = d([\xi_0], [\xi_u]) + 2\pi \,\mathrm{Re}(\mathrm{Res}(q_0 - q_u, \varphi'(v_1)) + O(e^{-2m}/u) = u + O(e^{-2m}/u).$$

The error term $O(e^{-2m}/u)$ uses a bound on the second derivative which does indeed hold but has never been spelt out. In any case the error term is $O(C(u)e^{-2m})$ for a constant $C(u)$ which is bounded when u is bounded from 0. Let $q'(z)dz^2$ be the quadratic differential for $d_{A'}([\xi_0'],[\xi_u'])$ at x', with stretch $p'(z)dz^2$ at $[\xi_u']$. Then since $\varphi'(v_1)=\xi_0(v_1)+O(e^{-m})$, $\varphi''(v_1)=\xi_u(v_1)+O(e^{-m})$, we deduce from 8.2 (for example) that

$$\mathrm{Res}(q',\varphi'(v_1))=1+O(e^{-m})=\mathrm{Res}(p',\varphi'(v_1)).$$

Then by the first derivative formula 8.4, remembering that $v_1\in A'$

$$\begin{aligned}d_{A'}([\varphi'],[\varphi''])&=d([\xi_0'],[\xi_u'])-2\pi\,\mathrm{Re}(\mathrm{Res}(p',\varphi'(v_1))\beta)+O(e^{-2m}/u)\\&=u-2\pi\,\mathrm{Re}(\beta)+O(e^{-2m}/u).\end{aligned}$$

Since $d([\varphi'],[\varphi''])\leqslant u$ we obtain the required result. □

24.7. Corollary. — *Any component W of*

$$\{([\varphi],z):\mathrm{Im}(z)=m,\ d([\varphi],\tau([\varphi]))\leqslant u,\ |\alpha([\varphi])-1|\leqslant Cu\log u,\\ \mathrm{Re}(\beta([\varphi],z))\geqslant -O(e^{-2m/u})\}$$

is disjoint from its iterates under the $G(B(A(\Delta_0),g_0)\times\mathbf{Z}$-action.

Proof. — If $d([\varphi],\tau([\varphi])\leqslant u$ then $[\varphi]$ is distance $O(u)$ from a point $[\varphi_\infty]$ fixed by τ. This is proved explicitly in 6.15 (and is basically Property 6 of 6.6) using iterated pullback. (It is also possible to prove it using the derivative of $[\varphi]\mapsto d([\varphi],\tau([\varphi])$.) Then we have $|\alpha([\varphi_\infty])-1|=O(u\log u)$. It follows that the rational map s_∞ with $[s_\infty^{-1}\circ\varphi_\infty\circ g_0]=[\varphi_\infty]$ is within $O(-u\log u)$ of a map with a parabolic fixed point. It follows that $\{[\varphi]:([\varphi],z)\in W\}$ for some $z\}$ has diameter $O(-u\log u)$. Then the set of z with $([\varphi],z)\in W$ for some $[\varphi]$ have $\mathrm{Im}(z)=m$ and $\mathrm{Re}(z)$ varying in an interval of length $\pi+O(-u\log u)$, by the relation $\beta([\varphi],z+c)=e^{ic}\beta([\varphi],z)$. So W is indeed disjoint from its iterates under the $G(B(A(\Delta_0),g_0)\times\mathbf{Z}$-action. □

24.8. Proof of 2 of 23.8. — We use 24.4, 24.5. From these, we deduce that if $([\varphi],z)$ is a point where $\|D_t\Phi([\varphi],z)\|=O(\zeta')$ for ζ' small, then $d(x',x'')=d([\varphi'],[\varphi''])\leqslant u-\delta e^{-m}$, and there is no large modulus annulus satisfying the Pole-Zero Condition (9.4) for q and homotopic to $\varphi'(\partial\Delta_0)$. Hence, for suitable C, arguing as in 24.7, $|\alpha([\varphi],u)-1|\leqslant Cu$. Since, by 9.5, Δ_0 is not of D_1-dominant area for q (for a suitable $D_1>0$) by 9.5, for a suitable $\delta>0$ independent of u,

$$d([\varphi],\tau([\varphi]))=d_{\Delta_0}([\varphi'],[\varphi''])\leqslant d([\varphi'],[\varphi''])(1-\delta)\leqslant(1-\delta)u.$$

Since $d([\varphi'],[\varphi''])\leqslant u-\delta e^{-m}$, arguing as in 24.7, we have

$$u-\delta e^{-m}\geqslant d_{A'}([\varphi'],\varphi''])=u-2\pi\,\mathrm{Re}(\beta)+O(e^{-2m}/u)$$

which gives $\mathrm{Re}(\beta)\leqslant -C^{-1}\delta e^{-m}$, as required.

24.9. Proof of 3 of 23.8. — The following lemma gives 3 of 23.8.

Lemma. — *The following holds for a suitable* $C_2 > 0$ *given* C_1. *Let* $\mathrm{Im}(z) = m$. *Let* $d([\varphi], \tau([\varphi])) \leqslant u(1-\delta)$. *Let* $|\alpha([\varphi], z) - 1| \leqslant C_1 u$ *and* $\mathrm{Re}(\beta) \leqslant -C_1^{-1} e^{-m}$. *Then* $\Phi([\varphi], z, u) \leqslant -C_2 e^{-m}$.

Proof. — As usual, write $x' = x'([\varphi], z) = [\varphi']$ and $x'' = x''([\varphi], z, u) = [\varphi'']$, and let $q(z)dz^2$ be the quadratic differential for $d_{Z'}(x', x'')$ at x'. By 8.2, it suffices to find χ with $[\chi \circ \varphi'] = [\varphi'']$ and

$$\int K(\chi)|q| \leqslant e^u - Ce^{-m}$$

for a suitable $C > 0$. We shall find χ using a different pair of quadratic differentials. Recall that, by definition of φ',

$$\sum_{y \in \Delta_0 \cap Y'} \varphi'(y) = 0.$$

For $y \in Z' \cap \Delta_0$,

$$\varphi''(y) = \varphi'(y) - \beta + O(ue^{-m}).$$

As in 24.8, we again have $|\beta| \geqslant C' e^{-m}$ for a suitable $C' > 0$, because the holomorphic map s used to define $[\varphi'']$ from $[\varphi']$ is close to a holomorphic map with an invariant orbit close to $\{\varphi'(g_0^j(v_1)) : j \geqslant 0\}$. Recall that $A' = \{v_1\} \cup Y' \smallsetminus \Delta_0 \subset Y'$ consists of 4 points. Let $[\xi_0]$, $[\xi_u] \in \mathcal{T}(A')$, $q_0(z)dz^2$, $q_u(z)dz^2$ be as in 22.4. Thus $[\varphi']_{A'}$ and $[\varphi'']_{A'}$ are near $[\xi_0]$, $[\xi_u]$, and $\xi_0(v_1) = \xi_u(v_1) = 0$. Let $[\xi'_u] \in \mathcal{T}(A')$ be near $[\xi_u]$ with $\xi'_u(y) = \xi_u(y) = \varphi''(y)$ for $y \in Y' \smallsetminus \Delta_0$, and $\xi'_u(v_1) = -\beta$. By the First Derivative Formula 8.4

$$d([\xi_0], [\xi'_u]) = d([\xi_0], [\xi_u]) + 2\pi \,\mathrm{Re}(\mathrm{Res}(q_0, 0)(-\beta)) + o(\beta) = u - 2\pi \,\mathrm{Re}(\beta) + o(\beta).$$

Let χ_1 be the map minimizing distortion with $[\chi_1 \circ \xi_0] = [\xi'_u]$. Then for z near 0,

$$\chi_1(z) = z - \beta + O(uz) + O(z^2).$$

So for $y \in \Delta_0 \cap Z'$,

$$\chi_1(\varphi'(y)) = \varphi''(y) + O(ue^{-m}) + O(e^{-2m}).$$

So we can choose χ_2 which is the identity except in discs $D(y)$ of radius $O(ue^{-m})$ round the points of $\chi_1 \circ \varphi'(y)$, $y \in Z' \cap \Delta_0$, with χ_2 of bounded distortion in these discs, and such that $[\chi_2 \circ \chi_1 \circ \varphi'] = [\varphi'']$. Suppose that each such disc has q measure $O(ue^{-m})$, that is, $a(D(y), q) = O(ue^{-m})$ for each $y \in Z' \cap \Delta_0$. Then

$$\begin{aligned}(1)\qquad \int K(\chi_2 \circ \chi_1)|q| &\leqslant \int K(\chi_1)|q| + O(ue^{-m}) = e^{u - \mathrm{Re}(\beta) + o(\beta)} + O(ue^{-m}) \\ &\leqslant \exp(u - C_2 e^{-m})\end{aligned}$$

for suitable $C_2 > 0$, as required. It remains to show that $a(D(y), q) = O(ue^{-m})$ for each $y \in Z' \cap \Delta_0$. Let $D(0, r) = \{z : |z| \leqslant r\}$, and let $A(0, r) = D(0, r) \smallsetminus D(0, r/2)$. For this, we need to show that $a(D(0, C_1 e^{-m}) \leqslant C_2 e^{-m}$ for C_2 sufficiently large given C_1.

There is no large modulus annulus homotopic to $\varphi'(\partial\Delta_0)$ satisfying the Pole-Zero Condition of 9.4, by the bound on $|\alpha - 1|$. So it suffices to show (in the notation of 9.4) that $a(\partial\Delta_0, q) \leqslant C_0 e^{-m}$ for some C_0. Suppose that this is not true. Then an annulus A homotopic to $\varphi'(\partial\Delta_0)$ of modulus 1 with $a(A, q)$ boundedly proportional to the minimum possible is distance $\geqslant C_1 e^{-m}$ from 0 for C_1 arbitrarily large given C_0. Then $D(0, C_1 e^{-m})$ has C_2-dominant area, where C_2 can be taken arbitrarily large for C_1 arbitrarily large. Then by 9.5, for a bounded constant C_3 independent of C_2,

$$d(x', x'') \leqslant d_{\Delta_0}(x', x'') + C_3 C_2^{-1} d(x', x''),$$

which gives a contradiction if C_2 is sufficiently large. □

24.12. Proof of 4 of 23.8. — We need to show that a component W'' of $W(u, \delta, C, m)$ is contractible within $W(u, C^{-1}\delta, C^2, m)$, where

$$\{([\varphi], z) : \mathrm{Im}(z) = m,\ d([\varphi], \tau([\varphi])) \leqslant (1-\delta)u,\ |\alpha([\varphi], z) - 1| \leqslant Cu,$$
$$\mathrm{Re}(\beta([\varphi], z)) \geqslant -C^{-1}\delta e^{-m}\} = W(u, \delta, C, m).$$

Using the derivative of $F([\varphi]) = d([\varphi], \tau([\varphi])$, since this is bounded from 0 by 2 of 20.14, we can find a vector field w with $DF(w) \leqslant -\|w\|/C_1$ near W'' and off $\{F([\varphi] = 0\}$ and hence can contract $\{[\varphi] : ([\varphi], z) \in W''$ for some $z\}$ within an $O(u)$ neighbourhood N and within $\{[\varphi] : F([\varphi]) \leqslant (1-\delta)u\}$ to $\{[\varphi] : F([\varphi]) = 0\}$. Then $|\alpha([\varphi]) - 1| \leqslant C_1 C u$ for suitable C_1 by the bound on N. Also $\mathrm{Re}\,\beta([\varphi], z)$ for $([\varphi], z) \in W''$ is within $O(u)$ of an interval of length π with centre β_0. Then we can choose a homotopy

$$h([\varphi], z, t) = (h_1([\varphi], z, t), h_2([\varphi], z, t)) \quad (t \in [0, 1])$$

with

$$h([\varphi], z, 0) = ([\varphi], z), \quad F(h_1([\varphi], z, 1)) = 0, \quad d([\varphi], h_1([\varphi], z, 1)) = O(u).$$

By varying the real part of z_t we can ensure that the argument of $\beta(h([\varphi], z, t)$ is contained in an interval of length $\pi + O(u)$ and $\beta(h([\varphi], z, 1)) = \beta_0$ for all $([\varphi], z) \in W''$. Then finally we can choose $h([\varphi], z, t)$ for $t \in [1, 2]$ so that $h([\varphi], z, 2)$ is constant on W''. □

PART IV

PROOF OF THE RESIDENT'S VIEW

CHAPTER 25

RESIDENT'S VIEW OF RATIONAL MAPS SPACE: OUTLINE PROOF

25.1. The maps with which the Resident views Rational Maps Space

The aim of the last few chapters is to prove the Resident's View of Rational Maps Space. This theorem was stated in 5.10, and slightly reduced in 7.9. The theorem concerns a component V_1 of the space of rational maps $V \subset B = B(Y, f_0)$, and $\widetilde{V}_1$ denotes the universal cover of V_1, which is biholomorphic to the open unit disc D. We now know (Injective on π_1) that the inclusion $V_1 \hookrightarrow B$ is injective on π_1, and, as a consequence, that $\widetilde{V}_1$ can be regarded as a subspace of $\mathcal{T} = \mathcal{T}(Y)$.

We start by extending the definition of ρ_2 — thus recalling the original definition in 1.12. We identify $\widetilde{V}_1$ with the open unit disc D, and write $\pi_2' : D \to V_1$ for the covering map. We also identify the universal cover of $\overline{\mathbf{C}} \setminus Z(f_0) = \overline{\mathbf{C}} \setminus Z$ with D, and write $\pi_2 : D \to \overline{\mathbf{C}} \setminus Z$ for the covering map. We assume without loss of generality that $\pi_2(0) = v_2$. The original map ρ_2 was defined on lifts in $\widetilde{V}_1$ of certain punctures in V_1 — since these were interpreted as paths from a basepoint f_0 to punctures in V_1, where the segment of all paths outside V_1 was the same. So we can assume without loss of generality that that the basepoint of $\mathcal{T}(Y(f_0))$ is $x_0 = [\text{identity}] \in \widetilde{V}_1$, corresponding to the polynomial $f_0 \in V_1$, and $\pi_2'(0) = f_0$. Now we shall define a continuous map $\rho_2 : \mathcal{T}(Y) \to D$, so that the map extends continuously to the original definition at the lifts of some punctures of V_1.

So let $[\varphi] \in \mathcal{T}(Y)$. Then there is a unique χ minimizing distortion up to isotopy constant on Z with $[\chi]_Z = [\varphi]_Z$. Then there is a path $\beta : [0,1] \to \overline{\mathbf{C}} \setminus Z$, with $\beta(0) = v_2$, uniquely determined up to homotopies preserving endpoints, such that

$$[\varphi]_Y = [\chi \circ \sigma_\beta]_Y$$

We then define

$$\rho_2([\varphi]) = \widetilde{\beta}(1),$$

where $\widetilde{\beta}$ is the lift of β with $\widetilde{\beta}(0) = 0$.

In particular, ρ_2 is defined on $\widetilde{V}_1 = D$. We claim that at any lift to $a \in \partial D$ of any puncture of V_1 which represents a degree two rational map, ρ_2 extends continuously to the original definition of $\rho_2(a)$ in 1.12. The reason is that, although this point is on the boundary of $\mathcal{T}(Y)$, it is only a finite distance in $\mathcal{T}(Z)$ from the basepoint. So we can write it in the form $[\chi \circ \sigma_\beta]$ where χ minimizes distortion up to isotopy constant on Z, and β is a path from v_2 to a point of $Z(f_0)$. Then it is clear that $\rho_2(a) = \widetilde{\beta}(1)$ coincides with the original definition. The map ρ_2 on $\mathcal{T}(Y)$ is not G-invariant. But since the definition on lifts of punctures is G-invariant, any continuous extension to ∂D with be $\pi_1(V_1)$-invariant.

The Resident's View of Rational Maps Space (5.10) is given by the following.

25.2. Theorem: View of $\widetilde{V}_1$ via ρ_2

(1) *The map ρ_2 has the following uniform continuity property on D. There is an at most countable set $A \subset \partial D$, and for each $a \in A$ there is a connected subset U_a of* $\mathrm{Int}(D)$ *whose closure intersects ∂D only in a, such that the following holds. Given $\varepsilon > 0$ there is a finite $A' \subset A$ and $\delta > 0$ such that if z, $z' \in D \setminus \cup_{a \in A'} U_a$ and $|z - z'| < \delta$ then $|\rho_2(z) - \rho_2(z')| < \varepsilon$.*

(2) *The set A is precisely the set of lifts of punctures of V_1, together with the end-points of lifts of closed geodesics in V_1, such that a generator g_a of the corresponding cyclic subgroup of $\pi_1(V_1)$ leaves almost invariant a set $\mathcal{T}(\partial P, \varepsilon)$, where P is the fixed set of (f_0, Γ) for a minimal isometric or pseudo-Anosov node $[f_0, \Gamma]$.*

Proof that the Resident's View of 5.10 follows from this. — We shall use several times in this proof the following classical result about pairs of fixed points of the action of a cofinite area discrete group of Möbius transformations on ∂D. The example we have in mind is $\pi_1(V_1)$, which acts by Möbius transformations on $D \cup \partial D$ when we identify D with the universal cover $\widetilde{V}_1$ of V_1. The result is simply that pairs of fixed points of hyperbolic elements of the group of Möbius transformations are dense in $\partial D \times \partial D$. This result is proved, for example, in [**G-H**].

Let $\varphi : \partial D \to \partial D$ be continuous monotone such that $\varphi^{-1}(x)$ is a nontrivial interval if and only if x is a discontinuity of ρ_2, that is, only if $x \in A$. Let

$$X = \{\varphi^{-1}(x) : \varphi^{-1}(x) \text{ is a point}\} = \{\varphi^{-1}(x) : \rho_2 \text{ is continuous at } x\}.$$

Then $\rho_2 \circ \varphi \mid X$ extends continuously to $\overline{X}$, by the uniform continuity property of ρ_2. Then we extend to a continuous map — which we again call $\rho_2 \circ \varphi$ — mapping ∂D into D, as follows. If $\varphi^{-1}(x)$ is singleton, that is, $x \notin A$, then we take $\rho_2 \circ \varphi(\varphi^{-1}(x)) = \rho_2 \circ \varphi(\partial \varphi^{-1}(x))$, since $\partial \varphi^{-1}(x) \subset \overline{X}$. If $\varphi^{-1}(x)$ is nonsingleton, we take $\rho_2 \circ \varphi(\varphi^{-1}(x))$ to be the geodesic in D joining the endpoints of $\rho_2 \circ \varphi(\varphi^{-1}(x))$. This happens only if $x = a$ for some $a \in A$. Only finitely many of these geodesics have over a given length by the uniform continuity property of ρ_2. So we can make $\rho_2 \circ \varphi$ continuous. We can define an action of $\pi_1(V_1)$ by homeomorphisms on ∂D so that $\varphi(g \cdot x) = g \cdot \varphi(x)$ for all $g \in \pi_1(V_1)$, $x \in \partial D$, the second action coming from the identification of D with $\widetilde{V}_1$. Then we also have $\rho_2 \circ \varphi(g \cdot x) = g \cdot \rho_2 \circ \varphi(x)$ for all $g \in \pi_1(V_1)$, $x \in \partial D$, the action on the range being the restriction of the G-action described in 1.13, 3.15. This commutativity follows from 1.13 and continuity of $\rho_2 \circ \varphi$.

For any fixed set P, we write $C(P)$ for the union of convex hulls $C(f_0, \Gamma)$, where (f_0, Γ) has fixed set P. All connected components of the union of Levy convex hulls are of this form. For $a \in A$ corresponding to P, we consider the action of $g_a \in \pi_1(V_1) \leqslant G$ on ∂D, regarding D as the universal cover of $\overline{\mathbf{C}} \setminus Z$. Thus, according to the definition of 1.13, the action of g_a is a lift of an element of the modular group. Note that, by 3.15, g_a fixes $C(P)$, but does not fix any other connected convex hull union. In the proof of 7.10, we saw that, replacing g_a by g_a^{-1} if necessary, for $x \in \partial D$, apart from one point in each component of $\partial D \setminus C(P)$ in the pseudo-Anosov case, $\lim_{n\to+\infty} g_a^n(x)$ exists and is in $\partial C(P)$. It can be shown similarly that for $x \notin \partial C(P)$, $\lim_{n\to\infty} g_a^{-n}(x)$ is the point fixed by g_a in the same component of $\partial D \setminus C(P)$ as x. This other fixed point must be in the boundary of a component of $D \setminus (\cup_{P'} C(P'))$ adjacent to $C(P)$, since each such component boundary does contain a fixed point of g_a outside $\partial C(P)$. We can also consider the action of g_a on ∂D obtained from regarding D as the universal cover $\widetilde{V}_1$ of V_1. Then a is a fixed point of g_a. Let a' be the other. Then $\rho_2 \circ \varphi(\partial(\varphi^{-1}(a)))$, $\rho_2 \circ \varphi(\partial(\varphi^{-1}(a')))$ are in $\partial C(P)$ or are fixed points in boundaries of adjacent components of $D \setminus (\cup_{P'} C(P'))$. The same is true for all points of A. So for any other point $a_1 \in A \setminus \{a, a'\}$, $\rho_2 \circ \varphi(\varphi^{-1}(a_1)) \cap \overline{C(P)} = \varnothing$. Moreover, if $\rho_2 \circ \varphi(x) \in \partial C(P)$ then $\rho_2 \circ \varphi(x) \in \rho_2 \circ \varphi(\varphi^{-1}(a) \cup \varphi^{-1}(a'))$, because otherwise $\rho_2 \circ \varphi(g_a^n x) = \rho_2 \circ \varphi(x)$ for all $n \in \mathbf{Z}$, which contradicts

$$\lim_{n\to\pm\infty} \rho_2 \circ \varphi(g_a^n(x)) \in \rho_2 \circ \varphi(\varphi^{-1}(a)) \cup \rho_2 \circ \varphi(\varphi^{-1}(a')).$$

Now we claim that $\rho_2 \circ \varphi(\varphi^{-1}(a)) \cup \rho_2 \circ \varphi(\varphi^{-1}(a'))$ lies in the closure of a single component of $D \setminus C(P)$. If this is not true, we can find intervals I_1, $I_2 \subset D$ such that $\rho \circ \varphi(I_j)$ lie in different components of $D \setminus C(P)$. But we can find $g \in \pi_1(V_1)$ with fixed points $a_j \in I_j$, $j = 1, 2$, with corresponding fixed set $P' \neq P$. This contradicts our previous finding on the positions of $\rho_2 \circ \varphi(a_j)$ relative to $C(P')$. So now we know that, interchanging a and a' if necessary, $\rho_2 \circ \varphi(\varphi^{-1}(a) \subset \partial C(P)$ and $\rho_2 \circ \varphi(\varphi^{-1}(a'))$ is a point in the boundary of an adjacent component of $D \setminus (\cup_{P'} C(P'))$. So far, it could be in $\partial C(P)$ also. It now follows that all points of $\rho_2 \circ \varphi(\partial D)$ are in the boundary of a single component of $D \setminus (\cup_{P'} C(P'))$ — for if not we can find a fixed set P', $g \in \pi_1(V_1)$ with fixed points a_1, $a_2 \in \partial D$ such that $\rho_2 \circ \varphi(\varphi^{-1}(a_1))$ is separated from $\rho_2 \circ \varphi(\varphi^{-1}(a_2))$ by $C(P')$, giving a contradiction.

Now suppose that $\rho_2 \circ \varphi$ is not monotone. Since the image is a circle, there must be nontrivial disjoint intervals I_1, $I_2 \subset \partial D$ such that $\rho_2 \circ \varphi(I_1) = \rho_2 \circ \varphi(I_2)$. Then we can find $g \in \pi_1(V_1)$ with attractive fixed point x_1 in I_1 and repelling fixed point x_2 in I_2. Then $\rho_2 \circ \varphi(g^n(I_2)) = \rho_2 \circ \varphi(g^n(I_1))$ for all n, hence $\rho_2 \circ \varphi(S^1) = \rho_2 \circ \varphi(x_1)$ is constant. This is not true, because ρ_2 is injective restricted to cusp points in ∂D (1.12). So $\rho_2 \circ \varphi$ must be monotone.

25.3. The Classical Proof. — One classical case in which a map on the open unit disc D extends continuously to the boundary is when the map in question is a quasi-isometry with respect to the Poincaré metric d_P on D. This was used by Mostow [**Mos**] in the proof of his rigidity theorem for most compact symmetric spaces. (Of course, for surfaces, there is no rigidity, but the part of the argument in which the quasi-isometry extends is valid.) Interestingly, in that case, as here, the map involved is a homomorphism for a group action. Our map $\rho_2 : D \to D$ is quite far from being a quasi-isometry, in all cases, and our proof is not particularly close to the classical one, but it is nevertheless worth highlighting some features of the classical proof.

So let $\rho : D \to D$ be a K-quasi-isometry, with $\rho(0) = 0$. For all x, $y \in D$, we have

$$\frac{d_P(x,y)}{K} \leqslant d_P(\rho(x), \rho(y)) \leqslant K d_P(x,y).$$

In particular,

$$d_P(\rho(x), 0) \geqslant \frac{d_P(x,0)}{K}.$$

Using the standard relationship between Poincaré and Euclidean distance, this means that

$$0 < 1 - |\rho(x)| \leqslant 2e^{-d_P(0,x)/K}.$$

This implies that $\rho(x)$ tends to ∂D as x tends to ∂D, uniformly in the Euclidean metric. This is obviously necessary if ρ is to extend (continuously) to map ∂D into ∂D. Now let x, $y \in D$ be such that $d_P(0,z) \geqslant d_P(0,x)$ for all z on the geodesic segment $[x,y]$ between x and y. Then $d_P(0,\rho(z)) \geqslant d_P(0,x)/K$ for all such z, and

$$|\rho(x) - \rho(y)| \leqslant e^{-d_P(x,0)/K} d_P(\rho(x), \rho(y)).$$

By breaking up $[x,y]$ into segments of bounded length, and adding, we deduce for a suitable C that

$$|\rho(x) - \rho(y)| \leqslant Ce^{-d_P(x,0)/K}.$$

The continuous extension of ρ is easily deduced from this inequality, and could clearly be deduced from much less. The essential point is that, if $\rho([x,y])$ is known to be near ∂D, then many different conditions will force it to have small Euclidean diameter, and the relationship between Euclidean and Poincaré metrics might well be helpful.

25.4. The Points of Discontinuity. — We shall formulate the conditions we shall use to prove continuity of ρ_2 in the next paragraph. First, we need to determine the set in which uniform continuity will be proved. We want to allow a countable set A of points of discontinuity on ∂D, at which right and left limits will exist. At each $a \in A$, we shall choose a set U_a, such that $\overline{U}_a \cap \partial D = \{a\}$. The sets U_a will be of two types, which we term *Stoltz angles* and *horoballs*. Given $a \in A \subset \partial D$, there is a cyclic subgroup in $\pi_1(V_1) \leqslant G$ which fixes a. We choose a generator g_a of the subgroup. Thus, a will obviously be in a countable set. As an element of $\pi_1(V_1)$, g_a is either hyperbolic or parabolic. As an element of G, $g_a \in G(f,\Gamma)$, where $[f,\Gamma]$ is either a pseudo-Anosov or an isometric minimal nonempty node. If P is the fixed set of (f,Γ), then, regarding g_a as an element of the modular group $\mathrm{MG}(\overline{\mathbf{C}},Y)$ (which contains G), g_a is the identity off P and is either a pseudo-Anosov on P (in the pseudo-Anosov case) or a Dehn twist round ∂P (in the isometric case).

In both cases, hyperbolic and parabolic, for M_a yet to be determined,

$$U_a = \{x \in D : d_P(x, g_a \cdot x) \leqslant M_a\}.$$

If g_a is hyperbolic then U_a is a *Stoltz angle*, and a is an endpoint of a (unique) geodesic ℓ_a which projects to a closed geodesic in V_1. An equivalent description of U_a in this case, for M_a' depending on M_a, is

$$U_a = \{z : d_P(z, \ell_a) \leqslant M_a'\}.$$

Then U_a is a geodesic-convex set such that ∂U_a meets a in two curves at equal angles $< \pi/2$ with ℓ_a, which is the reason for the term Stoltz angle.

If $a \in A$ is parabolic then U_a is a *horoball*, and a will lie in one of finitely many orbits under $\pi_1(V_1)$, corresponding to punctures in V_1. A *horoball* at a point $a \in \partial D$ is (as usual) a disc contained in D and tangent to D at a. As is well-known, horoballs are defined by the following property: points x and x' are on the boundary of some horoball if and only if the half-geodesics x_t, x_t' parametrised by length with ideal endpoints at a and $x_0 = x$, $x_0' = x'$, remain a bounded distance apart. In fact, $d(x_t, x_t') = O(e^{-t})$.

We shall choose our horoballs so that $U_{g\cdot a} = g \cdot U_a$ for $g \in \pi_1(V)$. This is enough to ensure that, given M, there is an integer N such that at most N horoballs U_a intersect any set of diameter $\leqslant M$, because the horoballs U_a are bounded distance neighbourhoods of smaller horoballs which are actually disjoint. We shall call this a *bounded intersection property*. It also ensures that, for some M, every point in D is distance $\leqslant M$ from some horoball U_a. In fact, large horoballs would cover D. We call this the *bounded covering property*.

25.5. The following reduces the proof of the Resident's View to proving two key conditions.

Continuity Condition Lemma. — *Let $D = \{z : \mid z \mid < 1\}$ and let $\rho_2 : D \to D$ be continuous with $0 = \rho_2(0)$. Let A be a countable set of points of ∂D, with a set $\{U_a : a \in A\}$, where each U_a is a Stoltz angle, or a horoball, such that the horoballs have the bounded intersection and bounded covering property. Then ρ_2 extends continuously to $\partial D \setminus A$, with right and left limits existing at a in $\overline{D} \setminus U_a$, and the Resident's View will be proved, if the following two properties hold.*

(1) Infinity Condition. *Given $\Delta > 0$ there exist $\Delta' > 0$, and a finite set $A' \subset A$ such that if $z \in D \setminus \cup_{a\in A'} U_a$ and $d(z, 0) \geqslant \Delta'$, then $d(\rho_2(z), 0) \geqslant \Delta$.*

(2) Eventually Close Condition. *There is a sequence $\{a_n\}$ with $\lim_{n\to\infty} a_n = 0$ such that the following holds for any $x \in D$. There is a path $\{x_u : u \in [0, T]\}$ in D*

with $x_0 = 0$, $x_T = x$, *such that the* Eventually Close Path Property *holds for the path. The property also holds for any path in* ∂U_a $(a \in A)$ *and any path of diameter* $\leqslant M$. *The limits along a path in* $\partial U_{g\cdot a}$ *are the images under* g *of the limits along a path in* ∂U_a $(g \in \pi_1(V_1))$.

(3) *The* Eventually Close Path Property *for a path* $\{x_t : t \in [0, T]\}$. *Let*

$$t_n = \sup\{t : d(0, \rho_2(x_t)) = n\}.$$

Then

ECPP $$|\rho_2(x_u) - \rho_2(x_{t_n})| \leqslant a_n \quad \text{for all } u \geqslant t_n.$$

Proof. — We need to prove uniform continuity of ρ_2 in the following sense. Given $\varepsilon > 0$, we need to find Δ' and finite $A' \subset A$ such that if z, z' are in the same component of

$$\{z : d_P(0, z) \geqslant \Delta'\} \setminus \bigcup_{a \in A'} U_a,$$

then

(4) $$|\rho_2(z) - \rho_2(z')| < \varepsilon.$$

Let $\{a_n\}$ be as in the statement of Eventually Close. Fix n, to be chosen later, so that a_n is small enough, depending on ε. Then by the Infinity Condition, there is a finite $A' \subset A$, and $\Delta' > 0$ such that if $d(0, z) \geqslant \Delta'$ and $z \in D \setminus \cup_{a \in A'} U_a$, then $d(0, \rho_2(z)) \geqslant n$.

By taking Δ' sufficiently large, we can assume that the sets $U_a \cap \{z : d_P(0, z) \geqslant \Delta'\}$, for $a \in A'$ and U_a a Stoltz angle, are disjoint. We can also assume that, for all $a \in A'$,

$$U_a \cap \{z : d_P(0, z) \leqslant \Delta'\} \neq \varnothing.$$

By adding extra horoballs if necessary, but keeping the above property, we can ensure that, for a fixed M independent of A', every point in $\{z : d_P(0, z) = \Delta'\}$ is distance $\leqslant M$ from some horoball U_a with $a \in A'$. This is because the horoballs in A have the bounded covering property. So then, for M_1 depending only on M whenever a, a' are adjacent points of A' (that is, adjacent on ∂D) we can find a path $\ell_{a,a'} \subset \{z : d_P(0, z) = \Delta'\}$ of length $\leqslant M_1$ (and possibly of length 0) joining points on ∂U_a, $\partial U_{a'}$. Now let R be any component of

$$\{z : d_P(0, z) \geqslant \Delta'\} \setminus \bigcup_{a \in A'} U_a.$$

By the Bounded Intersection Property of the horoballs, there is N_1, independent of A' and Δ', such that $\partial R \cap D$ is a union of $\leqslant N_1$ different connected sets, each of which is either in some $\ell_{a,a'}$ or in some ∂U_a.

So to prove (4), it suffices to show that, for a suitable integer N independent of n, if z, $z' \in R$, then

(5.N) $$|\rho_2(z) - \rho_2(z')| < N a_n.$$

Given $\varepsilon > 0$, we then choose n so that $Na_n \leqslant \varepsilon$. It suffices to assume that $z' \in \partial R \cap D$, since if $(5.N)$ holds in this case, then $(5.2N)$ will hold for any $z' \in R$. Consider the path $\{x_t : t \in [0,T]\}$ satisfying (ECPP) which joins z to 0, with $x_T = z$. Let

$$z'' = x_t \text{ where } t = \sup\{s : x_s \in \partial R\}.$$

Then (5.1) holds with z' replaced by z''. So it suffices to prove $(5.(N-1))$ with z, $z' \in \partial R$. But (ECPP) holds, and hence (5.1) holds, for each component of $\partial R \cap \partial U_a$ or $\partial R \cap \ell_{a,a'}$, any a, $a' \in A$. So then $(5.N_1)$ holds for any z, $z' \in \partial R$, and thus $(5.2N_1)$ holds for any z, $z' \in R$. □

25.6. Reduction to two theorems. — The Infinity Condition for ρ_2 will be deduced from Infinity Condition Theorem. We shall thus deduce the Resident's View of Rational Maps Space from the following two results.

Infinity Condition Theorem. — *Given M there is M' such that the following hold. Suppose that*

$$d_P(0, \rho_2(x)) \leqslant M.$$

Then all but at most length M' of ℓ is in $\mathcal{T}(\partial P, \varepsilon_0)$ for a fixed set P of some (f_0, Γ) satisfying the Invariance, Levy and Maximal Conditions.

Eventually Close Theorem. — *ECPP of the Continuity Lemma holds for our explicit choice of ρ_2: that is, ECPP holds for paths in ∂U_a, for paths of bounded length, and, for any $x \in \widetilde{V}_1$, for at least one path in $\widetilde{V}_1$ joining x to 0.*

25.7. Two Very Basic Lemmas. — Our first basic lemma is simply a very obvious estimate relating d_P, d_Y and d_Z — but is worth spelling out.

First Basic Lemma. — *For a constant $C > 0$ depending only on $\#(Y)$,*

$$d_Y(x_0, x) - d_Z(x_0, x) \leqslant C d_P(0, \rho_2(x)).$$

Proof. — As before, we have $x_0 = [\text{identity}]$, and $x = [\chi \circ \sigma_\beta]_Y$ where χ minimizes distortion up to isotopy constant on Z, and β is a path with initial point at v_2. Then

$$\begin{aligned} d_Z(x_0, x) &= d_Y([\sigma_\beta]_Y[\chi \circ \sigma_\beta]_Y). \\ d_Y(x_0, x) &\leqslant d_Y([\text{identity}]_Y, [\sigma_\beta]_Y) + d_Y([\sigma_\beta]_Y, [\chi \circ \sigma_\beta]_Y) \\ &\leqslant C d_P(0, \rho_2(x)) + d_Z(x_0, x). \end{aligned}$$

□

The following is fairly obvious, given our present state of knowledge, but it is worth spelling out, before we show how to obtain the Infinity Condition from the Infinity Condition Theorem.

Second Basic Lemma. — *Given $\Delta > 0$, there is $\Delta' > 0$ such that the following holds for any x, $y \in \widetilde{V}_1$. If $d_P(x, y) \geqslant \Delta'$, then $d_Y(x, y) \geqslant \Delta$.*

Proof. — Suppose the lemma is false. Then there are Δ and x_n and y_n with $d(x_n, y_n) \geqslant n$ but $d_Y(x_n, y_n) \leqslant \Delta$. Then if one of x_n or y_n is in $(\mathcal{T}(Y))_{<\varepsilon}$ for $\varepsilon > 0$ small enough depending on Δ, they must be in the same component of $\mathcal{T}(Y))_{<\varepsilon'}$. But in such a component, the comparison between d_Y and d_P is straightforward, and so $d_P(x_n, y_n)$ would be bounded. So there is $\varepsilon > 0$ such that x_n, $y_n \in \mathcal{T}(Y)_{\geqslant\varepsilon}$ for all n, and there is $\varepsilon' > 0$ such that x_n, $y_n \in \widetilde{V}_1$ project to $(V_1)_{\geqslant\varepsilon'}$ for all n. Then there are $M > 0$ $g_n, h_n \in \pi_1(V_1)$ and $x \in \widetilde{V}_1$ such that $g_n^{-1} \cdot x_n$ is a distance $\leqslant M$ from some x for all n, and $g_n^{-1} \cdot y_n$ is a distance $\leqslant M$ from $h_n \cdot x$. Then $d(x, h_n \cdot x) \to \infty$ but $d_Y(x, h_n \cdot x) \leqslant \Delta + 2M$ for all n. Then there exist $n \neq m$ with $h_n \neq h_m$ but $d_Y(h_n \cdot x, h_m \cdot x) = 0$, contradicting Injective on π_1. □

25.8. Towards the Infinity Condition. — The following is a step towards showing that the Infinity Condition Theorem implies the Infinity Condition for ρ_2. To get the complete Infinity Condition for ρ_2 we need M' in (1) below replaced by a constant M_a which does not depend on M (although it can depend on a). We shall obtain that refinement in 27.8.

Corollary. — *Suppose that the Infinity Condition Theorem holds. Let $x \in D$ with*

$$d_P(0, \rho_2(x)) \leqslant M.$$

Then there are M' and a finite $A' \subset A$, both depending on M, such that either $d_P(0, x) \leqslant M'$ or, for $g_a \in \pi_1(V_1)$ a generator of the subgroup fixing some $a \in A'$,

$$d_P(x, g_a \cdot x) \leqslant M'. \tag{1}$$

Proof. — Let ℓ denote the geodesic segment in $\mathcal{T}(Y)$ between x_0 and x. By the Infinity Condition Theorem, we deduce that all but length at most M_1 of ℓ lies in $\mathcal{T}(\partial P, \varepsilon_0)$ for the fixed set P of some (f_0, Γ) satisfying the Invariance and Levy, and Maximal Conditions. Because $\mathcal{T}(\partial P, \varepsilon_0)$ is entered by time M_1 (from x_0), there are only finitely many possibilities for ∂P, depending on M_1. In particular, we have a bound, depending on M_1, on $\kappa(P)$. Write $\kappa = \kappa(P)$ if $\kappa(P) > 0$. If $\kappa(P) = 0$, let $\kappa > 0$ be suitably small. Also, we have a bound on $d_Y(y, \tau(y))$ for $y \in \ell \smallsetminus \mathcal{T}(\partial P, \varepsilon_0)$ because $d_Y(x_0, \tau(x_0)) = d_Y(x, \tau(x) = 0$. Let $\alpha : [0,1] \times [0,1] \to \mathcal{T}$ be a homotopy between ℓ and a path in $\widetilde{V}_1$, with $\alpha([0,1] \times \{0\}) \subset \widetilde{V}_1$, $\alpha([0,1] \times \{1\}) \subset \ell$ and α constant on each of the sets $\{s\} \times [0,1]$, $s = 0$, 1. We apply the level κ tool of 7.7 to α, with ∂ equal to the complement in $\partial([0,1] \times [01])$ of two intervals in $[0,1] \times \{1\}$ containing $(0,1)$ and $(1,1)$, whose images under α are bounded with a bound depending only on a given ε_2, and such that $\alpha(\partial \cap ([0,1] \times \{1\})) \subset \mathcal{T}(\partial P, \varepsilon_2)$. Let α' be the new homotopy given by the Level κ tool. We can assume without loss of generality that α' is transverse to $\partial T'_\kappa(\varepsilon_1)$. Now consider the restriction of α' to the complement of the component of $(\alpha')^{-1}(T'_\kappa(\varepsilon_1))$ containing $\partial \cap ([0,1] \times \{1\})$. The restricted domain is topologically a square, and contains all of the original boundary apart from two subintervals of $[0,1] \times \{1\}$ containing $(0,1)$ and $(1,1)$. Reparametrise this restricted domain to a square, keeping the three original sides which are in the restricted domain, to obtain a new homotopy α'' between the original path in $\widetilde{V}_1$, and another path which, apart from two sets of diameter $\leqslant M(\varepsilon_2)$, is contained in some set $K_i(\mu, \varepsilon_1)$ (if $\kappa(P) > 0$) or in $\widetilde{V}_1 \cap T_{<\varepsilon_1}$ if $\kappa(P) = 0$. Write $\ell_1 = \alpha''([0,1] \times \{1\})$. Then

$$\ell_1 \cap \mathcal{T}(\partial P, \varepsilon_1) \subset (\widetilde{V})_{\leqslant \varepsilon_1} \quad \text{if } \kappa(P) = 0,$$
$$\ell_1 \cap \mathcal{T}(\partial P, \varepsilon_1) \subset C \subset K_1(\mu, \varepsilon_1)$$

for some minimal μ with fixed set P if $\kappa(P) > 0$, and component C of $K_1(\mu, \varepsilon_1)$. In both cases, ℓ_1 determines a cyclic subgroup of G with a generator g. If $\kappa(P) = 0$, we take g to be the element given by a simple path round the relevant puncture of V (not yet of V_1) and if $\kappa(P) > 0$ then we take g to be the generator of the cyclic group (whose existence is given by Descending Points in 7.7) which leaves C almost invariant.

Suppose we have the above situation for x_n for a certain number of n, where $d_P(0, x_n) \geqslant n$ but $d_P(0, \rho_2(x_n)) \leqslant M$. Because the intersection of ℓ_1 with $\mathcal{T}(\partial P, \varepsilon_1)$ starts within a bounded distance of the first point of intersection with $\mathcal{T}(\partial P, \varepsilon_1)$, there are only finitely many possibilities for C and g (depending on ε_2). So we can assume that we have the same $g \in G$ of the form above for all x_n. Let $\ell_{1,n}$ be the path as above corresponding to x_n, with second endpoint x_n. Moving the endpoints of $\ell_{1,n}$ a bounded distance (depending on ε_2), we can assume that the endpoints are $h_n \cdot x_0$ and $g^{m_n} h_n \cdot x_0 = x_n$ for a bounded $h_n \in G$. Then if there are sufficiently many n, we can assume after renumbering that $h_1 = h_2 = h$, $m_1 \neq m_2$. Then we have

$$x_2 = h g^{m_2 - m_1} h^{-1} \cdot x_1 \in \pi_1(V_1) \cdot x_1,$$

hence $h g^{m_2 - m_1} h^{-1} \in \pi_1(V_1)$, which is only possible if $hgh^{-1} \in \pi_1(V_1)$, by the Topographer's View 5.10: the detailed description of the group $G = \pi_1(B)$ shows that if $g_1 \in G$ with $g_1^n \in \pi_1(V_1)$ for some n then $g_1 \in \pi_1(V_1)$. Then let $a \in \partial D = \partial \widetilde{V}_1$ be a fixed point of g. Then for a suitable choice of the horoball or Stoltz angle U_a (depending only on M, and hence only on a) we have $x_n \in U_a$ for all n. □

CHAPTER 26

REDUCTIONS TO THE INFINITY CONDITION THEOREM

26.1. The most basic estimate. — In the next two chapters, we prove the Infinity Condition Theorem of 25.6. This chapter is devoted to reducing what has to be proved, to the First and Second Reductions of 26.4, 26.5. The Second Reduction will then be proved in the next chapter.

Throughout this Chapter, $x_0 = [\varphi_0]$, $x_1 = [\varphi_1]$ are points in $\widetilde{V}_1 \subset \mathcal{T}(Y)$, as in Chapter 25, $\ell = [x_0, x_1]$ is the geodesic between them in $\mathcal{T}(Y)$, and ℓ' is the geodesic in $\mathcal{T}(Z)$ joining $\pi_Z(x_0)$ and $\pi_Z(x_1)$. The point x_0 is the basepoint used in the definition of $\rho_2 : \widetilde{V}_1 \to D$ (21.1) with $\rho_2(x_0) = 0$. More generally, given $x_1 = [\varphi_1] \in \widetilde{V}_1$, we can define a function

$$x \longmapsto \rho_2(x, x_1) : \widetilde{V}_1 \longrightarrow D$$

with $\rho_2(x_1) = 0$. This function is uniquely determined up to a Möbius transformation of D.

The following is the most basic estimate in converting boundedness of ρ_2 into other properties. Mostly, we shall apply it with Condition (1), but we shall use Condition (2) to apply it to paths ending a bounded distance from some U_a.

Lemma. — *There is a constant C such that the following hold. Let*

$$d_P(0, \rho_2(x_1)) \leqslant M, \tag{1}$$

or suppose that for some $x_2 \in \mathcal{T}(Y)$ with $\pi_Z(x_2) \in \ell'$,

$$d_P(0, \rho_2(x_2, x_1)) \leqslant M, \quad d_Y(x_0, x_2) \leqslant d_Z(x_0, x_2) + M. \tag{2}$$

Then

$$d_Z(x_0, x_1) = \|\ell'\|_Z \leqslant \|\tau(\ell)\|_Z \leqslant d_Y(x_0, x_1) = \|\ell\|_Y \leqslant d_Z(x_0, x_1) + MC, \tag{3}$$

where $\|\cdot\|_Z$, $\|\cdot\|_Y$ denote length in $\mathcal{T}(Z)$, $\mathcal{T}(Y)$ respectively.

Proof. — To obtain the first two inequalities of (3): for any points x, y we have

$$d_Z(\tau(x), \tau(y)) \leqslant d_Y(x, y),$$

and hence, since the endpoints of $\tau(\ell)$ are $\pi_Z(x_0)$ and $\pi_Z(x_1)$,

$$d_Z(x_0, x_1) \leqslant \|\tau(\ell)\|_Z \leqslant \|\ell\|_Y.$$

Now assume that (1) holds. We can write $x_1 = [\varphi_1]_Y = [\chi_1 \circ \sigma_\beta \circ \varphi_0]_Y$, where $[\chi_1]$ minimizes distortion up to isotopy constant on Z and $\beta : [0,1] \to \overline{\mathbf{C}} \smallsetminus Z$ is a path with $\beta(0) = \varphi_0(v_2)$, $\beta(1) = \chi_1^{-1} \circ \varphi_1(v_2)$ and $\rho_2(x_1) = \widetilde{\beta}(1)$, where $\widetilde{\beta} : [0,1] \to D$ is the lift with $\widetilde{\beta}(0) = 0$. Then

$$\begin{aligned} d_Y(x_0, x_1) &\leqslant d_Y([\chi_1 \circ \sigma_\beta \circ \varphi_0]_Y, [\sigma_\beta \circ \varphi_0]_Y) + d_Y([\sigma_\beta \circ \varphi_0]_Y, [\varphi_0]_Y) \\ &\leqslant d_Z(x_0, x_1) + C d_P(0, \rho_2(x_1)), \end{aligned}$$

which gives the last inequality of (3). If (2) holds, then in a similar way we obtain

$$d_Y(x_2, x_1) \leqslant d_Z(x_2, x_1) + C_1 M,$$

which gives the last inequality of (3) with $C = C_1 + 1$. □

26.2. Lemma. — *Suppose that (1) or (2) of 26.1 holds. Given $M > 0$, $\delta_0 > 0$, there is M_2 such that the following holds, except for $x \in \ell$ in a union of $\leqslant M_2$ segments of length $\leqslant M_2$. Let μ be any segment of ℓ whose endpoints are both within 1 of x. Let y be the endpoint of μ separating μ and x_i ($i = 0$ or 1). Then*

$$\|\mu\| - \|\mu'\| \leqslant \delta_0 \tag{1}$$

$$\|\tau(\mu)\| - \|\mu''\| \leqslant \delta_0, \tag{2}$$

$$d_Z(x, y) + d_Z(y, x_i) - d_Z(x, x_i) \leqslant \delta_0, \tag{3}$$

$$d_Z(\tau(x), \tau(y)) + d_Z(\tau(y), x_i) - d_Z(\tau(x), x_i) \leqslant \delta_0, \tag{4}$$

where μ' is the geodesic in $\mathcal{T}(Z)$ whose endpoints are the projections of the endpoints of μ, and μ'' is the geodesic in $\mathcal{T}(Z)$ whose endpoints are the same as those of $\tau(\mu)$.

Proof. — To obtain (1) and (2), note that if μ_1, μ_2 are any two adjacent segments on ℓ (or $\tau(\ell)$), with $\mu = \mu_1 \cup \mu_2$, and μ', μ_1', μ_2' are the segments in $\mathcal{T}(Z)$ with endpoints the projections of the endpoints of μ, μ_1, μ_2, then

$$\|\mu\|_Y - \|\mu'\|_Z \geqslant (\|\mu_1\|_Y - \|\mu_1'\|_Z) + (\|\mu_2\|_Y - \|\mu_2'\|_Z).$$

Now write ℓ as a union of segments μ of length $\leqslant 3$ such that any $x \in \ell$ is in at most three segments, but, apart from length 1 at the endpoints of ℓ, any x lies in at least one such segment μ_1, distance $\geqslant 1$ from both endpoints of μ_1. Then

$$\sum_{\mu_1} (\|\mu_1\|_Y - \|\mu_1'\|_Z) \leqslant 3(\|\ell\|_Y - \|\ell\|_Z) \leqslant 3CM,$$

and similarly for the $\tau(\mu_1)$, μ_1''. So for all but $O(M/\delta_0)$ of these segments, any subsegment μ and corresponding $\tau(\mu)$ will satisfy (1) and (2). This gives (1) and (2) for μ within 1 of x, except for x in a union of $O(M/\delta_0)$ segments of total length $O(M/\delta_0)$.

The inequality (3) is really two inequalities, one for each x_i. We obtain each one except for a finite union of segments, bounded in terms of M, and δ_0. Without loss of generality, we consider the inequality for x_0. Then we write ℓ as a union of segments as before. Fix any segment μ_1, where μ_1 is divided into segments of length $< \delta_0/8$ and the successive points are y_i, $0 \leqslant i \leqslant n$. Then write

$$S(\mu_1) = \sum_{i=1}^{n} (d_Z(x_0, y_{i-1}) + d_Z(y_{i-1}, y_i) - d_Z(x_0, y_i)).$$

Then all the terms in each sum $S(\mu_1)$ are $\geqslant 0$, and

$$\sum_{\mu_1} S(\mu_1) \leqslant 3CM.$$

So for all but $\leqslant 12CM/\delta_0$ segments, we have $S(\mu) \leqslant \delta_0/4$ Then for y_p and y_q in a good μ, $p < q$, by summing from $p+1$ to q, we have

$$d_Z(x_0, y_p) + d_Z(y_p, y_q) - d_Z(x_0, y_q) \leqslant \delta_0/4.$$

Then for any x and $y \in \mu$ with x to the right of y we have

$$d_Z(x_0, y) + d_Z(y, x) - d_Z(x_0, x) \leqslant \delta_0.$$

Then (4) is proved in a similar way. □

26.3. Notation for the Reductions. — In what follows, we shall frequently refer to gaps without having introduced a loop set. A gap is simply a nonannular subsurface β of $\overline{\mathbf{C}}$ with $\partial\beta \cap Y = \varnothing$, such that all boundary components are nontrivial and nonperipheral. We allow the possibility $\beta = \overline{\mathbf{C}}$. We shall refer to periodic gaps and loops without having introduced an invariant loop set. A gap or loop is *periodic with orbit* $[\beta]$ if there exists a sequence β_i, $1 \leqslant i \leqslant n$, with $\beta = \beta_1 = \beta_{n+1}$, $\beta_i \cap \beta_j = \varnothing$ for $1 \leqslant i < j \leqslant n$, and β_i is a component of $f_0^{-1}(\beta_{i+1})$ up to Z preserving isotopy. We then write $[\beta] = \{\beta_i : 1 \leqslant i \leqslant n\}$, as in 17.3.

We use the notations $m_\beta(x)$ of 9.1, $a(\alpha, q)$ of 9.4, $|\varphi(\gamma)|_q$ of 14.5. Having specified a geodesic segment $\ell_1 \subset \mathcal{T}(Y)$, we shall write $I(\beta, C_1, \varepsilon)$ for the set of $x = [\varphi] \in \ell_1$ such that, if $q(z)dz^2$ is the quadratic differential at x for $d_Y(x_0, x_1)$, then

$$|\varphi(\partial\beta)|_q < \varepsilon,$$

$$a(\beta, q) \geqslant C_1 \quad \text{or} \quad a(\beta, q) \geqslant C_1/m_\beta(x),$$

depending on whether β is a gap or loop. (We have $\beta = \partial\beta$ is β is a loop.) In such circumstances (as in the past) we shall write $a(\beta, q) = a(\beta, x)$, where it is clear which quadratic differential is meant. If β is a gap, then we define $I(\beta, C_1, \varepsilon, \nu) \subset I(\beta, C_1, \varepsilon)$ to be the subset of x such that, in addition, $|\varphi(\gamma)|_q \geqslant \nu$ for all $\gamma \subset \text{Interior}(\beta)$. (This is an empty condition if β is a loop.) Note that if $x \in I(\beta, C_1, \varepsilon, \nu)$, then $x \in \mathcal{T}(\partial\beta, \delta)$, where $\delta = \delta(\varepsilon, C_1, \nu) \to 0$ as $\varepsilon \to 0$, for fixed C_1, ν, and $x \in (\mathcal{T}(A(\beta))_{\geqslant\eta}$, where $\eta = \eta(C_1, \nu)$ is bounded from 0 if ν and C_1 are.

We say that α is *workable* if

(i) α is a periodic loop or gap

(ii) if α is a loop then it generates a loop set satisfying the Levy Condition;

(iii) and if α is a gap then $v_2 \notin \alpha$ and it is irreducible and homeomorphic pseudo-Anosov or degree two nonrational.

26.4. The First Reduction. — *There is $C_1' = C_1'(M, \varepsilon_0, \eta_0) > 0$ such that the following hold. Let ℓ_1 be connected. Let (1) of 26.1 hold and $\ell_1 \subset \ell$, or let (2) of 26.1 hold and $x_1 \in \mathcal{T}_{\geqslant \varepsilon_0}$, $\ell_1 \subset [x_2, x_1]$ for x_2 as in (2) of 26.1. For $x \in \ell_1$, let (1)-(4) of 26.2 hold, for δ_0 sufficiently small given M, ε_0, η_0.*

Then there is a workable gap or loop α such that

$$\ell_1 \subset I(\alpha, C_1', \eta_0).$$

Remark. — The precise formulation of the First Reduction is made so as to imply the Infinity Condition both for paths for which $\|\ell\|_Y - \|\ell'\|_Z$ is large, and for paths in a set ∂U_a, for which $\|\ell\|_Y - \|\ell'\|_Z$ is bounded. From 26.2 it follows that if (1) of 26.1 holds then all but length $\leqslant M_2$ of ℓ is in $\leqslant M_2$ segments such as ℓ_1. We shall deduce the Infinity Condition from this in 27.7-8.

26.5. The Second Reduction. — *Let the same hypotheses hold for x_0, x_1, and ℓ_1 as in the First Reduction. Then there is a function $r : (0, \infty) \to (0, \infty)$ such that the following hold. Let $x \in I(\beta, C_1, r(\nu), \nu) \subset \ell_1$ for $\nu \geqslant \nu_1$ and periodic β. If δ_0 sufficiently small given M, C_1, ν_1, then β is workable.*

The following two lemmas explain why the Second Reduction implies the First.

26.6. Lemma. — *Let the same hypotheses hold for x_0, x_1, and ℓ_1 as in the First Reduction. Let $C_1 > 0$ and a function $r : (0, \infty) \to (0, \infty)$ be given. Then if $\delta_0 > 0$ is sufficiently small given M, C_1, r, there are $\nu_1 = \nu_1(M, C_1, \varepsilon_0, \eta_0, r) > 0$, $C_2 = C_2(M, C_1) > 0$ such that ℓ_1 is a union of sets $I(\beta, C_2, r(\nu), \nu)$ for periodic β (including $\beta = \overline{\mathbf{C}}$) and $\nu \geqslant \nu_1$.*

This will be proved in 26.12. It allows us to replace the hypotheses in the First Reduction by the hypotheses in the Second.

26.7. Lemma. — *Let the same hypotheses hold for x_0, x_1, and ℓ_1 as in the First Reduction. The following holds for a suitably chosen η_1 given η_0, M, and C_1' given M and C_1. Take any segment ℓ_2 of ℓ_1 such that ℓ_2 is a (connected) union of sets $I(\beta, C_1, \eta_1)$. Then if δ_0 is sufficiently small given M, ε_0, C_1', then there is a single workable α_0 such that*

$$\ell_2 \subset I(\alpha_0, C_1', \eta_0).$$

This will be proved in 26.13. It allows us to replace the conclusion of the Second Reduction by the conclusion of the First.

26.8. Area in the geodesics ℓ and ℓ'. — See 14.10 for the definition of d'_α. We write $d'_{\alpha,Z}$ when the underlying Teichmüller space is $\mathcal{T}(Z)$. The first step in analyzing subsurfaces α with $x \in \mathcal{T}(\partial\alpha, \varepsilon_0) \cap \ell$ and $a(\alpha, x)$ bounded from 0 is the following.

Lemma. — *Given $M > 0$, $C_1 > 0$, $\nu > 0$, there is $M_1 = M_1(M, \nu, C_1)$ such that the following hold. Let x_0, x_1 and ℓ_1 satisfy the hypotheses of the First Reduction. Let $x \in I(\alpha, C_1, \varepsilon, \nu) \subset \ell_1$ for some $\varepsilon \leqslant \nu$. Then*

$$d_{\alpha,Z}(x, \tau(x)) \leqslant M_1, \tag{1}$$

$$d'_{\alpha,Z}(x_i, x) \geqslant d_Y(x_i, x) - M_1, \tag{2}$$

for $i = 0$ or 1, depending on whether (1) or (2) of 26.1 holds.

Proof. — First, we want to apply 14.2 to the geodesic segment $\ell = [x_0, x_1]$ and a point x' on one of the following unions of geodesic segments in $\mathcal{T}(Y)$, depending on what hypotheses we are using:

$$[x_0, x'_0] \cup [x'_0, x_1]$$

where

$$x_0 = [\varphi_0], \quad x'_0 = [\sigma_\beta \circ \varphi_0]_Y, \quad x_1 = [\varphi_1]_Y = [\chi_1 \circ \sigma_\beta \circ \varphi_0]_Y$$

where β is a path starting at $\varphi_0(v)$ and χ_1 minimizes distortion up to isotopy constant on $\varphi_0(Z)$, or

$$[x_0, x_2] \cup [x_2, x'_1] \cup [x'_1, x_1]$$

where

$$x_0 = [\varphi_0], x_2 = [\chi_2 \circ \sigma_\alpha \circ \varphi_0], x'_1 = [\chi_1 \circ \chi_2 \circ \sigma_\alpha \circ \varphi_0], x_1 = [\sigma_\beta \circ \chi_1 \circ \chi_2 \circ \sigma_\alpha \circ \varphi_0].$$

Here, $[\varphi_0]$, $[\chi_2 \circ \varphi_0]$, $[\chi_1 \circ \chi_2 \circ \varphi_0]$ are on the same geodesic in $\mathcal{T}(Z)$, and χ_1, χ_2 minimize distortion up to isotopies constant on $\chi_2 \circ \varphi_0(Z)$, $\varphi_0(Z)$ respectively. The path β is bounded in both cases, giving

$$d_Y(x_0, x'_0) \leqslant CM, \quad d_Y(x_1, x'_1) \leqslant CM.$$

By the definitions, we have, under the first hypothesis

$$d_Y(x'_0, x_1) = d_Z(x_0, x_1),$$

and under the second hypothesis,

$$d_Y(x_0, x_2) \leqslant d_Z(x_0, x_2) + M, \; d_Y(x_2, x'_1) = d_Z(x_2, x_1).$$

Then for any $x' \in [x'_0, x_1]$ in case (1) of 26.1, or $x' \in [x_2, x_1]$ in case (2) of 26.1, we have

$$d_Y(x_0, x') + d_Y(x', x_1) \leqslant d_Z(x_0, x_1) + (1 + 2C)M.$$

Given $x \in \ell$ we can choose $x' \in [x'_0, x_1]$ or $x' \in [x_2, x'_1]$ with

$$|d_Y(x_0, x) - d_Y(x_0, x')| \leqslant CM.$$

So, since $a(\alpha, x') \geqslant C_1$, we can apply 14.2 to obtain, for suitable M_2,

$$d_{\alpha,Y}(x', x) \leqslant M_2.$$

Projecting, this gives

$$d_{\alpha,Z}(x', x) \leqslant M_2.$$

Then

$$d'_{\alpha,Z}(x', x_0) = d'_{\alpha,Y}(x', x'_0) \geqslant d'_{\alpha,Y}(x, x_0) - M_2 - CM \geqslant d_Y(x, x_0) - M_3 \tag{3}$$

for suitable M_3, under the first hypothesis. The last inequality uses $a(\alpha, q) \geqslant C_1$. Similarly, under the second hypothesis, if $x' \in [x'_2, x_1]$ we obtain

$$d'_{\alpha,Z}(x', x_1) \geqslant d_Y(x, x_1) - M_3. \tag{4}$$

It follows that there are $\eta = \eta(M, C_1, \nu)$, $C'_1 = C'_1(M, C_1, \nu)$, and α' with $\alpha \subset \alpha'$ (possibly $\alpha' = \overline{\mathbf{C}}$) such that $\pi_Z(x') \in \mathcal{T}(\partial\alpha', \eta)$, $\pi_Z(x') \in (\mathcal{T}(A(\alpha'))_{\geqslant \eta}$ if α' is a gap, and $a(\alpha', r) \geqslant C'_1$ or $a(\alpha', r) \geqslant C_1/m_{\alpha'}(x')$, where $r(z)dz^2$ is the quadratic differential for $d_Z(x_0, x')$ or $d_Z(x_1, x')$ at x'. These lower bounds on $a(\alpha', r)$ follow from (3) or (4), and $x_0 \in \mathcal{T}_{\geqslant \varepsilon_0}$ or $x_1 \in \mathcal{T}_{\geqslant \varepsilon_0}$, depending on whether the first or second hypothesis is used. For let $x' = [\varphi']$ and let $r_i(z)dz^2$ be the stretch of $r(z)dz^2$ at x_i. First let α' be a gap, and let $\gamma \subset \mathrm{int}(\alpha')$ with $|\varphi'(\gamma)|$ bounded. Then for $i = 0$ or 1 (depending on which hypothesis we are working under) both $|\varphi_i(\gamma)|_{r_i}$ and $|\varphi_i(\gamma)|_{r_i}/|\varphi'(\gamma)|_r$ are boundedly proportional to $e^{d_Z(x', x_i)}$, which means that $a(\alpha', r)$ must be bounded from 0. If α' is a loop, then $\alpha' = \alpha$, and we argue similarly, but replacing x' by $x'' = [\varphi''] \in [x', x_i]$ such that $\varphi''(\alpha)$ has length ε_0, and take $\gamma = \alpha$.

Then we can apply 14.2 to $[x'_0, x_1] \subset \mathcal{T}(Z)$, or $[x_2, x'_1] \subset \mathcal{T}(Z)$, with each of $\pi_Z(x)$, $\pi_Z(\tau(x))$ and with α', to obtain

$$d_{\alpha',Z}(x, x') \leqslant M_4,\ d_{\alpha',Z}(\tau(x), x') \leqslant M_4.$$

Then we have (1) and (2) of the statement of this lemma, as required. □

26.9. A good gap is a preimage of a good gap

Lemma. — *Take the same hypotheses as in 26.8. Then there are constants $\nu' = \nu'(M, C_1, \nu)$, $C'_1 = C'_1(M, C_1) > 0$, $\varepsilon' = \varepsilon'(M, C_1, \nu, \varepsilon)$ which $\to 0$ as $\varepsilon \to 0$, such that the following holds. There is $\alpha' \subset \overline{\mathbf{C}} \setminus Y$ such that α is contained in a component of $f_0^{-1}(\alpha')$ and $x \in I(\alpha', C'_1, \varepsilon', \nu')$.*

Proof. — In what follows, we work under the first hypothesis of 26.8. If working under the second hypothesis, we replace x_0 by x_1. By 26.8 we have, for suitable $M_1 = M_1(M, \nu, C_1)$,

$$d'_{\alpha,Z}(x_0, \tau(x)) = (d'_{\alpha,Z}(x_0, \tau(x)) - d'_{\alpha,Z}(x_0, x)) + d'_{\alpha,Z}(x_0, x) \geqslant d_Y(x_0, x) - 2M_1$$

If γ is a loop with $x \in \mathcal{T}(\gamma, \eta)$, then $\tau(x) \in \mathcal{T}(f_0^{-1}(\gamma), 4\eta)$ and if α is a gap, then $f_0^{-1}(\gamma) \cap \mathrm{Interior}(\alpha) = \varnothing$, if η is sufficiently small given ν, C_1 and M, because otherwise by 26.8 $x \in \mathcal{T}(f_0^{-1}(\gamma), \nu)$. So there are $\nu' = \nu'(M, C_1, \nu)$, α' with $\alpha \subset f_0^{-1}(\alpha')$, and ε' which $\to 0$ as $\varepsilon \to 0$ for fixed M, C_1, ν, such that $x \in I(\alpha', a(\alpha', q), \varepsilon', \nu')$.

It remains to show that $a(\alpha', q) \geqslant C_1' = C_1'(M, C_1)$ or $a(\alpha', q) \geqslant C_1'/m_{\alpha'}(x')$, depending on whether α' is a gap or a loop. To show that C_1' is independent of ν, note that we can subdivide α and hence could assume that $\nu = \varepsilon_0$, for some fixed ε_0. We claim that there is a constant $M_3 = M_3(M, C_1, \nu)$ such that

$$d'_{\alpha,Z}(x_0, \tau(x)) \leqslant d_{\alpha',Y}(x_0, x) + M_3 \tag{1}$$

We need to consider the definitions. Let s_0, s be the holomorphic branched coverings with branchpoints at $\varphi_0(v_j)$, $\varphi(v_j)$, $j = 1, 2$. Then, if α is a gap, $\exp(d'_{\alpha,Z}(x_0, \tau(x)))$ is given up to a bounded multiple by the maximum of $|s^{-1}(\varphi_0(\gamma))|'$ (see 14.10) over loops $\gamma \subset \alpha'$ where $1/D_1 \leqslant |\varphi(\gamma)|' \leqslant D_1$ for $D_1 = D_1(\nu') = D_1(M, C_1, \nu)$. But for a constant D_2,

$$|s^{-1} \circ \varphi_0(\gamma)|' \leqslant D_2 |\varphi_0(\gamma)|' \leqslant D_2 \exp(d'_{\alpha',Y}(x_0, x)) |\varphi(\gamma)|', \tag{2}$$

which gives (1) if α is a gap. If α is a loop, then so is α'. Choose $x' = [\varphi'] \in [x, x_0]$ such that $\varphi'(\alpha)$ has length ε_0. Then by 26.8, $d_\alpha(x, \tau(x))$ is bounded, so $\varphi'(\alpha')$ also has length boundedly proportional to ε_0. Then we obtain (2) above for $\gamma = \alpha'$. So we have, for a suitable $M_2 = M_2(M, C_1, \nu)$,

$$d'_{\alpha',Y}(x_0, x) \geqslant d_Y(x_0, x) - M_2.$$

Then since $x_0 \in \mathcal{T}_{\geqslant \varepsilon_0}$, arguing as in 26.8, we have $a(\alpha', q) \geqslant C_1'$ or $a(\alpha', q) \geqslant C_1'/m_{\alpha'}(x)$ for $C_1' = C_1'(M, C_1)$, as required. □

26.10. Corollary. — *Take the same hypotheses as in 26.8. Let a function $r' : (0, \infty) \to (0, \infty)$ be given. then the following holds for a suitable function r. Let $x \in I(\alpha, C_1, r(\nu), \nu)$. Then there are $C_1' = C_1'(M, C_1)$, $\nu_1' = \nu_1'(M, C_1, \nu, r') > 0$, $n \geqslant 0$, $k > 0$ and α_i $(1 \leqslant i \leqslant n < n + k)$ such that*

$$\alpha_1 = \alpha, \quad \alpha_n = \alpha_{n+k} \quad \text{but otherwise } \alpha_i \cap \alpha_j = \varnothing \text{ for } i \neq j,$$

α_i is contained in a component of $f_0^{-1}(\alpha_{i+1})$ for $1 \leqslant i < n+k$, and is such a component for $n \leqslant i < n + k$, and $x \in I(\alpha_i, C_1', r'(\nu_i), \nu_i)$ for some $\nu_i \geqslant \nu_1'(M, C_1, \nu, r')$.

Proof. — Given r', write $r' = r_N$ for a suitable $N \leqslant \#(Y)$, and we can construct functions r_i $(i \leqslant N)$, with r_i chosen given r_{i+1}, such that the following holds. We proceed inductively using 26.9, constructing α_{i+1} with $x \in I(\alpha_{i+1}, C_{i+1}, r_{i+1}(\nu_{i+1}), \nu_{i+1})$ for $C_{i+1} = C_{i+1}(M, C_i)$, $\nu_{i+1} \geqslant \nu'_{i+1}(M, C_i, \nu_i, r_{i+1})$, such that α_i is contained in a component of $f_0^{-1}(\alpha_{i+1})$. We can ensure that

$$r_i(\nu_i) < \nu_j \quad \text{for } i < j,$$

and hence that if $\alpha_i \cap \alpha_j \neq \varnothing$ then $\alpha_i = \alpha_j$. □

26.11. Lemma. — *Let the same hypotheses hold for x_0, x_1, and ℓ_1 as in the First Reduction. Let $C_1 > 0$ and $\eta_0 > 0$ be given. Then there is are constants $C_2 = C_2(M, C_1)$ and $\eta_1 = \eta_1(M, C_1, \varepsilon_0, \eta_0)$ such that the following holds. Let β be a periodic gap or loop with $v_2 \notin \beta'$ for $\beta' \in [\beta]$, and let $I(\beta, C_1, \eta_0) \subset \ell_1$ be sufficiently long. Then, for some workable α (see 26.3),*

$$I(\beta, C_1, \eta_1) \subset I(\alpha, C_2, \eta_0, \eta_0).$$

Proof. — Write $\{\beta_i : 1 \leqslant i \leqslant n\} = [\beta]$, $\beta = \beta_1 = \beta_{n+1}$, $\beta_i \subset f_0^{-1}(\beta_{i+1})$. First, suppose that β is a gap. Write $\pi_{[\beta]}(y) = (\pi_{\beta_i}(y))$. Then, as in 20.10, if $y \in \cap_i \mathcal{T}(\partial\beta_i, \varepsilon_0)$,

$$\pi_{[\beta]}(\pi_{\beta_i}(\tau(y)) = \tau_{[\beta]}(\pi_{\beta_i}(y)) + O(e^{-2\pi^2/\varepsilon_0}),$$

where

$$\tau_{[\beta]} : \prod_{i=1}^{n} \mathcal{T}(A(\beta_i)) \longrightarrow \prod_{i=1}^{n} \mathcal{T}(A(\beta_i)) : ([\varphi_i]) \longmapsto ([s_{i+1}^{-1} \circ \varphi_{i+1} \circ f_i])$$

is either the action of the appropriate element of the modular group or the pullback for the appropriate Teichmüller space for a critically finite branched covering on a finite disjoint union of spheres. (Compare with Chapter 20.) Here s_i is a holomorphic branched covering (of degree one for all but at most two i, when it is degree two) and

$$\psi_{[\beta]} : (z, i) \longmapsto (f_i(z), i+1),$$

that is

$$\psi_{[\beta]} : (\overline{\mathbf{C}} \times \{1 \cdots n\}, \cup_i(A(\alpha_i) \times \{i\}) \longrightarrow (\overline{\mathbf{C}} \times \{1 \cdots n\}, \cup_i(A(\alpha_i) \times \{i\})$$

is a critically finite branched covering of a union of spheres, which is, again, of degree one in all but at most two coordinates.

By 26.10, if ν_0 is sufficiently small given ν_1, M, C_1, and for $C_1' = C_1'(M, C_1)$ $I(\beta, C_1, \eta_0) \setminus I(\beta, C_1, \eta_0, \nu_0)$ is a union of sets $I(\alpha, C_1', \nu_1, \nu_1)$ for periodic α. By 26.8 we then have

$$d_{\alpha'}(x, \tau(x)) \leqslant M_1(M, \nu_1, C_1)$$

for $\alpha' \in [\alpha]$, which gives a bound on $d_{\alpha'}(\pi_{[\beta]}(x), \pi_{[\beta]}(\tau_{[\beta]}(x)))$. So $\partial[\alpha]$ would then have to be periodic under $\psi_{[\beta]}$, and α would have to be a union of irreducibles. In particular, there are only finitely many possibilities for α.

We need to show that if $\psi_{[\beta]} \mid [\beta]$ is reducible then any component of $I(\beta, C_1, \eta_0, \nu_0)$ has length bounded above in terms of C_1, M and ν_0. This argument can also be applied to any $[\beta'] \subset [\beta]$ with similar hypotheses. So let $y \in I(\beta, C_1, \eta_0, \nu_0)$ and let β be reducible. We have

$$F_{[\beta]}(y) = \sum_j d_{\beta_j}(\pi_{[\beta]}(y), \tau_{[\beta]}(\pi_{[\beta]}(y))) \leqslant M_1(M, C_1).$$

By 20.14, $F_{[\beta]} \mid I(\beta, C_1, \eta_0)$ is bounded only in $I(\beta, C_1, \eta_0, \nu_0)$ and the union of sets $I(\beta', C_1', \nu_1, \nu_1)$ for some $[\beta'] \subset [\beta]$, $[\beta'] \neq [\beta]$ periodic under $\psi_{[\beta]}$, and with ν_1 bounded from 0. So if a component of $I(\beta, C_1, \eta_0, \nu_0)$ is sufficiently long, we obtain points

$y = [\psi]$, $y' = [\psi']$ a long way apart with $|\psi(\partial[\beta'])|$ and $|\psi'(\partial[\beta'])|$ bounded for the same $[\beta']$. Then $[y, y']$ must enter $I(\beta', C_1', \nu_1)$ for some $[\beta']$.

It follows that for suitable η_1 depending on M, C_1, η_0, and C_2 depending on M, C_1, there is an irreducible α with $I(\beta, C_1, \eta_1) \subset I(\alpha, C_2, \eta_0)$. So, applying 26.8, $d_{\alpha_i}(y, \tau(y)) \leqslant M_1(M, \eta_0, C_1)$ for all $y \in I(\alpha, C_2, \eta_0)$ and all i. Assuming that $I(\beta, C_1, \eta_0)$ is long enough, $\psi_{[\alpha]}$ cannot be isotopic to an isometry, because then the set of $\pi_{[\alpha]}(y)$ for which $d(\pi_{[\alpha]}(y), \tau_{[\alpha]}(\pi_{[\alpha]}(y)))$ is $\leqslant M_2$ is bounded. For the same reason $\psi_{[\alpha]}$ cannot be Thurston equivalent to a rational map if it has degree > 1. So considering the classification of $\psi_{[\alpha]}$ (2.14-18), the degree of $\psi_{[\alpha]}$ is at most two. For all but at most one i, $f_i : (\overline{\mathbf{C}}, A(\alpha_i) \to (\overline{\mathbf{C}}, A(\alpha_{i+1}))$ is a homeomeorphism. So $\psi_{[\alpha]}$ is either pseudo-Anosov up to isotopy or a critically finite irreducible nonrational degree two branched covering up to Thurston equivalence.

Now suppose that α is a loop. Write $\mu = I(\alpha, C_2, \eta_0)$. Then each space $\mathcal{T}(A(\alpha_i))$ is the upper half-plane H. If α is not in the full orbit of a Levy cycle then a map $\tau_{[\alpha]} : H^n \to H^n$ is well-defined, such that $\pi_{[\alpha]}(\tau(x)$ is approximately $\tau_{[\alpha]}(\pi_{[\alpha]}(x))$, and $\tau_{[\alpha]}^n((x_i)) = (\lambda x_{i+1})$ where $\lambda = \frac{1}{2}$ or $\frac{1}{4}$, for a suitable identification of $\mathcal{T}(A(\alpha_i))$ with H. Then in order to achieve the bound $d(\pi_{[\alpha]}(x), \pi_{[\alpha]}(\tau(x))) \leqslant M_1$, we must have $\pi_{\alpha_i}(x)$ near the imaginary axis for all i. This contradicts $a(\alpha, y) \geqslant C_1$ for all $y \in \mu$, if μ is sufficiently long (that is, if $I(\beta, C_1\eta_1)$ is sufficiently long), because then the two endpoints of $\pi_\alpha(\mu)$ are far apart in H. Since these endpoints are both in $\{z : 1 \leqslant \mathrm{Im}(z) \leqslant \Delta_0\}$, for $\Delta_0 = \Delta_0(M, \eta_0, C_1)$, they cannot both be near the imaginary axis, giving the required contradiction. □

26.12. Proof of 26.6. — Fix $x \in \ell_1$, and let a function r_2 be given. We can find a gap or loop β_1 (possibly with $\beta_1 = \overline{\mathbf{C}}$) such that $x \in I(\beta_1, 1/\#(Z), r_2(\nu'), \nu')$ for $\nu \geqslant \nu_2 = \nu_2(M, r_2)$. We can assume without loss of generality that $r_2(\nu) \leqslant \eta_0$ for all ν where the existence of a loop of $q - d$ length $< \eta_0$ for x implies $x \in \mathcal{T}_{<\varepsilon_0}$ where ε_0 is the Margulis constant. Then we can apply 26.10 to find a periodic loop or gap so that $x \in I(\beta, C_1, r(\nu), \nu)$ if r_2 is suitably chosen given r, with ν bounded below in terms of ν_2. □

26.13. Proof of 26.7. — Take any segment ℓ_2 of ℓ_1 such that ℓ_2 is a (connected) union of sets $I(\beta, C_1, \eta_1)$. By 26.12, if η_1 is suitably chosen given η_1', and C_1' suitably chosen given M and C_1, ℓ_2 is a connected union of sets $I(\alpha, 2C_1', \eta_1)$ for workable α (see 26.3). We need to show that for a suitable workable α_0,

$$\ell_2 \subset I(\alpha, C_1', \eta_0)$$

Let $x \in I(\alpha_1, 2C_1', \eta_1) \cap I(\alpha_2, 2C_1', \eta_1)$. Let $\Gamma_i \supset [\alpha_i]$ be equivalent to the invariant loop set generated by $[\alpha_i]$ which satisfies the Levy condition, such that every periodic gap of Γ_i is irreducible. Let $\Delta([\alpha_i])$ be the component of $\overline{\mathbf{C}} \setminus (\cup[\alpha_i])$ containing v_1 and v_2. Then either $\Delta([\alpha_2]) \subset \Delta([\alpha_1])$ or $\Delta([\alpha_1]) \subset \Delta([\alpha_2])$. Assume without loss of generality that $\Delta([\alpha_2]) \subset \Delta([\alpha_1])$. Then Γ_2 can be extended to an invariant loop set containing $\partial[\alpha_1]$, without changing the gap containing v_2. As long as a gap boundary remains short along a geodesic, the area enclosed with respect to the quadratic differential along the geodesic does not change much. It follows that, for μ_1, μ_2 are segments of length $\leqslant M_1 = M(\eta_1)$ around the endpoints of the union,

$$I(\alpha_1, 2C_1', \eta_1') \cup I(\alpha_2, 2C_1', \eta_1') \setminus (\mu_1 \cup \mu_2) \subset I(\alpha_1, 3C_1'/2, \eta_1'),$$

and, if η_1' is small enough given η_0, and η_0 is small enough,

$$I(\alpha_1, 2C_1', \eta_1') \cup I(\alpha_2, 2C_1', \eta_1') \subset I(\alpha_1, C_1', \eta_0).$$

Then we can repeat the argument for adjacent segments to this union, obtaining a successively longer segment in $I(\alpha_1', 3C_1'/2, \eta_1')$ for varying workable α_1', and a slightly longer union of segments in $I(\alpha_1', C_1', \eta_0)$. The constants η_1', C_1' do not deteriorate because $M(\eta_1')$ can be used each time. So finally we obtain, for a suitable workable α_0,

$$\ell_2 \subset I(\alpha_0, C_1', \eta_0)$$

as required.

CHAPTER 27

PROOF OF THE INFINITY CONDITION THEOREM

27.1. The aim of this chapter is to complete the proof of the Infinity Condition for ρ_2. In 27.6 we prove the Second Reduction of 26.5, which, by 26.6-7, implies the First Reduction of 26.4. We deduce the Infinity Condition Theorem from this in 27.7. We complete the proof of the Infinity Condition for ρ_2 in 27.8.

27.2. The angle is small. — Throughout the first part of this chapter, until the end of 27.6, we assume the hypotheses of the First and Second Reduction for x_0, $x_1 \in \widetilde{V}_1 \subset \mathcal{T}(Y)$ and the segment ℓ_1 and $x \in \ell_1$. We take $x \in I(\beta, C_1, r(\nu), \nu) \subset \ell_1$ as in the Second Reduction, in particular, β is periodic. In fact, by 26.10, we can assume — and shall — that $x \in I(\beta', C_1, r(\nu), \nu)$ for all $\beta' \in [\beta]$. By 26.11, we only need to consider the case when β is a gap and $v_2 \in \beta$. Our aim is to obtain a contradiction. This will complete the proof of the Second Reduction. The method imitates the original proof that a pullback τ contracts distance [**T2**], [**D-H3**] in Thurston's geometrizing theorem for critically finite branched coverings: that is, we exploit the fact that a quadratic differential cannot be its own pullback — nor even "topologically equivalent" to its own pullback — except in very special circumstances.

We fix notation until the end of 27.6. Let $q(z)dz^2$ be the quadratic differential at x for $d_Y(x, x_i)$, and let $q_0(z)dz^2$, $p(z)dz^2$ be the quadratic differentials for $d_Z(x, x_i)$, $d_Z(\tau(x), x_i)$ at x, $\tau(x)$, where $i = 0$ or 1 depending on whether (1) or (2) of 26.1 holds. Let $s^*q(z)dz^2$ denote the pullback of $q(z)dz^2$ at $\tau(x)$.

Lemma. — *Let $\beta_1 \in [\beta]$. Let $S(\beta_1) = S(\beta_1, x)$ be as in 9.3. Let $\theta(z)$ be the angle between q and q_0. If δ_0 is sufficiently small given δ_0',*

$$\int_{S(\beta_1)} |\theta|^2 |q| \leqslant \delta_0',$$

and similarly for s^*q, p.

Proof. — Fix $y \in [x, x_i]$ with $d_Y(x,y) = 1$, and let $q_{\beta_1}(z)dz^2$, $q'(z)dz^2$ be the quadratic differentials at x for $d_{\beta_1}(x,y)$ and $d(x,y)$. Let $\theta_1(z)$, $\theta_2(z)$, $\theta_3(z)$ be the angle at z between $q_0(z)$ and $q'(z)$, $q(z)$ and $q_{\beta_1}(z)$, $q_{\beta_1}(z)$ and $q'(z)$ respectively. Then by 26.9 there is $C_0 = C_0(M, C_1, \nu) \leqslant C_1$ such that $a(\beta_1, q_0) \geqslant C_0$ and $a(\beta_1, p_0) \geqslant C_0$. By (4), (5) of 26.2 and 8.9,

$$\int |\theta_1|^2 |q_0| = O(\delta_0). \tag{1}$$

Then if δ_0 is small enough, by (1), (2) of 26.2 and 8.3,

$$\int |\theta_2|^2 |q_{\beta_1}| = O(\delta^3), \tag{2}$$

and by 9.5,

$$\int |\theta_3|^2 |q_{\beta_1}| = O(\delta^3). \tag{3}$$

From (2) and (3), we see that zeros of q_{β_1} and q in $S(\beta_1)$ can be paired so that no pair is separated by a large modulus annulus, and similarly for zeros of q_{β_1} and q' in $S(\beta_1)$. It follows that, except in discs round the zeros, $|q| = O(\delta^{-1}|q_{\beta_1}|)$ and

$$\int_{S(\beta')} |\theta_i|^2 |q| = O(\delta^2) \tag{4}$$

for $i = 2, 3$. Similarly we obtain (4) for $i = 1$ from (1). Then since $|\theta| \leqslant |\theta_1| + |\theta_2| + |\theta_3|$, we get the required result for q and q_0, if δ_0 is small enough given δ_0'. The proof of s^*q and p is similar. □

27.3. Equivalence Classes of Poles. — We are now going to modify the definition of pole of q on $S(\beta')$ for $\beta' \in [\beta]$. The idea is to define an equivalence so that a pole *equivalence class* on $\cup_{\beta'} S(\beta')$ is preserved by $\psi \circ \varphi^{-1}$ and also transforms suitably under s^{-1}. The key to this is the boundedness of the homeomorphism $\psi \circ \varphi^{-1}$ up to isotopy, at least on $S(\beta')$, where $\tau(x) = [\psi]$. By 26.8, up to homotopy, $\psi \circ \varphi^{-1}$ maps a path γ in $S(\beta')$ to a path γ' with

$$|\gamma'| \leqslant e^{M_1}(1 + |\gamma|)$$

for $M_1 = M_1(M, C_1, \nu)$ (remembering that M, as well as C_1 and ν, features in our hypotheses) and $|\gamma|$, as usual, denotes Poincaré length.

As usual (see 9.1), given a gap β' ($\beta' \in [\beta]$, in our case) we can choose a set $A(\beta') \subset Y$ consisting of all points in $\beta \cap Y$ and exactly one in each component of $\overline{\mathbf{C}} \setminus \beta$. We say that a point of $\varphi(A(\beta'))$ is a *possible pole* of $q(z)dz^2$ on $S(\beta')$ (for obvious reasons), and we define possible poles of $p(z)dz^2$ and $s^*q(z)dz^2$ similarly. Now we can define an equivalence relation on the set of possible poles of $q(z)dz^2$ on $S(\beta')$, and zeros in $S(\beta')$, and similarly for $p(z)dz^2$ and $s^*q(z)dz^2$ with the following properties, if δ_0 is small enough, and hence δ_0' of 27.2 is small enough.

Given any constant $C_0 > 1$, there are $\delta_4 > \delta_3 > \delta_2 > \delta_0'$ and $L_2 < L_3 < L_4$ with

$$(1) \qquad C_0\delta_0' < \delta_2, \quad C_0\delta_2 e^{2C_0M} < \delta_3, \quad C_0e^{2M_1} \leqslant L_2, \quad C_0e^{3M_1}L_2 < L_3,$$

$$(2) \qquad L_3 \leqslant C_0^{4\,\#(Z)}e^{12M_1\,\#(Z)},$$

$$(3) \qquad \delta_2 \leqslant C_0^{-8\,\#(Z)}e^{-24M_1\,\#(Z)}.$$

and similar inequalities relating δ_3, δ_4, L_3, L_4, and the following holds.

If a and b are zeros or possible poles of $q(z)dz^2$ on $S(\beta')$ then either they — or the components of $\overline{\mathbf{C}} \setminus S(\beta')$ containing them — are joined by a path η of length $\leqslant L_2$ in $S(\beta')$ with $|\eta|_{q,+} < \delta_2$, or there is no path ℓ joining them of length $\leqslant L_4$ with $|\ell|_{q,+} < \delta_4$, and similarly for $p(z)dz^2$, $s^*q(z)dz^2$.

If the constant C_0 is large enough, then there is an equivalence relation on the set of poles and zeros for $q(z)dz^2$ on $\cup_{\beta' \in [\beta]} S(\beta')$ defined as follows, and similarly for $p(z)dz^2$, $s^*q(z)dz^2$.

1. The possible poles and zeros a and b on $S(\beta')$ are equivalent if they are joined by a path η in $S(\beta')$ such that (1) holds for $(\delta, L) = (\delta_2, L_2)$. If they are not equivalent, then there is no path η joining them such that (4) holds for $(\delta, L) = (\delta_4, L_4)$:

$$(4) \qquad |\eta|_q \leqslant L, \ |\eta|_{q,+} < \delta.$$

Also the following hold, if δ_0 is small enough.

2. a and b are equivalent for $q(z)dz^2$ if and only if each point in $s^{-1}a$ is equivalent to a point in $s^{-1}b$ for $s^*q(z)dz^2$.

3. a and b are equivalent for $p(z)dz^2$ if and only if they are within δ_0' of points which are equivalent for $s^*q(z)dz^2$. Hence each equivalence class for $p(z)dz^2$ is within δ_0' of a unique equivalence class of $s^*q(z)dz^2$, and vice versa.

4. a and b are equivalent for $q(z)dz^2$ if and only if they are within δ_0' of points which are equivalent for $q_0((z)dz^2$. Hence each equivalence class for $q(z)dz^2$ is within δ_0' of a unique equivalence class of $q_0(z)dz^2$, and vice versa.

We can define the *index* N of an equivalence class $[a]$ to be the number of poles minus the number of zeros up to multiplicity, where the sum is taken over all poles and zeros of a component of $\overline{\mathbf{C}} \setminus S(\beta')$, if such a component is included in the equivalence class. The index of $[a]$ can only be strictly positive if $[a]$ contains at least one point of $\varphi(Z) \cap S(\beta')$ or at least one component of $\overline{\mathbf{C}} \setminus S(\beta')$ (since q and q_0 have close equivalence classes). A *pole equivalence class* is one of strictly positive index.

27.4. Choosing a Loop set. — We now introduce a tool — a loop set — which will enable us to map pole equivalence classes — and others. Write $x_i = [\varphi_i]$, $i = 0, 1$ depending on whether our hypothesis is that (1) or (2) of 26.1 holds. Our hypotheses ensure that $x_i \in \mathcal{T}_{\geqslant \varepsilon_0}$. We can choose a loop set $\Gamma_i \subset \overline{\mathbf{C}} \setminus Z$ with the following properties for $C_0 = C_0(\varepsilon_0)$:

(1) $|\varphi_i(\Gamma_i)| \leqslant C_0$.

(2) Every component of $\overline{\mathbf{C}} \setminus (Z \cup (\cup \Gamma_i))$ is a topological disc with at most one puncture.

Now let $x = [\varphi] \in \ell$. If C_0 is chosen large enough, the following holds. Let I be any segment of contracting foliation leaf of the quadratic differential for $d_Z(x_i, x)$ at x of length $\delta \geqslant C_0 e^{-d_Z(x_i,x)}$. Take $\varphi(\Gamma_i)$ in good position (14.5) with respect to this quadratic differential. Then the number of intersections of I by $\varphi(\Gamma_i)$ is between $C_0 \delta e^{d_Z(x_i,x)}$ and $(1/C_0)\delta e^{d_Z(x_i,x)}$. The reason is that, by 2, every stable leaf segment at x_i (for the quadratic differential for $d_Z(x_i, x)$ at x_i) of length C_0 must intersect $\varphi_i(\Gamma_i)$ at least once, but, by 1, not more than C_0^2 times. An exactly similar statement holds if I is a segment of stable foliation leaf for $d_Z(x_i, \tau(x))$ at $\tau(x)$. We can also assume that C_0 is large enough to act as C of 26.1. Then by 26.1, $d_Z(x_i, x)$ and $d_Z(x_i, \tau(x))$ differ by at most $C_0 M$. This will give us an effective way of comparing lengths of short segments of the stable foliation.

27.5. The surface $T([a], q, \delta, L)$. — For concreteness, take an equivalence class $[a]$ for $q(z)dz^2$, for a a possible pole or zero of q on $S(\beta')$ ($\beta' \in [\beta]$). Note that it can be encased in a surface $T([a], q, \delta, L) \subset S(\beta')$ of the following type for $(\delta, L) = (\delta_2, L_2)$ or (δ_3, L_3). The boundary $\partial_+ T([a], q, \delta, L) \cup \partial_- T([a], q, \delta, L)$ of $T([a], q, \delta, L)$ consists of finitely many expanding leaf segments $\partial_+ = \partial_+ T([a], q, \delta, L)$ of the expanding foliation of $q(z)dz^2$, and finitely many contracting leaf segments $\partial_- = \partial_- T([a], q, \delta, L)$. Any point in $T([a], q, \delta, L)$ can be joined to a pole or zero in $[a]$ by a path η in $T([a], q, \delta, L) \subset S(\beta', [\varphi], \nu)$ such that (4) of 27.3 holds, and

$$|\partial_+|_q = O(L), \quad |\partial_-|_q = O(\delta).$$

Conversely, $T([a], q, \delta, L)$ is a union of such paths. Note that our definitions ensure that any two surfaces $T([a], q, \delta_4, L_4)$ are equal or disjoint. Also, for fixed a, the surfaces $S([a], q, \delta, L)$ are isotopic for all $\delta_2 \leqslant \delta \leqslant \delta_4$ and $L_2 \leqslant L \leqslant L_4$.

Lemma. — *Let $y \in A(\beta')$, $S(\beta', [\varphi], \nu)$, $\beta' \in [\beta]$. Let $\delta_2 \leqslant \delta \leqslant \delta_4$ and $L_2 \leqslant L \leqslant L_4$. Then $T([a], q, \delta, L)$ is homeomorphic to $T([\psi(y)], p, \delta, L)$ under a homeomorphism isotopic to $\psi \circ \varphi^{-1}$ which preserves stable and unstable segments in the boundaries.*

Remark. — This does not mean that $T([\varphi(y)], q, \delta, L)$ and $T([\psi(y)], p, \delta, L)$ contain the same number of zeros: a rectangle in $T([\varphi(y)], q, \delta, L)$ bounded by stable and unstable leaf segments between poles and zeros of p might be collapsed by $\psi \circ \varphi^{-1}$, and vice versa. However, the number of zeros *up to multiplicity* will be the same.

Proof. — We use the loop set Γ_i constructed in 27.4. We assume that $\varphi(\Gamma_i)$ and $\psi(\Gamma_i)$ are in good position (14.5) with respect to $q(z)dz^2$ and $p(z)dz^2$ respectively. We take $T'([\varphi(y)]$ very close to $T([\varphi(y)], q, \delta_2, L_2)$, but containing it, with ∂_+ replaced by segments of $\varphi(\Gamma_i)$ and ∂_- as before. Take any two segments of $\varphi(\Gamma_i)$ in $T'([a])$ whose endpoints are joined by stable segments of length $< \delta_2$. Consider the images under $\psi \circ \varphi^{-1}$. The stable segments are mapped to paths of length $\leqslant C_0 e^{M_1}$ which are crossed by $\leqslant \delta_2 C_0 e^{d_Z(x,x_i)} \leqslant C_0\delta_2 e^{C_0 M + d_Z(\tau(x),x_i)}$ segments of $\psi(\Gamma_i)$ and hence have stable length $< \delta_3$. Also, all lengths in $S(\beta', [\varphi], \nu)$ ($\beta' \in [\beta]$) of at least 1 are multiplied by at most $O(e^{M_1})$ by application of $\psi \circ \varphi^{-1}$. So

$$\psi \circ \varphi^{-1}(T'([\varphi(y)]) \subset T([\psi(y)], p, \delta_3, L_3).$$

We also see that if two segments in $\varphi(\Gamma_i)$ are close to different unstable components of $\partial_+ T([\varphi(y)], p, \delta_2, L_2)$ which are not both closed loops, then they have points distance $\geqslant C_0 e^{M_1}$ apart (from the definition of L_2). So they cannot be mapped close to the same unstable component in $\partial_+ T([\psi(y)], p, \delta_3, L_3)$. It follows that $\psi \circ \varphi^{-1}$ can be perturbed to a homeomorphism of $T([\varphi(y)], p, \delta_2, L_2)$ into $T([\psi(y)], q, \delta_3, L_2)$ which maps across $\partial_+ T$ and $\partial_- T$ segments. Closed loops in $\partial_+(T)$ are obviously mapped across to closed loops. So $\psi \circ \varphi^{-1}$ can be perturbed to homeomorphism of $T_1 = T([\varphi(y)], q, \delta, L)$ to $T_2 = T([\psi(y)], p, \delta, L)$ mapping $\partial_+ T_1$-segments to $\partial_+ T_2$-segments, and $\partial_- T_1$-segments to $\partial_- T_2$-segments, for all $\delta_2 \leqslant \delta \leqslant \delta_4$, and $L_2 \leqslant L \leqslant L_4$. □

27.6. Proof of the Second Reduction. — Take δ_2 sufficiently small given ν (remembering that $\nu \geqslant \nu_1$ for some ν_1 depending on M) that $\delta_4 < \nu/4$. We also take δ_0 sufficiently small so that (as in 27.3) $C_0 \delta_0' < \delta_2$. We now show that it is impossible to have $x \in I(\beta, C_1, r(\nu), \nu)$ with β periodic degree 2, unless the gap map $[\psi_\beta]$ (2.13) is an irreducible critically finite nonrational degree two map. This will complete the proof of the second reduction.

If $y \in A(\beta')$, some $\beta' \in [\beta]$, and $\varphi(v_1)$, $\varphi(v_2) \notin [\varphi(y)]$ then the index of either component $[s^{-1}(\varphi(y)]$ is strictly less than that of $[\varphi(y)]$.The index of any equivalence class $[a]$ is $\leqslant 2$ by Euler's Theorem: because the index of $[\varphi(y)]$ is $2\chi(T) - N$ where $T = T([\varphi(y)], q, \delta_2, L_2)$, $\chi(T)$ is the Euler characteristic and N is the number of stable boundary components in N. So by 27.5, these are the same as the indices of $[\varphi(y_i)]$ for the points $y_i \in f_0^{-1}(y)$. If $[\varphi(y)]$ has index $\leqslant 1$ then any component of $[s^{-1}(\varphi(y)]$ has index strictly less than that $[\varphi(y)]$. If $[\varphi(y)]$ has index 2 and contains both critical values then the index of $[s^{-1}\varphi(y)]$ is strictly less than that of $[\varphi(y)]$. Now let β be periodic degree two. The total sum of indices for each $A(\beta')$ is 4. So the only possibilities are that:

(i) there are just two nontrivial equivalence classes, both of index 2, each containing one critical value, and either fixed or of degree two,

(ii) There are four nontrivial equivalence classes, all of index 1, with the following possible dynamics:

(iia) $\quad [\varphi(v_1)],\ [\varphi(v_2)] \mapsto [\varphi(f_0(v_1)] = [\varphi(f_0(v_2)] \mapsto [\varphi(f_0^2(v_1)] = [\varphi(f_0^3(v_1)],$

(iib) $\quad [\varphi(v_1)] \mapsto [\varphi(f_0(v_1))] \mapsto [\varphi(f_0^2(v_1))] = [\varphi(f_0(v_2))] \mapsto [\varphi(f_0^3(v_1))] = [\varphi(f_0(v_1))],$

(iib) $\quad [\varphi(v_2)] \mapsto [\varphi(f_0^2(v_1))].$

Now we claim that (i) is impossible. By (2) and (3) of 27.3, if (i) holds, then, ℓ enters $I(\gamma, C_1, e^{-12C_0M_1})$. By 26.11, this is impossible. So (ii) must hold. Then v_1, v_2

are nonperiodic. Then from earlier analysis of $[\psi_\beta]$ (2.18), v_1 must be eventually fixed. So (iia) is discounted, and we must have (iib). Also, $y = f_0^i(v_1)$ is the only element of $A(\beta')$ with $\varphi(y) \in [\varphi(f_0^i(v_1)]$, $0 \leqslant i \leqslant 2$. We claim that $\varphi(v_2) \notin S(\beta', [\varphi], \nu)$ (any $\beta' \in [\beta]$), and thus that $[\psi_\beta]$ is critically finite. Then each point $\varphi(y)$ of $\varphi(A(\beta')\setminus\{v_2\})$ is the only point of $\varphi(A(\beta')\setminus\{v_2\})$ in its equivalence class (for q). The same is true for the points $\psi(y)$, $y \in A(\beta')\setminus\{v_2\}$, for p. So for each equivalence class for s_*q, all poles in that equivalence class lie in a disc of radius δ_0'', where δ_0'' can be taken arbitrarily small by choice of δ_0', that is, by choice of δ_0. It follows that, for δ_0 sufficiently small, $\varphi(v_2) \notin S(\beta', [\varphi], \nu)$, as required. □

27.7. Proof of the Infinity Condition except for paths in ∂U_a. — We have now proved the First Reduction (which follows from the Second). The Infinity Condition Theorem then holds for any $x_1 \in \widetilde{V}_1$ for which $[x_0, x_1] \cap \mathcal{T}_{\geqslant\varepsilon_0}$ has length $\geqslant M_2$ for M_2 depending only on M and ε_0. If the sets U_a are suitably chosen (that is, M_a is taken large enough for each a), $[x_0, x_1] \cap \mathcal{T}_{\geqslant\varepsilon_0}$ has length $\leqslant M_2$ and $d_Y(x_0, x_1)$ is large enough given M_2, then the First Reduction, together with the following lemma, completes the proof of the Infinity Condition Theorem.

Lemma. — *The following holds for suitable C, Δ_2 given $N \geqslant 2$, α_i, M_1, Δ_1. Either let $x_0 \in \mathcal{T}_{\geqslant\varepsilon_0}$ and*

$$x_0 \in \mathcal{T}_{\geqslant\varepsilon_0},\ d_P(0, \rho_2(x_1)) \leqslant M, \tag{1}$$

or for some $x_2 \in \mathcal{T}(Y)$ with $\pi_Z(x_2)$ on the geodesic in $\mathcal{T}(Z)$ between $\pi_Z(x_0)$ and $\pi_Z(x_1)$, let

$$x_1 \in \mathcal{T}_{\geqslant\varepsilon_0},\ d_P(0, \rho_2(x_2, x_1)) \leqslant M,\ d_Y(x_0, x_2) \leqslant d_Z(x_0, x_2) + M. \tag{2}$$

Suppose also that $[x_0, x_1] = \ell$ is a union of $N+1$ segments of length $\leqslant \Delta_1$ and N segments in sets $I(\alpha_i, C_1, \nu_0, \nu_0)$ (in the notation of 26.3) for $1 \leqslant i \leqslant N$ with endpoints y_i and w_i, with $\alpha_i \neq \overline{\mathbf{C}}$ and $d_Y(y_i, w_i) \geqslant \Delta_2$. Then for a suitable $C > 0$,

$$d_P(0, \rho_2(x_1)) \geqslant C \sum_{i=1}^{N-1} \exp d_Z(y_i, x_0).$$

Remark. — The conclusion of the lemma is in contradiction with (1). Therefore (1) cannot hold in conjunction with the other hypotheses (excluding (2)). The alternative hypotheses (1) and (2) occur in the First and Second Reductions of Chapter 26. Alternative (1) is the hypothesis of the Infinity Condition Theorem 25.6. This is why this lemma implies the Infinity Condition Theorem. We shall use this lemma (with the hypothesis including (2)) to deduce the Infinity Condition for ρ_2 in 27.8.

Proof. — We choose α_i in its orbit to separate all other components of $[\alpha_i]$ from v_2. We can assume that $\partial\alpha_i$ and $\partial\alpha_{i+1}$ have more essential intersections in $\overline{\mathbf{C}} \setminus Y$ than in $\overline{\mathbf{C}} \setminus Z$ — otherwise, they have no essential intersections, and can be combined. This is an argument we have used before, in 3.14 and in 1.16 of [**R3**], and goes as follows. Assume all intersections are essential and replace f_0 by $f_0 \circ \chi = f_1$ for a homeomorphism χ which is isotopic to the identity relative to Z, and so that f_1 preserves each of $\partial[\alpha_i]$ and $\partial[\alpha_{i+1}]$ and the invariant loop sets Γ_i, Γ_{i+1} that these generate. We can even choose f_1 to be critically finite. Such a map cannot have two transversally intersecting loop sets satisfying the Invariance and Levy conditions. This completes the argument.

Write $x_1 = [\varphi_1]$, $x_0 = [\varphi_0]$. As usual we write ℓ for the geodesic in $\mathcal{T}(Y)$ between x_0 and x_1, and ℓ' for the geodesic in $\mathcal{T}(Z)$ between $\pi_Z(x_0)$ and $\pi_Z(x_1)$. By 14.2, for each $x' \in \ell'$ with $x' \in (\mathcal{T}(Z))_{\geqslant \varepsilon_0}$, there is a corresponding $x \in \ell$ with $d(x', x) \leqslant M_1$ — and then by 26.8 we also have $d(x', \tau(x)) \leqslant M'$. Then we can write

$$\begin{aligned}\varphi_1 &= \chi_{N,N-1} \circ \sigma_{\beta_{N-1}} \circ \cdots \circ \chi_{1,0} \circ \sigma_{\beta_0} \circ \varphi_0,\\ \chi_i &= \chi_{i,i-1} \circ \cdots \circ \chi_{1,0},\\ \psi_i &= \sigma_{\beta_i} \circ \chi_{i,i-1} \circ \sigma_{\beta_{i-1}} \cdots \sigma_{\beta_0} \circ \varphi_0,\end{aligned}$$

where β_i is a path with first endpoint at χ_i, χ_i minimizes distortion up to isotopy preserving $\varphi_0(Z)$, $\chi_{i+1,i}$ minimizes distortion up to isotopy preserving $\chi_i \circ \varphi_0(Z)$, and ψ_i is chosen to be a bounded distance (in terms of Δ_1) from w_i and y_{i+1}. So $[\chi_i \circ \varphi_0]_Z$ are all points on ℓ'. The point $[\chi_i \circ \varphi] \in \ell'$ is chosen in $(\mathcal{T}(Z))_{\geqslant \varepsilon_0}$ between the intersections of ℓ' with $\mathcal{T}(\partial[\alpha_{i-1}], \varepsilon_0)$ and $\mathcal{T}(\partial[\alpha_i], \varepsilon_0)$. So $|\chi_i \circ \varphi(\partial[\alpha_i])|$ and $|\chi_i \circ \varphi(\partial[\alpha_{i-1}])|$ are both bounded in terms of Δ_1. Also, since $\partial[\alpha_i]$ and $\partial[\alpha_{i-1}]$ have essential intersections, β_i has at least two essential intersections with $\chi_i \circ \varphi_0(\partial[\alpha_{i-1}])$. But the length of intersection of β_i with $\chi_i \circ \varphi_0([\alpha_{i-1}])$ is bounded in terms of Δ_1. It follows that if Δ_2 is sufficiently large given Δ_1 then $\chi_{i,i-1}^{-1}\beta_i$ has $\geqslant D_1 e^{d_Z(y_{i-1}, w_{i-1})}$ essential intersections with β_{i-1} for a suitable $D_1 > 0$. It follows that the path

$$\beta = \beta_0 * \chi_1^{-1}\beta_1 * \cdots * \chi_{N-1}^{-1}\beta_{N-1}$$

satisfies

$$|\beta| \geqslant C \sum_{i=1}^{N-1} \exp d_Z(x_0, y_i),$$

as required. □

27.8. Proof of the Infinity Condition for paths with one long component in the thin part. — Let M be given and let $x_1 \in \widetilde{V}_1$ satisfy

$$d(0, \rho_2(x_1)) \leqslant M.$$

By the Infinity Condition Theorem, and 25.8 there is M_1 depending on M such that all but length M_1 of the geodesic ℓ joining x_0 and x is in $\mathcal{T}(\partial P, \varepsilon_0)$ for some fixed set P of a minimal nonempty $[f_0, \Gamma]$ corresponding to some $a \in A$, with corresponding $g \in \pi_1(V_1)$ fixing $a \in A \subset \partial D$. It remains to show that for some Δ depending on M, either $d(x_0, x_1) \leqslant \Delta$ or $x \in U_a$ where U_a — or the constant M_a defining U_a (see 25.4) *does not depend on* M.

We can also consider g as an element of the modular group acting on $\mathcal{T}(Y)$ (as usual). In this case, we can write

$$x_1 = x_{1,n} = [\varphi'_n] = g^n x'_0 = [\varphi'_0 \circ \chi^n],$$

where x'_0 lies in a subset of $\widetilde{V}_1 \subset \mathcal{T}(Y)$ whose intersection with $\mathcal{T}_{\geqslant \varepsilon_0}$ is compact — which does, however, depend on M. Then (replacing g by g^{-1} if necessary), the components of $[x_0, g^n x'_0] \setminus \mathcal{T}(\partial P, \varepsilon_0)$ containing x_0, $g^n x'_0$ have lengths $\leqslant T_1 = T_1(x_0)$, $\leqslant T_2 = T_2(x_0, x'_0)$ respectively. This follows from 14.4, because $\varphi'_n(\partial P) = \varphi'_0(\partial P)$, and if γ is a loop with $\gamma \cap \partial P \neq \varnothing$ and $|\varphi_0(\gamma)| = O(|\varphi_0(\partial P)|)$, then $|\varphi'_n(\partial P)| = o(|\varphi'_n(\gamma)|)$. For $M_2 = M_2(x_0, \Gamma, x'_0)$, and for any γ with $\gamma \cap P = \varnothing$, we also have

$$\frac{|\varphi_0(\gamma)|'}{M_2} \leqslant |\varphi'_n(\gamma)|' \leqslant M_2 |\varphi_0(\gamma)|', \tag{1}$$

while if $\gamma \cap P \neq \varnothing$, then

$$\frac{e^{d_Z(x_{1,n}, x_0)}}{M_2} \leqslant |\varphi'_n(\gamma)|' \leqslant M_2 e^{d_Z(x_{1,n}, x_0)}. \tag{2}$$

We claim that we can write

$$x_{1,n} = [\chi_{2,n} \circ \sigma_{\beta_{2,n}} \circ \chi_{1,n} \circ \sigma_{\beta_{1,n}} \circ \varphi_0]_Y, \quad x'_{1,n} = [\chi_{1,n} \circ \sigma_{\beta_{1,n}} \circ \varphi_0]_Y.$$

where $\chi_{i,n}$ and $\beta_{i,n}$, $x_{1,n}$, $x'_{1,n}$ have the following properties:

$$|\beta_{1,n}| \leqslant M_3 = M_3(x_0, \Gamma), \tag{3}$$

$$d_Y(x'_{1,n}, x_{1,n}) \leqslant M_4 = M_4(x_0, \Gamma, x'_0), \tag{4}$$

and $\chi_{1,n}$ and $\chi_{2,n}$ minimize distortion up to isotopies constant on $\varphi_0(Z)$, $\chi_{1,n} \circ \varphi_0(Z)$ respectively, and $[\chi_{1,n} \circ \varphi_0]$ is on the geodesic between $[\varphi_0]$ and $[\chi_{2,n} \circ \chi_{1,n} \circ \varphi_0]$ in $\mathcal{T}(Z)$. We use 26.8: we can take $x'_{1,n} \in \mathcal{T}(Y)_{\geqslant \varepsilon_0}$ (that is, $\alpha = \overline{\mathbf{C}}$ in 26.8). We can achieve (3) by taking $\beta_{1,n}$ to be a loop up to intersection with ∂P, since we have already seen that $[x_0, x_{1,n}] \subset \mathcal{T}(Y)$ enters $\mathcal{T}(\partial P, \varepsilon_0)$ within distance T_1. We achieve (4), by using (1). Let $q_n(z)dz^2$ be the quadratic differential for $d_Z(x'_{1,n}, x_{1,n})$ at $x'_{1,n}$. Note that (4) implies, in particular, that

$$|\beta_{2,n}|_{q_n} < M_5 = M_5(x_0, \Gamma, x'_0), \tag{5}$$

To complete the proof it suffices to show that

$$\lim_{n \to \infty} |\beta_{1,n} * \chi_{1,n}^{-1} \beta_{2,n}| = +\infty.$$

Suppose this is not true. Then there is Δ' such that, for infinitely many n,

$$|\beta_{1,n} * \chi_{1,n}^{-1} \beta_{2,n}| \leqslant \Delta'. \tag{6}$$

Take such an n. Then $\beta_{2,n}$ does not have an essential intersection in $\overline{\mathbf{C}} \setminus \chi_{1,n} \circ \varphi_0(Z)$ with $\chi_{1,n} \circ \varphi_0(\partial P)$ apart from the first endpoint (if $[f_0, \Gamma]$ is isometric), or does not have an essential component of intersection with $\chi_{1,n} \circ \varphi_0(P)$ with both endpoints in $\chi_{1,n} \circ \varphi_0(\partial P)$ (if $[f_0, \Gamma]$ is pseudo-Anosov), because the resulting arc would be expanded by $\chi_{1,n}^{-1}$. It follows (as in 27.7) that $x_{1,n} \in \mathcal{T}_{\geqslant \varepsilon_0}$ for ε_0 independent of M. Therefore x_0' lies in a compact set, and the constants $M_i(x_0, \Gamma, x_0')$ depend only on M.

Now we claim that it is enough to obtain, for some $x_{2,n}$ with $\pi_Z(x_{2,n}) \in [\pi_Z(x_0), \pi_Z(x_{1,n})] \subset \mathcal{T}(Z)$,

$$d_Y(x_0, x_{2,n}) \leqslant d_Z(x_0, x_{2,n}) + M_6(x_0, \Gamma), \tag{7}$$

$$d_P(0, \rho_2(x_{2,n}, x_{1,n})) = 0. \tag{8}$$

If these hold, then the First Reduction 26.4 shows that the hypotheses of 27.7 are satisfied if M_a is large enough, depending only on x_0 — not on x_0'. So then we would contradict (6) for all but finitely many n, giving the required result.

In the isometric case, if $\beta_{2,n}$ is disjoint from $\chi_{1,n} \circ \varphi_0(\partial P)$, then, writing $x_{2,n} = [\sigma_{\beta_{2,n}} \circ \chi_{1,n} \circ \sigma_{\beta_{1,n}} \circ \varphi_0]_Y$, we obtain (7) and (8), as required, if n is sufficiently large. We are using, here, that the Teichmüller distance travelled in $\mathcal{T}(\partial P, \varepsilon_0)$ is, to within a bounded constant, the distance travelled in the maximal distance factor $\mathcal{T}(A(P))$.

Now let $[f_0, \Gamma]$ be pseudo-Anosov and write $\beta_{2,n} = \beta_{3,n} * \beta_{4,n}$ where $\beta_{4,n} \subset \chi_{1,n} \circ \varphi_0(P)$ and $\beta_{3,n} \cap \chi_{1,n} \circ \varphi_0(P) = \varnothing$. We have the same bound (5) on $\beta_{3,n}$, $\beta_{4,n}$ as on their union $\beta_{2,n}$. Given $\delta = M_5^{-2} > 0$, since (as we are assuming) (6) does not hold, we must have

$$|\beta_{4,n}|_{q_n,-} \leqslant \delta \quad \text{for all sufficiently large } n.$$

Then write

$$\chi_{1,n} = \chi_{4,n} \circ \chi_{3,n}, \quad \gamma_n = \beta_{3,n} * \chi_{4,n}^{-1}(\beta_{4,n}), \quad x_{2,n} = [\sigma_{\gamma_n} \circ \chi_{3,n} \circ \sigma_{\beta_{1,n}} \circ \varphi_0],$$

where $\chi_{3,n}$, $\chi_{4,n}$ minimize distortion up to isotopies constant on $\varphi_0(Z)$, $\chi_{3,n} \circ \varphi_0(Z)$, $[\chi_{4,n} \circ \varphi_0]$ is on the geodesic joining $[\varphi_0]$ and $[\chi_{1,n} \circ \varphi_0]$ in $\mathcal{T}(Z)$, and

$$|\chi_{4,n}^{-1}(\beta_{4,n})|_{p_n} \leqslant \sqrt{\delta} + M_5\sqrt{\delta},$$

where $p_n(z)dz^2$ is the quadratic differential at $x_{2,n}$ for $d_Z(x_{2,n}, x_{1,n})$. Then we again obtain (7) and (8), as required, again using that the distance travelled in $\mathcal{T}(\partial P, \varepsilon_0)$ is to within a bounded constant the distance travelled in the maximal distance factor $\mathcal{T}(A(P))$. □

CHAPTER 28

REDUCTIONS IN THE PROOF OF THE EVENTUALLY CLOSE THEOREM

28.1. To complete the proof of the Resident's View of Rational Maps Space, we need to show that the ECPP of 25.5 holds for paths of a number of different types, for ρ_2. We start with the following, for ρ_2. We use the fact that ρ_2 is defined on all of $\mathcal{T}(Y)$. We take $x_0 = [\text{identity}]$ as our basepoint in $\widetilde{V}_1 \subset \mathcal{T}(Y)$ with $\rho_2(x_0) = 0$. Then note that ρ_2 can be defined on all of $\mathcal{T}(Y)$. Then the first step in obtaining ECPP is the following, which gives ECPP for some geodesic segments.

Lemma. — *The following holds, given M, for all n sufficiently large. Let $[x, y]$ be a geodesic segment in $\mathcal{T}(Y)$ which is either of length $\leqslant M$ or $[x, y] \subset [x_0, y]$ such that*

$$n \leqslant d_P(0, \rho_2(z)) \text{ for all } z \in [x, y]. \quad d_P(0, \rho_2(x)), d_P(0, \rho_2(y)) \leqslant n + 1.$$

Then

$$|\rho_2(x) - \rho_2(z)| \leqslant e^{-\sqrt{n}} \quad \text{for all } z \in [x, y].$$

Proof. — Suppose for contradiction that

(1) $$|\rho_2(x) - \rho_2(y)| \geqslant e^{-\sqrt{n}}.$$

We are now going to obtain a contradiction to (1). We write $\beta_{u,v}$ for the geodesic with endpoints $\rho_2(u)$, $\rho_2(v)$. Then for some $[u, v] \subset [x, y]$, and suitable constants C_i, $(1 \leqslant i \leqslant 3)$ we shall show that (2) and (3) hold. We shall then obtain a contradiction.

(2) $$d_Z(x_0, u) + d_Y(u, v) - d_Z(x_0, v) \leqslant C_1.$$

(3) $\beta_{u,v}$ has length $\geqslant C_1^{-1}n$, and apart from a segment in the middle of length $\leqslant C_2 n^{2/3}$, any segment of $\beta_{u,v}$ of length $\geqslant C_3$ projects to cut the surface $\overline{\mathbf{C}} \setminus Z$ into (topological) discs or once-punctured discs.

Note that, by 26.1. and a suitable constant C_4,

$$d_Y(x_0, x) - d_Z(x_0, x) \leqslant C_4 n,$$

and similarly for y. So

$$d_Z(x_0, x) + d_Y(x, y) - d_Z(x_0, y) \leqslant d_Y(x_0, x) + d_Y(x, y) - d_Y(x_0, y) + 2C_4 n \leqslant 3C_4 n.$$

If $\{x_i : 1 \leqslant i \leqslant k\}$ are successive points on $[x, y]$ with $x_1 = x$, $x_k = y$, we have

$$d_Z(x_0, x) + d_Y(x, y) - d_Z(x_0, y) = \sum_{i=1}^{k-1} (d_Z(x_0, x_i) + d_Y(x_i, x_{i+1}) - d_Z(x_0, x_{i+1})).$$

In particular, for any $x' \in [x, y]$,

$$d_Y(x, x') - d_Z(x, x') \leqslant C_2 n.$$

So we can choose $O(n)$ successive points x_i in $[x, y]$ so that, for a constant C_1,

$$d_Z(x_0, x_i) + d_Y(x_i, x_{i+1}) - d_Z(x_0, x_{i+1})) \leqslant C_1,$$

and, for some i,

$$|\rho_2(x_i) - \rho_2(x_{i+1})| \geqslant \frac{e^{-\sqrt{n}}}{n}.$$

Then (2) holds for any $[u, v] \subset [x_i, x_{i+1}]$. To obtain (3) also, since ρ_2 is continuous on $[x_i, x_{i+1}]$, choose $[u, v] \subset [x_i, x_{i+1}]$ so that the (Poincaré) geodesic joining 0 and $\rho_2(u)$ — which is of length $\geqslant n$ — is such that, apart from the first portion of length $\leqslant n^{2/3}$, every segment of length $\geqslant C_3/2$ projects to cut the surface in $\overline{\mathbf{C}} \setminus Z$ into topological discs and once punctured discs. Then $\beta_{u,v}$ has the property 3. However, from 14.12 and 14.13 we obtain the following. We can write

$$u = [\varphi_u] = [\chi_u \circ \sigma_{\beta_u}], \quad v = [\varphi_v] = [\chi_v \circ \sigma_{\beta_v}],$$

where χ_u, χ_v minimize distortion up to isotopies constant on Z, β_u, β_v are paths starting from v_2,

$$v = [\varphi_v] = [\sigma_{\alpha_1} \circ \chi_{v,u} \circ \varphi_u],$$

where α_1 has the properties of $\gamma_3 * \gamma_2 * \gamma_1$ in 14.12, with [identity], $[\varphi_u]$, $\varphi_v]$ replacing $[\varphi_0]$, $[\varphi_1]$, $[\varphi_2]$, $\chi_{v,u}$ minimizes distortion up to isotopy constant on $\varphi_u(Z)$ with $[\varphi_v]_Z = [\chi_{v,u} \circ \varphi_u]_Z$, and then

$$[\chi_{v,u} \circ \varphi_u] = [\chi_{v,u} \circ \chi_u \circ \sigma_{\beta_u}] = [\sigma_{\alpha_2} \circ \chi_v \circ \sigma_{\beta_u}],$$

where α_2 has the properties of $\gamma_5 * \gamma_4 * \gamma_3 * \gamma_2 * \gamma_1$ in 14.13, again with $[\xi_0]$, $[\varphi_u]$, $[\varphi_v]$ replacing $[\varphi_0]$, $[\varphi_1]$, $[\varphi_2]$. Then

$$\chi_v^{-1}(\alpha_2 * \alpha_1) = \beta_{u,v}.$$

But by the properties of the paths γ_i of 14.12, 14.13 — leading to properties of the paths $\chi_v^{-1}(\alpha_1)$, $\chi_v^{-1}(\alpha_2)$ — $\beta_{u,v}$ cannot have property (3) above. So we obtain the required contradiction. □

28.2. ECPP for Paths in ∂U_a. — We now need to check the ECPP for paths in ∂U_a ($a \in A$) and ρ_2. We already know that the Infinity Condition holds for such paths and for ρ_2, by 27.8.

Lemma. — *ECPP holds for any path ℓ in ∂U_a and for ρ_2. In fact, if x, $y \in \ell$ and $d_P(0, \rho_2(z)) \geqslant n$ for $z = x$ or y, then*

$$|\rho_2(x) - \rho_2(y)| \leqslant e^{-\sqrt{n}}.$$

Proof. — For M depending only on a, we have

$$d_Y(z, x_0) \leqslant M + d_Z(z, x_0)$$

for all $z \in \ell$. Note that $\ell \subset \mathcal{T}_{\geqslant \varepsilon_0}$ for some ε_0 depending on M. Then by 14.12 we have $z = [\chi_z \circ \sigma_{\beta_z}]_Y$ where χ_z minimizes distortion up to isotopy constant on Z, and $|\beta_z|_+ \leqslant C$ for C depending on M, where $|\cdot|_+$ denotes length with respect to the expanding foliation of the quadratic differential for $d_Z(x_0, z)$ at x_0. Then the geodesic in the unit disc joining $\rho_2(x)$ and $\rho_2(y)$ is (by the definition of ρ_2) homotopic to $\gamma_1 * \gamma_2$, where γ_1 is homotopic to a suitable lift of β_y and γ_2 to a suitable lift of the reverse of β_x. Then the bound on distance follows exactly as in 28.1, that is, (3) of 28.1 cannot hold. □

28.3. A path $\Phi(\ell)$ and a Reduction in ECPP. — Let ℓ be a geodesic segment in $\mathcal{T}(Y)$ with endpoints in $\widetilde{V}_1$ for a component $\widetilde{V}_1$ of $\widetilde{V}$, and with initial endpoint x_0. We need to choose a path $\Phi(\ell) \subset \widetilde{V}_1$ with the same endpoints so that ECPP holds for $\Phi(\ell)$. As before, we define

$$F(x) = d_Z(x, \tau(x)) = d_Y(x, \tau(x)).$$

We shall choose

$$\Phi(x) = \lim_{m \to \infty} y_m(x),$$

where $y_0(x) = x$, and $y_m(x)$ has the following properties. We shall write $y_m(x) = y_m$ where possible, and also $y_m = [\varphi_m]$.

The properties depend on y_m: gaps or loops $\alpha_m = \alpha_m(x)$, (possibly $\alpha_m = \overline{\mathbf{C}}$); appropriate long, thick and dominant functions Δ, r, s (see 15.3) and a suitable constant m_0 in the Pole-Zero Condition (9.4, 15.8); closed subsets ℓ_m of ℓ with $\ell_{m+1} \subset \ell_m$ and $y_m(\ell_m) = y_m(\ell)$; constants D_1 M_1, M_1', Δ_0 with Δ_0 sufficiently large given D_1 and the long thick and dominant parameter functions; $\nu_0' > 0$; an integer k_0; a function $C : (0, \infty) \to (0, \infty)$ and constant K_0 as in Theorem 15.8. The following properties hold for $y_m = y_m(x)$, $x \in \ell$.

a) For each $x \in \ell$ and each m, either y_m or y_{m+1} is continuous at x, and the left and right limits $\lim_{x' \to x\pm} y_m(x')$ exist for all $x \in \ell$ and all m.

b) $d_Z(y_m, y_{m+1}) \leqslant D_1 F(y_m)$ and $\lim_{m \to \infty} F(y_m) = 0$.

c) For $t = t(m+1) \leqslant k_0$

$$[\varphi_{m+1}]_Y = [\psi'_{m+1,t}]_Y$$

where $[\psi'_{m+1,0}]_Y = [\varphi_m]_Y$ and

$$x^{i+1}_{m+1} = [\psi'_{m+1,i+1}] = [\xi_{m+1,i+1} \circ \psi_{m+1,i+1} \circ \psi_{m+1,i}]_Y$$

where $\psi_{m+1,i+1}$ minimizes distortion up to isotopy constant on $\psi'_{m+1,i}(Z)$, and $d_Y([\xi_{m+1,i}], [\text{identity}]) \leqslant M_2$ for all i and m. Moreover for every long ν-thick and dominant gap or m_0-Pole-Zero loop α along a segment of $[x^i_{m+1}, x^{i+1}_{m+1}]_Z$ containing some y, there is $y' \in [y_m, y_{m+1}]$ such that

$$d_{\alpha,Z}(y,y') \leqslant C(\nu) \quad \text{or} \quad |\operatorname{Re}(\pi_{\alpha,Z}(y) - \pi_{\alpha,Z}(y'))| \leqslant K_0,$$

depending on whether α is a gap or a loop.

d) For each m, one of the following holds for $y_m = y_m(x)$, $x \in \ell_m$.

d)(i) This property concerns only the projections of the y_m to $\mathcal{T}(Z)$. All distances are measured in $\mathcal{T}(Z)$, and all geodesic segments are in $\mathcal{T}(Z)$. The gap or loop α_m is long, ν_m-thick and dominant for some $\nu_m \geqslant \nu_0'$, or satisfies the Pole-Zero Condition, if a loop, on a segment centred on $y_m' \in [y_{m-1}, y_{m+1}]_Z$. The following hold.

(1)
$$\begin{gathered} d_{\alpha_m,\alpha_{m+1},Z}(y_m, y_{m+1}) \geqslant \Delta_0, \quad \alpha_m \cap \alpha_{m+1} \neq \varnothing, \\ d_{\alpha_m,Z}(y_m, y_m') \leqslant C(\nu_m), \quad \text{or} \quad |\operatorname{Re}(\pi_{\alpha_m,Z}(y_m) - \pi_{\alpha_m,Z}(y_m'))| \leqslant K_0 \end{gathered}$$

depending on whether α_m is a gap or a loop.

d)(ii) For all $k \geqslant m$, there is an invariant loop set Γ_k (not necessarily satisfying the Levy Condition 2.2, and possibly empty) with $y_k \in \mathcal{T}(\Gamma_k, M_1)$, and with the following properties. Let α_k denote the fixed set (2.8) of Γ_k. If Γ_k does not have a nonempty subset satisfying the Levy Condition, let $P_k = \alpha_k$. If Γ_k does have a nonempty subset satisfying the Levy Condition, let P_k be the irreducible fixed component of α_k. For all $k \geqslant m$, $x \in \ell_k$, $d_{\alpha_k}(y_k, \tau(y_k)) \leqslant M_1$, and $P_m \subset P_k$. Moreover, if $x \in \ell_k$, and $P_m = \alpha_m$ is not homeomorphic, then $d'_{P_m}(y_m, y_k) \leqslant M_1'$.

Note that Condition 28.3d)(ii) implies that, if $P_m \neq P_k$ for a least $k \neq m$, then P_k is not homeomorphic, P_j is not homeomorphic for all $j \geqslant k$ and $d'_{P_k}(y_k, y_j) \leqslant M_1'$ for all $j \geqslant k$.

28.4. Proposition. — *If Φ, y_m have the properties outlined in 28.3, and ℓ is the geodesic in $\mathcal{T}(Y)$ between x_0, $x \in \widetilde{V}$, then $\Phi(\ell)$ and ρ_2 satisfy ECPP of 25.5, for a suitable sequence $\{a_n\}$.*

This will be proved in 28.7 below. First, we need some preliminaries. We use the results of Chapters 14 and 15, and the various different notions of length given in Chapter 14.

28.5. Lemma. — *The following holds for some function $C_1 : (0,\infty) \to (0,\infty)$, and $K_1 > 0$ for a suitable choice of long thick and dominant functions. Let 28.3d)(i) hold for k replacing m, for all $j \leqslant k < m$ apart from (1) of 28.3d)(i), which is not needed. Then for all $j < m$, there is $y_{j,m,k} \in [y_j, y_m]_Z$ such that*

$$d_{\alpha_k,Z}(y_k, y_{j,m,k}) \leqslant C_1(\nu_k) \quad \textit{or} \quad |\operatorname{Re}(\pi_{\alpha_k,Z}(y_k) - \pi_{\alpha_k,Z}(y_{j,m,k}))| \leqslant K_1,$$

depending on whether α_k is a gap or a loop. Also, given new long thick and dominant parameter functions and a new Pole-Zero constant, the original long thick and dominant parameter functions and Pole-Zero constant can be chosen so that α_k is *long thick and dominant or satisfies the Pole-Zero Condition with respect to the new parameter functions and constant on a segment of* $[y_j, y_m]_Z$ *centred on* $y_{j,m,k}$.

Proof. — Fix m, j and k with $j < k < m$. The result is about $\mathcal{T}(Z)$. For the rest of the proof, we regard all points as being in $\mathcal{T}(Z)$, distances are measured in $\mathcal{T}(Z)$ and geodesic segments are in $\mathcal{T}(Z)$. We shall obtain the result from the criterion (3) of 15.9. The criterion 15.9 has two alternatives, involving a loop $\gamma \subset \alpha_k$ such that $|\varphi_k(\gamma)|$ is bounded, and two loops γ_0', γ_1' which both intersect γ transversally, but are mutually disjoint. We either have to prove some condition on y_m or on y_j. First we have to decide which. Let $y_k' \in [y_{k-1}, y_k]$ be as in 28.3d). Let $q_k(z)dz^2$ be the quadratic differential at $y_k' = [\varphi_k']$ for $d(y_k', y_{k+1})$. Then $|\varphi_k'(\gamma)|_{q_k}$ is also bounded. Now $|\varphi_k'(\gamma_0' \cap \alpha_k)|_{q_k}$ is boundedly proportional to either $|\varphi_k'(\gamma_0' \cap \alpha_k)|_{q_k,+}$ or $|\varphi_k'(\gamma_0' \cap \alpha_k)|_{q_k,-}$. We suppose without loss of generality that it is boundedly proportional to $|\varphi_k'(\gamma_0' \cap \alpha_k)|_{q_k,+}$. Then $|\varphi_k'(\gamma_1' \cap \alpha_k)|_{q_k}$ is also boundedly proportional to $|\varphi_k'(\gamma_1' \cap \alpha_k)|_{q_k,+}$, using 15.11 (as usual) because otherwise the loops would be disjoint. Now take any loop γ' intersecting γ transversally. Then, using (3) of 15.9 we need to show that, for a suitable constant M depending only on the choice of long thick and dominant parameter functions

$$(1_m) \qquad |\varphi_m(\gamma)|' \leqslant M|\varphi_m(\gamma_0')$$

The line of proof will then follow that of 15.8. We actually need to prove similar statements, not only for y_k but for $y = [\varphi]$ in a segment around y_k' in $[y_{k-1}, y_k]$ for γ replaced by any loop $\gamma_y \in \alpha_k$ such that $|\varphi(\gamma_y)|$ is bounded. The method will be the same. To prove (1_m) we use an induction. For $k \leqslant \ell < m$ we choose points $y_{\ell,1} \in [y_{\ell-1}, y_{\ell+1}]$ and for $k < \ell < m-1$ we also choose $y_{\ell,2} \in [y_\ell, y_{\ell+2}]$ with the following properties. The point $y_{\ell,1}$ is in the middle third segment centred on y_ℓ' along which α_ℓ is long thick and dominant or satisfies the Pole-Zero Condition, nearer $y_{\ell+1}$ than y_ℓ'. We can assume that the long thick and dominant functions and Pole-Zero constant are sufficiently good for it to be possible to apply 15.8 twice and obtain, for $C_2 : (0, \infty) \to (0, \infty)$ and $K_2 > 0$ depending only on the long thick and dominant functions and Pole-Zero constant,

$$d_{\alpha_\ell, Z}(y_{\ell,1}, y_{\ell,2}) \leqslant C_2(\nu_\ell) \quad \text{or} \quad |\operatorname{Re}(\pi_{\alpha_\ell, Z}(y_{\ell,1}) - \pi_{\alpha_\ell, Z}(y_{\ell,2})| \leqslant K_2,$$

depending on whether α_ℓ is a gap or a loop, and moreover we can assume the original long thick and dominant functions and Pole-Zero constant were sufficiently good that α_ℓ is also long thick and dominant or satisfies the Pole-Zero Constant along a segment of $[y_\ell, y_{\ell+2}]$ containing $y_{\ell,2}$ for (different) specified long thick and dominant parameter functions and Pole-Zero Condition which we shall need in a moment. Now write $y_{\ell,i} = [\varphi_{\ell,i}]$. Then we shall prove by induction on ℓ that

$$(1_{\ell,i}) \qquad |\varphi_{\ell,i}(\gamma)|' \leqslant M|\varphi_{\ell,i}(\gamma_0')$$

We shall obtain $(1_{\ell,i})$ inductively. First we shall obtain $(1_{k,1})$, and $(1_{\ell,1})$ will always imply $(1_{\ell,2})$ if $\ell < m-1$. Then $(1_{\ell,2})$ will imply $(1_{\ell+1,1})$ if $\ell < m-1$, while $(1_{m-1,1})$ will imply $(1_{m,1})$.

Let $q_{\ell,i}(z)dz^2$ be the quadratic differential at $y_{\ell,i}$ for $d(y_{\ell,i}, y_{\ell+i})$. Let $\mathcal{G}_{\ell,i,\pm}$ denote the expanding and contracting foliations of $q_{\ell,i}(z)dz^2$, that is, $\mathcal{G}_{\ell,i,+}$ expands when moving *forward* towards $y_{\ell+i}$. For $\ell > k$, $q_{\ell,1}(z)dz^2$ is the stretch of $q_{\ell-1,2}(z)dz^2$ at $y_{\ell,1}$.

The idea of the proof of $(1_{\ell,i})$ is to *lock* each point on $\varphi_{\ell,1}(\gamma)$ to a point on $\varphi_{\ell,i}(\gamma_0')$, where the locking is along bounded stable foliation segments, and there is a bound on the number of points on $\varphi_{\ell,1}(\gamma)$ locked to any particular point of $\varphi_{\ell,1}(\gamma')$. To prove the inductive step $(1_{\ell,i})$ we need a larger set of inductive properties, as follows. The properties involve an integer N which is bounded in terms of ν_k.

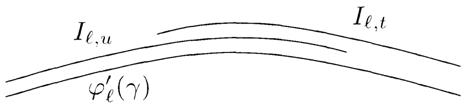
For each $k \leqslant \ell < m$, and $i = 1, 2$ (or just $i = 1$ if $\ell = m-1$) we shall find segments $I_{\ell,i,t} \subset \varphi_{\ell,i}(\gamma_0')$ for $1 \leqslant t \leqslant N' \leqslant N$, and maps

$$\sigma_{\ell,i,t} : I_{\ell,i,t} \longrightarrow \varphi'_{\ell,i}(\gamma) \cup (\cup_{u \neq t} I_{\ell,i,u})$$

with the following properties, with respect to a suitable function $\Delta_1 : (0,1) \to (1,\infty)$, which can be chosen suitably provided that the original long thick and dominant functions and Pole Zero function are suitably chosen.

$(2_{\ell,i})$ Each $I_{\ell,i,t}$ has, at each end a segment $I_{\ell,i,t,v}$ ($v = 0$ or 1) in $S(\alpha_\ell, y_{\ell,i}, \varepsilon_0)$ of length $\geqslant \Delta_1(\nu_{\ell,i,1})$.

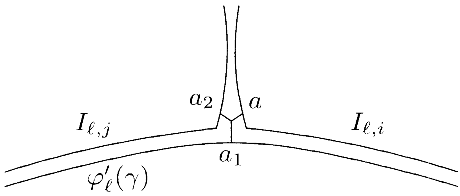
$(3_{\ell,i})$ The map $\sigma_{\ell,i,t}$ maps $I_{\ell,i,t} \smallsetminus (I_{\ell,i,t,0} \cup I_{\ell,i,t,1})$ to $\varphi_{\ell,i}(\gamma)$, maps $I_{\ell,t,i,0}$ either to $\varphi_{\ell,i}(\gamma)$ (case 1) or to $I_{\ell,i,u,1}$ for some u (case 2), and similarly for $I_{\ell,i,t,1}$. Each point on $\varphi_{\ell,i}(\gamma)$ is in $\sigma_{\ell,i,t}(I_{\ell i,t} \smallsetminus (I_{\ell,i,t,0} \cup I_{\ell,i,t,1}))$ for some t. Any points a and $\sigma_{\ell,i,t}(a)$ are joined along a $\mathcal{G}_{\ell,i,-}$-segment. For $x \in I_{\ell,i,t,v}$ ($v = 0$ or 1), this segment has length $\leqslant \Delta_1(\nu_\ell)$. The map $\sigma_{\ell,i,t}$ is continuous except possibly at a point $a \in \partial I_{\ell,i,t,v} \cap \partial(I_{\ell,i,t} \smallsetminus I_{\ell,i,t,v})$, for $v = 0$ or 1. At such a point, right and left limits $\sigma_{\ell,i,t}(a\pm)$ exist. If $v = 0$, then with $\sigma_{\ell,i,t}(a+) = a_1 \in \varphi_{\ell,i}(\gamma)$ and $\sigma_{\ell,i,t}(a-) = a_2 \in \partial I_{\ell,i,u} \cap \partial(I_{\ell,i,u} \smallsetminus I_{\ell i,u,1})$. Moreover, $\sigma_{\ell,i,u}(a_2-) = \sigma_{\ell,i,t}(a+) = a_1$ and $\sigma_{\ell,i,u}(a_2+) = \sigma_{\ell,i,t}(a-)$. See the diagrams.



Locking: Case 1.

$(4_{\ell,i})$ Any segment J of $I_{\ell,i,t,v}$ ($v = 0$ or 1) of length $\geqslant \Delta_1(\nu_\ell)$ has $|J|_{q_{\ell,i},+}| \geqslant \Delta_1(\nu_\ell)^{-1}|J|_{q_{\ell,i}}$.

The conditions $(2_{\ell,i})$ to $(4_{\ell,i})$ are sufficient to give $(1_{\ell,i})$. We start by obtaining $(2_{k,1})$ to $(4_{k,1})$, using 15.11. Since $|\varphi'_k(\gamma_0' \cap \alpha_k)|_{q_k}$ is boundedly proportional to $|\varphi'_k(\gamma_0' \cap \alpha_k)|_{q_k,+}$, $\varphi_{k,1}(\gamma_0')$ contains long segments almost tangent to the unstable foliation, if $d(y_{k,1}, y_k)$ is large enough, and, again if $d(y_{k,1}, y_k)$ is large enough given δ,



Locking: Case 2.

these long segments come within δ of every point in $S(\alpha_k, y_{k,1}, \varepsilon_0)$, by 15.11. Take $d(y_{k,1}, y'_k)$ just large enough for this, bounded in terms of some integer N, so that $|\varphi'_{k,i}(\gamma)|$ is bounded in terms of N. Then we can construct segments $I_{k,1,t} \subset \varphi_{k,1}(\gamma'_0)$ and maps $\sigma_{k,1}$ with properties $(2_{k,1})$ to $(4_{k,1})$

Now let $\ell > k$. Suppose inductively that we have $(2_{\ell-1,2})$ to $(4_{\ell-1,2})$. Then $(2_{\ell,1})$ to $(4_{\ell,1})$ also hold. In fact we claim that the estimates are better in some respects.

In $(2_{\ell,1})$: the segments of $I_{\ell,1,t,v}$ ($v = 0$ or 1) can be of length $\geqslant \Delta_2(\nu_\ell)$ where $\Delta_2(\nu_\ell)$ can be taken arbitrarily long by suitable the "long" function Δ with respect to which α_ℓ is long thick and dominant, or similarly if α_ℓ is a loop.

In $(3_{\ell,1})$: the length of the $\mathcal{G}_{\ell,1,-}$-segment joining x and $\sigma_{\ell,1,t}(x)$ for $x \in I_{\ell,1,t,v}$ can be taken of length $\leqslant \Delta_2(\nu_\ell)^{-1}$.

In (4_ℓ): any segment J of $I''_{\ell,1,t,v}$ ($v = 0$ or 1) of length $\geqslant \Delta_1(\nu_\ell)$ has

$$(4'_\ell) \qquad |J|_{q_{\ell,1},-} \leqslant \Delta_1(\nu_\ell)^{-1}|J|_{q_{\ell,1}}.$$

So we need to obtain the $I_{\ell,1,t}$, $\sigma_{\ell,1,t}$. Let χ be the map minimizing distortion with $[\chi \circ \varphi_{\ell-1,2}] = [\varphi_{\ell,1}]$. Then

$$\varphi_{\ell,1}(\gamma'_0) = \chi \circ \varphi_{\ell-1,2}(\gamma'_0), \quad \varphi_{\ell,1}(\gamma) = \chi \circ \varphi_{\ell-1,2}(\gamma).$$

We take $I_{\ell,1,t,v} \subset \chi(I_{\ell-1,2,t,v})$, and then take $I_{\ell,1,t}$ to be the subset of $\chi(I_{\ell-1,2,t})$ which ends in the segments $I'_{\ell,1,t,v}$. We need the fact that $\alpha_{\ell-1} \cap \alpha_{\ell,1} \neq \varnothing$. We then take $\sigma_{\ell,1,t} = \chi \circ \sigma_{\ell,2,t} \circ \chi^{-1}$. Thus $(2_{\ell-1,2})$ to $(4_{\ell-1,2})$ imply $(2_{\ell,1})$ to $(4_{\ell,1})$. In the same way, $(2_{k-1,2})$ to $(4_{k-1,2})$ imply (2_k) to (4_k).

It remains to show that $(2_{\ell,1})$ to $(4_{\ell,1})$ imply $(2_{\ell,2})$ to $(4_{\ell,2})$. We have a homeomorphism χ' which is bounded restricted to $S(\alpha_\ell, y_{\ell,1}, \varepsilon_0)$ in terms of $C(\nu_\ell)$ — or to a subannulus of this if α_ℓ is a loop — such that $[\chi' \circ \varphi_{\ell,1}] = [\varphi_{\ell,2}]$. We can also assume χ' is such that $\chi' \circ \varphi_{\ell,1}(\gamma)$, $\chi' \circ \varphi_{\ell,1}(\gamma'_0)$ are in good position. We claim that we can obtain $I_{\ell,2,t}$, $I_{\ell,2,t,v}$ with the required properties from $\chi'(I_{\ell,1,t})$, $\chi'(I_{\ell,1,t,v})$ by moving the endpoints a bounded distance in $S(\alpha_\ell, y_{\ell,2}, \varepsilon_0)$. The paths $I_{\ell,1,t}$, the path $\varphi_{\ell,1}(\gamma)$ and the $\mathcal{G}_{\ell,1,-}$-segments between them, give rise to a union of discs, foliated by $\mathcal{G}_{\ell,1,-}$-segments, although the discs can contain zeros of the foliation and the foliation is not wholly tangent, nor wholly transverse to the boundaries of the discs. These discs are mapped by χ' to discs, which are homotopically trivial in $\overline{\mathbf{C}} \setminus \varphi_{\ell,2}(Z)$. The foliation $\mathcal{G}_{\ell,1,-}$, is not, of course, mapped to $\mathcal{G}_{\ell,2,-}$ by χ'. But we still have a foliation of the new discs by $\mathcal{G}_{\ell,2,-}$-segments. We can obtain the $I_{\ell,2,t}$, $I_{\ell,2,t,v}$ with the required properties, provided that no segment J of $\chi'(I_{\ell,1,t,v})$ of length $\geqslant \Delta_1(\nu_\ell)$ in $S(\alpha_\ell,[\varphi'_{\ell,1}],\varepsilon_0)$ satisfies

$$|J|_{q_{\ell,2},+} \leqslant \Delta_1(\nu_{\ell,1})^{-1}|J|_{q_{\ell,2}}.$$

Suppose that there is such a segment. Remember that we also have $(4_{\ell,1})$. Then we can take such a J whose endpoints in $S(\alpha_\ell,[\varphi'_{\ell,1}],\varepsilon_0)$ can be joined by a bounded segment not crossing J. to form a closed simple loop $\varphi_{\ell,2}(\zeta)$ and such that

$$|\varphi_{\ell,2}(\zeta)|_{q_{\ell,2},+} \leqslant 2\Delta_1(\nu_{\ell,2})^{-1}|\varphi_{\ell,2}(\zeta)|_{q_{\ell,2}},$$
$$|\varphi_{\ell,1}(\zeta)|_{q_{\ell,1},-} \leqslant 2\Delta_1(\nu_{\ell,1})^{-1}|\varphi_{\ell,1}(\zeta)|_{q_{\ell,1}}.$$

Then the points $w=[\psi]\in[y_{\ell-1},y_{\ell+1}$ and $w'\in[t'_\ell,y_{\ell+2}]$ at which $\psi(\zeta)$, $\psi'(\zeta)$ achieve the minimum length along these geodesics (to within a bounded distance) are such that w is to the left of $y_{\ell,1}$ and w' to the right of $y_{\ell,2}$ — arbitrarily far from $y_{\ell,1}$, $y_{\ell,2}$ by choice of $\Delta_1(\nu_{\ell,1})$. This contradicts the bound on $d_{\alpha_\ell}(y_{\ell,1},y_{\ell,2})$ and the similar estimates throughout the corresponding segments on which α_ℓ is long thick and dominant, for suitable choice of $\Delta_1(\nu_\ell)$, if α_ℓ is a gap. The case of α_ℓ a loop is similar. So the inductive step is completed.

28.6. The No-cell-cutting Property. — A geodesic segment $\ell\subset\overline{\mathbf{C}}\setminus\varphi(Z)$ has the *no-cell-cutting property* with respect to a given $L_0>0$ and integer N if there is no set of successive subsegments $\ell_i\subset\ell$ $(1\leqslant i\leqslant n)$ such that each ℓ_i has length $\leqslant L_0$ and each component of $\overline{\mathbf{C}}\setminus(\varphi(Z)\cup\ell_i)$ is a topological disc with at most one puncture. This is a useful concept, because for any given L_0, N, with L_0 sufficiently long given $\varphi(Z)$, if R is sufficiently large, then the set of all geodesic segments of length R in $\overline{\mathbf{C}}\setminus\varphi_0(Z)$ with the no-cell-cutting property, and all with the same starting point, has relatively small measure.

Lemma. — *Let*

$$[\varphi_k]_Z=[\chi_{k,j}\circ\varphi_j]_Z=[\chi_k]_Z$$

where $\chi_{k,j}$ minimizes distortion up to isotopy constant on $\varphi_j(Z)$, χ_k minimizes distortion up to isotopy constant on Z. Let

$$[\varphi_0]_Y=[\chi_0\circ\sigma_{\beta'}]_Y$$

for a path β' starting from v_2. Let L_0 be given. Then for suitable choice of long thick and dominant functions in 28.3, and for N depending on L_0 and D_1 of 28.3, the following holds.

(1) $$[\varphi_m]_Y=[\chi_{m,j}\circ\sigma_{\beta_{m,j}}\circ\varphi_j]_Y=[\chi_m\circ\sigma_{\beta_m}\circ\sigma_{\beta'}]_Y,$$

(2) $$[\varphi_j]_Y=[\sigma_{\beta'_{0,j}}\circ\chi_{j,0}\circ\varphi_0]_Y$$

such that the geodesics homotopic to $\beta_{m,j}$, β_m via homotopies preserving endpoints have the no cell-cutting property (with respect to N and L_0). So does the path $\beta'_{0,j}$ if y_i satisfies 28.3(i) for all $i \leqslant j$.

Proof. — We first consider the case of $\beta'_{0,j}$. The case of $\beta_{m,j}$ will be similar. Let Γ_0 be a set of loops in $\varphi_0(\overline{\mathbf{C}} \setminus Z)$ such that $\overline{\mathbf{C}} \setminus (\cup\Gamma_0 \cup \varphi_0(Z))$ is a union of discs with at most one puncture. We assume that Γ_0 is in good position with respect to the quadratic differential for $d_Z(y_0, y_j)$. We start by choosing a simple path η from some intersection of two loops of Γ_0 to $\varphi_0(v_2)$ such that the intersection of γ with $S(y_0, \varepsilon)$ has bounded Poincaré length. Then define

$$\eta_1 = \chi_{j,0}(\eta), \quad \eta_2 = \varphi_j \circ \varphi_0^{-1}(\eta).$$

Then

$$[\varphi_j]_Y = [\sigma_{\eta_2 * \overline{\eta}_1} \circ \chi_{j,0} \circ \varphi_0]_Y.$$

Now η_1 is isotopic to a geodesic segment in $\overline{\mathbf{C}} \setminus \varphi_j(Z)$, under isotopy preserving endpoints, and hence, since it is simple, clearly has the no-cell-cutting property. But there is no reason why this should be true for η_2, which we now analyse more carefully. This proceeds in two steps. We shall write

$$[\varphi_k]_Y = [\sigma_{\zeta_k} \circ \chi_{k,k-1} \circ \varphi_{k-1}]_Y$$

for a path ζ_k which has the No-cell-cutting property (for suitable N and L_0) and a stronger property, to be specified — and then we shall obtain (2), where

(3) $$\beta'_{0,j} = \beta'_{0,j,1} * \cdots * \beta_{0,j,j}$$

where $\beta'_{0,j,k}$ has some connection to ζ_k, to be specified. Using the notation of 28.3c), and redefining the $\xi_{k,i}$ if necessary, we can assume either that the gap or loop $\omega_{k,i}$ containing $\psi'_{k,i}(v_2)$ is the end of a segment of $[x_k^{i-1}, x_k^i]$ and is long ν-thick and dominant or m_0-pole-zero (by 15.14), or $d_{\omega_{k,i}}(x_k^i, x_k^{i-1})$ is bounded. Then there is a corresponding point $y_k^i \in [y_{k-1}, y_k]_Z$ (as an element of $\mathcal{T}(Z)$) such that

(4) $$d_{\omega_{k,i},Z}(y_k^i, x_k^i) \leqslant C_1(\nu)$$

or

(5) $$|\operatorname{Re}(\pi_{\omega_{k,i},Z}(y_k^i) - \pi_{\omega_{k,i},Z}(x^i_{,k}))| \leqslant K_1$$

where the functions $C_1(\nu)$ and constant K_1 depend only on the function $C(\nu)$ and constant K_0 of 15.8 (and the long thick and dominant parameter functions and Pole-Zero constant) and the integer r_0 of 28.3c). We choose y_k^i as a element of $\mathcal{T}(Y)$ so that (3) or (4) holds — for possibly larger constants (although we could use the same ones) — with $d_{\omega_{k,i},Z}$ replaced by $d_{\omega_{k,i},Y}$ and $\pi_{\omega_{k,i},Z}$ replaced by $\pi_{\omega_{k,i},Y}$. Then we can write $y_k^i = [\varphi_k^i]_Y$ and

$$[\varphi_k^{i+1}]_Z = [\chi_k^{i+1} \circ \varphi_k^i]_Z$$

where χ_k^i minimises distortion up to isotopy constant on $\varphi_k^i(Z)$. Then by the $\mathcal{T}(Y)$ versions of (1) and (2), we can find bounded homeomorphisms θ_k^ℓ such that for all i, k,

$$[\varphi_k^i]_Y = [\theta_{0,j,k}^i \circ \psi'_{k,i}]_Y.$$

So then for $0 \leqslant k \leqslant j$,

$$[\chi_k^i \circ \varphi_k^{i-1}]_Y = [\theta_k^i) \circ \xi_k^i \circ \chi_k^i \circ (\theta_k^{i-1})^{-1} \circ \varphi_k^{i-1}]_Y$$

Then take a loop set Γ (more properly $\Gamma_{k,i-1}$) in $\overline{\mathbf{C}} \setminus \varphi_k^{i-1}(Z)$ such that the complement of $\cup\Gamma$ in $\overline{\mathbf{C}} \setminus \varphi_k^{i-1}(Z)$ is a union of at most once-punctured discs, in good position with respect of the quadratic differential for $d_Z(y_k^{i-1}, y_k^i)$ at $y_{0,j,k}^{i-1}$, and considering the homotopy between $(\theta_k^{i-1})^{-1}(\Gamma)$ and the good position of this with respect to the quadratic differential for $d_Z(x_k^{i-1}, x_k^i)$ at x_k^{i-1}, we obtain

$$[\varphi_k^i] = [\sigma_{\delta_{k,i}} \circ \chi_k^i \circ \varphi_k^{i-1}]$$

where

$$\delta_{k,i} = \delta'_{k,i} * (\theta_k^i \circ \xi_k^i)(\delta''_{k,i})$$

for bounded paths $\delta'_{k,i}$ and $\delta''_{k,i}$. Then write

$$\chi_k^{\ell,i} = \chi_k^\ell \circ \cdots \circ \chi_k^{i+1}, \quad y_{0,j,k} = [\varphi_{0,j,k}], \quad [\varphi_{0,j,\ell}]_Z = [\chi_{0,j,\ell,k} \circ \varphi_{0,j,k}]$$

if $0 \leqslant k \leqslant \ell \leqslant j$, where $\chi_{0,j,\ell,k}$ minimises distortion up to isotopy constant on $\varphi_{0,j,k}(Z)$, and

$$\zeta_k = \chi_k^{r,1}(\delta_{k,1}) * \cdots * \chi_k^{r,r-1}(\delta_{k,r-1}) * \delta_{k,r}.$$

Then

$$[\varphi_k]_Y = [\sigma_{\zeta_k} \circ \chi_{k,k-1} \circ \varphi_{k-1}]_Y.$$

Then for fixed k, for any given function $\varepsilon : (0,\infty) \to (0,\infty)$, assuming $d(y_k, y_{k+1})$ is sufficiently large, by choosing i suitably so that $d_Z(y_k, y_k^i)$ is bounded and much smaller than $d_Z(y_k, y_k^{i-1})$, we can write ζ_k in the form $\zeta_{k,1} * \zeta_{k,2}$ where $|\zeta_{k,2}| \leqslant L_2$ for a suitable constant L_2, and $|\zeta_{k,1}|_{q_{k,k-1},+} < \varepsilon(L_2)$, where $q_{k,k-1}(z)dz^2$ is the quadratic differential at y_k for $d_Z(y_{k-1}, y_{k-1})$, with stable foliation contracted when moving towards y_{k-1}. This property implies that any segment of $\zeta_{k,1}$ intersecting $S(\alpha_k, y_k, \varepsilon_0)$ has length bounded from 0. Also the number of self-intersections on ζ_k is bounded in terms of r and M_2.

Now we apply a similar technique, replacing the points x_k^i $(0 \leqslant i \leqslant r)$ by the points y_k $(0 \leqslant k \leqslant j)$ and the points y_k^i by the points $y_{0,j,k}$ — which are as in 28.5. This time, the homeomorphism θ_k^i are replaced by homeomorphisms $\theta_{0,j,k}$ with $\theta_{0,j,k}$ bounded only on $S(y_k, \alpha_k, \varepsilon_0)$. Then we obtain (2) and (3) with

$$\beta'_{0,j,k} = \chi_{j,k}(\zeta''_{k+1} * \zeta'_k) * \theta_{0,k,k}(\zeta_k)$$

with bounds only on $|\zeta'_k \cap S(y_{0,j,k}, \alpha_k, \varepsilon_0)|$, $|\zeta''_{k+1} \cap S(y_{0,j,k}, \alpha_k, \varepsilon_0)|$. But we do also have a bound on the number of self-intersections of each $\beta'_{0,j,k}$.

So $\beta'_{0,j,k}$ has the no-cell-cutting property for N depending only on r and M_2. To show that $\beta'_{0,j}$ does also (for suitable constants) it suffices to show that

$$\beta'_{0,j,1} * \cdots * \beta'_{0,j,j-1}$$

has the no-cell-cutting property, for suitable constants. In fact, it is more convenient to consider $\chi_j'^{-1}(\beta'_{0,j,1} * \cdots * \beta'_{0,j,j-1})$ where $[\varphi'_j] = [\chi_j'^{-1} \circ \varphi_j]$ is a point on $[y_0, y_j]$ at which α_j is long thick and dominant or satisfies the Pole-Zero Condition — but we assume that α_j is a gap. To simplify the notation, we assume that χ'_j is the identity. We also choose a loop $\gamma_i \subset \alpha_i$ such that $\varphi_i(\gamma_i)$ is bounded.

Suppose for contradiction that N successive segments of $\cup_{k\leqslant j-1}\beta'_{0,j,j-1}$ of total length $\leqslant L_0$ cut $\varphi'_{j,1}(\alpha_j)$ into cells. We can homotope this union of segments, via a homotopy preserving $\varphi_j(\partial\alpha_j)$, to a union ℓ of good position segments, moving endpoints of $\beta'_{0,j,i}$ in $\varphi_j(\alpha_j)$ along stable segments of the foliation for the quadratic differential $q_j(z)dz^2$ for $d_Z(y_0, y_j)$. We continue to call the respective images under homotopy $\beta'_{0,j,i} \cap \ell$, although some of these will now have zero length.

Now we claim that for each $\beta'_{0,j,k}$ lying entirely in ℓ, $\#(\beta'_{0,j,k} \cap \varphi_j(\gamma_k))$ is bounded, with a bound depending only on M_2, if $d_Z(y_i, y_{i+1})$ is big enough, for all $i \leqslant j$. For

$$\#(\beta'_{0,j,k} \cap \varphi_j(\gamma_i)) = \#(\chi_{0,j,k,j}(\beta'_{0,j,k}) \cap \varphi_{0,j,k}(\gamma_i)).$$

Now $\chi_{0,j,k,j}(\beta'_{0,j,k})$ is a union of 3 segments of length bounded by M_2 and of segments $\theta_{0,j,k}(\zeta_{k,i})$ for $i < r$ where the $\zeta_{k,i}$ are either bounded or are close to the unstable foliation of the quadratic differential for $d_Z(y_{k-1}, y_k)$ at y_k. The segments $\theta_{0,j,k}(\zeta_{k,i})$ which are not bounded then have unstable length bounded from 0 with respect to the quadratic differential for $d_Z(y_{0,j,k}, y_j)$ at $y_{0,j,k}$: because we can find long segments which are a bounded distance from $\varphi_{0,j,}(\gamma')$ for some loop $\gamma' \subset \alpha_k$ such that $\varphi'(\gamma')$ is bounded in the segment along which α_k is long thick and dominant but $[\varphi']$ is nearer $y_{0,j,k-1}$ than $y_{0,j,k}$. But if this happens then the contribution of $\zeta_{k,i}$ to $\beta_{0,j,k}$ gives length $> L_2$, assuming only that $d_Z(y_j, y_{0,j,k})$ is large enough. So $\chi_{0,j,k,j}(\beta'_{0,j,k})$ is bounded in terms of M_2, and $\#(\beta'_{0,j,k} \cap \varphi_j(\gamma_k))$ is indeed bounded. Since ℓ cuts $\varphi_j(\alpha_j)$ into cells, we can choose $\ell_1 \subset \varphi'_j(\alpha_j) \cap \ell$ with endpoints in $\varphi_j(\partial\alpha_j \cup \gamma_j)$ such that ℓ_1 is at constant angle to $q_j(z)dz^2$, and $|\ell_1|_{q_j,-} \geqslant C(L_0)|\ell_1|_{q_j}$ for $C(L_0) > 0$. Now any long segment of $\varphi_j(\gamma_i)$ $(i < j)$ is almost tangent to the unstable foliation. So by 15.11, for Δ_0 as in 28.3

$$\#(\ell_1 \cap \varphi_j(\gamma_{j-1})) \geqslant n(\Delta_0) \tag{6}$$

where

$$\lim_{\Delta_0 \to \infty} n(\Delta_0) = +\infty.$$

This follows from the fact that by 15.14 there must be a very long length of segments $\ell(\alpha_{j,i})$ between y_j and $y_{0,j-1,i}$ such that $\alpha_{j,i}$ is long thick and dominant or satisfies the Pole-zero Condition on $\ell(\alpha_{j,i})$ and $\alpha_{j,i+1} \cap \alpha_{j,i} \neq \varnothing$. In particular, (6) is true for $i = j-1$. So there is a segment $\ell_{1,j-1} \subset \ell_1$ with both endpoints on $\varphi_j(\gamma_{j-1})$ which lies in $\varphi'_j(\alpha_{j-1})$, and is not intersected by $\beta'_{0,j,j-1}$. Now we take preimages under $\chi'_{0,j,j,j-1}$ where $[\varphi_j]_Z = [\chi_{j,j-1} \circ \varphi_{0,j,j-1}]$. Eventually we obtain that for some k $\#(\beta'_{0,j,k} \cap \varphi_j(\gamma_k)) \geqslant n(\Delta_0)$, in contradiction to what was shown before.

The construction of $\beta_{m,j}$ is similar, with y_m taking the role of y_0. If y_i satisfies 28.3d)(ii) for some $i < m$ then this is also true for all k with $i \leqslant k \leqslant m$. Let i_1 be the least such i, or $i_1 = m$. If $i_1 < m$ then there are i_n for $n \leqslant r$, some r bounded in terms of $\#(Y)$ such that $i_r = m$, $\varphi_m(\partial P_{i_j})$ is bounded for $j \leqslant m$ and $\chi_{m,i_1}(\beta_{m,i_1,k})$ does not intersect the bounded loop set $\varphi_m(\partial P_{i_\ell})$ for $i_\ell \leqslant k < i_{\ell+1}$. Hence there is N_1 depending only on D_1, M_1, M'_1 and $\#(Y)$, $\chi_{m,i_1}(\beta_{m,i_1})$ can be written up to homotopy fixing the endpoints as a union of $\leqslant N_1$ geodesic segments such that if ℓ is one of these segments, not all components of $\overline{\mathbf{C}} \smallsetminus (\varphi_m(Z) \cup \ell)$ are discs with at most one puncture. This property is preserved under homeomorphism (unlike the weaker no-cell-cutting property). After extending the endpoints of β_{m,i_1} a bounded amount, $\beta_{m,j} = \beta_{j,i_1} * \chi^{-1}_{i_1,j}(\beta_{m,i_1})$ and hence also has the no-cell-cutting property, possibly for some N bounded in terms of C_1, M_1, M'_1. So we can allow y_i to satisfy 28.3d)(ii) for some i.

Now we want to adapt this argument to obtain similar results for the decomposition $[\varphi_m]_Y = [\chi_m \circ \sigma_{\beta_m} \circ \sigma_{\beta'}]_Y$. We want to use a similar argument to the above. So we need to consider the geodesics $[x_0, y_0]$, $[y_0, y_m]$ and $[x_0, y_m]$ for any given m. Remember that $x_0 =$ [identity]. Then by 15.8, we can find $p \leqslant m$ and $w \in [x_0, y_0]_Z$, $w' \in [x_0, y_p]_Z$ (if $p < m$) such that $d_{\alpha_p}(w, y_p) \leqslant C(\nu_p)$, $d_{\alpha_{p+1}}(w', y_{p+1}) \leqslant C(\nu_{p+1})$. If we drop the condition $\alpha_p \cap \alpha_{p+1} \neq \varnothing$ then, by inserting extra points y_j on $[y_p, y_{p+1}]$ if necessary, we can still assume that w, w' exist as above, and that there exists $w'' \in [y_p, y_{p+1}]_Z$ with no long thick and dominants on $[w'', y_{p+1}]$ intersecting α_{p+1} and not equal to α_{p+1}, and similarly for $[y_p, w'']$. Then by 15.14 there is K_1 depending on the long thick and dominant functions such that $d'_{\alpha_i}(y_i, w'') \leqslant K_1$, $i = p$, $p+1$. Now write

$$\chi_{p+1,p} = \chi^2 \circ \chi^1, \quad \pi_Z(w'') = [\chi'],$$

where $w'' = [\chi^1 \circ \varphi_p]$, and χ', χ^1, χ^2 minimize distortion up to isotopies constant on Z, $\varphi_p(Z)$, $\chi^1 \circ \varphi_p(Z)$. Hence, $[\chi_p] = [\xi_1 \circ \chi^1_0]_Z$, $[\chi_{p+1} = [\xi_2 \circ \chi^1_m]$, where ξ_1, ξ_2 are bounded on $\chi^1_0(\alpha_p)$, $\chi^1_m(\alpha_{p+1})$. Now by the same method as at the start of the proof,

$$\chi^1 \circ \chi_p = \chi' \circ \sigma_{\zeta_1}, \quad (\chi^2)^{-1} \circ \chi_{p+1} = \chi' \circ \sigma^{-1}_{\zeta_2},$$

where ζ_1, ζ_2 have $\leqslant C(K_1)$ intersections with $\chi_p(\alpha_p)$, $\chi_{p+1}(\alpha_{p+1})$. For any i, j, define $[\varphi_{j,i}]$ by

$$[\varphi_j]_Y = [\varphi_{j,i} \circ \varphi_i]_Y.$$

Therefore

$$[\varphi_m]_Y = [\varphi_{m,p+1} \circ \xi_{p+1} \circ \sigma_{\delta_{p+1}} \circ \chi_{p+1} \circ \chi^{-1}_{p+1} \circ \chi^2 \circ \chi^1 \circ \varphi_{p,0} \circ \chi_0 \circ \sigma_{\beta'}]_Y$$

Now apply the method of showing the no-cell-cutting property for $\beta_{m,j}$ to show

$$[\varphi_{0,p} \circ \chi_p]_Y = [\chi_0 \circ \sigma^{-1}_{\zeta_3}]_Y,$$

$$[\varphi_{m,p+1} \circ \xi_{p+1} \circ \sigma_{\delta_{p+1}} \circ \chi_{p+1}]_Y = [\chi_m \circ \sigma_{\zeta_4}]_Y,$$

where ζ_3,ζ_4 have the no-cell-cutting property. Then since $\varphi_{p,0} = \varphi_{0,p}^{-1}$ we have

$$\begin{aligned}[\varphi_m]_Y &= [\chi_m \circ \sigma_{\zeta_4} \circ \sigma_{\zeta_2} \circ \chi'^{-1} \circ \chi' \circ \sigma_{\zeta_1} \circ \sigma_{\zeta_3} \circ \chi_0^{-1} \circ \chi_0 \circ \sigma_{\beta'}]_Y \\ &= [\chi_m \circ \sigma_{\beta_m} \circ \sigma_{\beta'}]_Y,\end{aligned}$$

where $\beta_m = \zeta_3 * \zeta_1 * \zeta_2 * \zeta_4$ has the no-cell-cutting property as required. □

28.7. Proof of 28.4. — Let $\ell = [x_0, x]$ be the geodesic segment in $\mathcal{T}(Y)$ between x_0 and $x \in \widetilde{V}_1$, and fix n. We then have a natural ordering on ℓ, with $x_0 \leqslant t \leqslant x$ for all $t \in \ell$. Let

$$t_n = \sup\{t \in \ell : d_P(0, \rho_2(\xi(t)) \leqslant n\}.$$

Then we need to show that, for suitable a_n with $\lim_{n\to\infty} a_n = 0$, for any t, $t' > t_n$,

$$|\rho_2(\Phi(t)) - \rho_2(\Phi(t'))| \leqslant a_n.$$

In order to show this, it suffices to show that the set

$$\{\rho_2(\Phi(t)) : t > t_n\}$$

lies in a union of *disjoint* balls centred at points of S^1, each of Euclidean radius $\leqslant a_n$. Since the above set is connected, it must then lie in a single ball (if the set is nonempty) which contains the point $\rho_2(x)$, since $\Phi(x) = x$.

We first claim that, for suitable $a_n \geqslant e^{-\sqrt{n}}$, the set

$$\{\rho_2(t) : t > t_n\}$$

lies in the union of $\{z : d_P(0, z) < n/2\}$ and a union of disjoint balls centred on points of S^1, each of Euclidean radius $\leqslant a_n/2$, and such that the Euclidean distance between any two of these balls is $\geqslant e^{-\sqrt{n}}$. To see this, let

$$t'_n = \sup\{t \in \ell : \|\ell_t\|_Y - \|\ell_t\|_Z \leqslant \sqrt{\log n}\}.$$

Here, ℓ_t denotes the geodesic in $\mathcal{T}(Y)$ joining x_0 and t. Then, by the First Basic Lemma of 25.7, if $t > t'_n$, $d_P(0, \rho_2(t)) \geqslant \sqrt{\log n}/C_1$, and hence, by 28.1, if $t'_n < t \leqslant x$,

$$|\rho_2(t) - \rho_2(x)| \leqslant \frac{1}{\log n}.$$

However, if $t \leqslant t'_n$, then by 14.12, $t = [\psi_t \circ \sigma_{\beta_t}]_Y$, where β_t is a path based at v_2, ψ_t minimizes distortion up to Z-preserving isotopy and

$$|\beta_t|_{+,t} \leqslant Ce^{\sqrt{\log n}} < n^{1/4}.$$

Here, $|\cdot|_{+,t}$ denotes length with respect to the expanding foliation of ψ_t. We can take the paths β_t to be geodesic. Then the lifted geodesic with initial endpoint at 0 is the geodesic joining 0 and $\rho_2(t)$. By abuse of notation, we call this β_t also. It follows that on the circle $\{z : d_P(0, z) = 2\sqrt{n}\}$, there are intervals of Euclidean length $O(e^{-2\sqrt{n}})$, such that the successive Euclidean distance between any two is $e^{-2\sqrt{n}+O(n^{1/4})}$, which are not crossed by any β_t. It follows that the claim is proved with $a_n \geqslant 4/\log n$ (for example).

Now we need to consider $\rho_2(\Phi(t)) = \lim_{m\to\infty} \rho_2(y_m(t))$ instead of $\rho_2(t)$. Fix a large m depending on t so that $|\rho_2(\xi(t)) - \rho_2(y_m(t))| \leqslant e^{-n}$. Then we only need to consider $y_m(t)$ for $t \in \ell_m$ since $y_m(\ell) = y_m(\ell_m)$ and y_m maps components of $\ell \setminus \ell_m$ to points. By 28.6, we have a geodesic $\beta_m = \beta_{m,t}$, which we can consider either as a geodesic in $\overline{\mathbf{C}} \setminus Z$, or, by taking lifts, as geodesics in the unit disc, where β_m has initial endpoint at $\rho_2(t)$, and final endpoint at $\rho_2(y_m(t))$. We also have the no-cell-cutting property (28.6) for β_m. This means that, similarly to 28.1, for a suitable $\delta > 0$ determined by L_0 and N, regarding β_m as a geodesic in the unit disc, on any circle centred on 0 of Euclidean radius $1 - e^{-n}$, there are intervals of (Euclidean) length $\geqslant \delta e^{-n}$ intersecting every interval on this circle of length $\leqslant \delta^{-1} e^{-n}$, such that β_m does not cross any of these intervals.

Let $\mathcal{B}_0$ be the set of balls centred on points of S^1, each of Euclidean radius $\leqslant a_n/2$, such that

$$\rho_2((t_n, x]) \subset \cup\mathcal{B}_0 \cup \{z : d_P(0, z) < n/2\}.$$

Let $\mathcal{B}_1'$ be the balls of $\mathcal{B}_0$ with radii enlarged by $e^{-2\sqrt{n}}$. If the geodesic $\beta_{m,t}$ joining $\rho_2(t)$ and $\rho_2(\xi(t))$ does not intersect

$$\{z : d_P(z, 0 \leqslant 3)\sqrt{n}\},$$

then

$$|\rho_2(t) - \rho_2(\xi(t))| < e^{-2\sqrt{n}},$$

and $\rho_2(\Phi(t))$ is in a ball of $\mathcal{B}_1'$. If $\beta_{m,t}$ does intersect $\{z : d_P(0, z) \leqslant 3\sqrt{n}\}$, then $\beta_{m,t}$ crosses

$$\{z : d_P(0, z) = 4\sqrt{n}\},$$

and, by the no-cell-cutting property, avoids intervals on this circle of Euclidean diameter $\geqslant Ce^{-4\sqrt{n}}$ for C bounded from 0 and such that the distance between any two intervals is $O(e^{-4\sqrt{n}})$. It follows that $\rho_2((t_n, x])$ lies in a union of balls of diameter $< a_n/2 + e^{-\sqrt{n}} < a_n$ which are Euclidean distance $\geqslant e^{-5\sqrt{n}}$ apart, as required. □

CHAPTER 29

CHUNKS

29.1. In Chapter 28 (in particular in 28.3) we outlined a programme for moving a geodesic segment with endpoints in $\widetilde{V}$ to a path in $\widetilde{V}$, in a controlled way. The chief difficulty with this programme is dealing with geodesic segments which pass through $\mathcal{T}_{<\varepsilon}$ for a small $\varepsilon > 0$ but which do not lie in sets $\mathcal{T}(\Gamma, \varepsilon)$ for invariant (f_0, Γ) (because the geodesic segments are very long). This chapter is devoted to the study of loop sets Γ and the action on these by f_0, where it is not in general true that $\Gamma \subset f_0^{-1}\Gamma$, and to the study of long geodesic segments which pass through $\mathcal{T}_{<\varepsilon}$. In particular, we develop the concept of of *chunks*.

29.2. Definition of Chunks. — Let α be a homotopically essential nonperipheral subsurface of $\overline{\mathbf{C}} \smallsetminus Y$ and let $|\varphi(\partial\alpha)|$ be bounded for some $[\varphi] \in \mathcal{T}(Y)$. We shall say, somewhat loosely that α is *a gap* (or a loop if α is an annulus) at $[\varphi]$. One problem that we have to deal with is that although $f_0^{-1}(\alpha)$ is a gap at $\tau(y)$ whenever α is a gap at y, given y, $y' \in \mathcal{T}(Y)$, there is no obvious relation between gaps on $[\tau(y), \tau(y')]$ and gaps on $[y, y']$, because $[\tau(y), \tau(y')]$ may be a large distance from $\tau([y, y'])$. Because if this, we use *chunks* (of gaps and loops) along segments $[y, y']$ which are geodesic both in $\overline{\mathbf{C}} \smallsetminus Y$ and $\overline{\mathbf{C}} \smallsetminus Z$.

First we need a preliminary definition. For any $z \in [y, y']$, $z' \in [z, y']$ and any long thick and dominant α at z with $\alpha \cap f_0^{-1}(\alpha) \neq \varnothing$, we define $C(\alpha, z, z')$ to be the convex hull of thick and dominant gaps intersecting α along $[z, z']$. The convex hull is defined to be the union of the gaps, together with any homotopically trivial or peripheral complementary components. Thus $\alpha = C(\alpha, z, z)$ and $C(\alpha, z, z')$ increases as z' approaches y'. The finite type surface $C(\alpha, z, z')$ can obviously only increase finitely often. Also if $\alpha \cap \beta \neq \varnothing$ and α, β are at z, z' respectively then $C(\beta, z', z'') \subset C(\alpha, z, z'')$ for all $z'' \in [z', y']$.

In what follows, when we talk about a long thick and dominant gap or loop, we mean with respect to fixed parameter functions Δ, r, s for gaps and a constant C_1 for loops (see 15.3). Also, thick will mean ν-thick for some $\nu \geqslant \nu(\Delta, r.s, C_1)$, where

$\nu(\Delta, r, s, C_1)$ is the constant produced in 15.4, such that sufficiently long geodesic segments always contain such long, ν-thick and dominant gaps and loops. The constant $M = M(\Delta, r, s, C_1)$ is given by 15.14.

A *chunk* $[y_1, y_2] \times \alpha \subset [y, y'] \times \overline{\mathbf{C}}$ has the following property 1.

1. The subsurface α is a connected union of gaps and loops at each point of $[y_1, y_2]$. For any long thick and dominant gaps and loops (15.3) α_1, $\alpha_2 \subset \alpha$ at points y_1', y_2' in $[y_1, y_2]$ with $d'_{\alpha_i}(y_i, y_i') \leqslant M$, $i = 1, 2$, we have $\alpha = C(\alpha_1, y_1', y_2) = C(\alpha_2, y_2', y_1)$.

Note that a chunk might contain no long thick and dominant gaps, or it will be the convex hull of those it contains. A chunk which does contain long thick and dominant gaps is called *long thick and dominant.*

We say that $[y_1, y_2] \times \alpha$ is a *partial chunk* if we only have $C(\alpha_1, y_1', y_2) \subset \alpha$ and $C(\alpha_2, y_2', y_1) \subset \alpha$. It follows from 15.14 that, for a constant M depending on the parameter functions used to define long thick and dominant, if α is a partial chunk,

2. $$[y_1, y_2] \subset \mathcal{T}(\partial\alpha, M).$$

A *(partial) chunk system for* $y, y']$ is a collection Σ of disjoint (partial) chunks $[y_1, y_1'] \times \alpha$ such that:

3. $$[y, y'] \times \overline{\mathbf{C}} = \bigcup\{[y_1, y_1'] \times \overline{\alpha} : [y_1, y_1'] \times \alpha \in \Sigma\}.$$

29.3. Lemma. — *For some r bounded in terms of $\#(Y)$, a partial chunk can be partitioned into $\leqslant r$ chunks. Hence, given any y, $y' \in \mathcal{T}(Y)$, there is a chunk system for $[y, y']$ with $\leqslant r$ elements. Given a partial chunk system Σ_0 with p elements, there is a refinement of Σ_0 which is a chunk system with $\leqslant r'$ elements where r' is bounded in terms of p and $\#(Y)$.*

Proof. — Let $[y_0, y_0'] \times \alpha_0$ be a partial chunk which is not a chunk. Choose some $\alpha' \subset \alpha_0$ which is a long thick and dominant gap at y_0'' and as near as possible to y_0 (replacing y_0 by y_0' if necessary) with $C(\alpha', y_0, y_0') = C(\alpha', y_0'', y_0') \neq \alpha$. Then subdivide $[y_0, y_0']$ into segments $[z_i, z_i']$ along which $C(\alpha', y_0, z)$ is constant. Then consider the set Σ_1 of partial chunks $[z_i, z_i'] \times C(\alpha', y_0, z_i')$ and $[z_i, z_i'] \times (\alpha \smallsetminus C(\alpha', y_0, z_i'))$. Inductively, suppose not all elements of Σ_j are chunks. Then those which are not can be properly subdivided to obtain a new system Σ_{j+1} of partial chunks such that if $[z, z'] \times \beta \in \Sigma_j$, $[w, w'] \times \zeta \in \Sigma_{j+1}$ with $[w, w'] \times \zeta \subset [z, z'] \times \beta$ and $[w, w'] \times \zeta \neq [z, z'] \times \beta$ then $\zeta \neq \beta$. If Σ_j is not a set of chunks for $0 \leqslant j \leqslant s$ then we have a strictly decreasing sequence $\alpha_j \smallsetminus Y$ of nontrivial nonperipheral subsurfaces of $\overline{\mathbf{C}} \smallsetminus Y$. So for some s bounded in terms of $\#(Y)$, Σ_s must be a set of chunks. Since the number of elements of Σ_{j+1} is bounded in terms of j and $\#(Y)$, the number of elements of Σ_s is bounded in terms of $\#(Y)$.

The last two statements of the lemma are obtained by applying this to the partial chunk $[y, y'] \times \overline{\mathbf{C}}$, and to the elements of a partial chunk system Σ_0. $\square$

29.4. Lemma. — *Given a sequence* $\{(\Delta_n, r_n, s_n, C_{1,n})\}$ *of long thick and dominant parameter functions and constants, and given an integer* p, *and given a partial chunk system, we can find a chunk system* Σ *for* p *successive* $(\Delta_n, r_n, s_n, C_{1,n})$, $N \leqslant n \leqslant N + p - 1$ *with* N *bounded in terms of* p *and* $\#(Y)$, *and* $\#(\Sigma)$ *bounded in terms of* $\#(Y)$.

Proof. — We use exactly the same method of the previous lemma. Assume without loss of generality that $(\Delta_{n+1}, r_{n+1}, s_{n+1}, C_{1,n+1})$ is a better set of parameter functions and constants than $(\Delta_n, r_n, s_n, C_{1,n})$. If Σ_n is a chunk system for $(\Delta_n, r_n, s_n, C_{1,n})$, then it is a partial chunk system for $(\Delta_{n+1}, r_{n+1}, s_{n+1}, C_{1,n+1})$ and any partial chunk $[y_1, y_2] \times \alpha$ which is not a chunk for the better set can be properly subdivided — in particular α is properly subdivided — so that it is. But any sequence of proper subdivisions is bounded in terms of $\#(Y)$. So there is a bound on the number of n such that some chunk needs subdividing. □

29.5. r**-fold chunks.** — We shall say α is an *r-fold chunk along* $[y_1, y_1']$ if we can find successive points z_i, $0 \leqslant i \leqslant r$ with $z_0 = y_1$, $z_r = y_2$, such that α is a chunk along each $[z_i, z_{i+1}]$. An *r-fold chunk system* is a chunk system in which every chunk is r-fold.

Lemma. — *Any partial chunk can be written as a disjoint union of* $\leqslant s$ *r-fold chunks, where* s *is bounded in terms of* r *and* $\#(Y)$. *Hence, any chunk system* Σ_0 *can be refined to an r-fold chunk system* Σ_r, *where the number of elements of* Σ_r *is bounded in terms of* $\#(Y)$, r, *and the number of elements of* Σ_0.

Proof. — We only need to prove the first statement. The key to this is the fact that a chunk $[y_1, y_2] \times \alpha$ which is a single long thick and dominant gap is automatically an r-fold chunk: either α is long thick and dominant on each of r disjoint subintervals of $[y_1, y_2]$ or $d_\alpha(y_1, y_2) \leqslant M$. Let $[y_0, y_0'] \times \alpha_0$ be a partial chunk. Suppose inductively that any partial chunk $[z, z'] \times \beta$ with $\beta \subset \alpha$, $\beta \neq \alpha$ can be written as a disjoint union of $r+1$-fold chunks, and that $[y_0, y_0'] \times \alpha$ can be written as a disjoint union of r-fold chunks. Then we can assume without loss of generality that $[y_0, y_0'] \times \alpha$ is an r-fold chunk, with subdivision of $[y_0, y_0']$ into $[z_i, z_{i+1}]$, $0 \leqslant i < r$. If $[z_i, z_{i+1}] \times \alpha_0$ is a 2-fold chunk for at least one i, then $[y_0, y_0'] \times \alpha_0$ is an $r+1$-fold chunk. So suppose $[z_i, z_{i+1}] \times \alpha_0$ is not a 2-fold chunk for any i. To simplify the notation, take $i = 0$. Fix some arbitrary point $z_0' \in (z_0, z_1]$. Assume without loss of generality that $[z_0, z_0'] \times \alpha_0$ is not a chunk. (Otherwise, interchange z_0 and z_1.) If $[z_0', z_1] \times \alpha_0$ is also not a chunk, then each of $[z_0, z_0'] \times \alpha_0$, $[z_0', z_1] \times \alpha_0$ can be divided into chunks $[z, z] \times \beta$, always with $\beta \neq \alpha$, and these can be further divided into $r+1$-fold chunks by the inductive hypothesis. If $[z_0', z_1] \times \alpha_0$ is a chunk, then take the first point z_0'' to the right of z_0' such that one of $[z_0, z_0''] \times \alpha_0$, $[z_0'', z_1] \times \alpha_0$ is not chunk (since we are assuming that $[z_0, z_1] \times \alpha_0$ is not a 2-fold chunk). Then the difference is made by the long thick and dominant gaps and loops at z_0''. Each of these is an $r+1$-fold chunk along subinterval of $[z_0, z_1]$, giving a decomposition of $[z_0, z_1] \times \alpha$ into sets $[z, z'] \times \beta$ such that either β is an $r+1$-fold chunk or $\beta \neq \alpha$ is a partial chunk. So, again, we can divide $[z_0, z_1] \times \alpha$ into $r+1$-fold chunks. □

As in 29.4, this result extends to work for a finite sequence of parameter functions and constants.

29.6. Chunk systems on adjacent segments

Lemma. *Let Σ_1, Σ_2 be chunk systems for $[y_0, y_1]$, $[y_1, y_2]$. Then there is a chunk system Σ for $[y_0, y_2]$ such that every chunk is contained in a chunk of $\Sigma_1 \cup \Sigma_2$. The number of elements of Σ is bounded in terms of $\#(Y)$ and the number of elements of $\Sigma_1 \cup \Sigma_2$.*

Proof. — By 15.8, if α is long ν-thick and dominant at $y \in [y_0, y_1]$ then $d_\alpha(y, y') \leqslant C_1(\nu)$ for some $y' \in [y_0, y_2] \cup [y_1, y_2]$. If $y' \in [y_0, y_2]$ then for all $z \in [y_0, y]$ and long thick and dominant β at z with $\beta \cap \alpha \neq \varnothing$ — that is, $\beta \subset C(\alpha, y, z)$ — the corresponding point $z' \in [y_0, y_2]$ also. Then if $[z, z'] \times \beta$ is a chunk for Σ_1 and $y' \in [y_0, y_2]$, for all $y \in [z, z']$ and long thick and dominant $\alpha \subset \beta$ at y, we take $[z, z'] \times \beta$ to be a chunk in Σ. If $y' \in [y_1, y_2]$ for all such y, α, we do not take $[z, z'] \times \beta$ to be a chunk of Σ. If $y' \in [y_0, y_2]$ for some y, α and $y' \in [y_1, y_2]$ for others, then we want to subdivide $[z, z'] \times \beta$ into chunks, some of which will be in Σ and some will not. This is the rough idea. We shall carry out a similar procedure for Σ_2. Since every long η-thick and dominant ζ at a point $w \in [y_0, y_2]$ satisfies $d_\zeta(w, w') \leqslant C_1(\eta)$ for some $w' \in [y_0, y_1] \cup [y_1, y_2]$, we shall have the required chunk system Σ, after possibly adding some chunks disjoint from all long thick and dominants.

So fix $[z_0, z_1] \times \beta \in \Sigma_1$ such that for some (y, α) we have $y' \in [y_0, y_2]$ and for some other (y, α) we have $y' \in [y_1, y_2]$. We say that (α, y) is a *cutting point* if $y' \in [y_0, y_2]$, and for any other (ζ, w) with $\beta \cap \alpha \neq \varnothing$, $\beta \setminus \alpha \neq \varnothing$, $w \in [y, y_1]$, we have the corresponding point $w' \in [y_1, y_2]$. We can choose a set P of cutting points (y, α) bounded in terms of $\#(Y)$ such that any other cutting point is (w, α) for some $(y, \alpha) \in P$. Take $y' \in [y_0, y_2]$ with $d_\alpha(y, y') \leqslant C_1(\nu)$ (where α is ν-thick) Consider the convex hulls $C(\alpha, y', z')$ for $d_\zeta(z, z') \leqslant C_1(\eta)$ for $z \in [z_0, z_1]$, $\zeta \subset \beta$ η-thick at z and $\zeta \cap \alpha \neq \varnothing$. There are boundedly many intervals $[w_i, w_{i+1}]$ such that $C(\alpha, y', z')$ is constant for $z' \in [w_i, w_{i+1}]$. Then we consider the partial chunks $C(\alpha, y', w_i) \times [w_i, w_{i+1}]$ and take the union of these over the cutting points (α, y). Then by taking intersections and differences, we can obtain a set of disjoint partial chunks whose union is the union of all the $C(\alpha, y, w_i) \times [w_i, w_{i+1}]$, the size of this set being bounded in terms of $\#(Y)$. Then we can refine this to a union of chunks. We can assume that each point y' arises from a cutting point from $[y_0, y_1]$ and another from $[y_1, y_2]$. Since this can be done for all chunks of $\Sigma_1 \cup \Sigma_2$, we obtain the result. □

29.7. Pullback of chunks. — If $[y_0, y_0'] \times \alpha$ is a chunk, and $S\alpha$ is a component of $f_0^{-1}\alpha$, then $[\tau(y_0), \tau(y_0')] \times S\alpha$ is a partial chunk. We now deal with pullbacks of chunk systems.

Lemma. — *Let Σ be a partial chunk system for $[y, y']$. Then there is a chunk system Σ' for $[\tau(y), \tau(y')]$ such that every chunk $\beta \times [z, z']$ of Σ' satisfies $\beta \subset S\alpha$ for some chunk $\alpha \times [y_0, y_0'] \in \Sigma$ and some component $S\alpha$ of $f_0^{-1}\alpha$, with $\#(\Sigma)$ bounded in terms of $\#(Y)$ and $\#(\Sigma)$.*

Proof. — We can assume without loss of generality that there is a decomposition $[y, y'] = \cup_{i=0}^{r-1}[y_i, y_{i+1}]$ such that the partial chunks of Σ are all of the form $[y_i, y_{i+1}] \times \beta$ for some i, since we can always refine a partial chunk system (although not a chunk system) to be of this form. Then we have partial chunks $[\tau(y_i), \tau(y_{i+1})] \times S\beta$. Then applying the previous lemma repeatedly, we obtain a chunk system Σ_n for $[\tau(y_0), \tau(y_n)]$ from the chunk system Σ_{n-1} for $[\tau(y_0), \tau(y_{n-1})]$ and the partial chunk system consisting of the chunks $[\tau(y_{n-1}), \tau(y_n)] \times \beta$. Finally we put $\Sigma' = \Sigma_{r-1}$. □

It makes sense to describe the chunks of Σ' as *preimages* of the chunks of Σ, even though Σ' is not canonically defined from Σ and although, for $[y_1', y_2'] \times \alpha' \in \Sigma'$, α' is only a subset of a component of $f_0^{-1}(\alpha)$ for some $[y_1, y_2] \times \alpha \in \Sigma$. Similarly if z is a bounded distance from $[y, \tau(y)]$we can use a chunk system Σ for $[y, \tau(y)]$ to define a chunk system Σ' for $[z, \tau(z)]$ with chunks of Σ' in the *backward orbit* of chunks of Σ. In future we shall often simply write α for a chunk $[y_1, y_2] \times \alpha$, and shall write $\alpha' \subset f_0^{-1}\alpha$.

CHAPTER 30

OUTLINE CONSTRUCTION OF A GOOD SEQUENCE

30.1. We need to complete the task outlined in Chapter 28. This means that, given a geodesic ℓ in $\mathcal{T}(Y)$ with endpoints in a component $\widetilde{V}_1$ of $\widetilde{V}$, we need to construct the path $\Phi(\ell)$ of 28.3 in $\widetilde{V}_1$ with the same endpoints as ℓ. This means that, given $x \in \ell$, we need to construct the sequence $\{y_m(x)\}$ with the properties given in 28.3. The rough idea is that this will make the union of geodesic segments $[y_m(x), y_{m+1}(x)]$ in $\mathcal{T}(Z)$ be as near as possible to a geodesic in $\mathcal{T}(Z)$. The hard properties of 28.3 to obtain are c) and d). In order to construct y_{m+1} from y_m, we find in general that we need a supplementary sequence w_j.

30.2. A sequence $\{w_j\}$. — The pullback map $\tau : \mathcal{T}(Y) \to \mathcal{T}(Y)$ has the property that if α is a gap at y, then each nontrivial nonperipheral component of $f_0^{-1}(\alpha)$ is homotopic in $\overline{\mathbf{C}} \setminus Z$ to a gap α_1 at $\tau(y)$. By the definition of τ (that is, $d_Z(y, \tau(y)) = d_Y(y, \tau(y))$ up to homotopy all intersections between α and α_1 are essential in $\overline{\mathbf{C}} \setminus Z$, and if $\alpha \cap \alpha_1 = \varnothing$ up to Z-preserving isotopy then $F_0^{-1}\alpha_1 \cap f_0^{-1}\alpha = \varnothing$ also. The same argument does not work if we replace α_1 by α_i, where α_j is defined inductively by: α_{j+1} is a gap at $\tau^{j+1}(y)$ which is homotopic in $\overline{\mathbf{C}} \setminus Z$ to a component of $f_0^{-1}(\alpha_j)$. That is, if $\alpha \cap \alpha_i = \varnothing$ up to Z-preserving isotopy then this may not be true up to Y-preserving isotopy, and hence we cannot necessarily deduce that $f_0^{-1}(\alpha) \cap f_0^{-1}(\alpha_i) = \varnothing$. For this reason, among others, instead of considering sequences $\{\tau^i(y)\}_{i \geqslant 0}$, we consider a sequence $\{w_j\}_{j \geqslant 0} \subset \mathcal{T}(Y)$, $w_j = w_j(y)$, with $w_0 = y$ and chunk systems Σ_j, Σ'_j for $[w_j, \tau(w_j)]$, $[w_j, w_{j+1}]$ and with the following properties. We abbreviate "long thick and dominant" to ltd, where possible. The Basic property depends on a suitable constant Δ_1 which is chosen just large enough. The properties depend on a fixed choice of long thick and dominant parameter functions.

Basic properties for $\{w_j\}$. Every ltd chunk of Σ_q is, up to bounded distance, either in $f_0^{-1}(\Sigma_{q-1})$ or in $f_0^{-2}(\Sigma_{q-1})$ (but the latter happens only on a connected union of chunks of length $\leqslant \Delta_1$). For a bounded integer r_0, and all j, every ltd chunk of Σ'_j is a subset of a chunk in $\cup_{j-r_0 \leqslant i \leqslant j}\Sigma_i$.

Here, the term "up to bounded distance" for ltd chunks $I \times \zeta$ means that for all long ν-thick and dominant gaps or loops $\beta \subset \zeta$ and $z \in I$ there is $z' \in I'$ for some chunk $I' \times \zeta'$ on $f_0^{-r}(\Sigma_{q-r})$, with $d_\beta(z, z') \leqslant C_2(\nu)$, for a suitable function $C_2 : (0, \infty) \to (0, \infty)$ which depends only on the function C of 15.8. For chunks which are not ltd we simply replace $C_2(\nu)$ by M_2 for a suitable constant $M_2 > 0$.

The most important property of the sequence w_j is the following.

Straightening property. — a) The following holds for an integer r_0. Take any $r \geqslant r_0$ and any $q > r$. If β is long ν-thick and dominant or m_0-Pole-Zero along $I \subset [w_{q-r}, w_q]_Z \cup [w_q, \tau(w_q]$ and $w \in I$ then there is $w' \in [w_{q-r}, \tau(w_q)]$ such that

$$d_\beta(w, w') \leqslant C(\nu) \quad \text{or} \quad |\operatorname{Re}(\pi_\beta(w) - \pi_\beta(w'))| \leqslant K_0$$

depending on whether β is a gap or a loop.

Another way of saying this is that every chunk of Σ_q is not a bounded distance from any chunk along $[w_{q-r}, w_q]$. The following part of the Straightening Property will not be referred to so frequently, but it will still be used.

b) Given $\varepsilon > 0$, the following holds for suitable choice of the long thick and dominant parameter functions and Pole-Zero constant. If β is a loop which is in the convex hull of long thick and dominants along both $[w_0, w_q]$ and $[w_q, \tau(w_q)]$, and $w_{q+1} = [\varphi]$, then $|\varphi(\beta)| < \varepsilon$.

Now we list the other properties of the sequence w_j.

The following is for suitable functions C_2, $C_3 : (0, \infty) \to (0, \infty)$, depending on q.

v_2-tracing property. If β is long, ν-thick and dominant along segments of $[w_i, w_{i+1}]$ and $[w_j, w_{j+1}]$ containing w and w' respectively which are both within a bounded distance of segments of $[w_0, \tau(w_q)]$ $(i, j \leqslant q)$ and $d_{\beta,Z}(w, w') \leqslant C_2(\nu)$ then $d_{\beta,Y}(w, w') \leqslant C_3(\nu)$. and similarly if β is a loop satisfying the Pole-Zero Condition along the segments.

Note that if this holds for long thick and dominant gaps ad Pole-Zero loops β, then we have a similar result for *all* gaps and loops β, with $C_2(\nu)$ and $C_3(\nu)$ replaced by suitable constants M_2, M_3.

The mostly Z property. — The following holds for a constant M_2 and an integer r_0. For $w_n = [\psi_0]$, and $r \leqslant r_0$,

$$w_{n+1} = [\xi_r \circ \chi_r \circ \cdots \circ \xi_1 \circ \chi_1 \circ \xi_0 \circ \psi],$$

where, if $[\psi_i] = [\xi_i] \circ \chi_i \circ \xi_0 \circ \psi_0]$, χ_i minimises distortion up to isotopy constant on $\psi_{i-1}(Z)$ and $d_Y([\xi_i], [\text{identity}]) \leqslant M_2$ for all i. Moreover if

$$w_{n+1}^i = [\xi_i \circ \cdots \circ \chi_1 \circ \xi_0 \circ \psi]_Y,$$

then for every long ν-thick and dominant gap or m_0-Pole-Zero loop α along a segment of $[w_{n+1}^{i-1}, w_{n+1}^i]_Z$ containing some w, there is $w' \in [w_n, w_{n+1}]$ such that

$$d_{\alpha,Z}(w, w') \leqslant C(\nu) \quad \text{or} \quad |\operatorname{Re}(\pi_{\alpha,Z}(w) - \pi_{\alpha,Z}(w'))| \leqslant K_0$$

for $C(\nu)$ and K_0 as in 15.8, depending on whether α is a gap or a loop.

The mostly Z property will imply the similar property 28.3c) for the y_m.

F-decreasing property

$$F(w_i) \leqslant F(w_{i-1}).$$

Furthermore given ε_1, there is $\eta_1 > 0$, bounded below in terms of $F(y_m)$, ε_1 such that either (i) or (ii) holds.

(i) $$d_Y(w_i, \tau(w_{i-1})) \leqslant \varepsilon_1.$$

(ii) $$F(w_j) \leqslant F(w_{j-1}) - \eta_1.$$

The next (and last) chapter will be devoted to proving the following.

30.3. Theorem. — *Given y, it is possible to construct a sequence $\{w_j\}$ with $w_0 = y$ and the properties above.*

30.4. Consequences of the properties. — By 29.4, it is possible to construct boundedly finite chunk systems with respect to sets of long thick and dominant parameter functions, one given in terms of the other. If a chunk $[y_1, y_2] \times \alpha$ is not long thick and dominant, then, for M depending on the long thick and dominant functions, $d_\alpha(y_1, y_2) \leqslant M$. So then for either component α_1 of $f_0^{-1}(\alpha)$,

$$d_{\alpha_1}(\tau(y_1), \tau(y_2)) \leqslant M + C_2$$

for a suitable constant C_2 (see 26.9). If α is a loop we also have

$$|\operatorname{Re}(\pi_{\alpha_1}(\tau(y_1)) - \pi_{\alpha_1}(\tau(y_2)))| \leqslant |\operatorname{Re}(\pi_\alpha(y_1) - \pi_\alpha(y_2)| + C_2$$

It follows that if $I_1 \times \alpha_1 \subset f_0^{-1}(I \times \alpha)$ is long thick and dominant with respect to $(\Delta_1, r_1, s_1, m_1)$ and also with respect to $(\Delta_2, r_2, s_2, m_2)$ defined suitably in terms of $(\Delta_1, r_1, s_1, m_1)$, then α must also be long thick and dominant with respect to $(\Delta_2, r_2, s_2, m_2)$.

Write $\Sigma_{i,q} = \Sigma_{i,q,0}$ for the set of ltd chunks along $[w_0, \tau(w_q)]$ which are derived from Σ_i. We could also write $\Sigma_{i,q,k}$ for the set of ltd chunks along $[w_k, \tau(w_q)]$ which are derived from Σ_i, if $k \leqslant i$. But in fact, by the Straightening Property, $\Sigma_{i,q,k} = \Sigma_{i,q}$ if $i - k > 0$ is sufficiently large.

Lemma. — *Let $\zeta_r \times I_r \in \Sigma_q$ be part of a sequence of ltd chunks $I_i \times \zeta_i \in \Sigma_{q-r+i}$ $(0 \leqslant i \leqslant r)$ with $\zeta_{i+1} \subset f_0^{-1}(\zeta_i)$, with I_r not within distance Δ_1 of $\tau(w_q)$.*

(1) *Then $I_i \times \zeta_i \in \Sigma_{q-r+i,j}$ for all $q - r + i \leqslant j \leqslant q$.*

(2) *All ltd chunks of Σ_{q-r+i} nearer w_{q-r+i} than $I_i \times \zeta_i$ are also in $\Sigma_{q-r+i,j}$ for all $q - r + i \leqslant j \leqslant q$, and all chunks of $\Sigma_{k,j-r+i}$ which are nearer w_0 than $I_i \times \zeta_i$ are in $\Sigma_{q-r+i,j}$ for all $q - r + i \leqslant j \leqslant q$ and $k \leqslant j - r + i$.*

All these inclusions are up to bounded distance.

Proof

(1) We prove this by induction on $r-i$. It is trivially true for $r-i=0$. Renaming $r-i+1$ as r if necessary, assume that $I_i\times\zeta_i\in\Sigma_{q-r+i,j}$ for all $q-r+i\leqslant j\leqslant q$ and $i\geqslant 1$. Then the inductive hypothesis also gives $I_0\times\zeta_0\in\Sigma_{q-r,j}$ for $q-r\leqslant j\leqslant q-1$. Now suppose that $I_0\times\zeta_0\in\Sigma'$ where Σ' is the set of chunks in the chunk system for $[w_0,\tau(w_{q-1})]$, which are a bounded distance from $f_0^{-1}(\Sigma_{q-1})$, and hence not in the chunk system for $[w_0,\tau(w_q)]$. Let Σ'' denote the complement of Σ' in $f_0^{-1}(\Sigma_{q-1})$, up to bounded distance. Now chunks from Σ_q must be a bounded distance from chunks in $f_0^{-1}(\Sigma_{q-1})$ or distance $\leqslant\Delta_1$ from $\tau^2(w_{q-1})$, by the Basic Property. So $I_1\times\zeta_1$ comes within a bounded distance of $f_0^{-1}(\Sigma')=f_0^{-1}(\Sigma_{q-1})\setminus\Sigma_q$ (this uses the v_2-tracing property), contradicting $I_1\times\zeta_1\in\Sigma_{q-r+1,q}$.

(2) The first statement is true for $j=q-r+i$ by the straightening property. Then use induction. The second statement is similar. □

30.5. Periodic Chunks. — The definition of a periodic chunk is a little involved, but, in particular, we say that a ltd chunk along $[w_{q-r},\tau(w_{q-r})]$ is *of period* r if it is of the form $I_0\times\zeta_0$ and there are ltd chunks $I_i\times\zeta_i$ along $[w_{q-r+i},\tau(w_{q-r+i})]$ $(0\leqslant i\leqslant r)$ with $I_{i+1}\times\zeta_{i+1}\subset f_0^{-1}(I_i\times\zeta_i)$ if $i<r$ and $\zeta_0\cap\zeta_r\neq\varnothing$, and r is the least integer >0 for which this holds for any choice of $I_i\times\zeta_i$, $i>0$. A chunk of period one is also called a *fixed chunk*.

But we need a more general definition that this. We want to say that a ltd chunk $I_0\times\zeta_0$ along $[w_0,\tau(w_q)]$ which is, up to bounded distance, a subset of a chunk of Σ_i for $i\leqslant q-r$, is of period r in certain circumstances, even if there may not exist a sequence of ltd chunks $I_i\times\zeta_i$ as above. For a chunk $I\times\zeta$ on $[w_0,\tau(w_q)]$ which is, up to bounded distance, a subset of a chunk of Σ_j, $T(I\times\zeta)$ will be a finite set of chunks which are, up to bounded distance subsets of chunks of $\cup_{k\geqslant 0}\Sigma_{j+1-k}$. If there is are ltd chunks on $[w_0,\tau(w_{q+1})]$ a bounded distance from each component of $f_0^{-1}(I\times\zeta)$ then this set of ltd chunks is taken to be $T(I\times\zeta)$. Otherwise we take the maximal ltd chunks which are a bounded distance from chunks in $f_0^{-1}(\Sigma(I\times\zeta,<))$ where $\Sigma(I\times\zeta,<)$ denotes the set of ltd chunks $<I\times\zeta$ in $[w_0,\tau(w_q)]$ where $J\times\omega<I\times\zeta$ if J separates I from w_0 and $\omega\times\zeta\neq\varnothing$. This relation $<$ is transitive, because if $J_2\times\omega_2$ is between $J_1\times\omega_1$ and $J_3\times\omega_3$ and ω_2 intersects both ω_1 and ω_3 and γ_i are loops in ω_i with $|\varphi_i(\gamma_i)$ bounded at some fixed points $[\varphi_i]\in J_i$ then $\varphi_2(\gamma_1\cap\omega_2)$ and $\varphi_2(\gamma_3\cap\omega_2)$ are approximately stable and unstable leaves of the quadratic differential foliation and hence intersect by 15.11.

Note that if $T(I_1\times\zeta_1)\cap T(I_2\times\zeta_2)\neq\varnothing$ then either $I_1\times\zeta_1\cap I_2\times\zeta_2\neq\varnothing$ or, for both $i=1$ and 2, $T(I_i\times\zeta_i)$ is a bounded distance from subsets of chunks of $f_0^{-1}(\Sigma(I_i\times\zeta_i,<)$.

Then we say that $I\times\zeta$ is *of period* p (for $[w_0,\tau(w_q)]$) if $I\times\zeta$ is derived from Σ_i, $i\leqslant q-p$, $T^p(I\times\zeta)\cap\zeta)\neq\varnothing$. We do not insist that p is the least integer >0 for which this holds.

30.6. Lemma. — *Let $J_0 \times \omega_0$ be a ltd period p chunk along $[w_0, \tau(w_q)]$ which is contained in a chunk of Σ_i, $i \leqslant q - 2p$, and let $T^{2p}(J_0 \times \omega_0) \cap \omega_0 \neq \varnothing$. Then either a bounded refinement of Σ_i gives no period p chunks within $J_0 \times \omega_0$ or $J_0 \times \omega_0$ contains a period p chunk for $[w_0, \tau(w_{q+1})]$ up to bounded distance.*

Remark. — The condition $T^{2p}(J_0 \times \omega_0) \cap \omega_0 \neq \varnothing$ is automatically satisfied if there is a sequence of ltd chunks $J_i \times \omega_i$ for $[w_0, \tau(w_q)$, $0 \leqslant i \leqslant p$ with $J_{i+1} \times \omega_{i+1}$ a bounded distance from a subset of $f_0^{-1}(J_i \times \omega_i)$, $\omega_p \cap \omega_0 \neq \varnothing$ and ω_p is itself periodic of period p — which would happen if this sequence $J_i \times \omega_i$ was defined for $i \leqslant 2p - 1$ and $\omega_p \cap f_0^{-1}(\omega_{2p-1}) \neq \varnothing$

Proof. — Suppose for contradiction that $J_0 \times \omega_0$ is contained in Σ_{q+1} up to bounded distance. Then either the refinement of $J_0 \times \omega_0$ given by intersecting with Σ_{q+1} splits it into chunks of which none have period p, or at least one has period p. We can assume that it is contained in ζ_{q+1} for a chunk $I_{q+1} \times \zeta_{q+1}$ of Σ_{q+1}. Now let $J_k \times \omega_k$ be the orbit of $J_0 \times \omega_0$, that is, $J_{k+1} \times \omega_{k+1} \subset f_0^{-1}(J_k \times \omega_k)$. Let $I_j \times \zeta_j \in \Sigma_j$ with $I_{j+1} \times \zeta_{j+1} \subset f_0^{-1}(I_j \times \zeta_j)$ up to bounded distance. We have $T^p(J_0 \times \omega_0) \cap \omega_0 \neq \varnothing$, and hence, $T^p(J_0 \times \omega_0) \cap \zeta_{q+1} \neq \varnothing$. Then by induction $T^{p-\ell}(J_0 \times \omega_0) \cap \zeta_{q+1-\ell} \neq \varnothing$ and $\omega_{p-\ell-1}$ and $\zeta_{q+1-\ell}$ are contained in the images of $T^{p-\ell+1}(J_0 \times \omega_0)$, $\zeta_{q+2-\ell}$ under consistent branches of f_0^{-1}. This uses the v_2-tracing property, since all chunks are along $[w_0, \tau(w_q)]$. We deduce that $\omega_0 \cap \zeta_{q+1-p} \neq \varnothing$, and hence $\zeta_{q+1} \cap \zeta_{q+1-p} \neq \varnothing$. So $I_{q+1-p} \times \zeta_{q+1-p}$ is period p also, and intersects $J_0 \times \omega_0$, and $q + 1 - p > q - p \geqslant i$. Similarly using $T^{2p}(J_0 \times \omega_0) \cap \omega_0 \neq \varnothing$, we deduce that $\zeta_{q+1-p} \cap T^p(J_0 \times \omega_0) \neq \varnothing$. So $I_{q+1-p} \times \zeta_{q+1-p}$it must be in $f_0^{-1}(\Sigma_q)$ up to bounded distance. So by the same argument as above, either there is a boundedly finite refinement of ζ_{q+1-p} into chunks none of which are period $\leqslant p$ — in which case we can go back and refine $J_0 \times \omega_0$ — or there is $I_{q+1}^2 \times \zeta_{q+1}^2 \in \Sigma_{q+1}$ which is a bounded distance from $I_{q+1-p} \times \zeta_{q+1-p}$, after reducing a bit if necessary, and then this chunk also gives rise to a period p chunk $I_{q+1-p}^2 \times \zeta_{q+1-p}^2$ which intersects $I_{q+1-p} \times \zeta_{q+1-p}$. It also intersects $J_0 \times \omega_0$ because $I_{q+1-p} \times \zeta_{q+1-p}$ intersects $T^p(J_0 \times \omega_0)$. So all ltd chunks between $J_0 \times \omega_0$ and $I_{q+1-p} \times \zeta_{q+1-p}$ are a bounded distance from ltd chunks between $I_{q+1} \times \zeta_{q+1}$ and $I_{q+1}^2 \times \zeta_{q+1}^2$. We have

$$d'_{\omega_0, \zeta_{q+1-p}}(J_0, I_{q+1-p}) \leqslant d'_{\zeta_{q+1}, \zeta_{q+1}^2}(I_{q+1}, I_{q+1}^2) + O(1)$$

$$d'_{\zeta_{q+1-p}, \zeta_{q+1-p}^2}(I_{q+1-p}, I_{q+1-p}^2) + O(1) < d'_{\omega_0, \zeta_{q+1-p}}(J_0, I_{q+1-p})$$

giving the required contradiction. □

30.7. Corollary. — *Let $J_0 \times \omega_0$ be ltd period p for $[w_0, \tau(w_q)]$ and in $\Sigma_{i,q}$, up to bounded distance. Let $J_j \times \omega_j$ be defined for $-p \leqslant j \leqslant 0$ and $\Sigma_{i+j,q}$ up to bounded distance, with $J_{j+1} \times \omega_{j+1}$ a subset of $f_0^{-1}(J_j \times \omega_j)$, up to bounded distance. Then $J_{-p} \times \omega_{-p}$ is period p on $[w_0, \tau(w_r)]$ for all $r \geqslant q$.*

Proof. — With respect to $[w_0, \tau(w_q)]$ we have $T^p(J_{-p} \times \omega_{-p}) \cap \omega_{-p} \neq \varnothing$, since

$$T^p(J_{-p} \times \omega_{-p}) = \omega_0 \quad \text{and} \quad T^{2p}(J_{-p} \times \omega_{-p}) = T^p(J_0 \times \omega_0).$$

Then $T^{2p}(J_{-p} \times \omega_{-p}) \cap \omega_{-p} \neq \varnothing$, since $J_0 \times \omega_0$ is ltd. By induction on t this remains true with respect to $[w_0, \tau(w_t)]$ for $t \geqslant q$. In fact if there is any change the sets, $T^p(J_0 \times \omega_0)$ has to change first and move closer to ω_0, and they may coalesce in the future. So for $q+1$ replaced by $t+1$ in the previous lemma we deduce from $J_{-p} \times \omega_{-p}$ being period p on $[w_0, \tau(w_t)]$ that it is also period p on $[w_0, \tau(w_{t+1})]$. □

30.8. Lemma. — *For r_0, r_1 bounded in terms of $vv\#(Y)$, the following holds. Every chunk $I_0 \times \zeta_0$ in $\Sigma_{i,q}$ either has some period $r \leqslant r_0$ or there is no sequence $\{I_j \times \zeta_j\}$ defined for $0 \leqslant j \leqslant r_0$ with $I_{j+1} \times \zeta_{j+1} \subset f_0^{-1}(I_j \times \zeta_j)$, up to bounded distance, and $I_j \times \zeta_j \in \Sigma_{i+j,q}$ for $0 \leqslant j$, $i+j \leqslant q$.*

Proof. — Suppose that such a sequence $I_j \times \zeta_j$ does exist. If all the ζ_j are nontrivial nonperipheral then because they all lie in a finite tye surface, for some $j \geqslant 0$ and $r > 0$ both bounded in terms of $\#(Y)$ we must have $\zeta_j \cap \zeta_{j+r} \neq \varnothing$ — or if they are both loops we might have $\zeta_j = \zeta_k$ up to Z-preserving isotopy. If ζ_j is a gap, or a loop with ζ_j and ζ_{j+r} intersecting transversally then, as we have seen, we must have $\zeta_0 \cap \zeta_r \neq \varnothing$. If ζ_j and ζ_{j+r} are nontrivial nonperipheral loops which are equal up to Z-preserving isotopy, then we still deduce that ζ_0 is periodic. For if ζ_0 is disjoint from all ζ_j for $j \leqslant r_0$, for some suitable r_0 depending only on $\#(Y)$, and no component of $\partial\zeta_0$ coincides up to Z-preserving isotopy with any component of $\partial\zeta_j$ for any $j \leqslant r_0$, then the backward orbit of $\partial\zeta_0$ gives a loop set Γ with $f_0^{-1}(\Gamma) \subset \Gamma$ and $f_0^{-n}(\Gamma)$ nontrivial nonperipheral for all $n \geqslant 0$. If this happens, $\Gamma = f_0^{-1}(\Gamma)$, and hence ζ_0 must be periodic after all. □

30.9. Some special chunks of chunks. — A *chunk of chunks* along a geodesic is a union A of chunks along this geodesic towithin bounded distance, with the following properties.

(1) If $I_j \times \zeta_j$ are ltd chunks with $I_1 \times \zeta_1$, $I_3 \times \zeta_3 \in A$ and I_2 is between I_1 and I_3 and $\zeta_1 \cap \zeta_2 \neq \varnothing$, $\zeta_3 \cap \zeta_2 \neq \varnothing$, then $I_2 \times \zeta_2 \in A$.

In such circumstances, we say that $I_1 \times \zeta_1 < I_2 \times \zeta_2$, and $I_2 \times \zeta_2 < I_3 \times \zeta_3$. We also define $C(A)$ to be the convex hull of all ζ with $I \times \zeta \in A$ and $I \times \zeta$ ltd (for varying ζ). We also define $C(< I \times \zeta, A)$ to be the convex hull of all ζ' with $I' \times \zeta' \in A$, $I' \times \zeta'$ ltd and $I' \times \zeta' < I \times \zeta$. Surfaces $C(> I \times \zeta, A)$, $C(\leqslant I \times \zeta, A)$, $C(\geqslant I \times \zeta, A)$ are similarly defined. Then the second condition is as follows

(2) If $I \times \zeta \in A$ is ltd and such that ζ is not contained in $C(< I \times \zeta, A)$ then $C(A) = C(\geqslant I \times \zeta, A)$, and if ζ is not contained in $C(> I \times \zeta, A)$ then $C(A) = C(\leqslant I \times \zeta, A)$.

A chunk of chunks A is *r-fold* if it can be written as a union of r chunks of chunks A_i, $1 \leqslant i \leqslant r$ with $C(A_i) = C(A)$ for all i.

Lemma. — *The following holds, given r_0, for some bounded integers r_1, t_1 depending only on $\#(Y)$, r_0 and M_1 depending on the long thick and dominant parameter functions. Fix ltd parameter functions and a Pole-zero constant. For $i \leqslant j \leqslant j+t_1 \leqslant q$, let $\Omega_{i,j,q}$ be the convex hull of all ltd chunks in $\cup_{i\leqslant k\leqslant j}\Sigma_{k,q}$.*

Then given i there is some $\Omega = \Omega_{i_1,j_1,q}$ with $0 \leqslant i_1 - i \leqslant r_1$, $r_1 \geqslant j_1 - i_1 \geqslant r_0$ such that $\Omega \subset f_0^{-1}(\Omega)$ modulo trivial and peripheral components. Either there is a connected component Ω_1 of Ω with $\Omega_1 \subset f_0^{-1}(\Omega_1$ and $f_0^{-s}(\Omega_1) = \overline{\mathbf{C}}$ such that the set $\Sigma(\Omega_1)$ of chunks contributing to Ω_1 is a three-fold chunk of chunks, or for some $s \leqslant \#Y$ there is $\Delta \supset f_0^{-s}(\overline{\mathbf{C}} \smallsetminus \Omega$ which is nonempty and satisfies $\Delta = f_0^{-1}(\Delta)$ modulo trivial and peripheral loops with

$$d_\Delta(w_{i_1+t+s}, \tau(w_{i_1+t+s})) \leqslant M_1,$$

and for $[\varphi] = w_{i_1+t+s}$,

$$|\varphi(\partial\Delta)| \leqslant M_1.$$

Remark. — By 30.7, if t is sufficiently large depending only on $\#(Y)$ then for $I_0 \times \beta_0 \in \Omega_{k,k,t,q}$ we have $I_0 \times \beta_0 \in \Sigma_{k,r}$ for all $r \geqslant q$.

Proof. — Choose a sequence $(\Delta_j, r_j, s_j, m_j)$ of ltd parameter functions for $0 \leqslant j < \infty$ such that if a chunk $I \times \alpha$ contains no ltd gaps with respect to (Δ_j, r_j, s_j) and no m_j-Pole-Zero loops then the same is true for components of $f_0^{-1}(I \times \alpha)$ for $j+1$ replacing j. By 29.4 we can find a bounded i_1 and $(\Delta_j, r_j, s_j, m_j)$ and Σ_i such that every chunk of Σ_i is either ltd with respect to all $(\Delta_j, r_j, s_j, m_j)$.for all $i_1 \leqslant j \leqslant i_1+n_1$, or none of them. To simplify the notation, take $i_1 = 0$. Now define $\Omega'_{i+k,j,q,t}$ to be the union of all β_0 with $I_0 \times \beta_0 \in \cup_{i+k\leqslant n\leqslant j}\Sigma_{k,q}$ so that for $\ell \leqslant t$ there is $I_\ell \times \beta_\ell$ which is ltd with respect to $(\Delta_{k+\ell}, r_{k+\ell}, s_{k+\ell}, m_{k+\ell})$ and $I_{\ell+1} \times \beta_{\ell+1} \subset f_0^{-1}(I_\ell \times \beta_\ell)$ for $\ell < t$. It follows from the properties of the sequence of long thick and dominant parameter functions that we can extend the sequence $I_\ell \times \beta_\ell$ to $-k \leqslant \ell \leqslant t$ so that $I_\ell \times \beta_\ell$ is ltd with respect to $(\Delta_{k+\ell}, r_{k+\ell}, s_{k+\ell}, m_{k+\ell})$ for all $\ell \geqslant -k$.

Then we have we have $\Omega'_{i+k+1,j+1,q+1,t-1} \subset f_0^{-1}(\Omega'_{i+k,j,q,t})$ by the definition. We have seen in 30.7 that $\Omega'_{i+k,j,q+1,t} = \Omega'_{i+k,j,q,t}$. So we have $\Omega'_{i+k+1,j+1,q,t} \subset f_0^{-1}(\Omega'_{i+k,j,q,t})$. Also $\Omega'_{i',j',q,t} \subset \Omega_{i'',j'',q',t'}$ if $i'' \leqslant i' \leqslant j' \leqslant j''$, $q' < q$, $t' > t$. Now we claim that we can find $i+k$, j such that the set $\Sigma(\Omega'_{i+k,j,q,t})$ of chunks contributing to $\Omega'_{i+k,j,q,t}$ satisfies condition 2 for being a 9-fold chunk of chunks. It may not satisfy Condition 1 because there may be extra ltd chunks in between chunks of $\Sigma(\Omega'_{i+k,j,q,t})$.

This is similar to the construction of chunk systems in 29.3-5. For if $\Sigma(\Omega'_{i+k,j,q,t})$ does not satisfy condition 2 for being a 9-fold chunk of chunks, choose an interval $[i', j'] \subset [i+k, j+k]$ with $\Omega'_{i',j',q,t}$ properly contained in $\Omega_{i+k,j,q,t}$ and repeat the process. If we end with $\Omega_{i+k,j,q,t} = \varnothing$, so much the better, provided $j-(i+k)$ is sufficiently large as we can assume.

So then we have, for $\Omega = \Omega'_{i+k,j,q,t}$, $\Omega \subset f_0^{-1}(\Omega)$, and $\Omega'_{i_1,i_2,q,t} = \Omega'_{i+k,j,q,t}$ where $[i_1, i_2]$ is about the middle third and $\Sigma(\Omega'_{i_1,i_2,q,t})$ satisfies condition 2 for being 3-fold (although possibly not condition 1). We clearly have $\Omega'_{i+k,j,q,t} \subset \Omega_{i+k,j,q}$ for any $i+k \leqslant j \leqslant q$. But if $I \times \zeta \in \Sigma_{\ell,q}$ is between two chunks $I_0 \times \beta_0$ and $I'_0 \times \beta'_0$ which contribute to $\Omega_{i+k,j,q,t}$, in particular so that ζ intersects the interior of Ω, then we must have $\zeta \subset \Omega$. For if not, take $[\varphi] \in I$ and $\varphi(\partial\Omega)$ in good position with respect to the quadratic differential determined by I. If (without loss of generality) the good position of $\varphi(\partial\Omega \cap \zeta)$ contains segments bounded from the unstable manifold, then by 15.11 it contains segments intersecting $\varphi(\alpha' \cap \Omega)$, which contains segments close to the stable manifold for any ltd gap or Pole -Zero loop $\alpha' \subset \beta'_0$.

So we obtain $\Omega = \Omega_{i_1,i_2,q} \cap \partial\Omega'_{i+k,j,q,t} = \varnothing$. So each component of $\Omega'_{i_1,i_2,q,t}$ is a component of Ω and similarly if we split into thirds. So if $\Omega' = \Omega'_{i_1,i_2,q,t}$ has a component Ω_1 satisfying $\Omega_1 \subset f_0^{-1}(\Omega_1)$ then Ω_1 is also a component of Ω, and $\Sigma(\Omega_1)$ — the set of all ltd chunks in $\cup_{i_1 \leqslant k \leqslant i_2} \Sigma_{k,q}$ contributing to Ω_1 — is three-fold. If $\Omega' \neq \varnothing$ then $f_0^{-s}(\Omega')$ is a union of components of $f_0^{-s}(\Omega)$ and if $\Omega_1 \neq \varnothing$ then $f_0^{-s}(\Omega_1)$ is a union of components of $f_0^{-s}(\Omega')$. So we can only have $f_0^{-s}(\Omega) = \overline{\mathbf{C}}$ if $f_0^{-s}(\Omega_1) = \overline{\mathbf{C}}$. Otherwise take Δ to be the complement of $f_0^{-s-t}(\Omega')$ which is nonempty, invariant. Suppose that $I \times \zeta$ is a ltd chunk in Σ_{i_1+s+t} which intersects Δ and is ltd for $(\Delta_{i_1+s+t}, r_{i_1+s+t}, s_{i_1+s+t}, m_{i_1+s+t})$. Then we can write $I \times \zeta = I_p \times \zeta_p$ with $I_{\ell+1} \times \zeta_{\ell+1} \subset f_0^{-1} I_\ell \times \zeta_\ell$ with $I_\ell \times \zeta_\ell$ ltd for $(\Delta_\ell, r_\ell, s_\ell, m_\ell)$. In particular, assuming t is sufficiently large ζ_0 is periodic and hence $I_0 \times \zeta_0 \in \Sigma_{i_1,q}$. So $\zeta_0 \in \Omega'$ and $\zeta \in f_0^{-s}(\Omega')$, contradicting $\zeta \cap \Delta \neq \varnothing$. So no chunks in Σ_{i_1+s+t} which are ltd for $(\Delta_{i_1+s+t}, r_{i_1+s+t}, s_{i_1+s+t}, m_{i_1+s+t})$ intersect Δ and we have

$$d_\Delta(w_{i_1+s+t}, \tau(w_{i_1+s+t})) \leqslant M_1,$$

and for $[\varphi] = w_{i_1+t+s}$,

$$|\varphi(\partial\Delta)| \leqslant M_1,$$

as required. □

30.10. Lemma. *Suppose that Ω_j is a fixed component of $\Omega_{i_n+k_n,j_n,q}$ in the notation of the previous lemma for $j = 1, 2$, with $i_2 > i_1 + k_1$. Suppose also that $f_0^{-s}(\Omega_j)$ is the same surface for $j = 1, 2$. Then for any β_0 in the middle third of Ω_1 as in 30.9 and any β' correspondingly in the middle third of Ω_2, $\beta \cap \beta' \neq \varnothing$.*

Proof. Each $\Omega'_{i_n+k_n,j_n,q,t}$ satisfies the second condition for being 3-fold, from the proof of 30.9. Then $\beta \cap \beta'' \neq \varnothing$ for any β'' in the last third of $\Omega'_{i_1+k_1,j_1,q,t}$ and $\beta' \cap \beta''' \neq \varnothing$ for any β''' in the first third of $\Omega'_{i_2+k_2,j_2,q,t}$. The only way we can have $\beta_0 \cap \beta'_0 = \varnothing$ is if $\beta'' \cap \beta''' = \varnothing$ for all β'', β'', that is, $\Omega_2 \subset f_0^{-s}(\Omega_1) \setminus \Omega_1$. But this would imply that Ω_2 is aperiodic, which is impossible because in the proof of 30.9 we have $\Omega_j = \Omega'_{i_n,j_n,q,t}$, and the chunks used to define this are periodic. □

30.11. Construction of y_m, α_m. — We now assume the construction of the w_j with the required properties. Let $w_j = w_j(y_0)$. Fix suitable integers q and t. We use the construction of 30.9. Take a sequence of sets $\Omega_{i_j,i_j+k_j,i_j+q}$ for $r_0 \leqslant k_j \leqslant r_1$ and $0 < i_{j+1} - (i_j + k_j) \leqslant r_1$. From the proof of 30.9, we have

$$\partial\Omega_{i_j,i_j+k_j,i_j+q} \supset \partial\Omega'_{i_j,i_j+k_j,i_j+q,t}, \tag{1}$$

where $\Omega'_{i_j,i_j+k_j,i_j+q,t}$ can be defined using different — stronger — ltd parameter functions and pole-zero constant. But we can choose the ltd parameter functions and pole-zero constants defining the $\Omega'_{i_j,i_j+k_j,i_j+q,t}$ differently for different j so that

$$\Omega'_{i_{j+1},i_{j+1}+k_{j+1},i_{j+1}+q,t} \subset \Omega'_{i_j,i_j+k_j,i_j+q,t}.$$

Whatever parameter functions we choose, we always have (1), and hence, for $w_k = [\varphi]$, $i_j + t + s \leqslant k \leqslant i_{j+1} + t + s$, and $\gamma \subset \partial f_0^{-s}\Omega'_{i_j,i_j+k_j,i_j+q,t}$, or, more generally $\gamma \subset \partial f_0^{-s}\Omega_{i_j,i_j+k_j,i_j+q}$,

$$|\varphi(\gamma)| \leqslant M_1,$$

where M_1 depends only on $\#(Y)$, r_1 and the ltd parameter functions and pole-zero constant used to define the w_i's, not on the parameter functions used to define the different $\Omega'_{i_j,i_j+k_j,i_j+q,t}$. Write

$$\Omega_j = f_0^{-s}\Omega'_{i_j,i_j+k_j,i_j+q,t},$$

and Δ_j for the complement of Ω_j. Then we have $\Omega_{j+1} \subset \Omega_j$ and $\Delta_j \subset \Delta_{j+1}$. We also have $f_0^{-1}(\Omega_j) = \Omega_j$ and $f_0^{-1}(\Delta_j) = \Delta_j$. So $\partial\Omega_j = \partial\Delta_j$ is an invariant loop set and exactly one of Ω_j, Δ_j has a fixed component. The sets Ω_j and Δ_j are defined using ltd parameter functions and pole-zero constants which vary with j — although, as we have seen, their boundaries are independent of these choices. Each component of Δ_j (or Ω_j) is eventually periodic, and each periodic component is mapped either homeomorphically or not. We claim that the constant M_1 can be chosen, independent of j so that if P is a homeomorphic periodic component of Δ_j, for any j, then, for all $k \geqslant i_j + t + s$,

$$d_P(w_k, \tau(w_k)) \leqslant M_1. \tag{2}$$

In fact, this follows from the Straightening Property, as we shall see. We shall also see that we can choose the w_k and M_1 so that (2) holds when P is periodic nonhomeomorphic, for $k \geqslant i_j + t + s$. It then follows that (2) holds also for all components P of Δ_j, for $k \geqslant i_j + 2s + t$, for s large enough given $\#(Y)$. Note that if Δ_j has a fixed component then so does Δ_k for all $k > j$, because we then have $\Delta_j \subset \Delta_k$.

So it remains to choose the y_m and α_m so that the properties of 28.3 are satisfied. We always choose y_m by modifying $y'_m \in [w_k, \tau(w_k)]$ for some $k = k(m)$, strictly increasing with m, and with $F(y_m) \leqslant F(y'_m)$, which gives the F-decreasing property 28.3b). We shall also have $k(m+1) - k(m)$ bounded above. We take $y_m = y'_m$ except in two cases.

Case 1. — Δ_j has a nonhomeomorphic fixed component for $i_j \leqslant k(m) \leqslant i_j + k_j$, when y_m is chosen to be a modification of y'_m as described in 30.15 below.

Case 2. — $y'_m \in \mathcal{T}(\Gamma, \varepsilon'_1)$ for some (f_0, Γ) satisfying the Invariance and Levy Conditions, and ε'_2 is sufficiently small given $\mathrm{Max}\{F(x) : x \in \ell\}$, where the $w_j(x)$ and $y_m(x)$ are being constructed for $x \in \ell$, for some geodesic ℓ in $\mathcal{T}(Y)$ with endpoints in $\widetilde{V}_1$. Then for some $\varepsilon_2 < \varepsilon'_2$ again depending on $\mathrm{Max}\{F(x) : x \in \ell\}$, and $y'_m \in \mathcal{T}(\Gamma, \varepsilon_2)$, we take $y_m = \sigma(y'_m)$ for σ as in Chapters 18-21, and for $y'_m \in \mathcal{T}(\Gamma, \varepsilon'_2) \smallsetminus \mathcal{T}(\Gamma, \varepsilon_2)$ we use the vector fields involved in the definition of σ to define y_m from y'_m. We shall later choose $\varepsilon_1 < \varepsilon_2$ sufficiently small given ℓ. This is similar to how ε_1 was chosen in Chapter 7: small enough given $F(x)$ for x in some set.

Suppose inductively that $y_m \in [w_k, \tau(w_k)]$ has been chosen, and that $i_j < k < i_{j+1}$. Write $j = j(m)$. If $\Delta_{j(m)}$ has a fixed component then we take α_m to be this fixed component, and then take $y_{m+n} = w_{k+n}$ for $n > 0$ and α_{m+n} to be the fixed component of $\Delta_{j(m+n)}$ where $j(m+n)$ is such that $i_{j(m+n)} \leqslant k+n < i_{j(m+n+1)}$. If $\Omega_{j(m)}$ has a fixed component, then we choose $\alpha_m \subset \Omega(j(m))$ so that $\{y_m\} \times \alpha_m$ is a bounded distance from the middle third of the 3-fold chunk of chunks along $[w_{i_m}, \tau(w_{i_m+k_m})]$ (up to bounded distance), with $y_m \in \cup_{i_m \leqslant k \leqslant i_m+k_m} [w_k, \tau(w_k)]$. By 30.10, we do then have $\alpha_m \cap \alpha_{m+1} \neq \varnothing$, as required by 28.3d)(i). By choosing y_m at an appropriate place, in the middle of some segment $[w_k, \tau(w_k)]$ corresponding to a segment of $[w_{i_{j(m)}}, \tau(w_{i_{j(m)}+k_{j(m)}})]$ along which α_m is ltd or satisfies the Pole-zero condition, that the rest of 28.3d)(i) is satisfied. We can also ensure that for each x and m, either $x' \mapsto y_m(x')$ or $x' \mapsto y_{m+1}(x')$ is continuous at x, as also required by 28.3.

The following lemma gives the mostly Z property for $\{w_n\}$, and will also give the related property 28.3c) of $\{y_m\}$.

30.12. Lemma. — *The following holds for a constant M depending only on a choice of long thick and dominant parameter functions. Let $[\varphi] \in \mathcal{T}(Y)$ and let $[w, w']_Z = [[\varphi], [\chi \circ \varphi]]_Z$ be a geodesic segment in $\mathcal{T}(Z)$, with χ minimising distortion up to isotopy constant on $\varphi(Z)$. Let x, $x' \in \mathcal{T}(Y)$ and $[x, x']_Z$ be such that there is a chunk system for $[x, x']_Z$ with chunks a bounded distance from subsets of ltd chunks of a chunk system for $[w, w']_Z$.*

Then we can write $x = [\varphi']$ and $x' = [\xi \circ \chi' \xi' \circ \varphi']_Y$ where $d_Y([\xi], [\mathrm{identity}]) \leqslant M$ and $d_Y([\xi'], [\mathrm{identity}]) \leqslant M$, χ' minimises distortion up to isotopy constant on $\xi' \circ \varphi'(Z)$.

Proof. — By hypothesis, there is a decomposition of $\varphi'(\overline{\mathbf{C}} \smallsetminus Y)$ into subsurfaces α with $|\varphi'(\partial\alpha)|$ bounded and $d_{\alpha,Y}(z, z_\alpha)$ bounded for some $x_\alpha \in [w, w']_Z$, if α is a gap, or $|\operatorname{Re}(\pi_{\alpha,Y}(x) - \pi_{\alpha,Y}(x_\alpha)|$ bounded if α is a loop. The quadratic differentials $q_\alpha(z)dz^2$ at $z_\alpha = [\varphi_\alpha])$ for $d_Z(z_\alpha, w')$, for varying α, are all stretches, by varying factors, of the quadratic differential at w for $d_Z(w, w')$. If γ is a common boundary component of adjacent surfaces α and β then the good position of $\varphi_\alpha(\gamma)$ with respect to $q_\alpha(z)dz^2$ is a stretch of the good position of $\varphi_\beta(\gamma)$ with respect to $q_\beta(z)dz^2$. Then we can paste these together to obtain a *fake* quadratic differential on some surface $\overline{\mathbf{C}} \smallsetminus \varphi_1(Y)$, and in particular, a fake unstable foliation. If $d_{\alpha,Z}(x, x')$ is bounded, then we may paste in something else: any suitable surface will do. Then take any loop set

$\Gamma \subset \overline{\mathbf{C}} \setminus Z$ such that $|\varphi''(\Gamma)|$ is bounded, for $[\varphi''] = x'$ and such that $\overline{\mathbf{C}} \setminus (\Gamma \cup Z)$ is a union of at most once punctured discs. Then there is a homeomorphism ξ_1 with $d_Y([\xi_1], [\text{identity}])$ bounded, between the good positions of $\varphi_1(\Gamma)$ with respect to the fake quadratic differential, and the good position of $\varphi'(\Gamma)$ with respect to the quadratic differential $p(z)dz^2$ at x for $d_Z(x, x')$. Composing $[\xi_1]$ on the right with another bounded homeomorphism if necessary (because x may not be exactly $[\varphi_1]$) we have that $d_Y([\chi' \circ \xi' \circ \varphi'], x')$ is bounded. So we can find $[\xi]$ with $d_Y([\xi], [\text{identity}])$ bounded with $x' = [\xi\chi' \circ \xi' \circ \varphi']$, as required. □

This lemma gives the mostly Z Property for $\{w_n\}$, because by the Basic Property there is a chunk system for $[w_n, w_{n+1}]_Y$ which is, for ltd chunks up to bounded distance a subset of a chunk system for $\cup_{j \leqslant r_0}[w_{n-j}, \tau(w_{n-j})]$. So we obtain, for $w_{n,0} = w_n$, $w_{n+1} = w_{n,r}$, $w_{n,i}$, $w'_{n,i}$ $(1 \leqslant i < r)$ with $d_Y(w_{n,i}, w'_{n,i})$ bounded and $w_{n,i}$, $w'_{n,i+1}$, $\tau(w_{n-i})$, (w_{n-i}) having the properties of x, x', w, w' in the lemma, for $0 \leqslant i < r$.

The lemma also gives similar properties to the mostly Z property for $y \in [w_n, w_{n+1}]$ and $y' \in [w_{n+p}, w_{n+p+1}]$, for r (in the mostly Z Property) bounded in terms of p, and hence gives property 28.3c) of y_{m+1} in terms of y_m.

So to complete the construction of the y_m, it suffices to prove (2) for homeomorphic periodic components of Δ_j, and show that we can ensure (2) for periodic nonhomeomorphic components of Δ_j, and deal with exceptional indices.

30.13. The homeomorphic periodic components of Δ_j. — At this stage, we can forget about the precise definition of Δ_j, but the set Δ_j is increasing with j and therefore changes only boundedly finitely often, with bound given by $\#(Y)$. So now suppose that P is a (not necessarily connected) set with $f_0(P) = P$ and $f_0 \mid P$ homeomorphic. and suppose that for $[\varphi] \in [w_k, \tau(w_k)]$ for $n_1 \leqslant k \leqslant n_2$,

$$|\varphi(\partial P)| \leqslant M_1. \tag{1}$$

Suppose also that

$$d_P(w_{n_1}, \tau(w_{n_1})) \leqslant M_1.$$

Then we shall show that the Straightening Property implies that for a suitable constant M_2, and all $n_1 \leqslant k \leqslant n_2$,

$$d_P(w_k, \tau(w_k) \leqslant M_2. \tag{2}$$

This will complete what we need for homeomorphic periodic components of the Δ_j, as stated in 30.13. Recall that d_P is actually $\mathrm{Max}_i\, d_{P_i}$ where P_i are the components of P and d_{P_i} the Teichmüller distance in $\mathcal{T}(A(P_i))$, where $A(P_i) \subset Y$ is chosen relative to P_i (see 9.1). Then we write $\mathcal{T}(P) = \Pi_i \mathcal{T}(A(P_i))$ as a set of isotopy classes $[\varphi]$ where φ is a homeomorphism defined on a disjoint union of sets $(\overline{\mathbf{C}}, A(P_i))$. Then $\mathcal{T}(Y)$ projects naturally to $\mathcal{T}(P)$. We define $[\varphi] \cdot [\chi_P] = [\varphi \circ \chi_P]$ for all $\varphi] \in \mathcal{T}(P)$, and this also defines a map from $\mathcal{T}(Y)$ to $\mathcal{T}(P)$. Then for M_1' depending only on M_1 in (1), $[\varphi] \in [w_k, \tau(w_k)]$ for $n_1 \leqslant k \leqslant n_2$,

$$d_P(\tau([\varphi]), [\varphi] \cdot [\chi_P]) \leqslant M_1'.$$

So in order to show (2) it suffices to show that for a suitable M_2' for all $n_1 \leqslant k \leqslant n_2$,

$$d_P(w_k, w_k \cdot [\chi_P]) \leqslant M_2', \tag{3}$$

just from the Straightening Property. We can assume that the Straightening Property is stated for $\tau(w_j)$ replaced by $\pi_P(w_j) \cdot [\chi_P]$, that is, there are chunk systems for $[\pi_P(w_{n_1}), \pi_P(w_k)]$ and $[\pi_P(w_k), \pi_P(w_k) \cdot [\chi_P]]$ such that every chunk is, up to bounded distance, a union of chunks for a chunk system for $[\pi_P(w_{n_1}, \pi_P(w_k) \cdot [\chi_P]]$, for $n_1 + n_0 \leqslant k \leqslant n_2$. (Here, n_0 is bounded in terms of $\#(Y)$.) Now $\pi_P(w_{k+1})$ and $[\pi_P(w_{k+1}), \pi_P(w_{k+1}) \cdot [\chi_P]]$ are obtained from $[\pi_P(w_k) \cdot [\chi_P], \pi_P(w_k) \cdot [\chi_P^2]]$ by removing chunks. Since χ_P is a homeomorphism, every ltd chunk is periodic — not just eventually periodic, as happens with τ. If an orbit of chunks is going to need removing it will happen for $n_1 \leqslant k \leqslant n_1 + n_0'$ for some n_0' bounded in terms of the topological type of P (and hence in terms of $\#(Y)$). So there is n_0' such that for $n_1 + n_0' \leqslant k \leqslant n_2$, $\pi_P(w_{k+1}) = \pi_P(w_k) \cdot [\chi_P]$, up to bounded distance. So for $n_1 + n_0' \leqslant k < n_2$, every segment of

$$[\pi_P(w_k), \pi_P(w_k) \cdot [\chi_P]] \cup [\pi_P(w_k) \cdot [\chi_P], \pi_P(w_k) \cdot [\chi_P^2]],$$

along which some gap is ltd, or some loop is Pole-zero, is a bounded distance from another such segment on $[\pi_P(w_k), \pi_P(w_k) \cdot [\chi_P^2]]$. This can only be true if $\pi_P(w_k)$ is a bounded distance from the set on which $d_P(w, w \cdot [\chi_P])$ is within a bounded distance of the minimum or infimum possible. If $[\chi_P]$ is a pseudo-Anosov class then this means that $\pi_P(w_k)$ must be a bounded distance from the geodesic along which the minimum of $d(w, w \cdot [\chi_P])$ is achieved. If $[\chi_P]$ is an isometric isotopy class then $\pi_P(w_k)$ must be a bounded distance from the set on which $d(w, w \cdot [\chi_P])$ takes the value 0. If $[\chi_P]$ is a reducible isotopy class then for any γ in the loop set Γ of loops in P which is left invariant by $[\chi_P]$, and of which complementary components are irreducible, then for $w_k = [\varphi]$, $|\varphi(\gamma)$ is bounded, the projections $\pi_\gamma(w_k)$ are such that $d_\gamma(w_k, w_k \cdot [\chi_P])$ are bounded. Also for any periodic orbit Q of $P \setminus (\cup\Gamma)$, $\pi_Q(w_k)$ satisfies similar conditions to the above for $\pi_P(w_k)$, depending on whether $[\chi_Q]$ is pseudo-Anosov or isometric. Altogether, this gives a bound

$$d_P(w_k, w_k \cdot [\chi_P]) \leqslant \kappa_P$$

where κ_P depends only on the isotopy class $[\chi_P]$, which gives (3).

30.14. A nonhomeomorphic fixed component of Δ_j. — Let P be a nonhomeomorphic fixed component of Δ_j. Again, we can forget the precise definition of Δ_j. We simply assume that for some integers $n_1 < n_2$,

$$d_P(w_{n_1}, \tau(w_{n_1})) \leqslant M_1,$$

and for $w_k = [\varphi]$, any $n_1 \leqslant k \leqslant n_2$,

$$|\varphi(\partial P)| \leqslant M_1.$$

Now we want to show that if w_k can be chosen to satisfy all the properties of 30.2 then it can also be chosen to satisfy

$$d_P(w_k, \tau(w_k)) \leqslant M_2.$$

for a constant M_2 depending only on M_1 and $\#(Y)$ and all $n_1 \leqslant k \leqslant n_2$. In order to do this we shall show that given a constant M_4, for k_0 depending only on $\#(Y)$ and some $0 < i \leqslant k_0$,

$$d(w_{k+i}, \tau(w_{k+i})) \leqslant d(w_k, \tau(w_k)) - M_4. \tag{1}$$

We obtain (1) as follows. As in Chapter 27, we basically need to use the argument of 6.6, where the key fact used is: given a quadratic differential at a point in $\mathcal{T}(Y)$, the pullback under a holomorphic map of sufficiently high degree is not a quadratic differential at any point in $\mathcal{T}(Y)$.

Now v_2 and v_1 are not both in the same complementary component of P. Then

$$F(\tau(w_k)) \leqslant \sum_{i=0}^{N-1} d_Z(\tau(w_{k,i}), \tau(w_{k,i+1}))$$

and $F(w_{k+1})$ is not much bigger. If

$$F(w_k) - F(\tau(w_k)) \geqslant 2M_4 \tag{2}$$

then we already have the required bound on $F(w_{k+1}) - F(w_k)$. In general, to obtain (1), we only need, for some $j < k_0$,

$$F(\tau^{j+1}(w_k)) - F(\tau^j(w_k)) \geqslant 2M_4 \tag{3}$$

Now we fix long thick and dominant parameter functions (r, Δ, s) and a pole-zero constant C and $\nu_1 = \nu_1(r, \Delta, s)$ such that ltd loops or long ν-thick and dominant gaps always exist on sufficiently long geodesic segments for some $\nu \geqslant \nu_1(r, \Delta, s)$ (see 15.4), and so that, for all ν, $\Delta(\nu)$ is sufficiently large (in a way specified below) given M_4. We break $[w_k, \tau(w_k)]_Z = [w_k, \tau(w_k)]_Y$ into segments $[w_{k,i}, w_{k,i+1}]$ ($0 \leqslant i \leqslant N$, some N depending on $F(w_k)$) of length between $\Delta_0/2$ and Δ_0, for a suitable Δ_0 so that each $[w_{k,i}, w_{k,i+1}]$ is sufficiently long to contain a segment along which some gap or loop is ltd or pole-zero for (Δ, r, s, C).

Now for any sequence $\beta_{k,i}$ ($1 \leqslant 0 \leqslant N-1$) of gaps and loops such that $\beta_{k,i}$ is ltd or pole-zero along a segment $[z_{k,i,1}, z_{k,i,2}]$ of $[w_{k,i}, w_{k,i+1}]$, define

$$\begin{aligned} &S(z_{k,0,1}, z_{k,0,2} \cdots z_{k,N-1,2}, \beta_{k,0}, \dots, \beta_{k,N-1}) \\ &\quad = d'_{\beta_{k,0},\beta_{k,1}}(w_{k,i}, z_{k,0,1}) + \sum_{i=0}^{N-1} \left(d_{\beta_{k,i}}(z_{k,i,1}, z_{k,i,2}) + d'_{\beta_{k,i},\beta_{k,i+1}}(z_{k,i,2}, z_{k,i+1,1})\right) \\ &\qquad + d'_{\beta_{k,N-1},\beta_{k,N}}(z_{k,N-1,2}, w_{k,N}). \end{aligned}$$

where $d'_{\beta,\zeta}$ is as in 14.10. Now for a suitable constant C_1,

$$\begin{aligned}\mathrm{Max}\{S(z_{k,0,1}, z_{k,0,2} \cdots z_{k,N-1,2}, \beta_{k,0}, \ldots, \beta_{k,N-1}) - NC_1 =\leqslant d(w_k, \tau(w_k))\\ \leqslant \mathrm{Max}\{S(z_{k,0,1}, z_{k,0,2} \cdots z_{k,N-1,2}, \beta_{k,0}, \ldots, \beta_{k,N-1}) + NC_1.\end{aligned}$$

We have (2) unless, for any choice of sequence $\beta_{k,i}$, for all i,

$$d_{f_0^{-1}(\beta_{k,i})}(\tau(z_{k,i,1}, \tau(z_{k,i,2})) \geqslant d_{\beta_{k,i}}(z_{k,i,1}, z_{k,i,2}) - 2M_4$$

and there are components $\beta_{k,i,1}$ of $f_0^{-1}(\beta_{k,i})$ with $\beta_{k,i,1} \cap \beta_{k,i+1,1} \neq \varnothing$ which are $\nu_{k,i}^{(1)}$-thick (for suitable numbers $\nu_{k,i}^{(1)}$) or pole-zero along all but bounded distance of $[\tau(z_{k,i,1}), \tau(z_{k,i,2})]$ and $d_{\beta_{k,i,1},Z}$-distance $\leqslant C_1(\nu_{k,i,1}, \leqslant C_1(\nu_{k,i+1,1})$ from a similar segment on $[\tau(z_{k,i,1}), \tau(z_{k,i+1,2})]$, $[\tau(z_{k,i-1,1}), \tau(z_{k,i,2})]$, or similarly if $\beta_{k,i}$ is a loop. Then by 28.5, each segment $[\tau(z_{k,i,1}), \tau(z_{k,i,2})]$ is also a bounded $d_{\beta_{k,i,1},Z}$-distance from a similar segment along $[w_k, \tau(w_k)]$. Now similarly, we have for some j, either

$$F(\tau^{j+1}(w_k)) \leqslant F(\tau^j(w_k)) - 2M_4$$

or for $1 \leqslant \ell \leqslant j+1$ there are $\beta_{k,i,\ell} \subset f_0^{-1}(\beta_{k,i,\ell-1})$ (with $\beta_{k,i,0} = \beta_{k,i}$) and

$$[\tau^\ell(z_{k,i}), \tau^\ell(z_{k,i})]$$

is $d_{\beta_{k,i,\ell}}$-distance $\leqslant C_1^(\nu_{k,i}^{(\ell)})$ from a similar segment on $[\tau^\ell(z_{k,i,1}), \tau^\ell(z_{k,i,2})]$, or similarly if $\beta_{k,i,\ell}$ is a loop and

$$d_{\beta_{k,i,\ell}}(\tau^\ell(z_{k,i,1}), \tau^\ell(z_{k,i,2})) \geqslant d_{\beta_{k,i,\ell-1}}(\tau^{\ell-1}(z_{k,i,1}), \tau^{\ell-1}(z_{k,i,2})) - 2M_4.$$

But if this happens for all $\ell \leqslant k_0$ for a suitable k_0 depending only on $\#(Y)$, we obtain a contradiction. We use the concept of pole equivalence class of 27.3. By the same argument as in 27.4-5, the quadratic differentials at points of

$$\pi_{\beta_{k,i,\ell}}([\tau^\ell(z_{k,i,1}), \tau^\ell(z_{k,i,2})] \cap \mathcal{T}(\partial\beta_{k,i,\ell}, \varepsilon_0))$$

must have pole equivalence classes of the same types as at points of

$$\pi_{\beta_{k,i,\ell}}(\tau^\ell([z_{k,i,1}, z_{k,i,2}]) \cap \mathcal{T}(\partial\zeta_i, \varepsilon_0)).$$

This is impossible, for ℓ sufficiently large depending only on $\#(Y)$, because the degree of f_0^ℓ increases strictly at regular intervals, since $\beta_{k,i,\ell} \cap P = \varnothing$ and v_1 and v_2 are not in the same component of the complement of P — and hence the same is true for $\beta_{k,i,\ell}$, for some $\ell \leqslant k_0$.

30.15. Exceptional Indices. — Now as in the proof of the Level κ-tool in Chapters 18-21, we modify the definition of y_m for a sequence of exceptional indices m_i. We write κ_i (in decreasing order) for the possible (discrete set of) values taken by $\kappa(f_0, \Gamma)$ which are $\leqslant \max\{F(x) : x \in \ell\}$. We choose ε_1 so that $\varepsilon_1/(1 - E_i) < \varepsilon_2$ for $i = 0$, 1 2 where ε_2 is as in 30.13, and the constants E_i are as in 7.7 for $i = 0, 1, 2$. Suppose that m_j and $y_{m_j}(x)$ have been chosen for $j \leqslant i - 1$.We choose $m_i > m_{i-1}$ sufficiently large that if $y_m(x) \in \mathcal{T}(\Gamma, \varepsilon_1/(1 - E_j\varepsilon_1))$ for $\kappa(f_0, \Gamma) = \kappa_i$) and $j = 0$, 1 or 2 and $m < m_i$ then $y_m(x) \in K_j(\mu, \varepsilon_1)$ where μ is minimal nonempty with $\mu \leqslant [f_0, \Gamma]$ (see 7.7 for the notation). Then we define $y_{m_i}(x) = y_{m_i-1}(x)$ except when $y_{m_i-1}(x) \in K_1([f_0, \Gamma], \varepsilon_1)$, when we take $y_{m_i}(x) \in K_2([f_0, \Gamma], \varepsilon_1)$, as in 7.8.

CHAPTER 31

STRAIGHTENING

31.1. Canonical Straightening. — We now give a canonical construction of w_{q+1} satisfying the Basic, Straightening and v_2-tracing properties, under the assumptions that, for $i \leqslant q$, w_i satisfies the Basic, Straightening and v_2-tracing properties. The mostly Z property then follows from 30.12. In fact, if we obtain w_{q+1} satisfying the Basic and Straightening properties, which only concern the projection to $\mathcal{T}(Z)$, we can satisfy the v_2-tracing property also. In order to do this we construct $w_{q+1,k}$ for $0 \leqslant k \leqslant \min(N, q+1)$ for some N bounded in terms of $\#(Y)$ and take $w_{q+1} = w_{q+1,N}$, where $w_{q+1,k}$ satisfies the Basic property and the Straightening Property with respect to $[w_{q+1-k}, \tau(w_{q+1,k})]$ rather than $[w_0, \tau(w_{q+1,k})]$. We start with $w_{q+1,0} = \tau(w_q)$. Having constructed $w_{q+1,k}$ we fix a chunk system $\Sigma_{q+1,k}$ for $[w_{q+1,k}, \tau(w_{q+1,k})]$ and let $\Sigma^{\rm ltd}_{q+1,k}$ denote the ltd chunks in $\Sigma_{q+1,k}$. Inductively we shall choose $\Sigma_{q+1,k+1}$ so that each chunk of $\Sigma^{\rm ltd}_{q+1,k+1}$ is a subset of a chunk of $\Sigma^{\rm ltd}_{q+1,k}$. We start from $\Sigma_{q+1,0} = f_0^{-1}(\Sigma_q)$. (We already have a family of chunk systems $\Sigma_{n,m}$ with $n \leqslant m$, with $\Sigma_{n,n} = \Sigma_n$. This is consistent with what we want to do now because for some $N \leqslant q+1$ we shall define $\Sigma_{q+1} = \Sigma_{q+1,N}$, once we have defined $\Sigma_{q+1,N}$.) For $0 \leqslant i \leqslant k$, we then fix a chunk system $\Sigma_{q-i,q+1,k}$ for which each chunk is a subset of $\Sigma_{q-i,q+1,k-1}$ if $i \leqslant k-1$, or of $\Sigma_{q-k+1,q}$ if $i = k$, and such that

$$\cup_{i \leqslant k-1} \Sigma_{q-i,q+1,k} \cup \Sigma_{q+1,k}$$

is a chunk system for $[w_{q+1-k}, \tau(w_{q+1,k})]$. We define $w_{q,k}$ differently for $k \leqslant N-1$ and for $k = N$.

So now suppose that $w_{q+1,k}$, $\Sigma_{q+1,k}$ and $\Sigma_{q-i,q+1,k}$ have been constructed for $i \leqslant k$. We can assume after refinement of $\Sigma_{q+1,k}$ if necessary that every chunk of $\Sigma_{q+1,k}$ is either a bounded distance from a subset of some chunk of $\Sigma_{q-k,q}$ or has no subset which is a bounded distance from a subset of any such chunk. Then let Ω_1 be the set of ltd chunks $\Sigma_{q+1,k}$ a bounded distance from a subset of some chunk of $\Sigma_{q-k,q}$, with Ω_2 the corresponding set of ltd chunks in a refinement of $\Sigma_{q-k,q}$. Let Ω_3 be the set of ltd chunks $I_{k+1} \times \zeta_{k+1}$ in a chunk system for $[w_{q+1,k}, \tau(w_{w+1,k})]$ for which there exists a sequence $I_j \times \zeta_j$ with $I_{j+1} \times \zeta_{j+1} \subset I_j \times \zeta_j$ up to bounded distance, $I_0 \times \zeta_0 \in \Omega_2$ and $I_j \times \zeta_j \in \Sigma_{q-k+j,q}$ for $j \leqslant k$. If $\Omega_1 \cap \Omega_3 = \varnothing$ then we define

$$\Sigma^{\rm ltd}_{q+1,k+1} = \Sigma^{\rm ltd}_{q+1,k} \setminus (\Omega_1 \cup \Omega_3)$$

If $\Omega_1 \cap \Omega_3 \neq \varnothing$ then we need to be more careful.

Lemma. – *We can find $\Omega_4 \subset \Omega_1$ such that, if Ω_5 and Ω_6 are defined relative to Ω_4 in the same way as Ω_2, Ω_3, respectively relative to Ω_1, then $\Omega_1 \subset \Omega_4 \cup \Omega_6$ and $\Omega_4 \cap \Omega_6 = \varnothing$.*

Proof. — Let $\Omega_1(z_1)$ be a chunk system which is a refinement of $\Omega_1 \mid [w_{q+1,k}, z_1]$, for any $z_1 \in [w_{q+1,k}, \tau(w_{q+1,k})]$. Let $\Omega_2(w)$ and $\Omega_3(z_1)$ be defined in the same way as Ω_2 and Ω_3 but relative to $\Omega_1(z_1)$. Of course, $\Omega_1(w_{q+1,k}) \cap \Omega_3(w_{q+1,k}) = \varnothing$. Take the first z_1 for which the intersection is nonempty. Then $\Omega_1(z_1) \cup \Omega_3(z_1)$ divides at least one ltd chunk of $\Sigma_{q+1,k}$. Then define $\Omega(z_1, z_2)$ to be a refinement of $\Omega_1 \mid [z_1, z_2]$ for any $z_2 \in [z_1, \tau(w_{q+1,k})]$, and let $\Omega_2(z_1, z_2)$ and $\Omega_3(z_1, z_2)$ be similarly defined to before — but looking only at ltd chunks in $\Sigma_{q+1,k} \setminus (\Omega_1(z_1) \cup \Omega_3(z_1))$ Take the first z_2 such that $\Omega_1(z_1, z_2) \cap \Omega_3(z_1, z_2) \cap \neq \varnothing$. Similarly define $\Omega_k(z_j, z_{j+1})$ for $k = 1$, 2, 3, and $j < m$, where m is the first integer such that

$$\Omega_1 \subset \Omega_1(z_1) \cup \left(\cup_{j=1}^{m-1} \Omega_1(z_j, z_{j+1})\right) \cup \Omega_3(z_1) \cup \left(\cup_{j=1}^{m-1} \Omega_3(z_j, z_{j+1})\right).$$

Then we take

$$\Omega_4 = \Omega_1(z_1) \cup \left(\cup_{j=1}^{m-1} \Omega_1(z_j, z_{j+1})\right), \quad \Omega_6 = \Omega_3(z_1) \cup \left(\cup_{j=1}^{m-1} \Omega_3(z_j, z_{j+1})\right). \qquad \square$$

Having done this we then define

$$\Sigma^{\mathrm{ltd}}_{q+1,k+1} = \Sigma^{\mathrm{ltd}}_{q+1,k} \setminus (\Omega_4 \cup \Omega_6).$$

Then $\Sigma^{\mathrm{ltd}}_{q+1,k+1}$ is a subset of $\Sigma_{q+1,k}$ which is a chunk system for some $[z, z']$, not necessarily with $z' = \tau(z)$. But this can fail only if $k+1 > 1$ and z' is in the boundary of a boundary chunk $I \times \zeta$ of Ω_5 with $\zeta \cap \alpha \neq \varnothing$ or if z is a bounded distance from the boundary of a boundary chunk $I \times \zeta$ of Ω_4 such that $z' \times \alpha$ intersects a chunk derived from $I \times \zeta$.

To ensure that $z' = \tau(z)$ and to define the non-ltd chunks of $\Sigma_{q+1,k+1}$ we consider the chunks of $\cup_{0 \leqslant i < k} \Sigma_{q-i,q+1,k}$ nearer $w_{q+1,k}$ on $[w_{q-k+1}, w_{q+1,k}]$ than the chunks of Ω_5, which are all ltd. Because the Straightening Property holds for $w_{q+1,k}$ relative to $[w_{q-k+1}, \tau(w_{q+1,k})]$, these chunks are not ltd. Now take any chunk $I \times \zeta$ of $\Omega_4 \setminus \Omega_6$, and write $I \times \zeta = I_{q+1} \times \zeta_{q+1}$ with $I_{i+1} \times \zeta_{i+1} \subset f_0^{-1}(I_i \times \zeta_i)$ and $I_{q-i} \times \zeta_{q-i} \in \Sigma_{q-i,q+1,k}$. Then $\zeta_i \cap \zeta_{q+1} = \varnothing$ for $q - k < i \leqslant q+1$, because $I_j \times \zeta_j$ has to be ltd, and for any ltd chunk $J \times \omega \in \Sigma_{j,q+1,k}$ for $j > q-k$, $\omega \cap \zeta_{q+1} = \varnothing$. Also if ζ_{q+1} is a loop, $\zeta_j \neq \zeta_{q+1}$ up to Z-preserving isotopy. So pulling back we have $\zeta_i \cap \zeta_j = \varnothing$ for $q - k < i < j \leqslant q+1$ or $q - k \leqslant i < j \leqslant q$, where this includes, for such i and j, $\zeta_i \neq \zeta_j$ up to Z-preserving isotopy. This puts a bound on k in terms of $\#(Y)$, if $\Omega_4 \neq \varnothing$, because there is a bound depending only on $\#(Y)$ on the number of disjoint nontrivial nonperiheral gasp and loops ζ_j. But $I \times \zeta$ is also a bounded distance from a chunk $I'_{q-k} \times \zeta_{q-k}$ of Ω_5. Write $\zeta'_{q-k} = \zeta_{q-k}$. So pulling back under f_0^{j-i} — which is permissible under the v_2-tracing property — we have $\zeta'_i \cap \zeta_j = \varnothing$ for any $q - k \leqslant i < j \leqslant q$, for any sequence $I'_{i+1} \times \zeta'_{i+1} \subset f_0^{-1}(I'_i \times \zeta'_i)$. In particular, $\zeta'_i \cap \zeta = \varnothing$. that is, $\zeta'_{q-k} \cap \zeta'_i = \varnothing$ for $q-k < i \leqslant q$. Pulling back further, note that this implies $\zeta'_i \cap \zeta'_j = \varnothing$ for $q-k \leqslant i < j \leqslant q$ or $q-k < i < j \leqslant q+1$. This puts a bound on k in terms of $\#(Y)$, if $\Omega_4 \neq \varnothing$ — and if $\Omega_4 = \varnothing$, $\Omega_5 = \Omega_6 = \varnothing$. For there is a bound on the number of disjoint surfaces ζ'_i, unless some ζ'_i is loop with $\zeta'_i = \zeta'_j$ up to Z-preserving isotopy. Then we can choose non-ltd chunks $J_i \times \zeta'_i$ to replace the chunks $I_i \times \zeta'_i$, and we take $\Sigma_{q+1,k+1}$ to be the chunk system with these replacements. Then $\Sigma_{q+1,k+1}$ is indeed a chunk system for $[w_{q+1,k+1}, \tau(w_{q+1,k+1})]$ for some $w_{q+1,k+1}$.

If $\Omega_4 = \varnothing$ then $\Omega_5 = \Omega_6 = \varnothing$. We have seen that there is a bound on the k for which $\Omega_4 \neq \varnothing$. So there is a least N, bounded in terms of $\#(Y)$, such that $\Sigma_{q+1,j} = \Sigma_{q+1,N}$ for all $j \geqslant N$. We then define $\Sigma_{q+1} = \Sigma_{q+1,N}$ and $w_{q+1,N} = w_{q+1}$. □

31.2. Corollary. — *The following holds for an integer r_1 depending only on $\#(Y)$ and a constant M_1 depending only on the long thick and dominant parameter functions and the constant in the v_2 tracing property. Suppose the w_i are defined using canonical straightening. Then*

$$(1) \qquad d_Y(w_q, w_{q+1}) \leqslant \sum_{j=i-r_1}^{i-1} F(w_j) + M_1,$$

Proof. — This is immediate from 31.1: (1) follows because length can be computed using just the ltd chunks and in canonical straightening every ltd chunk of Σ'_{q+1} is derived from a chunk of Σ_{q-k} for some $k \leqslant r_1$, for a suitable r_1. □

31.3. Construction of w_n via successive geodesics: $[\psi^i]$ and $[\xi^i]$. — Suppose that w_{n-1} has been constructed. We shall define w_n by constructing a sequence $[\psi^i]$ in $\mathcal{T}(Y)$ for $0 \leqslant i \leqslant N$, for some N bounded in terms of $\#(Y)$, and taking $w_n = [\psi^N]$. We take $[\psi^0] = \tau(w_{n-1})$. For all i we define $[\xi_i] = \tau([\psi^i]) \in \mathcal{T}(Y)$- remembering that, from the definition of τ, $d_Y([\psi^i],[\xi^i]) = d_Z([\psi^i],[\xi^i])$. For all i, we either choose $[\psi^{i+1}] \in [[\psi^i],[\xi^i]]_Z = [[\psi^i],[\xi^i]]_Y$ with $d_Z([\psi^i],[\xi^i])$ just sufficiently large (where what "sufficiently large" means is yet to be determined), or we choose $[\psi^{i+1}]$ so that the ltd chunks of chunk system for $[[\psi^{i+1}],[\xi^{i+1}]]_Z$ are, up to bounded distance, ltd chunks of some chunk system for $[[\psi^i],[\xi^i]]$. It will be chosen using Canonical Straightening 30.11, and with the v_2-tracing property in mind. In fact, for $i \geqslant 2$, the ltd chunks of chunk system for $[[\psi^{i+1}],[\xi^{i+1}]]_Z$ are, up to bounded distance, precisely all the ltd chunks of some chunk system for $[[\psi^i],[\xi^i]]$, and only the nonltd chunks are changed. The first case will occur for $i=1$ and for at least every other i. For $[\psi^{i+1}]$ as in the first case we shall have $F([\psi^{i+1}]) \leqslant F([\psi^i])$. For $[\psi^{i+1}]$ as in the second case we shall have

$$(1) \qquad F([\psi^{i+1}]) < F([\psi^{i-1}]).$$

These two cases will give

$$F(w_n) = F([\psi^N]) < F(w_{n-1}).$$

We shall obtain the stronger F-decreasing property by stronger estimates on $F([\psi^{i+1}]) - F([\psi^i])$, $F([\psi^{i+1}]) - F([\psi^{i-1}])$. To obtain $F([\psi^{i+1}]) \leqslant F([\psi^i])$ in the first case of $[\psi^{i+1}]$:

$$\begin{aligned} F([\psi^{i+1}]) &\leqslant d_Z([\psi^{i+1}],[\xi^i]) + d_Z([\xi^i],[\xi^{i+1}]) = d_Y([\psi^{i+1}],[\xi^i]) + d_Z(\tau([\psi^i]),\tau([\psi^{i+1}])) \\ &\leqslant d_Y([\psi^{i+1}],[\xi^i]) + d_Y([\psi^i],[\xi^i]) = d_Y([\psi^i],[\xi^i]) = F([\psi^i]). \end{aligned}$$

We shall concentrate on the construction of $[\psi^2]$ from $[\psi^1]$ — assuming that $[\psi^2]$ is as in the second case, which is usually so — and $[\psi^0]$. Let $q_2(z)dz^2$ be the quadratic differential for $d_Y([\psi^2],[\xi^2])$ at $[\psi^2]$ with stretch $p_2(z)dz^2$ at $[\xi^2]$. Write $[\xi^2] = [\chi^2 \circ \psi^2]$. Of course, χ^2 is only defined up to isotopy constant on $\psi^2(Y)$. Let $\theta_2(z)$ denote the angle between the direction of maximal distortion of χ_2^{-1} at z and the stable foliation of $p_2(z)dz^2$. We shall show that, for a suitable choice of χ^2, for L as in 8.3:

$$\int K(\chi^2)^{-1}|p_2| \leqslant \exp(2F([\psi^0]))\Big(1 + \tfrac{1}{2}L\int \theta_2^2|p_2|\Big) \tag{2}$$

Then by 8.3 we have (1) as required for $i = 1$. We shall proceed in the same way for $[\psi^{i+2}]$, $[\psi^{i+1}]$, $[\psi^i]$ replacing $[\psi^2]$, $[\psi^1]$, $[\psi^0]$, if $[\psi^{i+2}]$ is defined as in the second case from $[\psi^{i+1}]$, that is, so that the ltd chunks of chunk system for $[[\psi^{i+2}],[\xi^{i+2}]]_Z$ are, up to bounded distance, ltd chunks of some chunk system for $[[\psi^{i+1}],[\xi^{i+1}]]$, with $[\psi^i]$ replacing $[\psi^0]$ in (2).

31.4. Definition of $[\psi^2]$: the rough idea. — Largely for ease of notation, we concentrate on the construction of $[\psi^2]$. In fact, as we shall see, the construction of $[\psi^{i+1}]$ from $[\psi^i]$ is somewhat simpler for $i \geqslant 2$, when the ltd chunks for $[[\psi^{i+1}],[\xi^{i+1}]]$ are the same as those for $[[\psi^i],[\xi^i]]$ (up to bounded distance). It is possible that this might also be true for $i = 1$, but in general it is not. It is only some non-ltd chunks which change. We defined $[\psi^1]$, $[\xi^1]$ in 31.2. We also write

$$\tau(w_{n-1}) = [\xi^0] = [\chi^{t_1,1} \circ \psi^1]$$

where $\chi^{t_1,1}$ minimizes distortion up to isotopy constant on $\psi^1(Y)$ and $t_1 = d_Y([\psi^1],[\xi^0])$. We write

$$\tau([\psi^1]) = [\chi^{t_2+t_1,t_1,1} \circ \chi^{t_1,1} \circ \psi^1] = [\chi^{t_2+t_1,1} \circ \psi^1]$$

where $\chi^{t_2,t_1,1}$ minimizes distortion up to isotopy constant on $\psi^1(Z)$ and

$$d_Z(\tau([\psi]),\tau([\psi^1]) = t_2.$$

The reason for this notation is that we shall write $[\chi^{t,1} \circ \psi^1]$ for the point distance t along $[[\psi^1],[\xi^0]]_Y$ if $0 \leqslant t \leqslant t_1$, and for the point distance $t - t_2$ along $[\tau([\psi]),\tau([\psi^1]]_Z$ if $t_1 \leqslant t \leqslant t_1 + t_2$. Here $\chi^{t,1}$ minimizes distortion modulo isotopy constant on $\psi^1(Z)$ if $t \leqslant t_1$, and $\chi^{t,t_1,1}$ minimizes distortion modulo isotopy constant on $\chi^{t_1,1} \circ \psi(Z)$ if

$t \geqslant t_1$. (This notation is reasonably consistent with that used in 28.6.) In general, we define

$$\chi^{t,u,1} = \chi^{t,1} \circ (\chi^{u,1})^{-1}$$

for $u \leqslant t$, which is consistent with the above. So then

$$\chi^{t,u,1} \circ \chi^{u,1} = \chi^{t,1}$$

for $u \leqslant t$.

We have already said that $[\psi^2]$ is defined up to bounded distortion to be the canonical straightening on non-ltd chunks. To define $[\psi^2]$ pointwise, we should like to define

$$\psi^2(z) = \chi^{t(z),1} \circ \psi^1(z),$$

where $t(z)$ is piecewise constant — but obviously there is some work needed to make sense of this, because ψ^2 needs to be continuous. We want to do this in such a way that $\tau([\psi^2]) = [\xi^2$ where

$$\xi^2 = \eta \circ \chi^{u(z),1} \psi^1(z)$$

for some $u(z) - t(z) \leqslant t_1$ and some homeomorphism η of bounded distortion. If η is the identity near a point then of course the corresponding local distortion is $\leqslant e^{2t_1}$. The map η is not globally the identity, but we shall be able to compensate this by using 8.3: the factor to which $\chi^{u(z),t(z),1}$ is not the best distortion.

31.5. Pointwise definitions. — Now we define precisely the surface $\overline{\mathbf{C}} \setminus \psi^2(Y)$ and a homeomorphism χ^2 with $[\xi^2] = [\chi^2 \circ \psi^2]$. To define the surface we use the quadratic differential $q_{t,1}(z)dz^2$ at $[\chi^{t,1} \circ \psi^1]$ for $d_Z([psi^1], [\chi^{t,1} \circ \psi^1])$ (if $t < t_1$) and for $d_Z([\chi^{t_1,1} \circ \psi^1], [\chi^{t,1} \circ \psi^1])$ (if $t_1 < t \leqslant t_1 + t_2$). The Canonical Straightening of w_{n-1} gives a chunk system Σ for $[[\psi^2], [\xi^2]]$ whose ltd chunks are a subset of a chunk system for $[w_{n-1}, \tau(w_{n-1})]$ and hence the ltd chunks are a subset of a chunk system for $[[\psi^1], [\xi^1]]$, up to bounded distance. For the moment we say that a ltd chunk $I \times \omega \in \Sigma$ is *semiminimal* if ω is not contained in the union of surfaces ζ with $J \times \zeta < I \times \omega$, that is, with J to the left of I in $[[\psi^1], [\xi^1]]$ and $J \times \zeta \in \Sigma$. As usual we say that $J \times \omega$ is *minimal* if $\omega \cap \zeta = \varnothing$ for all such $J \times \zeta$. For a semiminimal chunk the *visible lower boundary* is the complement in ω of the union of all the ζ with $J \times \zeta$ ltd and $J \times \zeta < I \times \omega$. The visible boundary gives a decomposition of the surface $\overline{\mathbf{C}} \setminus \psi^2(Y)$. The surfaces has not yet been defined precisely, but will be built up using the decomposition. So take any $I \times \omega$ such that ω contributes to the visible lower boundary. We can choose $\psi^1(\partial\omega)$ so that $\chi^{t,1} \circ \psi^1(\partial\omega)$ is in good position relative to $q_{t,1}(z)dz^2$ for all $0 \leqslant t < t_1$. If $I \times \omega$ is ltd then $\chi^{t,1} \circ \psi^1(\partial\omega)$ is embedded. If $I \times \omega$ is not ltd then we choose a small perturbation of the good position so that $\chi^{t,1} \circ \psi^1(\partial\omega)$ is embedded for all $t < t_1$. So now we form the surface $\overline{\mathbf{C}} \setminus \psi^2(Y)$ by gluing together surfaces $\chi^{s_1(\omega),1} \circ \psi^1(\omega)$ for minimal $I \times \omega$ and surfaces

$$\chi^{s_1(\omega),1} \circ \psi^1(\omega) \setminus (\cup \chi^{s_1(\zeta),1}(\zeta))$$

for semiminimal $I \times \omega$. Any loop $\chi^{t,1} \circ \psi^1(\partial\omega)$ has a natural identification with $\chi^{t,1} \circ \psi^1(\partial\omega)$ for any $t \leqslant u < t_1$, using $\chi^{u,t,1} = \chi^{u,1} \circ (\chi^{t,1})^{-1}$. This enables us to glue surfaces together. For the ltd chunks, we do simply glue the surfaces together. For some of the other chunks, we shall glue in something slightly different. For the moment we leave that aside, and consider the visible upper boundary for Σ, which is defined in the same way as the visible lower boundary. Gluing together is not quite so simple because there is a discontinuity in the map $t \mapsto \chi^{t,1} \circ \psi^1(\partial\omega)$ at $t = t_1$. When gluing together $\chi^{t,1} \circ \psi^1(\partial\omega)$ and $\chi^{u,1} \circ \psi^1(\partial\omega)$ for $t < t_1$ and $u > t_1$ we simply use length parametrisation. In this way we get a surface which is $\overline{\mathbf{C}} \setminus \xi^2(Y)$ up to bounded distortion only, which will coincide with $\overline{\mathbf{C}} \setminus \xi^2(Y)$ on any part of the surface of the form $\chi^{t_1+t_2,1} \circ \psi^1(f_0^{-1}\omega)$ for which $\psi^1(\omega)$ is part of the surface $\overline{\mathbf{C}} \setminus \psi^2(Y)$. On other parts of the surface, we have to compose with a bounded distortion η to get to $[\xi^2] = \tau([\psi^2])$.

As for the remaining non-ltd chunks, where the surfaces for $\overline{\mathbf{C}} \setminus \psi^2(Y)$ have not been glued in: the chunk system given by Canonical Straightening tells us what surface we would like to glue in up to bounded distortion, but at this point we may not be able to get all the projections right in one step, if we want to keep $F([\psi^2]) \leqslant F([\psi^0])$, and it is for this reason only that we may need $[\psi^i]$ for $i \geqslant 2$. We would like to define $\pi_{\zeta_j}([\psi^2])$ differently from $\pi_{\zeta_j}([\psi^0])$ for $0 \leqslant j \leqslant r$, where $\zeta_{j+1} \subset f_0^{-1}(\zeta_j)$ and $I_0 \times \zeta_0$ comes within a bounded distance of chunks on $[w_{n-r-1}, w_{n-r}]$. But we may be constrained to define just some, not all of the $\pi_{\zeta_j}([\psi^2])$ differently from $\pi_{\zeta_j}([\psi^0])$, and then define some $\pi_{\zeta_j}([\psi^{i+2}])$ differently from $\pi_{\zeta_j}([\psi^i])$ for $i \leqslant N-2$ until all the $\pi_{\zeta_j}([\psi^N])$ are correct. There is also another choice, for $\pi_{\partial\zeta_j}([\psi^2])$ and $\pi_{\partial\zeta_j}([\psi^i])$. We leave this choice open for the moment — to be specified in 31.6 — but it will give use the property $F([\psi^{i+2}]) < F([\psi^i])$ which we need.

As for the v_2-tracing property: we have so far defined $[\psi^2]$ only as an element of $\mathcal{T}(Z)$, using these surfaces. We choose $[\psi^2]$ so that for any gap β at $[\psi^2]$ if $d_{\beta,Z}([\psi^2], w) \leqslant M_1$ for some $w \in [w_i, w_{i+1}]$ within a bounded distance of $[w_0, w_{q-1}]$ or $w \in [w_{q-1}, \tau(w_{q-1})]$ then $d_{\beta,Y}([\psi^2], w) \leqslant M_2$. We do not need to modify χ^2 for the v_2-tracing property, however, because $[\psi^2]$ is chosen so that the Straightening Property holds, that is, $[[\psi^2], [\xi^2]]$ and $[w_0, [\psi^2]]$ do not come within a bounded distance of each other on ltd chunks — or equivalently (as we saw in Canonical Straightening) $[[\psi^2], [\xi^2]]$ and $[w_{n-r}, [\psi^2]]$ do not come within a bounded distance of each other on ltd chunks, for some (or equivalently any) sufficiently large r.

31.6. The F-decrease property. — We now need to show the F-decrease property, that is, as suggested in 31.2, we need to show that

$$\int K(\chi^2)^{-1}|p_2| \leqslant \exp(2F([\psi^0])\Big(1 + \tfrac{1}{2}L \int \theta_2^2|p_2|\Big) \tag{1}$$

where L is as in 8.3, which, by 8.3, then implies that

$$F([\psi^2]) \leqslant F([\psi^0]) = R.$$

The F-decreasing property is actually a negative upper bound on $F([\psi^2]) - F([\psi^0])$ and we shall actually obtain this. So we need to examine the set of z on which we do not have $K((\chi^2)^{-1})(z) \leqslant \exp 2F([\psi^0])$. This is obviously contained in the set where η is not the identity. From 31.5, this set is the union of two subsets U_1 and U_2.

The subset U_1 is contained in the union of sets $S_1(\omega_1)$ which are bounded neighbourhoods of $S(\omega_1, \varepsilon_0)$ where $u \times \omega_1$ is part of the visible upper boundary for Σ with $u < t_1 + t_2$- and in fact we shall then have $u < t_1$.

The subset U_2 is contained in the union of sets $S_1(\omega_1)$ which are bounded neighbourhoods of $S(\omega_1, \varepsilon_0)$ for $\omega_1 = \zeta_j$ for some $1 \leqslant j \leqslant r$, for some $r \geqslant 2$, where r and the ζ_j are as in 31.6. Again, in this case, we shall want to estimate θ_2 on a larger set, and to choose the modulus of annuli at $[\psi^2]$, $[\xi^2]$ homotopic to $\psi^2(\partial\zeta_j)$, $\xi^2(\partial\zeta_j)$ appropriately — but in some cases, as we shall see, we might have to leave modifications until $[\psi^i]$ for some $i > 2$.

In the case of both U_1 and U_2, we shall show that for any of the sets $S_1(\omega_1)$ contributing to U_1 or U_2, there is a larger set $T(\omega_1)$ such that

$$\int_{T(\omega_1)} K(\chi^2)^{-1}|p_2| \leqslant \exp(2R)\Big(1 + \tfrac{1}{2}L\int \theta_2^2|p_2|\Big) \tag{2}$$

To do this, we shall choose $T(\omega_1)$ containing $S_1(\omega_1)$ such that the p_2-measure $a(T(\omega_1), p_2)$ is much larger than $a(S_1(\omega_1), p_2)$, such that either $K(\chi^2)\exp(-2F([\psi^0])$ is bounded from 1 on a positive proportion of $T(\omega_1)$, or θ_2 is bounded from 0 on a positive proportion of $T(\omega_1)$. Let $q_{t_1,-,1}(z)dz^2$ and $q_{t_1,+,1}(z)dz^2$ denote the quadratic differentials at $[\xi^0]$ for $d_Z([\psi^1],[\xi^0])$ and $d_Z([\xi^0],[\xi^1])$ respectively, which extends the earlier notation $q_{t,1}(z)dz^2$.

So now we consider $S(\omega_1) \subset U_1$. If ω_1 is dominant at $[\xi^2]$ for $p_2(z)dz^2$, then we can change χ^2 on a positive p_2-measure-proportion to reduce the distortion substantially as follows. By 15.8, we can find points $[\chi^{s_1,1} \circ \psi^1]$ and $[\chi^{s_2,1} \circ \psi^1]$, with $s_1 < t)_1$ and $t_1 < s_2 < t_1 + t_2$, such that (if ω_1 is a gap) ω_1 is long ν-thick and dominant at $[[\chi^{s_2,1} \circ \psi^1]$ for $q_{s_2,1}(z)dz^2$ and long ν'-thick and dominant at $[\chi^{s_1,1} \circ \psi^1]$ for $q_{s_1,1}(z)dz^2$ and with

$$d_{\omega_1}([\chi^{s_1,1} \circ \psi^1], [\chi^{s_2,1} \circ \psi^1]) \leqslant C(\nu),$$

but with $s_2 - s_1$ much larger than $C(\nu)$, and much larger than $(t_1 + t_2) - s_2$. Then we can replace $(\chi^{s_2,s_1,1})^{-1}$ on a positive proportion of $(\chi^{t_1+t_1,s_2,1})^{-1}(S_1(\omega_1))$, to get distortion $O(C(\nu))$, thereby changing the definition of χ^2 and much reducing distortion and obtaining

$$\int_{S_1(\omega_1)} K((\chi^2)^{-1})|p_2| < e^{2R}. \tag{3}$$

The case when ω_1 is a loop is similar. If ω_1 is dominant for either of these quadratic differentials at $[\xi^0] = \tau(w_{n-1})$, we can employ a similar argument, finding s_1 and s_2 as above, using them to change the definition of χ^2 although this time $t_1 + t_2 - s_2$ is not necessarily small in comparison to $s_2 - s_1$. So we again obtain (3).

If ω_1 is not dominant for $q_{t_1,+,1}(z)dz^2$ at $[\xi^0]$ or $p_2(z)dz^2$ at $[\xi^2]$ then for $[\psi^0] = \tau(w_{n-1})$, $|\chi^{t,1}(\partial\omega_1)|_{q_{t,1}} = O(e^{-|t_1-t|}|\psi^0(\partial\omega_1)|_q)$ for $q = q_{t_1,s}$ and $t > t_1$. Then we take $T(\omega_1) = (\chi^{t_1+t_2,t_1,1})^{-1}S_2$ where S_2 is a bounded neighbourhood of the good position of $\psi^0(\partial\omega_1)$ with respect to $p_{t_1,+,1}(z)dz^2$. Then $T(\omega_1)$ has much larger diameter than $S_1(\omega_1)$ and hence also much larger p_2-measure, since ω_1 is not dominant for p_2. Then on T_1, either θ_2 is bounded from 0 or $K((\chi^{t_1+t_2,1}) \circ \chi^{t_1,1})^{-1}$ is boundedly less than $K((\chi^{t_1+t_2,1})^{-1})K((\chi^{t_1,1})^{-1})$, because the angle between the unstable foliations of $q_{t_1+t_2,1}(z)dz^2$ and $p_2(z)dz^2$ is bounded from 0 on a good proportion of $T(\omega_1)$.

Now we consider U_2. and consider sets $S(\zeta_j)$ for $1 \leqslant j \leqslant r$ for some $r \geqslant 2$ and sets ζ_j as in 31.5. Once we have chosen $|\xi^2(\partial\zeta_0)|$ and $|\psi^2(\partial\zeta_r)|$, we choose a function g of the form

$$g(t) = C\exp(|t - t_0|)$$

such that for points $t_3 < t_4$ with $t_4 - t_3 = (r-1)(t_2 - t_1)$,

$$g(t_3) = |\xi^2(\partial\zeta_0)|,$$

$$g(t_4) = |\psi^2(\partial\zeta_r)|.$$

For each j we would like to define

$$|\psi^2(\partial\zeta_j)| = |\xi^2(\partial\zeta_{j+1})| = g(t_3 + j(t_2 - t_1)),$$

and to define $\pi_{\zeta_j}([\psi^2])$ as dictated by Canonical Straightening. This is possible if

$$|\xi^2(\partial\zeta_j)| = o(|\xi^1(\partial\zeta_j)|), \tag{4}$$

which is what will give θ_2 bounded from 0 on a set $T(\omega_1)$ of much bigger p_2-measure that $S_1(\omega_1)$. This is certainly true if $|\xi^1(\partial\zeta_0)| = o(|\xi^0\partial\zeta_0|)$. We do obtain (4) for $j = 0$, if ζ_0 is not dominant for $q_{t_1,+,1}(z)dz^2$. We can then make $\pi_{\zeta_0}([\psi^2])$ and $\pi_{\zeta_1}([\xi^2])$ different from $\pi_{\zeta_0}([\psi^1])$ and $\pi_{\zeta_1}([\xi^1])$, and keep the increased distortion below

$$Le^{2F([\psi^0])}\int_{T(\zeta_1)} \theta_2^2|p_2|.$$

If ζ_0 is dominant for $q_{t_1,+,1}(z)dz^2$ then we do not attempt any change on this but take $[\psi^2]$ to be defined using just the changes made elsewhere, choose $[\psi^3] \in [[\psi^2], [\xi^2]]$ with $d_Z([\psi^2], [\psi^3])$ just large enough, and for $[\psi^3]$ either ζ_0 is not dominant for the corresponding quadratic differential or ζ_1 is dominant along a stretch of $[[\psi^3], [\xi^3]]$ and $S_1(\zeta_1)$ has moved into the U_1-set. So in finitely many steps we reach $[\psi^N]$, which is a canonical straightening.

BIBLIOGRAPHY

[Abi] W. Abikoff – *The real analytic theory of Teichmüller Space*, Lecture Notes in Math., vol. 820, Springer, 1980.

[A-B] L. Ahlfors & L. Bers – The Riemann Mapping Theorem for Variable Metrics, *Ann. of Math.* **72** (1960), p. 385–404.

[Bers] L. Bers – An Extremal Problem for quasi-conformal mappings and a Theorem of Thurston, *Acta Math.* **141** (1978), p. 73–98.

[Cas] A.J. Casson & S.A. Bleiler – *Automorphisms of Surfaces after Nielson and Thurston*, LMS Student Texts, vol. 9, Cambridge U.P., 1988.

[D-H1] A. Douady & J.H. Hubbard – Itération des polynômes quadratiques complexes, *C. R. Acad. Sci. Paris Sér. I Math.* **294** (1982), p. 123–126.

[D-H2] A. Douady & J.H. Hubbard – *Études dynamiques des polynômes complexes I, II*, Publications Mathématiques d'Orsay, 1985, avec la collaboration de P. Lavaurs, Tan Lei, P. Sentenac.

[D-H3] ______, A proof of Thurston's topological characterization of rational functions, *Acta Math.* **171** (1993), p. 263–297.

[Ear] C.J. Earle – The Teichmüller Distance is Differentiable, *Duke Math. J.* **44** (1977), p. 389–397.

[F-B] E. Fadell & J. van Buskirk – On the braid groups of E^2 and S^2, *Duke Math. J.* **29** (1962), p. 243–257.

[F-L-P] A. Fahti, F. Laudenbach & V. Poénaru – *Travaux de Thurston sur les Surfaces*, Astérisque, vol. 66-67, Société Mathématique de France, 1989.

[Gar1] F. Gardiner – On the variation of Teichmüller's metric, *Proc. Edinburgh Math. Soc. (2)* (1980), p. 143–152.

[Gar2] ———, *Teichmüller Theory and Quadratic Differentials*, Wiley-Interscience, New York, 1987.

[G-H] W.H. GOTTSCHALK & G.A. HEDLUND – *Topological Dynamics*, AMS Publications, vol. 36, American Mathematical Society, 1955.

[Gun] R.C. GUNNING – *Lectures on Riemann Surfaces*, Princeton University Press, 1966.

[L] S.V.F. LEVY – Critically finite rational maps, Ph.D. Thesis, Princeton University, 1985.

[Mag] W. MAGNUS – Uber Automorphismen von Fundementalengruppen berandeter Flachen, *Math. Ann.* **109** (1934), p. 617–646.

[Mi] J.W. MILNOR – On the Homotopy Class of a CW-Complex, *Trans. Amer. Math. Soc.* **90** (1959), p. 272–80.

[Mos] G.D. MOSTOW – *Strong Rigidity of Locally Symmetric Spaces*, Annals of Math. Studies, vol. 78, Princeton U.P., Princeton N.J., University of Tokyo Press, Tokyo, 1973.

[R1] M. REES – An alternative approach to the ergodic theory of measured foliations on sufaces, *Ergod. Th. & Dynam. Sys.* **1** (1981), p. 461–488.

[R2] ———, Components of degree two hyperbolic rational maps, *Invent. Math.* **100** (1990), p. 357–382.

[R3] ———, A Partial Description of the Parameter space of Rational Maps of Degree Two, I, *Acta Math.* **168** (1992), p. 11–87.

[R4] ———, A Partial Description of the Parameter space of Rational Maps of Degree Two, II, *Proc. London Math. Soc. (3)* **70** (1995), p. 644–690.

[Spa] E.H. SPANIER – *Algebraic Topology*, McGraw Hill, 1966.

[Sti] J. STIMSON – Degree two rational maps with a periodic critical point, Ph.D. Thesis, University of Liverpool, July 1993.

[TL] TAN LEI – Matings of Quadratic Polynomials, *Ergod. Th. & Dynam. Sys.* **12** (1992), p. 589–620.

[T1] W.P. THURSTON – The Geometry and Topology of Three-manifolds, Notes, Princeton University, 1978.

[T2] ———, On the Combinatorics of Iterated Rational Maps, Preprint, Princteton University and I.A.S., 1985.

[W] B. WITTNER – On the bifurcation loci of rational maps of degree two, Ph.D. Thesis, Cornell University, 1988.