Teichmüller Distance, Moduli Spaces of Surfaces, and Complex Dynamical Systems

Mary Rees

1 Introduction

1.1

This article arises, strange as it may seem, out of a project in complex dynamics. Complex dynamics is the study of the dynamics of rational maps of the Riemann sphere: that is, given a rational map f of $\hat{\mathbf{C}}$, one studies the behaviour of the sequence $\{f^n(z)\}$ for varying z. (Alternatively, people study the dynamics of entire or meromorphic functions of the complex plane.) As one varies a rational map through some natural parameter space - such as the space of all maps of some fixed degree $d > 1$ - dynamical behaviour varies widely - even wildly. An important aspect of study in complex dynamics is to understand variation of dynamics within parameter spaces. I shall attempt to give some indication of how this led me to develop a calculus of Teichmüller distance, in the hope of striking one or two chords. Two types of moduli spaces (at least) are involved here: the parameter spaces of rational maps (the original objects of study); the classical moduli space arising as the quotient of Teichmüller space by the mapping class group also plays a rôle. Teichmüller distance arises naturally in the classical problem of isotopy classification of surfaces homeomorphisms, as will be described later. The way in which Teichmüller distance arises in complex dynamics is related to this. .

1.2 Periodic Points

In any area of dynamics, periodic points are important, and this is especially true in complex dynamics. A point z is *periodic* under a map f, of period n, if $f^n(z) = z$, and n is the least integer > 0 for which this is true. Periodic points come in various types. A periodic point z of period n is *attractive* if the derivative $(f^n)'(z)$ is < 1 in modulus. Here, I am restricting to holomorphic maps. This definition is independent of choice of local coordinates, and of the point in the forward orbit of z, that is, for a periodic point z, z is attractive if and only if $f(z)$ is. Attractive periodic points influence the behaviour of nearby points: for any point w sufficiently near an attractive periodic point z, the sequence $f^m(w)$ converges to the periodic orbit of z. Further, attractive periodic points influence the behaviour of nearby maps: if z is attractive periodic under f then for g near

f, g has an attractive periodic point near z. This is a simple consequence of the Implicit Function Theorem, and, although elementary, is an important tool in dynamics in general.

1.3 Hyperbolic Maps

It can happen, for a rational map f, that for almost all points w (in a topological or measure-theoretic sense) the sequence $f^m(w)$ converges to one of finitely many attractive periodic orbits. This happens, for example, if such convergence holds for all *critical* points w. A point w is *critical* for a (holomorphic) map f if $f'(w) = 0$. This definition is independent of choice of local coordinates. Since there are only finitely many critical points (in fact, at most $2d-2$ for a rational map of degree d), it is possible - at least in theory, and often in practice too - to check whether their orbits converge to atttractive peirodic orbits. If this does happen for all critical points of a map f then f is *hyperbolic.* This is a very important concept in dynamical systems. The formal definition of hyperbolic is somewhat different from the above in general, but the general consequences are much the same. For hyperbolic maps, dynamics can be analysed: as stated above, for a hyperbolic rational map, a generic orbit converges to an attractive periodic orbit. This is considered rather dull behaviour, and the set of points with orbits not converging to some attractive periodic orbit is called the *Julia set.* Dynamics on this set can be analysed thoroughly, using standard techniques available in dynamical systems.

1.4 Hyperbolicity and Stability

In summary, there are two (dynamical) kinds of rational maps: hyperbolic and nonhyperbolic. The hyperbolic maps can be analysed. They are *J-stable* - stable is generally a good word in mathematics, but apparently with many different meanings. In dynamics, a map f is stable if all g near f are conjugate to f, that is, of the form $\phi \circ f \circ \phi^{-1}$ for some homeomorphism ϕ. For rational maps, J- stable means that such a conjugacy holds in a neighbourhood of the Julia set. Thus, each hyperbolic map is in an open connected set of maps which are all hyperbolic and have the same dynamics, at least on their Julia sets (and with very minor variations elsewhere). It is conjectured that hyperbolic maps are generic - open and dense - in natural parameter spaces of rational maps, for example the space of all maps of some fixed degree $d > 1$. It is known that stable maps are dense [7]. So the conjecture is that all stable maps are hyperbolic: a result which is true in some other categories of dynamical systems. (However, in other categories, stable systems are not usually dense.) In contrast to some other situations in mathematics, there are infinitely many hyperbolic (and hence stable) components in any reasonable parameter space of rational maps, so even if the conjecture holds and hyperbolic maps are dense, the variation of dynamcis in the parameter space will necessarily be complicated.

1.5 Versions of homotopy equivalence for maps

A rational map is obviously hyperbolic if every critical point w is periodic, because then w is attractive with $(f^n)'(w) = 0$ (by the Chain Rule). Many hyperbolic components will contain such a map, which is called *critically finite*. The total dynamics of such a map is determined by its critical orbits, or, more precisely, by the appropriate type of homotopy class of the map with respect to these orbits. This type of homotopy equivalence for critically finite branched coverings is known as *Thurston equivalence*. If f is a branched covering then the *postcritical set* $P(f)$ is the set $\{f^n(c) : n > 0, c \text{ critical}\}$. Then critically finite branched coverings f_0 and f_1 are *Thurston equivalent* if there is a homotopy f_t through critically finite branched coverings from f_0 to f_1 such that $P(f_t)$ varies isotopically. Thus Thurston equivalence is stronger than homotopy equivalence and can be regarded as the appropriate analogue of isotopy for homeomorphisms: we recall that two homeomorphisms are isotopic if they are homotopic through homeomorphisms.

1.6 The importance of homotopy-type information

Another common theme in dynamics is that isotopy, or sometimes even homotopy, determines a map up to some kind of semiconjugacy, for certain homotopy (or isotopy) classes. There are many results of this type. For example [5], any continuous map f of a torus (of any dimension) which is homotopic to a *hyperbolic* toral automorphism g is semiconjugate to g, that is, there is a continuous map φ such that $\varphi \circ f = g \circ \varphi$. There is a result of this type [6] for other surface homeomophisms (with a somewhat weaker semiconjugacy statement), and another such for critically finite branched coverings ([8] , 4.1). This is one rough reason for dynamicists to be interested in the problems of: classification of surface homeomorphisms up to isotopy; classification of critically finite branched coverings up to Thurston equivalence; and so on.

1.7 Classical Problems

The problem of isotopy classification of surface homeomorphisms has been well worked over during the past seventy years or so. There have been varying approaches to the problem by (among others) Nielsen, Thurston [4], and (following Thurston to some extent) Bers [1]. It is Bers' approach that I shall be highlighting shortly. First, I want to try and indicate certain aspects of all the proofs which seem to me to give important clues as to how investigations into dynamical structure should develop. All proofs point towards a "best map", or at least, a small class of best maps, within an isotopy, or Thurston equivalence, class. These best maps have strong geometric structure and strong dynamical properties. For example, the best map in a Thurston equivalence class is often a rational map, usually unique up to Möbius conjugacy. The family of best maps in an isotopy class of surface homemorphism is, for some choices of isotopy class (the *pseudo-Anosov* classes), a family of maps each of which preserves two transverse measured foliations (with singularities) on the surface, expanding leaves of one foliation and contracting leaves of the other. The other important point about these proofs is that they turn out to give information about the topology of the isotopy class of homeomorphisms (or Thurston equivalence class of critically finite branched coverings) itself. Let us consider the case of isotopy classes of surface homeomorphisms for the moment. Any isotopy class itself (at least for a surface of negative Euler characteristic) is easily seen to be contractible. However, the proof of isotopy classification gives, further, the topology of the connected components of the space of pair (f, S), where S is homeomorphic to a fixed compact surface S_0 and f is a homeomorphism of S. A component of this space corresponding to a pseudo-Anosov isotopy class is homotopy equivalent to the circle. Now let us consider critically finite branched coverings. Thurston ([10], [2]) showed that certain Thurston equivalence clases of critically finite branched coverings (satisfying a certain condition concerning sets of disjoint simple closed loops) contained a unique rational map up to Möbius conjugation. The proof, almost incidentally, also showed that such Thurston equivalence classes (after quotienting by Möbius conjugation) were contractible - in this case, a nontrivial fact.

1.8 Generalizing the Classical Problems

I shall shortly come to some concrete definitions, statements, and even proofs. However, the point of the work I shall attempt to describe is not to reprove the results stated above but to carry out other investigations which were suggested by the above. We have seen that the dynamics of a rational map is heavily influenced by what happens to the critical orbits - and, in fact, critically finite maps are relatively easy to analyse. It is therefore natural to consider slices in parameters space in which various critical points are constrained to have finite forward orbits. Such slices will intersect many hyperbolic components, for example, and will contain many critically finite maps. By constraining all but one of the critical points, one obtains slices of complex dimension one, which are good for initial consideration. Such spaces of rational maps are natural subspaces of spaces of branched coverings with specific dynamics on certain finite sets, which include some critical points. These can be regarded as "one dimension up" from (for example) the isotopy class of a surface homeomorphism, or of a homeomorphism of $(\hat{\mathbf{C}}, A)$, where $A \subset \hat{\mathbf{C}}$ is a finite set. I have tried to show the evidence that investigation of dynamical structure should go hand-in-hand with a study with the space of maps itself. It seems likely that the structure of a moduli space often reflects structure of the objects in the modui space in some way. Also, (perhaps less clear in the above) although one might want to study a relatively small parameter space, in order to study it comprehensively, it might be necessary to consider a somewhat larger space. This is because of the fact that dynamical structure is often implied by very simple homotopy-type information. It is also because information about varying dynamics within a parameter space is inextricably linked to information about the structure - especially topological and geometric structure - of the parameters space itself, including the toplogical structure of inclusions into certain much larger spaces (but which are nevertheless quite simply defined), topologically. The larger spaces can also be regarded as moduli spaces, although not in the strict mathematical sense described in Frances Kirwan's talk (of this meeting), because they are not in general algebraic varieties. I am afraid it is not possible to justify this completely here. However, there might possibly be an analogy with other situations in mathematics - in particular, involving moduli spaces - such as blow-ups of singularities. Some of the enlargement (but not all) of rational map parameter spaces mentioned above is indeed connected with singularities in these spaces.

2 Calculus

2.1 The programme from now on

For the rest of this article, I shall restrict to developing some calculus of Teichmüller distance, and applying it to an adaptation of Bers' proof of isotopy classification of surface homeomorphisms. It is not possible to give much detail - for further detail see [9], especially Chapters 10-13. I think this calculus is of interest in its own right. As we shall see, I have developed it only in a restricted case - for marked spheres rather than even general finite type surfaces. I do believe that it can be developed for general finite type surfaces though, a belief that was strengthened by some discussion at the EWM meeting, and subsequently with a colleague. The calculus is not that easy to use - nor formidably difficult. It is certainly totally unnecessary to burden Bers' elegant proof with it. However, the isotopy classification is useful as an illustration, and the calculus seems to be essential to the problem for which it was developed.

This specific problem (for which the calculus was developed) is part of the general project - indicated above - of understanding variation of dynamics in parameter spaces of rational maps. The specific problem would take a very long time to state, but is, in fact, quite closely related to the problem of isotopy classification of surface homeomorphisms. For further details, see [9].

2.2 Teichmüller Space

Let S_0 be a compact surface (of two real dimensions) and $Y_0 \subset S_0$ finite. Then homeomorphisms $\varphi_0, \varphi_1 : (S_0, Y_0) \to (S, Y)$ are *isotopic* if there is a continuous family φ_t of homeomorphisms for $t \in [0, 1]$.

Now let S be a Riemann surface (or a two-dimension hyperbolic manifold). Then $\phi : (S_0, Y_0) \to (S, Y)$ and $\varphi' : (S_0, Y_0) \to (S', Y')$ are *equivalent* (as elements of Teichmüller space) if there is a biholomorphic map (or hyperbolic isometry) $\tau : (S, Y) \to (S', Y')$ such that φ' and $\tau \circ \varphi$ are isotopic. If $[\varphi]$ denotes the equivalence class of φ then the *Teichmüller space of* (S_0, Y_0), which we denote by $\mathcal{T}(S_0, Y_0)$, is

$$\{[\phi] : \varphi : (S_0, Y_0) \to (S, Y) \text{ an orientation} - \text{preserving homeomorphism}\}.$$

Roughly speaking, $\mathcal{T}(S_0, Y_0)$ is the space of framed geometric structures homeomorphic to (S_0, Y_0).

2.3 Example

Let $S_0 = \hat{\mathbf{C}}$. Then $\mathcal{T}(S_0, Y_0)$ is a point if $\#(Y_0) \leq 3$, because given any two sets of three points in $\hat{\mathbf{C}}$, there is a Möbius transformation taking the first set to the second. If $\#(Y_0) = n \geq 3$ then, applying a Möbius transformation, we can assume that $0, 1, \infty \in Y_0$. Then for any $[\varphi] \in \mathcal{T}(S_0, Y_0)$ we can choose φ in its equivalence class to fix $0, 1, \infty$, and we do this from now on. From now on, we write

$$Y_0 = \{0, 1, \infty\} \cup \{y_i : 1 \leq i \leq n-3\}.$$

Then the map

$$[\phi] \mapsto (\phi(y_i)) \in \mathbf{C}^{n-3}$$

is a local homeomorphism, and gives $\mathcal{T}(S_0, Y_0)$ the structure of a complex manifold. In fact, it can be shown that $\mathcal{T}(S_0, Y_0)$ is homeomorphic to $\mathbf{R}^{2n-6}$. If $\#(Y_0) = 4$ then $\mathcal{T}(S_0, Y_0)$ is biholomorphic to the unit disc.

2.4 The Mapping Class Group

Teichmüller space was used by Thurston (and Bers) to analyse isotopy classes of homeomorphisms $\psi : (S_0, Y_0) \to (S_0, Y_0)$. The *mapping class group* $M(S_0, Y_0)$ is the group of orientation-preserving homeomorphisms modulo isotopy. The mapping class group acts on $\mathcal{T}(S_0, Y_0)$ on the right by

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

The quotient of Teichmüller space by the mapping class group is often called the *moduli space of the surface.*

2.5 Teichmüller Distance

Let $\chi : U \to \mathbf{C}$ be a local diffeomorphism, where $U \subset \mathbf{C}$ is open. Then one can define the (pointwise) *distortion* $K(\chi)(z)$ at a point $z \in U$ by

$$K(\chi)(z) = \sqrt{\frac{|\lambda_1|}{|\lambda_2|}},$$

where $\lambda_1 \geq \lambda_2 > 0$ are the eigenvalues of $D\chi_z^t D\chi_z$, where $D\chi_z$ denotes the derivative of χ at z (considered as a 2×2 matrix). Note that $K(\chi)(z) = 1$ if χ is holomorphic near z. We can define $K(\chi)(z)$ similarly for $z \in S_1$ if $\chi : S_1 \to S_2$ is a local diffeomorphism and S_1, S_2 are any Riemann surfaces: the definition is independent of the choice of local coordinates. Then we can define

$$||\chi||_{qc} = ||K(\chi)||_\infty,$$

where $\|.\|_\infty$ denotes the L_∞ norm. This norm is finite if χ is *quasi-conformal*, which is true, for example, if χ is a C^1 diffeomorphism between compact spaces. It is also finite more generally: since we are taking the L_∞ norm, $\|\chi\|_{qc}$ might be finite even if χ is nondifferentiable on a set of zero measure.

The *Teichmüller distance*

$$d_{\mathcal{T}} : \mathcal{T} \times \mathcal{T} \to (0, \infty)$$

is defined by

$$d_{\mathcal{T}}([\varphi_1], [\varphi_2]) = \frac{1}{2} \inf\{\log \|\chi\|_{qc} : [\chi \circ \varphi_1] = [\varphi_2]\}.$$

The Teichmüller distance coincides with $\frac{1}{2}$ the Poincaré distance if $\mathcal{T} = \mathcal{T}(\hat{\mathbf{C}}, Y_0)$ and $\#(Y_0) = 4$, so that $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$ is the unit disc. The Teichmüller distance is a genuine metric: if χ is biholomorphic then $\log \|\chi\|_{qc} = 0$. Moreover, the action of $M(S_0, Y_0)$ on $(T)(S_0, Y_0)$ preserves Teichmüller distance, that is,

$$d_{\mathcal{T}}([\varphi_1].[\psi], [\varphi_2].[\psi]) = d_{\mathcal{T}}([\varphi_1], [\varphi_2])$$

for all $[\varphi_1]$, $[\varphi_2] \in \mathcal{T}(S_0, Y_0)$ and for all $[\psi] \in M(S_0, Y_0)$. This ensures that the Teichmüller distance metric descends to a metric on moduli space. The infimum in the definition of $d_{\mathcal{T}}$ is attained for a unique χ. This χ is defined by the following properties. Except at finitely many points on $\varphi_1(S_0)$ and $\varphi_2(S_0)$, there are local coordinates $x + iy$ on $\varphi_1(S_0)$ and $\varphi_2(S_0)$ with respect to which χ has the formula

$$\chi(x + iy) = \lambda x + i\frac{1}{\lambda}y$$

for some $\lambda \geq 1$. We then have

$$\log \lambda = \frac{1}{2} \log \|\chi\|_{qc} = d_{\mathcal{T}}([\varphi_1], [\varphi_2]).$$

Moreover, the local coordinates (with singularities) are given by *quadratic differentials* on $\varphi_1(S_0)$, $\varphi_2(S_0)$. A quadratic differential on $(\varphi_j(S_0), \varphi_j(Y_0))$ has the following property with respect to (any system of nonsingular) charts (U, ξ_U), (V, ξ_V) on $\varphi_j(S_0)$, that is, with U, $V \subset \varphi_j(S_0)$. Suppose that $U \cap V \neq \emptyset$ and that the quadratic differential is given in $\xi_U(U)$, $\xi_V(V)$ by $q_U(z)dz^2$, $q_V(z)dz^2$. Then q_U, q_V are meromorphic with at worst simple poles, with these occurring only (at most) at points of $\xi_U \circ \varphi_j(Y_0)$, $\xi_V \circ \varphi_j(Y_0)$, and on $\xi_U(U \cap V)$ we have

$$q_U(z) = ((\xi_V \circ \xi_U^{-1})'(z))^2.q_V \circ \xi_V \circ \xi_U^{-1}(z).$$

The singular local coordinates with respect to which χ has its special form are then given on U by

$$(x + iy)(\xi_U^{-1}(z)) = \int_{z_0}^{z} \sqrt{q_U(\zeta)}d\zeta.$$

Up to addition of a constant and plus or minus sign, these local coordinates are independent of choice of chart U: the way the quadratic differential transforms under change of chart, and the change of variable formula for integrals, ensure that.

Now let $S_0 = \hat{\mathbf{C}} = \varphi_j(S_0)$ and consider the chart $\mathbf{C}$. Then a quadratic differential on $(\hat{\mathbf{C}}, \varphi(Y_0))$ is given in this chart by $q(z)dz^2$, where q is a rational function wiith at worst simple poles, occurring at most at points of $\varphi(Y_0)$, and with at least three more poles than zeros. This last condition assumes that ∞ is a point of $\varphi(Y_0)$: otherwise we need at least four more poles than zeros. It arises from considering the local coordinate $1/z$ at ∞: we need $q(1/z)z^{-4}$ to have at most a simple pole at 0.

2.6 Bers' Approach to Isotopy Classification of Surface Homeomorphisms

Let an isotopy class $[\varphi]$ in $M(S_0, Y_0)$ be given. Then consider the map

$$F : \mathcal{T}(S_0, Y_0) \to \mathbf{R}_+$$

given by

$$F([\varphi]) = d_{\mathcal{T}}([\varphi].[\psi], [\varphi]).$$

Consider where the infimum of this function is attained. There are three possibilities, corresponding to three different types of isotopy classes.

1. The minimum value 0 is attained at a point in $\mathcal{T}$. In that case, it is attained at a unique point in $\mathcal{T}$ at a point $[\varphi]$, and $\varphi \circ \psi \circ \varphi^{-1}$ is isotopic to a biholomorphism (or hyperbolic isometry) of $(\varphi(S_0), \varphi(Y_0))$, which is necessarily of finite order.

2. The infimum is 0, but this is not attained at any point of $\mathcal{T}$. This is the *reducible* case. In this case, it is possible to show that the infimum is only attained by going to infinity in $\mathcal{T}$ in specific directions. We omit the details.

3. The infimum is strictly positive, and is attained uniquely on a curve of one real dimension in $\mathcal{T}$ which is invariant under the action of $[\psi]$: it has to be, since right action by $[\psi]$ preserves Teichmüller distance. This curve is known as a *geodesic* (and distance between any two points on the geodesic is indeed attined uniquely along paths in the geodesic). Such a minimizing geodesic is uniquely determined by any fixed point $[\varphi]$ on it, and a unique quadratic differential (up to multiplication by a nonzero real number) for $(\varphi(S_0), \varphi(Y_0))$. If we take the singular local coordinates $x + iy$ on $\varphi(S_0)$ given by this quadratic differential, then all other points on the geodesic are $[\chi_\lambda \circ \varphi]$ $(\lambda > 0)$, where, if $S'_\lambda = \chi_\lambda \circ \varphi(S_0)$ then $\chi_\lambda : S'_0 = \varphi(S_0) \to S'_\lambda$ is given in singular local coordinates by

$$\chi_\lambda(x + iy) = \lambda x + i\frac{y}{\lambda}.$$

This is known as the *pseudo-Anosov* case. If $[\varphi]$ is a point on the geodesic, then so is $[\varphi \circ \psi]$, and it is apparent from the characterization above of points on the geodesic that $\varphi \circ \psi \circ \varphi^{-1}$ has very special properties. In fact, the singular local coordinates $x+iy$ give the two transverse measured foliations (with singularities) mentioned earlier in this article. Leaves of these foliations are given locally by the curves $y=$ constant and $x=$ constant.

Bers [1] was able to do his analysis of the function F without any reference to its derivative, which means without any reference to the derivative of $d_{\mathcal{T}}$. (Note that, with respect to the natural local coordinates on $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$ that we described earlier, the right action of $[\psi]$ is given by the identity map.) However, at the same time that Bers was carrying out his analysis, Earle [3] was developing a formula for the derivative of $d_{\mathcal{T}}$. I shall not give the precise formula. It is not difficult, but involves the introduction of Beltrami differentials associated to elements of Teichmüller space. Anyway, the nub of the formula is that the derivative of $d_{\mathcal{T}}([\varphi_1],[\varphi_2])$ at $([\varphi_1],[\varphi_2])$ with $[\varphi_1]\neq[\varphi_2]$ is given by the pair of quadratic differentials at $(\varphi_1(S_0),\varphi_1(Y_0))$, $(\varphi_2(S_0),\varphi_2(Y_0))$ used to define the map χ with $[\chi\circ\varphi_1]=[\varphi_2]$ which minimises distortion. I shall give a formula below in the special case of $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$. Of course, when one is thinking of infima of functions - which often turn out to be minima - then one thinks of the second derivative. If F is twice differentiable at any minimum value, the first derivative must be zero at such a point, and the second derivative must be positive at such a point. This is indeed the case, at least in the cases for which I have been able to compute second derivative of distance so far, that is, for marked spheres.

2.7 The derivative formula for $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$

This particularly simple formula (which is a special case of Earle's) is possible because of the simple local coordinates on $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$. Let $\#(Y_0)=n$, $n\geq 3$. Let $[\varphi]$, $[\varphi']\in\mathcal{T}(\hat{\mathbf{C}}, Y_0)$, so that φ, φ' are orientation-preserving homeomorphisms of $\hat{\mathbf{C}}$, and assume without loss of generality that they fix 0, 1, ∞. Then if $h=(h_i)\in\mathbf{C}^{n-3}$ is small enough given $[\varphi]$, the element $[\varphi]+h$ of $\mathcal{T}(\hat{\mathbf{C}}, Y_0)$ is well-defined by taking this to be the isotopy class of the homeomorphism near φ which fixes 0, 1, ∞ and sends y_i to $\varphi(y_i)+h_i$. Let $q(z)dz^2$ and $-p(z)dz^2$ be the quadratic differentials at $[\varphi]$, $[\varphi']$ for $d_{\mathcal{T}}([\varphi],[\varphi'])$. Thus, $p(z)dz^2$ is the *stretch* of $q(z)dz^2$, that is, with respect to the local coordinates given by $q(z)dz^2$, $p(z)dz^2$, the homeomorphism minimizing distortion is in the form

$$x+iy\mapsto x+i\frac{y}{\lambda}$$

where $\log\lambda=d_{\mathcal{T}}([\varphi],[\varphi'])$. We recall from the definition of quadratic differentials that q and p have at most simple poles. Then the formula for the first derivative in this special case ([9] Chapter 9) is given by

$$d_{\mathcal{T}}([\varphi]+h,[\varphi']+h')=d_{\mathcal{T}}([\varphi],[\varphi'])+2\pi\sum_{i=1}^{n-3}\Re(\mathrm{Res}(q,\varphi(y_i))h_i-\mathrm{Res}(p,\varphi'(y_i)h_i'))$$

$$+o(h)+o(h').$$

2.8 Quadratic Differentials and Hyperelliptic Curves

We continue with the conventions on Y_0, φ established above. We recall that if $q(z)dz^2$ is a quadratic differential at $(\hat{\mathbf{C}}, \varphi(Y_0))$ then q is a rational function with at most n simple poles, at most occurring at the points $\varphi(Y_0)$, and at most 3 less zeros than poles, up to multiplicity. We suppose for the moment that all zeros are simple, that all points of $\varphi(y_i)$ are simple poles of q, and that q has exactly three more poles than zeros. (This is essentially the definition of ∞ being a simple pole of q.) We consider the Riemann surface

$$S_q = \overline{\{(z, w) \in \mathbf{C}^2 : q(z) = w^2\}}.$$

Here, we use $\overline{\}}$ to denote the following possibly nonstandard closure: without closure, the set described above is biholomorphically a compact surface minus finitely many points, corresponding to taking z or $w = \infty$. We define the closure to be the union with these finitely many points, thus giving a compact Riemann surface S_q, such that $\pi : S_q \to \hat{\mathbf{C}}$ given by $\pi(z, w) = z$ is a branched double cover, branched over the zeros and poles of q - including ∞. Thus, S_q is a compact surface of genus $n - 3$ which I shall call a *hyperelliptic curve*. (This is the standard definition, modulo possible quibbles about the closure operation.)

Then $wdz = \pi^*(\sqrt{q(z)}dz)$ is a holomorphic 1-form on S_q. Some checking in local coordinates near points where $w = 0$ or ∞, or where $z = \infty$, is necessary to confirm this, but it is so. The standard theory tells us that the holomorphic 1-forms on S_q form a vector space of complex dimension $n - 3$. The general formula for a holomorphic 1-form on S_q is

$$r(z)wdz = \pi^*(r(z)\sqrt{q(z)}dz),$$

where r is a rational function whose denominator is the numerator of q and the numerator is a polynomial of degree $\leq n - 3$. The real and imaginary parts of holomorphic 1-forms are *harmonic 1-forms*. The real and imaginary parts of a complex valued form given locally by $(a + ib)(dx + idy)$ (where a and b are functions) are $adx - bdy$ and $bdy + adx$. A harmonic 1-form is *closed* - that is, $\partial a/\partial y = \partial b/\partial x$ - and the *harmonic* condition is $\partial a/\partial x + \partial b/\partial y = 0$. Let ω be a holomorphic or harmonic 1-form on S_q, and γ a loop on S_q. Then

$$\int_\gamma \omega$$

depends only on the homology class of γ. Thus each holomorphic 1-form defines an element of $H^1(S_q, \mathbf{C})$ and each harmonic 1-form defines an element of $H^1(S_q, \mathbf{R})$. In fact, the space of harmonic 1-forms is isomorphic to $H^1(S_q, \mathbf{R})$.

2.9 Connection with the Second Derivative of Teichmüller Distance

In this section I shall indicate very briefly how to use harmonic 1-forms on S_q, S_p to obtain the second derivative of the Teichmüller distance function. See [9] for further details, and also on the function F. (For reasons of space, I shall not attempt to indicate, here, the precise form of $D^2 d_{\mathcal{T}}$, nor why certain terms in this are positive, nor why $D^2 F$ is positive.)

We recall that if $\log\lambda = d_{\mathcal{T}}([\varphi],[\varphi'])$, then λ is the distortion of the best map $\chi : (\hat{\mathbf{C}}, \varphi(Y_0)) \to (\hat{\mathbf{C}}, \varphi'(Y_0))$ with $[\chi \circ \varphi] = [\varphi']$. The key point is that λ can be characterised in terms of a linear map between $H^1(S_q, \mathbf{R})$ and $H^1(S_p, \mathbf{R})$, as follows. Note that the map $\chi : \hat{\mathbf{C}} \to \hat{\mathbf{C}}$ lifts to a map $\tilde{\chi} : S_q \to S_p$. Then $\tilde{\chi}$ induces a map from $H_1(S_q, \mathbf{Z})$ to $H_1(S_p, \mathbf{Z})$. Then the following holds, where we identify complex numbers with 2-dimensional real column vectors, so that left multiplication by real 2×2 matrices makes sense. For all $\gamma \in H_1(S_q, \mathbf{Z})$,

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \int_\gamma \pi_q^*(\sqrt{q(z)}dz) = \int_{\widetilde{\chi}(\gamma)} \pi_p^*(\sqrt{p(z)}dz).$$

That is, in terms of $H^1(S_q, \mathbf{R})$, $H^1(S_p, \mathbf{R})$,

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \pi_q^*(\sqrt{q(z)}dz) = \widetilde{\chi}^* \pi_p^*(\sqrt{p(z)})dz. \tag{1}$$

Now suppose that $[\varphi]$ and $[\varphi']$ are changed to $[\varphi]+h$ and $[\varphi']+h'$, and we want to find the perturbations q_1 and p_1 which satisfy the equation corresponding to (1). Write

$$\varphi(y_i) = b_i, \text{ and } a_i = \mathrm{Res}(q, b_i).$$

Then q_1 has possible poles at 0, 1, ∞ and $b_i + h_i$, $1 \le i \le n-3$, where the residues are k_i, $1 \le i \le n-3$. Then

$$q(z) = \sum_{i=1}^{n-3} \frac{b_j(b_j-1)a_j}{z(z-1)(z-b_j)}.$$

Then expanding, $\sqrt{q_1}$ this gives

$$\sqrt{q_1(z)}dz = \sqrt{q(z)}dz + \frac{1}{2}\sum_{j=1}^{n-3} \frac{b_j(b_j-1)k_j}{z(z-1)(z-b_j)\sqrt{q}}dz +$$

$$\sum_{j=1}^{n-3} \frac{b_j(b_j-1)a_j h_j}{z(z-1)(z-b_j)^2\sqrt{q(z)}}dz + o(h) + o(k).$$

Then we note that

$$\pi_q^*\left(\frac{b_j(b_j-1)}{z(z-1)z-b_j\sqrt{q(z)}}dz\right) \tag{2}$$

form a basis of holomorphic 1-forms on $H^1(S_q, \mathbf{C})$. The real and imaginary parts then form a basis of harmonic 1-forms in $H^1(S_q, \mathbf{R})$. Meanwhile, the

$$\pi_q^*\left(\frac{b_j(b_j-1)a_j}{z(z-1)(z-b_j)^2\sqrt{q(z)}}dz\right)$$

are meromorphic 1-forms *of the first kind*, that is, the residues at singularities are zero. Thus these forms, also, define elements of $H^1(S_q, \mathbf{C})$. In fact, each of these 1-forms has exactly one singularity, which is a double pole. It is possible to write these forms in terms of the basis of harmonic 1-forms given by (2). The coefficient matrix is a sum of products of matrices of improper intergrals and inverses of such. This looks like a very classical calculation, apart from the nonstandard form of hyperelliptic curve (that is, using q rather than a polynomial). Nevertheless, details can be found in [9] Chapters 13-14.

In exactly the same way, an expresson can be found for the perturbation $\pi_p^*(\sqrt{p_1(z)}dz)$ of $\pi_p^*(\sqrt{p(z)}dz)$. Then the perturbation of the equation (1) can be written down to first order, and the vectors k and k', and the perturbation of the distance, can be computed to first order in terms of h and h'. This, then, gives the formula for the second derivative of Teichmüller distance, at least generically. The formula extends continuously at other points. For further details see [9] Chapters 10-13.

The formula for the second derivative of $d_{\mathcal{T}}$ then gives a formula for the second derivative of the function F of 2.6. It is possible to show that, as expected, D^2F is positive, although not, in general, positive definite. The whole procedure can be regarded as an exercise in Morse theory, that is, using level and critical sets of a function to analyze the topology of a space.

References

[1] Bers, L.: An Extremal Problem for quasi-conformal mappings and a Theorem of Thurston. *Acta Math. 141* (1978) 73-98.

[2] Douady, A. and Hubbard, J.H.: A proof of Thurston's topological classification of rational functions. *Acta Math. 171* (1993) 263-297.

[3] Earle, C.: The Teichmüller Distance is Differentiable. *Duke Math. J. 44* (1977) 389-397.

[4] Fahti, A., Laudenbach, F., Poënaru, V.: Travaux de Thurston sur les Surfaces. *Astérisque 66-67* (1989).

[5] Franks, J.:Anosov Diffeomorphisms.*Proc. Symposia in Pure Math. 14* (1968), 61-93.

[6] Handel, M: Global Shadowing of pseudo-Anosov homeomorphisms. *Ergod. Theor. Dynam. Sys. 5* (1985), 373-377.

[7] Mañé, R., Sad., P., and Sullivan, D.: On the Dynamics of Rational Maps. *Ann. Sci. Ecole. Norm. Sup. 4e, Série 16* (1983) 193-217.

[8] Rees, M.: A Partial Description of the Parameter Space of Rational Maps of Degree Two: Part 1.*Acta Math. 168* (1992) 11-87.

[9] Rees, M.: Views of Parameter Space: Topographer and Resident. *To be submitted to Astérisque in 1999.*

[10] Thurston, W.P.: On the Combinatorics of Iterated Rational Maps. *Preprint, Princeton University and I.A.S,* (1985).