J. DIFFERENTIAL GEOMETRY
52 (1999) 407-452

SYSTOLES AND TOPOLOGICAL MORSE FUNCTIONS FOR RIEMANN SURFACES

PAUL SCHMUTZ SCHALLER

1. Introduction

Let $\mathcal{M}$ be a (g,n)-surface, this is a Riemann surface of genus g with n cusps, equipped with a complete metric of constant curvature -1. Let $T(g,n)$ be the Teichmüller space of $\mathcal{M}$, and $\Gamma(g,n)$ the corresponding mapping class group.

The aim of this paper is the construction of topological Morse functions on $T(g,n)$ which are invariant with respect to $\Gamma(g,n)$ and proper on the moduli space $T(g,n)/\Gamma(g,n)$. Such functions are useful for the study of the topology of the mapping class groups and the moduli spaces. Until now, no such functions were known; the functions of this paper are the first examples.

The conception of topological Morse functions has been introduced by Morse [18] himself. It generalizes the traditionally smooth Morse functions. For example, the function $f : \mathsf{R} \to \mathsf{R}$, $f(x) = |x|$, is a topological Morse function. In order to be more precise, let X be a connected topological manifold and

$$f : X \longrightarrow \mathsf{R}_+$$

a continuous function. Then $x_0 \in X$ is an ordinary point of f if there exists an open neighbourhood U of x_0 in X and a homeomorphic parametrization of U by $N = \dim X$ parameters such that one of them is f. Otherwise, x_0 is a critical point of f. In the latter case, x_0 is non-degenerate if there exists an open neighbourhood U of x_0 in X, a

Received November 11, 1997, and, in revised form, September 16, 1999.

homeomorphic parametrization of U by parameters $y_1, \dots, y_N$, and an integer J ($0 \le J \le N$) such that, for all $x \in U$,

$$f(x) - f(x_0) = \sum_{i=1}^{J} y_i^2 - \sum_{i=J+1}^{N} y_i^2 .$$

If all critical points of f are non-degenerate, then f is a topological Morse function.

In order to introduce our Morse functions, some preparation is needed. Let $F = \{u_1, \dots, u_m\}$ be a set of m marked, simple closed geodesics of $\mathcal{M}$ which fills up. For $M \in T(g,n)$ let $Tan(M)$ be the tangent space of $T(g,n)$ in M. For $\xi \in Tan(M)$ and $u \in F$ let $\xi(u) \in \mathsf{R}$ be the derivation, induced by ξ, of the length function $L(u)$. Define the *set of minima* of F as

$$\begin{aligned} Min(F) = \{M \in T(g,n) : \forall \xi \in Tan(M) \text{ either } \xi(u) = 0,\ \forall u \in F,& \\ \text{or } \exists u, v \in F,\ \xi(u) > 0,\ \xi(v) < 0\}&. \end{aligned}$$

For $M \in T(g,n)$ define the vector space

$$W_F(M) = \{ \, (\xi(u_1), \dots, \xi(u_m)) \in \mathsf{R}^m : \xi \in Tan(M) \, \} .$$

Then $M \in Min(F)$ is called *F-regular* if $\dim W_F$ is locally constant on $Min(F)$ in M.

Let now

$$syst : T(g,n) \longrightarrow \mathsf{R}_+$$

be the function which associates to $M \in T(g,n)$ the length of a systole, a shortest simple closed geodesic; $syst$ is continuous on $T(g,n)$, but has plenty of corners. Let $S(M)$ be the set of systoles of M.

Theorem A. *The function* syst *is a topological Morse function on $T(g,n)$ if the following conditions hold:*

(i) For all $M \in T(g,n)$ with $M \in Min(F)$ (where $F = S(M)$), M is F-regular.

(ii) For all $M \in T(g,n)$ with $M \in \partial Min(F)$ (where $F = S(M)$), M is strongly F-regular (see Section 2 for definitions).

If condition (ii) holds, then $M \in T(g,n)$ is a critical point of syst *if and only if $M \in Min(S(M))$.*

How strong are these restrictions? Note first (see Theorem 23 in Section 3) that, up to isometry, there are only finitely many surfaces

$M \in T(g,n)$ with $M \in Min(S(M)) \cup \partial Min(S(M))$. Moreover, surfaces $M \in \partial Min(S(M))$ are very rare (if they exist at all, compare the remark at the end of Section 3) so that condition (ii) seems to be no serious restriction. Concerning condition (i) there is some good and some bad news. The bad news is that there exist rather often surfaces $M \in T(g,n)$ with $M \in Min(F)$ $(F = S(M))$ which are not F-regular; see Theorem 47 in Section 6. The good news is that such surfaces are often (always?) nevertheless non-degenerate critical points of *syst*; see again Theorem 47. In all studied cases for small values of g and n, *syst* is in fact a topological Morse function; see Section 6; and it seems quite probable that *syst* is a topological Morse function on all Teichmüller spaces $T(g,n)$.

Let now $\mathcal{M} \in T(g,n)$, $n > 0$. Let O be a cusp of $\mathcal{M}$ and let $G_O(g,n) \subset \Gamma(g,n)$ be the subgroup (of index n) of elements which fix O. Let $\mathcal{R}$ be a certain set of closed geodesics in $\mathcal{M}$ (see Section 4 for details). For $M \in T(g,n)$, let $syst_{\mathcal{R}}(M)$ be the length of a shortest element of $\mathcal{R}$ in M.

Theorem B. *On $T(g,n)$, $n > 0$, $syst_{\mathcal{R}}$ is a topological Morse function, invariant with respect to $G_O(g,n)$.*

Once again, we have a stronger result for surfaces with cusps than for closed surfaces, as it is usually the case in the study of the topology of $\Gamma(g,n)$ (see for example [7]). Note that the methods and results of this paper apply to many further functions which are similar to *syst*, an example is contained in the proof of Theorem 41 in Section 6.

The difficulty in the proof of Theorem A (and of Theorem B) lies in the fact that we are not in a Euclidean linear situation. This is why we need the sets of minima and their convexity properties; in some sense they provide a local cell decomposition of $T(g,n)$ which replaces the Euclidean linearity. On the sets of minima we have to construct flows, this is why some regularity of these sets is required.

Note that the sets of minima have other interesting applications; they can be used for a (global) cell decomposition of $T(g,n)$; see [27], [28], [29].

The paper is organized as follows. Section 2 introduces and studies the sets of minima $Min(F)$. Section 3 contains the proof of Theorem A while Section 4 gives the proof of Theorem B. Section 3 also contains a result which is an analogue of Voronoï's theorem [37] in the theory of positive definite quadratic forms; namely, $M \in T(g,n)$ is a local maximum for the function *syst* if and only if M is a critical point of *syst* of index 0. In Section 6 it is shown that for $(g,n) \in \{(0,4),(0,5),(0,6),(1,1),(1,2),(2,0)\}$ *syst* is a topological Morse function and all critical points are classified with their index. The important Theorem 47 is proved with another method for a local cell decomposition of $T(g,n)$. It is further shown (Theorem 51) that many of the critical points of *syst* correspond to arithmetic Fuchsian groups, and a surprising stability property (derived from Theorem D below) of some critical points of *syst*, which correspond to subgroups of Fuchsian triangle groups, is described:

Theorem C. *There exist surfaces in certain $T(g,n)$ which are critical points for all (topological) Morse functions on $T(g,n)$, invariant with respect to $\Gamma(g,n)$. For some of these surfaces, the index as a critical point has always the same parity, but not for other.*

Section 5 contains two applications which we are going to discuss now.

If *syst* is a topological Morse function on $T(g,n)$, one obtains a mass formula. Let $\mathcal{C}$ be the set of the mutually non–isometric critical points of *syst* in $T(g,n)$. Denote by $J(M)$ the index of a critical point M of *syst*. Denote by $|Isom(M)|$ the order of the automorphism group of M containing the orientation preserving automorphisms. Then

Theorem D.

$$\sum_{M \in \mathcal{C}} \frac{(-1)^{J(M)} n!}{|Isom(M)|} = \chi(\Gamma(g,n)),$$

where $\chi(\Gamma(g,n))$ is the virtual Euler characteristic of $\Gamma(g,n)$.

This (mass) formula follows by results from the cohomology of groups; see in particular [5], [10], [30].

In Section 5 the following result is also proved (in a constructive way).

Theorem E. *For every $g \geq 2$ there exists $M_g \in T(g,0)$ which is a non-degenerate critical point of* syst *of index $4g-5$.*

It is possible that $4g-5$ is the maximal index of critical points of *syst* in $T(g,0)$. If this is the case (and if *syst* is a topological Morse function on $T(g,0)$), then *syst* would induce a contractible simplicial complex on $T(g,0)$, invariant with respect to $\Gamma(g,0)$, of the best possible dimension, namely $4g-5$ (since the virtual cohomological dimension *vcd* of $\Gamma(g,0)$ is also $4g-5$ by Harer [6]). Until now, such complexes with the best possible dimension are known only for $T(g,n)$ with $n > 0$; they have been constructed in a combinatorial way; see [7] and [10] for overviews; with these complexes, the virtual Euler characteristic as well as the virtual cohomological dimension of the mapping class groups had been calculated. Note that $syst$ has in general much less critical points, compared with these combinatorial complexes; a striking example is Theorem B in connection with [27] (see the remark at the end of Section 4).

In the original version of this paper I made the mistake to think that all $M \in Min(F)$ are always F-regular ("proving" so Theorem A without the conditions). This mistake is also contained in the survey paper [26].

Here are some additional remarks concerning related literature. In the context of the Euclidean geometry of numbers it has been shown that the packing function (for lattice sphere packings) is a topological Morse function; see Ash [1],[2], and also [17]. For related problems in different geometries of numbers; see Bavard [3],[4]. Sarnak [23] conjectures, based on [20],[21], that the function

$$-\log \det \Delta(M)\,,$$

which is the (zeta regularised) product of the non-zero eigenvalues of the Laplacian of a closed surface M of genus g, is a Morse function on the space of all smooth Riemannian metrics on M. Thurston ([35], see also [7]) studies a function related to our function $syst$, which however is critical for every $M \in T(g,n)$ where $S(M)$ fills up. Kerckhoff [13] introduces lines of minima in $T(g,n)$; see Section 2 for some more comments. Finally, for more on systoles of (g,n)-surfaces and on hyperbolic geometry of numbers see Schmutz Schaller [26].

2. Sets of minima in Teichmüller space

First are stated the basic notation and definitions. I have tried to collect them in three separate definitions; the first is about surfaces, Teichmüller spaces and length functions, the second is about tangent spaces and tangent vectors, and the third introduces the set $Min(F)$ and a corresponding function ϕ_F.

Definition. (i) A (g,n)-*surface* $\mathcal{M}$ is a Riemann surface, of genus g with n cusps, equipped with a complete metric of constant curvature

-1. The case $(g, n) = (0, 3)$ will always be excluded without further comment.

A *surface* is a (g, n)-surface where g and n are not specified.

(ii) The *Teichmüller space* of a (g, n)-surface $\mathcal{M}$ is denoted by $T(\mathcal{M})$ or $T(g, n)$.

(iii) Let u be a closed geodesic of a surface $\mathcal{M}$. Then u is considered as *marked*. The *length* of u in $M \in T(\mathcal{M})$ is denoted by $L_M(u)$. $L_M(u)$ or $L(u)$ is called a *length function*.

(iv) Let $F = \{u_1, \ldots, u_m\}$ be a set of m closed geodesics in a surface $\mathcal{M}$. Then, for $M \in T(\mathcal{M})$, define

$$L_M(F) = (L_M(u_1), L_M(u_2), \ldots, L_M(u_m)) \in \mathsf{R}^m.$$

(v) A set F of closed geodesics of a surface $\mathcal{M}$ is said to *fill up* if every closed geodesic of $\mathcal{M}$ is intersected by at least one element of F.

Remark. (i) Let $\mathcal{M}$ be a surface and let $M \in T(\mathcal{M})$. I do not make a difference between $T(\mathcal{M})$ and $T(M)$. I shall denote by $\mathcal{M}$ a ("base") surface which will determine the Teichmüller space $T(\mathcal{M})$, and by M, M' points of $T(\mathcal{M})$.

(ii) The topology and the differential structure on $T(\mathcal{M})$ are given by a (finite) set of length functions which are real analytic functions.

Definition. (i) The *tangent space* of $T(\mathcal{M})$ in M is denoted by $Tan(M)$. Let $\xi \in Tan(M)$. Let u be a closed geodesic of $\mathcal{M}$. Then $\xi(u) \in \mathsf{R}$ denotes the real number which is the derivation, induced by ξ, of the length function $L_M(u)$. Let $F = \{u_1, \ldots, u_m\}$ be a set of m closed geodesics of $\mathcal{M}$. Define

$$\xi(F) = (\xi(u_1), \xi(u_2), \ldots, \xi(u_m)) \in \mathsf{R}^m.$$

Denote by $W_F(M)$ the vector space

$$W_F(M) = \{\xi(F) : \xi \in Tan(M)\}.$$

(ii) Let $x = (x_1, \ldots, x_m) \in \mathsf{R}^m$. Then $x > 0$ means that $x_i > 0$, $i = 1, \ldots, m$, while $x \geq 0$ means $x \neq 0$ and $x_i \geq 0$, $i = 1, \ldots, m$. R^m_+ is the subset of R^m containing all points $x > 0$.

(iii) Let $x = (x_1, \ldots, x_m) \in \mathsf{R}^m$ have k components which are zero. Then $\check{x}$ denotes the point in R^{m-k} where the zero-components have been omitted.

Definition. Let $\mathcal{M}$ be a surface with a set F of m closed geodesics which fills up.

(i) Fix $a \in \mathsf{R}^m$. Then

$$q_a(M) = \langle a, L_M(F)\rangle, \quad M \in T(\mathcal{M}),$$

defines a (real-valued) function on $T(\mathcal{M})$ by the Euclidean inner product of a and $L_M(F)$.

(ii) Let $a \in \mathsf{R}^m$, $a \geq 0$. To $\check{a}$ corresponds a subset of F, denoted by $F(\check{a})$. If $F(\check{a})$ fills up and if $a \neq \check{a}$, then a is called an *admissible boundary point*. By $A(F)$ is denoted the union of the set R^m_+ and the set of the admissible boundary points.

Let $a \in \mathsf{R}^m$. If the function $q_a(M)$ has a unique minimum in $T(\mathcal{M})$, in M_0 say, then define

$$\phi_F(a) = M_0.$$

This defines a function ϕ_F from a subset of R^m to $T(\mathcal{M})$ (Lemma 1 will show that ϕ_F is well defined on the whole set $A(F)$).

(iii) Put

$$Pos(F) := \{M \in T(\mathcal{M}) : \exists \xi \in Tan(M) \text{ with } \xi(F) > 0\},$$

where Pos stands for positive.

(iv) Put

$$Min(F) := \{M \in T(\mathcal{M}) : \forall \xi \in Tan(M) \text{ either } \xi(F) = 0 \text{ or } \exists\, u, v \in F \\ \text{with } \xi(u) > 0, \xi(v) < 0\},$$

where Min stands for minimum.

(v) Define $\quad \partial Min(F) := T(\mathcal{M}) \setminus (Pos(F) \cup Min(F))$.

Notation. For the rest of this section, F will denote a set of m simple closed geodesics of a surface $\mathcal{M}$ which fills up.

Remark. For $m = 2$ the sets $Min(F) \subset T(\mathcal{M})$ have been introduced by Kerckhoff [13] who called them *lines of minima*. His main aim was the proof (in [13]) that for every two different points M, M' in $T(\mathcal{M})$ there exists a set $F = \{u, v\}$ such that $M, M' \in Min(F)$. In order to do so one has however to extend the set of simple closed geodesics to the set of geodesic laminations (so that $u, v \in F$ are geodesic laminations). This generalization would also be possible for sets F with more than two elements, but we shall not need it (we will however need another generalization; see the remark after Corollary 16 at the end of this section).

Extending Kerckhoff's terminology I call $Min(F)$ a *set of minima.* The reason for this terminology will become clear from Lemma 1 and Theorem 4.

Lemma 1. *Let $a \in A(F)$. Then q_a has a unique critical point which is a minimum.*

Proof. (i) Let M_0 be a critical point of q_a. Let $\xi \in Tan(M_0)$. Then $\langle a, \xi(F)\rangle = 0$. By [11],[13], there exists an earthquake path δ passing through M_0 which induces ξ; moreover, q_a is strictly convex along δ and has thus a minimum in M_0 (with respect to δ). Since this is true for every $\xi \in Tan(M_0)$, it follows that q_a has a minimum in M_0. This shows that every critical point of q_a is a minimum.

(ii) If $M_1 \neq M_0$ is another critical point of q_a, then relate M_0 and M_1 by an earthquake path γ (which exists by Thurston [34]). Since again q_a is strictly convex along γ, this yields a contradiction which proves the uniqueness.

(iii) As for existence, since q_a is continuous, it is enough to prove that the sets

$$K(C) = \{M \in T(\mathcal{M}) : q_a(M) \leq C\}\,, \quad C \in \mathsf{R},$$

are bounded. Let $M \in K(C)$ (in the sequel we allow M to vary in $K(C)$). Since $F(\check{a})$ fills up,

$$M \setminus \bigcup_{u \in F(\check{a})} u$$

is a finite union of polygons and it follows (since $M \in K(C)$) that the diameter of each polygon is bounded (the diameter is the maximal distance between two points of the polygon). Let G be a finite set of simple closed geodesics of $\mathcal{M}$ such that the length functions of G parametrize $T(\mathcal{M})$ (in the sense that $L_M(G) = L_{M'}(G)$, $M, M' \in T(\mathcal{M})$, implies $M = M'$). Then each element $v \in G$ intersects only a finite number of polygons so that the length of v is bounded. This proves that $K(C)$ is bounded. q.e.d.

Remark. In the sequel, I shall use without comment the fact that geodesic length functions are convex along earthquake paths and that earthquake paths exist between two different points in the same Teichmüller space.

Lemma 2. *There exists an open set $B(F)$ in R^m which contains $A(F)$ such that q_a has a unique minimum for each $a \in B(F)$.*

Proof. Let $a_0 \in A(F)$ be an admissible boundary point. By Lemma 1 it is enough to show that there exists an open neighbourhood U of a_0 in R^m such that q_a has a unique minimum for each $a \in U$. Since q_a is smooth with respect to a, there exists an open neighbourhood of a_0 such that the sets (compare the proof of Lemma 1)

$$K_a(C) = \{M \in T(\mathcal{M}) : q_a(M) \leq C\}$$

are still compact (if $C < \epsilon + \min q_{a_0}$, for an $\epsilon > 0$, say). This proves that q_a has a minimum for every $a \in U$. Let M_i, $i = 1, 2$, be two different points on the boundary of $K_a(C)$. Let δ be an earthquake path from M_1 to M_2. Since q_{a_0} is strictly convex along δ and since q_a is smooth with respect to a it follows that q_a is still strictly convex along δ. By this argument we can choose U so that the uniqueness of the minimum of q_a is assured for every $a \in U$. q.e.d.

Definition. Denote by $B(F)$ a contractible open set in R^m containing $A(F)$ for which Lemma 2 holds. The definition of the function ϕ_F will be restricted to $B(F)$.

Corollary 3. *Let $a \in B(F)$. Then the following hold:*

(i) M_0 is the unique minimum of q_a if and only if $\langle a, \xi(F)\rangle = 0$ for every $\xi \in Tan(M_0)$.

(ii) If $a \in A(F)$, then $\phi_F(a) \in Min(F) \cup \partial Min(F)$ and $\phi_F(a) \in Min(F)$ if $a > 0$.

Proof. Part (i) was shown for $a \in A(F)$ during the proof of Lemma 1. By the proof of Lemma 2, the argument also holds for $a \in B(F)$. Part (ii) is an immediate consequence of (i) and the definitions. q.e.d.

Theorem 4. *(i) Let $M \in Min(F)$. Then there exists an $a \in \mathsf{R}^m_+$ such that $\phi_F(a) = M$.*

(ii) The image $\phi_F(A(F))$ is the set $Min(F) \cup \partial Min(F)$.

(iii) ϕ_F is continuous on $B(F)$.

Proof. (i) Let $M \in Min(F)$. Then, as a consequence of the definition, there exists an x in the orthogonal complement of $W_F(M)$ in R^m with $x > 0$. From Corollary 3 it follows that $\phi_F(x) = M$.

(ii) By Corollary 3 we have $\phi_F(A(F)) \subset Min(F) \cup \partial Min(F)$. By a similar argument as in part (i), for $M \in \partial Min(F)$, there exists an $a \in A(F)$, which is an admissible boundary point with $\phi_F(a) = M$.

(iii) Let $a_i \longrightarrow a$ be a convergent sequence in $B(F)$. By the proofs of Lemmas 1 and 2, $\phi_F(a_i)$ has a convergent subsequence which we denote by $\phi_F(a_j)$ and its limit by M_0. Put

$$\{a_i\}^\perp := \{\lambda a_i : \lambda \in \mathsf{R}\}^\perp \subset \mathsf{R}^m.$$

Then, as point sets,

$$\{a_i\}^\perp \longrightarrow \{a\}^\perp. \tag{1}$$

It follows from Corollary 3 that

$$W_F(\phi_F(a_i)) \subset \{a_i\}^\perp. \tag{2}$$

For the subsequence a_j, since the dimension of $W_F(M_0)$ cannot decrease locally, we have

$$W_F(\phi_F(a_j)) \longrightarrow W_F(M_0),$$

which implies, due to (1) and (2), that $W_F(M_0) \subset \{a\}^\perp$. It then follows from Corollary 3 that $\phi_F(a) = M_0$. We therefore have proved that every convergent subsequence of $\phi_F(a_i)$ has the same limit $\phi_F(a)$. The continuity of ϕ_F is an immediate consequence. q.e.d.

Notation. (i) Let ϕ_F be defined on $B(F)$. Then the image $\phi_F(B(F))$ in $T(\mathcal{M})$ is denoted by $Conv(F)$, where $Conv$ stands for convex.

(ii) Let $M \in Conv(F)$. Then by Corollary 3, $\phi_F^{-1}(M)$ is the intersection of $B(F)$ with a linear subspace of R^m. The dimension of $\phi_F^{-1}(M)$ is therefore well defined and is denoted by $j_F(M)$.

Lemma 5. *Let $M \in Min(F) \cup \partial Min(F)$. Then there exists $F' \subset F$ such that $M \in Min(F') \cup \partial Min(F')$, $j_{F'}(M) = 1$, and* $\dim W_{F'}(M) = \dim W_F(M)$.

Proof. If $j_F(M) = 1$ take $F' = F$. So we assume $j_F(M) \geq 2$. Since $\phi_F^{-1}(M)$ is the intersection of $B(F)$ with a linear subspace, it follows that $\phi_F^{-1}(M)$ contains a sequence a_i which converges to a point $b \geq 0$ in the boundary of R^m_+. By continuity,

$$\inf q_b = \lim_{a_i \longrightarrow b} \langle a_i, L_M(F)\rangle = \langle b, L_M(F)\rangle,$$

which shows that $b \in A(F)$ (the uniqueness of the minimum of q_b follows as in the proof of Lemma 1). Recall that $F(\check{b})$ denotes the subset of F corresponding to the non-zero components of b. By construction, we have $M \in Min(F(\check{b}))$ and $j_{F(\check{b})}(M) < j_F(M)$ which proves by induction that F has a subset $F(1)$ such that $M \in Min(F(1))$ and $j_{F(1)}(M) = 1$. Let m_1 be the order of $F(1)$. Then $\dim W_{F(1)}(M) = m_1 - 1 \leq \dim W_F(M) = m - j_F(M)$ so that $m_1 \leq m + 1 - j_F(M)$. Assume that $m_1 - 1 < \dim W_F(M)$. Then there exists a $u \in F \setminus F(1)$ so that, putting $F(2) = F(1) \cup \{u\}$, we have

$$m_1 = \dim W_{F(1)}(M) + 1 = \dim W_{F(2)}(M) \text{ and } j_{F(2)}(M) = 1.$$

It is clear that $M \in Min(F(2)) \cup \partial Min(F(2))$. Since we can repeat this argument, the lemma follows. q.e.d.

Definition. $M \in Conv(F)$ is called *regular* or *F-regular* if j_F is locally constant in M. More precisely, if M is regular then there exists an open neighbourhood U of M in $T(\mathcal{M})$ such that j_F is constant on $U \cap Conv(F)$.

Lemma 6. *Let $M \in Conv(F)$ with $\dim W_F(M) = \dim T(\mathcal{M})$ or $j_F(M) = 1$. Then M is regular.*

Proof. By Corollary 3 we have

$$|F| = j_F(M) + \dim W_F(M)\,.$$

Clearly, $\dim W_F(M)$ cannot decrease locally and therefore, $j_F(M)$ cannot increase locally. q.e.d.

Lemma 7. *Let $M \in Conv(F)$ be F-regular. Then there exists an open neighbourhood U' of M in $T(\mathcal{M})$ such that $U := U' \cap Conv(F)$ is homeomorphic to a k-ball with $k = m - j_F(M)$ and such that j_F is constant on U.*

Definition. Let $M \in Conv(F)$ be F-regular and let U be defined as in Lemma 7. Then U is called a *regular neighbourhood* (or an *F-regular neighbourhood*) of M in $Conv(F)$.

Proof of Lemma 7. (i) Let $j_F(M) = 1$. We restrict the definition of ϕ_F to $a \in B(F)$ with $\|a\| = 1$. Then $\phi_F^{-1}(M)$ is unique and it is clear that the lemma holds in this case.

(ii) Let now $j_F(M) \geq 2$. Note that if $F' \subset F$, then $Conv(F') \subset Conv(F)$ by definition. It thus follows from Lemma 5 and its proof that there exists $F' \subset F$ such that $M \in Conv(F')$, $j_{F'}(M) = 1$, and

$\dim W_{F'}(M) = \dim W_F(M)$. By Lemma 6, M is F'-regular and therefore, by (i), the lemma holds for F'. Let thus U' be an open neighbourhood of M in $T(\mathcal{M})$ such that $U = U' \cap Conv(F')$ is an F'-regular neighbourhood of M in $Conv(F')$. Since M is F-regular, it follows that U' can be chosen such that $U' \cap Conv(F) = U$. q.e.d.

Definition. Put

$$\pi_F : T(\mathcal{M}) \longrightarrow \mathsf{R}^m\,, \quad M \mapsto L_M(F)\,.$$

Lemma 8. *Let $M \in Conv(F)$ be regular, and U be a regular neighbourhood of M in $Conv(F)$. Then π_F restricted to U is a diffeomorphism onto its image.*

Proof. Let $M \in U$. Since, by Lemma 2, M is the unique minimum for a function q_a, it follows that π_F is injective on U. Obviously, π_F is smooth and, since F fills up, proper, $\pi_F(U)$ is a diffeomorphism onto its image. q.e.d.

The next lemma will be needed in Section 5.

Lemma 9. *Let F be such that every subset $F' \subset F$, $F' \neq F$, does not fill up. Then $j_F(M) = 1$ for every $M \in Min(F)$, $\partial Min(F) = \emptyset$, and for every $a \in \mathsf{R}^m_+$ there exists a unique $M \in Min(F)$ such that $L_M(F)$ is a multiple of a.*

Proof. By hypothesis $A(F)$ has no admissible boundary point. This implies by Corollary 3 that $\partial Min(F) = \emptyset$ and by the argument in the proof of Lemma 5 that $j_F(M) = 1$ for all $M \in Min(F)$. Restrict ϕ_F to the set $S^{m-1}_+ = \{a \in \mathsf{R}^m_+ : \|a\| = 1\}$, denote still by ϕ_F this restriction; by Theorem 4, ϕ_F is then a homeomorphism onto $Min(F)$. We may identify $Min(F)$ with $\pi_F(Min(F))$; see Lemma 8. Moreover, the map p

$$p : L_M(F) \longrightarrow \frac{L_M(F)}{\|L_M(F)\|}\,,\ M \in Min(F),$$

is injective (if $L_M(F) = \lambda L_{M'}(F)$ for $\lambda > 1$, then an earthquake path from M' to M induces $\xi \in Tan(M)$ with $\xi(F) > 0$) and obviously continuous and proper so that p is a homeomorphism onto its image. Thus the map $p \circ \pi_F \circ \phi_F$ is a homeomorphism of S^{m-1}_+ into S^{m-1}_+. Since $\partial Min(F) = \emptyset$, it follows that the image of $p \circ \pi_F \circ \phi_F$ is closed in S^{m-1}_+ proving that this image is the whole S^{m-1}_+. q.e.d.

Remark. In the situation of Lemma 9 and its proof where ϕ_F is a homeomorphism when restricted to the points of norm 1, ϕ_F is the inverse of the Gauss map.

Theorem 10. *Let $M_0 \in Min(F)$ (or $M_0 \in Conv(F)$) be regular and let U be a regular neighbourhood of M_0 in $Conv(F)$. Then U is a differentiable submanifold of $T(\mathcal{M})$ (of type C^1) of dimension $k := m - j_F(M_0)$.*

Proof. By Lemma 5 and its proof U is covered by charts $U \cap Conv(F')$ where the $F' \subset F$ are subsets with the properties that $M_0 \in Conv(F')$, $j_{F'}(M_0) = 1$, and $\dim W_{F'}(M_0) = \dim W_F(M_0)$.

We have to prove that the differentiable structure of $T(\mathcal{M})$ induces a differentiable structure on $U' := U \cap Conv(F')$. By Lemma 8 we may identify U' with its image with respect to $\pi_{F'}$ in R^{k+1}. Then U' is contained in the boundary of a strictly convex body in R^{k+1} such that through each point $L_M(F') \in U'$, there exists a unique supporting hyperplane (parallel to $W_{F'}(M)$). It follows that the length functions of F' induce a differentiable structure on U', induced by the differentiable structure of $T(\mathcal{M})$. q.e.d.

Corollary 11. *In regular points, ϕ_F is a C^1 map.*

Proof. Let $M \in Conv(F)$ be regular and U a regular neighbourhood of M. We have the following situation:

$$W \subset \mathsf{R}^m \quad \overset{\phi_F}{\longrightarrow} \quad U \quad \overset{\pi_F}{\longrightarrow} \quad \mathsf{R}^m$$

By Lemma 8, π_F is a diffeomorphism onto the image. By Corollary 3 and Theorem 4 the inverse image $\phi_F^{-1}(\pi_F^{-1}(x))$ associates to a point $x \in \pi_F(Conv(F))$ its normal vector space, and the corollary follows by Theorem 10. q.e.d.

Corollary 12. *Let $M \in Conv(F)$ be regular. Then $W_F(M)$ is the tangent space of $\pi_F(Conv(F))$ in $\pi_F(M)$.*

Proof. Clear by Theorem 10. q.e.d.

Corollary 13. *If every $M \in Min(F)$ (or in $Conv(F)$) is regular, then $Min(F)$ (respectively $Conv(F)$) is a differentiable submanifold of $T(\mathcal{M})$ (of type C^1) of dimension $m - j_F(M)$.*

Proof. Clear by Theorem 10 since $Min(F)$ and $Conv(F)$ are connected by Theorem 4. q.e.d.

Lemma 14. *Let $M \in \partial Min(F)$. Then there exists a unique subset $F_0 \subset F$ such that $M \in Min(F_0))$ and $F' \subset F_0$ for all $F' \subset F$ with $M \in Min(F')$.*

Proof. Since $M \in \partial Min(F)$ there exists $\xi \in Tan(M)$ with $\xi(F) \geq 0$. Among all such ξ let $\xi_0 \in Tan(M)$ be one such that the number of zero components of $\xi_0(F)$ is minimal. Let $F_0 \subset F$ correspond to the zero components of $\xi_0(F)$. Assume that there exists $\zeta \in Tan(M)$ such that $\zeta(F_0) \geq 0$. Then, for $\lambda \in \mathrm{R}$ big enough, $\xi'(F) = (\zeta + \lambda\xi_0)(F) \geq 0$ has less zero components than ξ_0, a contradiction. It follows that $M \in Min(F_0)$.

Let $F' \subset F$ such that $M \in Min(F')$. Then $\xi_0(F') = 0$ which implies $F' \subset F_0$. q.e.d.

Theorem 15. *Let $M_0 \in \partial Min(F)$, and define $F_0 \subset F$ as in Lemma 14. Let*

$$\begin{aligned}\mathcal{G} = \{F\} \cup \{G : F_0 \subset G \subset F \textit{ and } \exists \xi \in Tan(M_0) \\ \textit{such that } \xi(G) = 0,\, \xi(F \setminus G) > 0\,\}.\end{aligned}$$

Let M_0 be G-regular for every $G \in \mathcal{G}$. Then there exists an F-regular neighbourhood U of M_0 in $Conv(F)$ such that, in U,

$$Min(F) \cup \partial Min(F) = \bigcup_{G \in \mathcal{G}} Min(G),$$

and this is a disjoint union.

Proof. (i) It follows from Lemma 14 that there exists an F-regular neighbourhood U of M_0 in $Conv(F)$ such that

$$U \subset \bigcup_{G \in \mathcal{G}} Min(G)\,. \tag{3}$$

Let $k = \dim W_F(M_0)$. Let $k_0 = \dim W_{F_0}(M_0)$ and let $q = k - k_0$. If $q = 1$ then $\mathcal{G} = \{F_0, F\}$ and the theorem clearly holds. So assume $q \geq 2$. By (3) there then exists $G \in \mathcal{G}$, such that $\dim W_G(M_0) = k - 1$. It follows that there exists an, up to a scalar factor, unique $\xi \in Tan(M_0)$ with $\xi(G) = 0$ and $\xi(F \setminus G) > 0$. Since M_0 is G-regular, we can choose U such that $U \cap Conv(G)$ is a G-regular neighbourhood of M_0 and such that, for every $M \in U \cap Conv(G)$, there exists $\xi \in Tan(M)$ with $\xi(G) = 0$ and $\xi(F \setminus G) > 0$.

(ii) Let $G_1, G_2 \in \mathcal{G}$ such that $\dim W_{G_i}(M_0) = k-1$, $i = 1, 2$. Let $M \in U \cap Min(G_1) \cap Min(G_2)$. By (i), it follows that there exists $\xi_i \in Tan(M)$ such that $\xi_i(G_i) = 0$ and $\xi_i(F \setminus G_i) > 0$, $i = 1, 2$. Thus $G_1 = G_2$. In other words, if G_1 and G_2 are two different elements of $\mathcal{G}$ with $\dim W_{G_i}(M_0) = k-1$, then $Min(G_1)$ and $Min(G_2)$ are disjoint in U.

The theorem now follows by (i) and by induction with respect to q since the hypothesis of the theorem is fulfilled for every $G \in \mathcal{G}$ with $\dim W_G(M_0) = k-1$. q.e.d.

Definition. Let $M \in \partial Min(F)$. Let $k = \dim W_F(M)$. Then M is called *strongly F-regular* if M is F-regular and if M has an F-regular neighbourhood U in $Conv(F)$ such that $U \cap \partial Min(F)$ is homeomorphic to a $(k-1)$-ball which separates U into $U \cap Min(F)$ and $U \cap Pos(F)$ which both are homeomorphic to a k-ball.

Remark. (i) Let $M \in \partial Min(F)$ and let $j_F(M) = 1$. By Lemma 6, M is then F-regular. Moreover, by the map ϕ_F and Theorem 4, M is strongly F-regular.

(ii) Let $M \in \partial Min(F)$. It is quite probable that M is strongly F-regular if and only if M is F-regular.

Corollary 16. *Let $M \in \partial Min(F)$, let $F_0 \subset F$ and $\mathcal{G}$ be defined as in Theorem 15. Let M be G-regular for every $G \in \mathcal{G}$. Then M is strongly F-regular.*

Proof. Clear by Theorem 15. q.e.d.

Remark. Let F be a set of m simple closed geodesics of a surface $\mathcal{M}$. Let $v_1, \ldots, v_d$ be d further simple closed geodesics of $\mathcal{M}$. Let $\lambda_1, \ldots, \lambda_d$ be d fixed non-negative reals. For $M \in T(\mathcal{M})$ put

$$L_M(U) = \sum_{i=1}^{d} \lambda_i L_M(v_i),$$

where U is treated like a closed geodesic. Let $F' = \{U\} \cup F$. Then the theory developed in this section applies for F' since $L_M(U)$ is a well-defined length function which is convex along earthquake paths (since $\lambda_i \geq 0$, $i = 1, \ldots, d$).

This generalization will be used in the next section.

I add some results concerning the possible size of $Min(F)$ which however will not be used in this paper.

Proposition 17. *(i) $Pos(F)$ is never empty.*

(ii) Let $M, M' \in T(\mathcal{M})$. If $L_{M'}(u) \le L_M(u)$, for every simple closed geodesic u in $\mathcal{M}$, then $M = M'$.

Proof. (i) Let u be a simple closed geodesic of $\mathcal{M}$. Let $F(u)$ be the subset of F of geodesics which intersect u. Let $M_0 \in T(\mathcal{M})$. Execute a twist deformation (starting in M_0) along u so that in the resulting surface M we have $\xi_u(F(u)) > 0$ (this is possible since $F(u)$ is finite). Further, a Weil-Petersson geodesic from M_0 to M induces $\zeta \in Tan(M)$ such that $\zeta(u') > 0$ for all $u' \in F \setminus F(u)$ by the strict convexity of the length functions along Weil-Petersson geodesics (Wolpert [39]). It follows that there is a positive α such that $(\alpha\xi_u + \zeta)(F) > 0$ and therefore, $M \in Pos(F)$.

(ii) Assume that $M' \neq M$. Let γ be a Weil-Petersson geodesic from M' to M and prolong γ beyond M. Let N be on this prolongation. Since the length functions are strictly convex along Weil-Petersson geodesics, we have $L_M(u) < L_N(u)$, for every simple closed geodesic u of $\mathcal{M}$. Let γ' be an earthquake path from M to N and prolong γ' beyond N. By the convexity of the length functions along earthquake paths and since earthquake paths can always be prolonged, it follows that along γ', the length of all simple closed geodesics grows to infinity which contradicts the fact that, for all elements of $T(\mathcal{M})$, the length of a shortest closed geodesic has a (finite) common upper bound. q.e.d.

Proposition 18. *Let K be a compact subset of $T(\mathcal{M})$. Then there exists a finite set F such that $K \subset Min(F)$.*

Proof. (i) Let $M \in T(\mathcal{M})$. For every simple closed geodesic u of $\mathcal{M}$ put

$$B_M(u) = \{M' \in T(\mathcal{M}) : L_{M'}(u) \le L_M(u)\}.$$

Define further

$$B_M(F) = \bigcup_{u \in F} B_M(u).$$

If $B_M(F)$ contains an open neighbourhood U of M in $T(\mathcal{M})$, then $M \in Min(F)$ (since if there exists $\xi \in Tan(M)$ with $\xi(F) \ge 0$, then, by convexity, the points on the Weil-Petersson geodesic starting in M induced by ξ are not in $B_M(F)$). Moreover, $\dim W_F(M) = \dim T(\mathcal{M}) =: N$ since the same argument shows that if $\xi \in Tan(M)$ with $\xi(F) = 0$, then $\xi \equiv 0$.

(ii) Let U be an open (topological) ball around M in $T(\mathcal{M})$; let ∂U be its boundary. Choose U such that the Weil-Petersson geodesics from M to the points of ∂U fill U (this is possible by [39]). By Proposition 17 (ii) (and since ∂U is compact) there exists a finite set $F(M)$ of simple closed geodesics such that $\partial U \subset B_M(F(M))$. From the strict convexity of the geodesic length functions along Weil-Petersson geodesics it follows that $U \subset B_M(F(M))$ which, by (i), proves that $M \in Min(F(M))$ and that $\dim W_{F(M)}(M) = N$. This implies by Lemma 6 that M is $F(M)$-regular so that M has an open neighbourhood $V(M)$ in $T(\mathcal{M})$ with $V(M) \subset Min(F(M))$.

Therefore,

$$M \longrightarrow V(M), \quad M \in K,$$

provides an open covering of K which has a finite subcovering (K is compact) and the proposition follows. q.e.d.

3. Morse functions and systoles

Definition. Let $\mathcal{M}$ be a surface.

(i) Let $M \in T(\mathcal{M})$. Then $S(M)$ is the *set of the systoles* of M where a *systole* of M is a shortest closed geodesic.

(ii) Define the function

$$syst : T(\mathcal{M}) \longrightarrow \mathsf{R}_+,$$

where $syst(M)$ is the length of a systole of $M \in T(\mathcal{M})$.

Remark. A systole u in $S(M)$ is, as always, also considered as a marked geodesic, but, of course, in another surface M', u is in general not a systole.

Definition. Let X be a compact connected manifold. Let

$$f : X \longrightarrow \mathsf{R}_+$$

be a continuous function. Then $x \in X$ is called an *ordinary point* of f if there exist an open neighbourhood U of x in X and a homeomorphic parametrization of U by $N = \dim X$ parameters such that one of them is the function f. One calls x a *critical point* of f if x is not an ordinary point.

If x is critical, then x is called a *non-degenerate critical point* if there exists an open neighbourhood U of x in X, a homeomorphic parametrization of U by N parameters $y_1, \dots, y_N$, and an integer J $(0 \leq J \leq N)$ such that

$$f(x') - f(x) = \sum_{i=1}^{J} y_i^2 - \sum_{i=J+1}^{N} y_i^2$$

for all $x' \in U$. The integer J is called the *index* of x.

If all critical points of f are non-degenerate, then f is called a *topological Morse function.*

Remark. (i) Usually, the index of a critical point is defined as the number of the negative squares. But since the dimension of the Teichmüller spaces used in this paper is always even, the definition of the index chosen here works well for the function *syst*.

(ii) General Morse functions on $T(\mathcal{M})$ are well-known. As an example, take the functions q_a, defined in Section 2. By Lemma 1, q_a has a unique critical point which is a minimum (this point clearly is non-degenerate); therefore, q_a is a Morse function (note that q_a is proper by the proof of Lemma 1). Similar functions have been used for example in [11] and [39]. Further examples (the energy of a harmonic map) can be found in [38] and [36].

(iii) Of particular interest are Morse functions on $T(\mathcal{M})$ which are invariant with respect to the mapping class group $\Gamma(\mathcal{M})$; they can be used for analysing the topology of the *moduli space*

$$Mod(\mathcal{M}) = T(\mathcal{M})/\Gamma(\mathcal{M}),$$

which is much more complicated than that of $T(\mathcal{M})$ (the latter space being a cell). The first examples of such functions are those of this paper.

Note that *syst* is invariant with respect to $\Gamma(\mathcal{M})$. It will be proved that, under certain restrictions, *syst* is a topological Morse function. In Section 4 we shall construct topological Morse functions on $T(g,n)$, $n > 0$, which are invariant with respect to a subgroup of $\Gamma(g,n)$ of index n.

(iv) Traditionally, Morse functions are defined for compact spaces. Of course, $T(\mathcal{M})$ is not compact, but the function *syst* is invariant with respect to $\Gamma(\mathcal{M})$ and is proper on $Mod(\mathcal{M})$ (by a result of Mumford [19]; see the proof of Theorem 23 below) so that the fact that $T(\mathcal{M})$ is not compact does not cause problems.

Note that *syst* is not proper on $T(\mathcal{M})$. Namely, let $M \in T(\mathcal{M})$ be such that a systole u of M is very small. Then a twist deformation along u does not change $syst = L(u)$ and $syst^{-1}(L(u))$ cannot be compact.

Remark and Definition. (i) Let $R \subset T(g,n)$ be a differentiable submanifold. Let $\tilde{\xi}$ be a tangent vector field on R. Then ξ_M denotes the induced vector in $Tan(M)$, $M \in R$.

(ii) Let $M_0 \in T(g,n)$, and let u be a simple closed geodesic of M_0. Let $\xi_u \in Tan(M_0)$ be the tangent vector induced by a twist deformation of unit speed along u (we also fix the direction in which the twist is executed). Since u is considered as a marked geodesic, we can define $\xi_u \in Tan(M)$ for every $M \in T(g,n)$. This defines a vector field denoted by $\tilde{\xi}_u$ or simply by $\tilde{\xi}$ if we do not want to mention u. This construction is also possible for generalized twist deformations, namely the earthquakes. I shall say that these vector fields are *derived from an earthquake.*

The vector fields derived from an earthquake are known to be real analytic (see [12]).

Theorem 19. *Let $M_0 \in T(g,n)$ and $F = S(M_0)$. If $M_0 \in Pos(F)$, then M_0 is an ordinary point of* syst.

Proof. Let $\xi_0 \in Tan(M_0)$ such that $\xi_0(F) > 0$. By [13] there then exists a vector field $\tilde{\xi}$ (derived from an earthquake) such that $\xi_{M_0} = \xi_0$. We choose an open neighbourhood U of M_0 in $T(g,n)$ such that $\xi_M(F) > 0$ and $S(M) \subset F$ for all $M \in U$. It follows that along the integral curves of the flow induced by $\tilde{\xi}$, *syst* is strictly monotonic in U. Therefore, these integral curves can be parametrized in U by *syst*. Since the integral curves give a disjoint partition of U (by the general theory of ordinary differential equations and since $\tilde{\xi}$ is differentiable), the theorem follows. q.e.d.

Corollary 20. *Let $M \in T(g,n)$. If $S(M)$ does not fill up, then M is an ordinary point of* syst.

Proof. Assume that $F = S(M)$ does not fill up. By Thurston [35] (discussed in [7]) there then exists $\xi \in Tan(M)$ such that $\xi(F) > 0$. The corollary now follows from Theorem 19. q.e.d.

Remark. Let F be a set of simple closed geodesics of a surface $\mathcal{M}$. It follows from the proof of Corollary 20 that $Min(F) \cup \partial Min(F)$

is empty if F does not fill up.

Definition. Let $M, M' \in T(g,n)$. Let F and F' be finite sets of simple closed geodesics in M and M', respectively. Then F and F' are called *equivalent* if there exists a homeomorphism from M to M' which maps the homotopy classes of the elements of F to those of F'.

Corollary 21. *Let $M, M' \in T(g,n)$ so that M and M' are not isometric. If $S(M)$ and $S(M')$ are equivalent, then M and M' cannot both be critical points of* syst.

Proof. Assume that $S(M)$ and $S(M')$ are equivalent. Then there exists an $N \in T(g,n)$ which is isometric to M' such that $S(N) = S(M) =: F$ in the sense of marked geodesics. Since M and M' are not isometric, we have $M \neq N$. Assume without restriction that $syst(M) \geq syst(M')$. Then there exists an earthquake path γ from N to M which induces $\xi \in Tan(M)$ with $\xi(F) \geq 0$. Let $F_0 \subset F$ be the maximal subset of F such that $\xi(F_0) = 0$. Since this means that $L(u)$ is constant along γ for all $u \in F_0$, it follows that F_0 does not fill up. As in the proof of Corollary 20 there then exists $\xi' \in Tan(M)$ such that $\xi'(F_0) > 0$. An appropriate linear combination of ξ and ξ' thus produces $\zeta \in Tan(M)$ with $\zeta(F) > 0$. This hence proves by Theorem 19 that M is an ordinary point of *syst*. q.e.d.

Definition. Let $M \in T(g,n)$ such that $M \in \partial Min(S(M))$. Then M is called a *boundary point* of *syst*.

Corollary 22. *Critical points and boundary points of* syst *are isolated in $T(g,n)$.*

Proof. Let $M \in T(g,n)$. Then there exists an open neighbourhood U of M in $T(g,n)$ such that $S(M') \subset S(M)$ for all $M' \in U$ (the elements of $S(M)$ and of $S(M')$ taken as marked geodesics). By Corollary 21 and its proof, *syst* can only have a finite number of critical points and boundary points in U. q.e.d.

Theorem 23. *Up to isometry, there exists only a finite number of critical points and boundary points of* syst *in $T(g,n)$.*

Proof. Let $M \in T(g,n)$ and let u and v be two simple closed geodesics of M of the same length which intersect. Then, by hyperbolic trigonometry, their length is bounded from below by an $\epsilon > 0$. By a result of Mumford [19], there exists a compact subset $T_\epsilon(g,n)$ of $T(g,n)$ such that for every $M \in T(g,n)$ with $syst(M) \geq \epsilon$, there exists $M' \in$

$T_\epsilon(g,n)$ which is isometric to M. By Corollary 20 we have $syst(M_0) \geq \epsilon$ if M_0 is a critical or a boundary point of $syst$. The theorem therefore follows by Corollary 22. q.e.d.

Lemma 24. *Let $M_0 \in T(g,n)$, let $F = S(M_0)$. Let $M_0 \in Min(F)$ be regular. Then there exist a regular neighbourhood V of M_0 in $Min(F)$ and a differentiable tangent vector field $\tilde{\theta}$ on $V \setminus \{M_0\}$ such that the integral curves of $\tilde{\theta}$ all start on ∂V (the boundary of V in $Min(F)$) and end in M_0, and* syst *is strictly growing along these integral curves.*

Proof. Note first that $syst$ restricted to $Min(F)$ has a unique maximum in M_0 since if $syst(M) \geq syst(M_0)$ for $M \neq M_0$, then an earthquake path from M_0 to M induces a vector $\xi \in Tan(M)$ with $\xi(F) \geq 0$ which implies $M \notin Min(F)$. We have thus to construct an analogue of the gradient flow.

Let U be a regular neighbourhood of M_0 in $Min(F)$. For $\delta > 0$, let $V(\delta)$ be the connected component, containing M_0, of the set

$$\{M \in U : syst(M) > syst(M_0) - \delta\}$$

which is thus an open neighbourhood of M_0 in U. We need some conditions for the choice of δ. A non-empty subset Σ of F is called admissible if $M_0 \in Pos(\Sigma)$. For every admissible Σ choose $\sigma \in Tan(M_0)$ such that $\sigma(\Sigma) > 0$. Denote by $\tilde{\sigma}$ the induced vector field, derived from the earthquake associated to σ (different subsets Σ will induce different vector fields which, by abuse of notation, will all be denoted by $\tilde{\sigma}$). Choose $\delta > 0$ so small that:

(1) $S(M) \subset F$, $\forall M \in V(\delta)$,

(2) for every admissible Σ, we have $\sigma_M(\Sigma) > 0$, for all $M \in V(\delta)$.

We fix such a $\delta > 0$ and put $V_0 = V(\delta) \setminus \{M_0\}$. Since $syst|Min(F)$ attains its maximum in M_0, and an earthquake path from M to M_0 induces a vector $\xi \in Tan(M_0)$ with $\xi(S(M)) > 0$, we note that if $M \in V_0$, then $S(M)$ is admissible.

Let $\epsilon > 0$ and put, for every admissible Σ,

$$U_\epsilon(\Sigma) = \{M \in V_0 : S(M) \subset \Sigma \text{ and } L_M(u) < syst(M) + \epsilon, \forall u \in \Sigma\}.$$

This defines a finite open covering of V_0 since $U_\epsilon(\Sigma)$ is obviously open and for every $M \in V_0$, we have $M \in U_\epsilon(\Sigma)$ for $S(M) = \Sigma$. Choose a differentiable partition of unity

$$\{f_\Sigma : \Sigma \text{ admissible}\}$$

of this covering and put, for every $M \in V_0$,

$$\theta_M = \sum_{\Sigma} f_{\Sigma}\sigma_M \,, \quad \Sigma \text{ admissible}.$$

This defines a differentiable vector field $\tilde{\theta}$ and thus a differentiable flow on V_0. Along the integral curves of this flow, *syst* is strictly increasing; therefore, all integral curves start on $\partial V(\delta)$, the boundary of $V(\delta)$ in U, and end in M_0. This proves the lemma. q.e.d.

Theorem 25. *Let $M_0 \in T(g,n)$, and let $F = S(M_0)$. Then the following conditions are equivalent:*
(a) $M_0 \in Min(F)$ and $\dim W_F(M_0) = \dim T(g,n)$.
(b) M_0 is a non-degenerate critical point of syst *of index zero.*
(c) M_0 is a local maximum of syst.

Proof. (i) We assume that (a) holds. Since $\dim W_F(M_0) = \dim T(g,n)$, it follows that M_0 is regular (by Lemma 6) and has a regular neighbourhood $U \subset Min(F)$ which is open in $T(g,n)$. Let V be an open neighbourhood of M_0 as in Lemma 24 with the vector field $\tilde{\theta}$. Then by Lemma 24, we have a homeomorphic parametrization of V by

$$M \longrightarrow \sqrt{syst(M_0) - syst(M)}\, x, \quad M \in V,$$

where $x = (x_1, \ldots, x_N) \in \mathsf{R}^N$ $(N = \dim T(g,n))$ is a point of unit norm corresponding to a point in ∂V. This yields, for $M \in V$,

$$(4) \qquad syst(M) - syst(M_0) = -\sum_{i=1}^{N} \left(\sqrt{syst(M_0) - syst(M)}\right)^2 (x_i)^2,$$

which shows that M_0 is a non-degenerate critical point of index zero. We thus have proved $(a) \Longrightarrow (b)$.

(ii) If M_0 is a non-degenerate critical point of *syst* of index zero, then there exists an open neighbourhood U of M_0 in $T(g,n)$ such that, for $M \in U$, (compare (4))

$$syst(M) - syst(M_0) \leq 0,$$

which implies that M_0 is a local maximum of *syst*. This proves $(b) \Longrightarrow (c)$.

(iii) If M_0 is a local maximum of *syst*, then M_0 is a critical point of *syst* and it follows from Corollary 22 that M_0 is an isolated maximum. Therefore, if $\xi \in Tan(M_0)$ with $\xi(u) \geq 0$, $\forall u \in F$, then ξ is identical zero, since otherwise, an earthquake path induced by ξ would, by the convexity of the length functions along earthquake paths, contradict that M_0 is an isolated maximum. Hence $M_0 \in Min(F)$ and $\dim W_F(M_0) = \dim T(g,n)$. q.e.d.

Remark. This theorem, first appeared in Schmutz [24] in a different form without the notion of critical points, is an analogue of Voronoï's theorem in [37] stating that positive definite quadratic forms are extremal if and only if they are critical of index zero (or, in an often used language, if and only if they are eutactic and perfect).

Lemma 26. *Let $M_0 \in T(g,n)$, and let $F = S(M_0)$. Let $M_0 \in Min(F)$ be F-regular, let $k = \dim W_F(M_0)$, and let $N = \dim T(g,n) > k$. Let v be a simple closed geodesic of M_0 such that, putting $G = F \cup \{v\}$, we have that $M_0 \in Conv(G)$ is G-regular with $\dim W_G(M_0) = k+1$. Then there exist a G-regular neighbourhood V of M_0 in $Conv(G)$ and a Lipschitz continuous tangent vector field $\tilde{\eta}$ on V with $\eta_M(v) > 0$, $\forall M \in V$, such that on each integral curve* c *of $\tilde{\eta}$ there is a unique point $c(M) \in Min(F)$. Moreover, this point $c(M)$ is the minimum of* syst *restricted to* c, *and on each connected component of $c \setminus \{c(M)\}$,* syst *is strictly monotonic.*

Proof. (i) By hypothesis we have $Min(F) \subset \partial Min(G)$. Therefore, we can choose V homeomorphic to a $(k+1)$-ball such that $V_1 := V \cap Min(G)$ and $V_2 := V \cap Pos(G)$ are both homeomorphic to a $(k+1)$-ball, with the common boundary $V_0 := V \cap Min(F)$ (which is homeomorphic to a k-ball). We further choose V such that $S(M) \subset F$ for all $M \in V$.

(ii) Let $M \in V_1$. Then $M \in Pos(F)$ so that there exists $\eta \in Tan(M)$ with $\eta(F) < 0$. Since $M \in Min(G)$, it follows that $\eta(v) > 0$. By Theorem 10 and a partition of unity we can therefore choose a differentiable vector field $\tilde{\eta}$ such that $\eta_M(F) < 0$, $\eta_M(v) > 0$, and $\|\eta_M(G)\| = 1$ for all $M \in V_1$.

Let $M \in V_2$. Then there exists $\eta \in Tan(M)$ with $\eta(G) > 0$. By the same argument as above we obtain a differentiable vector field $\tilde{\eta}$ such that $\eta_M(G) > 0$, and $\|\eta_M(G)\| = 1$ for all $M \in V_2$.

Let $M \in V_0$ and let $M_i \longrightarrow M$ be a convergent sequence, $M_i \in V_1$. Then, by Theorem 10, $\eta_{M_i}(G)$ has a convergent subsequence; for the limit vector $\eta_M(G) \in W_G(M)$ we have $\|\eta_M(G)\| = 1$ and $\eta_M(F) \leq 0$. It follows that $\eta_M(F) = 0$ and $\eta_M(v) = 1$.

Let $M \in V_0$ and let $M_i \longrightarrow M$ be a convergent sequence, $M_i \in V_2$. By the same argument, each convergent subsequence of $\eta_{M_i}(G)$

converges to the vector $\eta_M(G) \in W_G(M)$ with $\eta_M(F) = 0$ and $\eta_M(v) = 1$.

We therefore have defined a continuous vector field $\tilde{\eta}$ on V which is differentiable in $V \setminus V_0$.

(iii) Let $\xi = \eta_{M_0}$. To ξ is associated an earthquake, let $\tilde{\xi}$ be the derived vector field. The vector fields derived from earthquakes vary continuously with $M \in T(g,n)$ and $\xi \in Tan(M_0)$ (see [12]). We therefore can choose the vector field $\tilde{\eta}$ and the neighbourhood V such that, for each $M \in V$ and each $u \in G$, the norm of the gradient of $\eta_M(u)$ remains bounded. Thus $\tilde{\eta}$ is Lipschitz continuous and gives rise to a flow.

By construction, the integral curves of this flow are as it is claimed by the lemma. q.e.d.

Lemma 27. *Let F be a set of m simple closed geodesics of a surface $\mathcal{M}$. Let $M_0 \in Min(F) \cup \partial Min(F)$ be F-regular. Let* $\dim W_F(M_0) = k < N = \dim T(\mathcal{M})$. *Then there exists a simple closed geodesic v in $\mathcal{M}$ such that $M_0 \in Conv(G)$ is G-regular with* $\dim W_G(M_0) = k+1$, *where $G = F \cup \{v\}$.*

Proof. By Lemma 5 there exists $F' \subset F$ such that

$$M_0 \in Min(F') \cup \partial Min(F')$$

with $j_{F'}(M_0) = 1$ and $\dim W_{F'}(M_0) = \dim W_F(M_0) = k$. Since $|F'| = k+1$ there exists $w \in F'$ with the following property for $F(w) = F' \setminus \{w\}$: If $\xi \in Tan(M_0)$ with $\xi(F(w)) = 0$, then $\xi(w) = 0$. This implies that M_0 has an F-regular neighbourhood U in $Conv(F)$ such that $\dim W_{F(w)}(M) = k$ for all $M \in U$. Since $F(w)$ has k elements, there exists simple closed geodesics $v_1, \dots, v_{N-k}$ in $\mathcal{M}$ such that, putting $H = F(w) \cup \{v_1, \dots, v_{N-k}\}$, we have $\dim W_H(M) = N$ for all M in an open neighbourhood of M_0 in $T(\mathcal{M})$. Choosing $v = v_1$ the lemma follows. q.e.d.

Theorem 28. *Let $M_0 \in T(g,n)$, let $F = S(M_0)$, let $|F| = m$, and let* $N = \dim T(g,n)$. *If $M_0 \in Min(F)$ is regular, then M_0 is a non–degenerate critical point of* syst *of index $J = N - (m - j_F(M_0))$.*

Proof. Put $k := \dim W_F(M_0) = m - j_F(M_0)$. If $N-k = 0$, then the theorem was already proved by Theorem 25. Therefore, we can assume in the sequel that $N - k > 0$.

We assume that $M_0 \in Min(F)$ is regular.

We shall see that there exist an open neighbourhood V of $M_0 \in T(g,n)$ and a homeomorphic parametrization of V by N parameters $y_1, \dots, y_N$ such that

$$syst(M) - syst(M_0) = \sum_{i=1}^{J} y_i^2 - \sum_{i=J+1}^{N} y_i^2 .$$

This is sufficient in order to prove the theorem since it then follows by Morse [18] that M_0 is a critical point of $syst$.

(i) We choose an F-regular neighbourhood V' of M_0 in $Min(F)$ with a flow as in Lemma 24. We then have a homeomorphic parametrization of $V_0 = V' \setminus \{M_0\}$ by

$$M \longrightarrow \sqrt{syst(M_0) - syst(M)}\, x, \quad M \in V_0,$$

where $x \in \mathsf{R}^k$ is a vector of unit norm. (Recall that V' is chosen such that $S(M) \subset F$, $\forall M \in V'$.)

(ii) Since $N - k > 0$, there exists, by Lemma 27, a simple closed geodesic v of M_0 such that, putting $G = F \cup \{v\}$ we have that $M_0 \in Conv(G)$ is G-regular with $\dim W_G(M_0) = k+1$. We apply Lemma 26. Let V and $\tilde{\eta}$ be defined as in this lemma and choose V so small that $V \cap Min(F) \subset V'$. For $M \in V$ let $c(M)$ be the point in $V \cap Min(F)$ which lies on the same integral curve of $\tilde{\eta}$ as M.

If $N - k = 1$, then we have the desired homeomorphic parametrization of V by

$$M \longrightarrow \left(\sqrt{syst(M) - syst(c(M))}\, y, \sqrt{syst(M_0) - syst(c(M))}\, x\right),$$

where $x \in \mathsf{R}^k$ was defined in (i), and y is a vector of unit norm in R^1.

(iii) Assume that $h := N - k > 1$. Then, by Lemma 27, there exist simple closed geodesics $v_1, \dots, v_h$ of M_0 such that $\dim W_H(M) = N$ for every M in an open neighbourhood W of M_0 in $T(g,n)$, where $H = F \cup \{v_1, \dots, v_h\}$. Let

$$\Lambda = \{\lambda \in \mathsf{R}^h : \lambda \geq 0,\ \|\lambda\| = 1\}.$$

For $\lambda = (\lambda_1, \dots, \lambda_h) \in \Lambda$ and $M \in W$ put

$$L_M(U(\lambda)) = \sum_{i=1}^{h} \lambda_i L_M(v_i),$$

and $F(\lambda) = F \cup \{U(\lambda)\}$ (compare with the remark after Corollary 16 at the end of Section 2).

Let $\lambda, \lambda' \in \Lambda$, $\lambda \neq \lambda'$. For $M \in W \cap Conv(F(\lambda))$ there then exists (by construction) $\xi \in Tan(M)$ with $\xi(F(\lambda)) = 0$ and $\xi(\lambda') > 0$. It follows that $W \cap Conv(F(\lambda)) \cap Conv(F(\lambda')) = W \cap Min(F)$.

(iv) Apply Lemma 26 to F and $F(\lambda)$. Let $V(\lambda)$ be the corresponding $F(\lambda)$-regular neighbourhood of M_0 in $W \cap Conv(F(\lambda))$. Then by (iii) and the invariance of domain we have

$$W \setminus Min(F) = \bigcup_{\lambda \in \Lambda} (W \cap V(\lambda) \setminus Min(F)) \,,$$

and this is a disjoint union. Therefore, if W is chosen accordingly, every $M \in W \setminus Min(F)$ lies in $Conv(F(\lambda))$ for a unique $\lambda \in \Lambda$ and there exists a unique $c(M) \in W \cap Min(F)$ such that M and $c(M)$ lie on the same integral curve induced by the flow on $Conv(F(\lambda))$. We then obtain the desired homeomorphic parametrization of W by

$$M \longrightarrow \left(\sqrt{syst(M) - syst(c(M))}\; y, \sqrt{syst(M_0) - syst(c(M))}\; x \right),$$

where $x \in \mathsf{R}^k$ was defined in (i), and y is in Λ. q.e.d.

Theorem 29. *Let $M_0 \in T(g,n)$, and let $F = S(M_0)$. Let $M_0 \in \partial Min(F)$ be strongly F-regular. Then M_0 is an ordinary point of* syst.

Proof. Let $N = \dim T(g,n)$, let $k = \dim W_F(M_0)$ and $q = N - k$. Let U be an open neighbourhood of M_0 in $T(g,n)$ such that $S(M) \subset F$ for all $M \in U$ and such that $M \in Pos(S(M))$ for every $M \in U \setminus \{M_0\}$ (the latter is possible by Corollary 22). Let $s_0 = syst(M_0)$. Let

$$\mathcal{S}(t) = \{M \in U : syst(M) = s_0 + t\}\,.$$

(i) Assume that M_0 is an ordinary point of $syst$ restricted to $Conv(F)$. If $q = 0$ we are done. So let $q \geq 1$. Let v be a simple closed geodesic of M_0 such that, putting $G = F \cup \{v\}$, we have that M_0 is G-regular with $\dim W_G(M_0) = k+1$ (v exists by Lemma 27). We can choose a G-regular neighbourhood $W \subset U$ of M_0 in $Conv(G)$ such that $W_0 := W \cap Conv(F)$ is homeomorphic to a k-ball which separates W into $W_1 := W \cap Min(G)$ and $W_2 := W \cap Pos(G)$ which both are homeomorphic to a $(k+1)$-ball. On W_2 construct a differentiable vector field $\tilde{\eta}$ (as in the proof of Lemma 26) such that $syst$ is strictly monotonic along the integral curves of the induced flow. Then $\tilde{\eta}$ induces a homotopy from $W \cap \mathcal{S}(t)$ to W_0 if $t \geq 0$ is small enough. Therefore, for $t \geq 0$ small enough, $W \cap \mathcal{S}(t)$ is contractible. Since every $M \in W \setminus \{M_0\}$ is an ordinary point of $syst$, it follows that, for $t \geq 0$ small enough, $W \cap \mathcal{S}(t)$ is homeomorphic to a k-ball (if W is chosen accordingly).

It follows from the proof of Lemma 24 that $W_1 \cap \mathcal{S}(t)$ is homeomorphic to a k-ball for $t < 0$, $|t|$ small enough. Together with the vector field $\tilde{\eta}$ on W_2 this implies that $W \cap \mathcal{S}(t)$ is homeomorphic to a k-ball for $t < 0$, $|t|$ small enough (if W is chosen accordingly). This proves that M_0 is an ordinary point of $syst$ restricted to $Conv(G)$.

It then follows by induction with respect to q that M_0 is an ordinary point of $syst$.

(ii) It remains to prove that M_0 is an ordinary point of $syst$ restricted to $Conv(F)$.

By hypothesis we can choose an F-regular neighbourhood $V \subset U$ of M_0 in $Conv(F)$ such that $V_0 := V \cap \partial Min(F)$ is homeomorphic to a $(k-1)$-ball which separates V into $V_1 := V \cap Min(F)$ and $V_2 := V \cap Pos(F)$ which both are homeomorphic to a k-ball. It follows as in (i) that V can be chosen such that, for $t \geq 0$ small enough, $V \cap \mathcal{S}(t)$ is homeomorphic to a $(k-1)$-ball.

Let $\tilde{\theta}$ be the tangent vector field on V_1 constructed in Lemma 24. By its definition it follows that $\tilde{\theta}$ can also be defined on an open set $Y \subset V$ containing every $M \in V \setminus \{M_0\}$ with $syst(M) \leq s_0$; we continue to denote by $\tilde{\theta}$ this extended vector field. Recall that $syst$ is strictly growing along the integral curves of $\tilde{\theta}$. Since $\mathcal{S}(0)$ is homeomorphic to a $(k-1)$-ball in V, it follows that, for $t < 0$, $|t|$ small enough, $V \cap \mathcal{S}(t)$ is contractible, since $\tilde{\theta}$ induces a homotopy from $V \cap \mathcal{S}(t)$ to $V \cap \mathcal{S}(0)$. Thus as in (i), for every $t < 0$, $|t|$ small enough, $V \cap \mathcal{S}(t)$ is homeomorphic to a $(k-1)$-ball. We have therefore shown that M_0 is an ordinary point of $syst$ restricted to $Conv(F)$. q.e.d.

Theorem 30. *Assume in $T(g,n)$ every M with $M \in Min(S(M))$ is $S(M)$-regular, and every M with $M \in \partial Min(S(M))$ is strongly $S(M)$-regular. Then* syst *is a topological Morse function on $T(g,n)$. Moreover, $M \in T(g,n)$ is a critical point for* syst *if and only if $M \in Min(S(M))$.*

Proof. This follows from Theorems 19, 28, and 29. q.e.d.

Remark. (i) For $n > 0$ and $L > 0$, let $T(g,n;L)$ be the Teichmüller space of all surfaces of genus g which have n boundary components which all are simple closed geodesics of length L. Then Theorem 30 also holds for $T(g,n;L)$ since it is still true (by [34]) that for $M, M' \in T(g,n;L)$, $M \neq M'$, there is an earthquake path from M to M'.

(ii) It is shown in [28] that for every $L \geq 0$, there exists $M(C) \in T(2,1;L)$ such that $F := S(M(C))$ is a set of 9 elements (in [28], this F is called "of type C"). One can show that $M(C)$ is an ordinary point of *syst* for L small while for large L, $M(C)$ is a (non-degenerate) critical point of *syst*. Therefore, by continuity, there exists an L_0 such that $M(C) \in T(2,1;L_0)$ is a boundary point of *syst*. This is the unique example of a boundary point of *syst* which I know.

It is quite probable that for $n > 0$ we can always eliminate boundary points of *syst* on $T(g,n;L)$ by varying L.

4. A Morse function for surfaces with cusps

Definition. (i) Let $\mathcal{M}$ be a (g,n)-surface, $n > 0$. If $n = 1$, denote by O the cusp of $\mathcal{M}$. If $n > 1$, denote by $O, A_1, \dots, A_{n-1}$ the cusps of $\mathcal{M}$.

Let u be a simple closed geodesic of $\mathcal{M}$ which is the boundary of an embedded $(0,3)$-subsurface $Y(u)$ of $\mathcal{M}$ which contains two cusps $\neq O$, the *associated cusps* of u. Denote by R the set of all simple closed geodesics of $\mathcal{M}$ of this type, and by $t(u)$ the unique simple geodesic, contained in $Y(u)$, between the two associated cusps of u.

Let t be a simple geodesic in $\mathcal{M}$, which starts and ends on the same cusp A, with $A \neq O$ if $n > 1$ (and $A = O$ if $n = 1$). Let $Y(t) \subset \mathcal{M}$ be the unique embedded $(0,3)$-subsurface which contains t. Let t' be the unique common orthogonal (in $Y(t)$) between the two boundary components different from A. Put $t(v) = t$ where v is the unique closed geodesic, contained in $Y(t)$, with two self-intersections, which is symmetric with respect to t', and intersects each of t and t' twice. Denote by R' the set of all closed geodesics of $\mathcal{M}$ of this type. Put

$$\mathcal{R} = \mathcal{R}(\mathcal{M}) := R \cup R' .$$

(ii) Define the function $syst_{\mathcal{R}}$ by

$$syst_{\mathcal{R}} : T(\mathcal{M}) \longrightarrow \mathsf{R}_+, \quad M \mapsto \min\{L_M(u) : u \in \mathcal{R}\} .$$

(iii) A subset $F = \{u_1, \dots, u_k\} \subset \mathcal{R}$ is called *admissible* if F fills up and if $t(u_i)$ and $t(u_j)$ have no common inner point, $1 \leq i < j \leq k$.

(iv) Let $\Gamma(g,n)$ be the mapping class group of $\mathcal{M}$ and let $G_O(g,n) \subset \Gamma(g,n)$ be the subgroup (of index n) of elements leaving the cusp O invariant.

Theorem 31. *Let $\mathcal{M}$ be a (g,n)-surface, $n > 0$. Let $F \subset \mathcal{R}(\mathcal{M})$ be an admissible subset of maximal possible order. Then the following hold:*

(i) $|F| = \dim T(g,n) + 1 = 6g + 2n - 5$;

(ii) the length functions of the elements of F parametrize $T(g,n)$;

(iii) $j_F(M) = 1$ *for all* $M \in Min(F)$.

Proof. (i) and (ii) are proved in [27], the proof is straightforward. (iii) then follows from (i) and (ii). q.e.d.

Corollary 32. *Let $\mathcal{M}$ be a (g,n)-surface, $n > 0$. Let $F \subset \mathcal{R}(\mathcal{M})$ be admissible. Then $j_F(M) = 1$ for all $M \in Min(F)$.*

Proof. Clear by Theorem 31. q.e.d.

Theorem 33. *For $n > 0$, $syst_{\mathcal{R}}$ is a topological Morse function on $T(g,n)$, invariant with respect to the group $G_O(g,n)$.*

Proof. Clearly, Theorem 30 holds for $syst_{\mathcal{R}}$ instead of $syst$. Note that if $S_{\mathcal{R}}(M)$ is the set of the shortest elements of $\mathcal{R}$ in $M \in T(g,n)$, then $S_{\mathcal{R}}(M)$ is admissible if $S_{\mathcal{R}}(M)$ fills up. Therefore, the needed regularity holds by Corollary 32, and the theorem follows from Theorem 30. q.e.d.

Remark. It is proved in [27] that, for $n > 0$,

$$T(g,n) = \bigcup_F Min(F)\,, \quad F \subset \mathcal{R} \text{ admissible}\,, \tag{5}$$

and this is a disjoint union. This gives rise to a contractible simplicial complex, invariant with respect to $G_O(g,n)$. The "critical points" of this complex are thus all admissible sets in $\mathcal{R}$. On the other hand, for general pairs (g,n), not for every admissible set $F \subset \mathcal{R}$ there exists $M \in T(g,n)$ such that F is the set of the shortest elements of $\mathcal{R}$ in M. Therefore, $syst_{\mathcal{R}}$ has less critical points than the complex induced by (5).

5. A mass formula and other applications

Definition. Let $\mathcal{M}$ be a (g,n)-surface.

(i) The *mapping class group* $\Gamma(\mathcal{M})$ or $\Gamma(g,n)$ is the group of the isotopy classes of orientation preserving self–homeomorphisms of $\mathcal{M}$ (permutation of the cusps is allowed).

(ii) Denote by $\mathcal{C}$ the set of mutually non–isometric critical points of *syst* in $T(g,n)$. For $M \in \mathcal{C}$, denote by $J(M)$ the index of M (as a non-degenerate critical point of *syst*).

(iii) Let $Isom(M)$ be the automorphism group of $M \in T(g,n)$ containing the orientation preserving isometries. Define

$$|Aut(M)| = \frac{|Isom(M)|}{n!}.$$

(iv) Let Γ be a torsion free subgroup of finite index j in $\Gamma(\mathcal{M})$. Define the *virtual Euler characteristic* of $\Gamma(\mathcal{M})$ by

$$\chi(\Gamma(\mathcal{M})) = \frac{1}{j}\, e(T(\mathcal{M})/\Gamma),$$

where $e(T(\mathcal{M})/\Gamma)$ is the ordinary Euler characteristic of $T(\mathcal{M})/\Gamma$.

Remark. The rational number $\chi(\Gamma(\mathcal{M}))$ does not depend on the choice of Γ; see for example [5].

Theorem 34. *Assume that* syst *is a Morse function of* $T(g,n)$. *Then*

$$\sum_{M\in\mathcal{C}} \frac{(-1)^{J(M)}}{|Aut(M)|} = \chi(\Gamma(g,n)).$$

Proof. By Morse [18], *syst* gives rise to a complex. The theorem now follows by results from the cohomology of groups; see for example Brown [5], Ivanov [10], Serre [30]. q.e.d.

Remark. (i) The formula of Theorem 34 is called a mass formula (in the literature also the term weight formula appears) in analogy to a formula by Siegel [31] in the theory of quadratic forms. This latter formula containing only terms of the same sign; the formula in Theorem 34 is sometimes called a mass formula with sign.

(ii) The virtual Euler characteristic of $\Gamma(g,n)$ is given in Theorem 35; it has been first proved by, independently, Harer/Zagier [8] and Penner [22] (see also Harer [7] and Kontsevich [14]).

Theorem 35. *Let B_k be the k-th Bernoulli number. Then*

(i) $\chi(\Gamma(g,0)) = \frac{B_{2g}}{4g(g-1)}$.

(ii) For $g \geq 2$ we have

$$\chi(\Gamma(g,n)) = (-1)^n B_{2g} \frac{(2g+n-3)!}{2g(2g-2)!}.$$

(iii) $\chi(\Gamma(1,n)) = \frac{(-1)^n (n-1)!}{12}$.

(iv) $\chi(\Gamma(0,n)) = (-1)^{n-1}(n-3)!$.

In Section 6, I shall determine the set $\mathcal{C}$ for some small values of (g,n) calculating thus directly (applying Theorem 34) the number $\chi(\Gamma(g,n))$.

Remark. The denominator of the Bernoulli numbers is well know (see for example [9]). By Theorems 34 and 35, this can give some information about possible orders of automorphism groups of a (g,n)-surfaces $\mathcal{M}$ in general, and of critical points in $T(\mathcal{M})$ in particular (references for related questions are [15],[16]). However, in the formula of Theorem 34, two critical points with the same order of the automorphism group mutually annihilate themselves if their index is different modulo 2.

Theorem 36. *Let M_g be a $(g,0)$-surface such that $Isom(M_g)$ is a quotient of the $(2,4,2g+2)$ triangle group, $g \geq 2$. Then M_g is a non–degenerate critical point of* syst *of index $4g-5$.*

Proof. Fix g for the whole proof.

(i) It is well-known that M_g exists and is unique up to isometry. A fundamental domain P of M_g (respectively of a Fuchsian group associated to M_g) can be constructed as follows. Let P_g be a regular right-angled polygon with $2g+2$ sides. Let X be a vertex of P_g. Take four copies C_i, $i = 1, \ldots, 4$, of P_g and glue them so that they have disjoint interior and that X is a common vertex of them. We thereby obtain a right-angled polygon P with $8g-4$ sides. In order to obtain M_g from P the identifications of the sides of P must be such that for every side u of a copy C_i, $i = 1, \ldots, 4$, there exists a side v in a C_j, $j \in \{1,2,3,4\} \setminus \{i\}$, such that $u \cup v$ is a simple closed geodesic in M_g. Let F be the set of the $2g+2$ simple closed geodesics obtained in this way; they all have the same length $2s$ where s is the length of a side of P_g.

(ii) We now show that F is the set of systoles of M_g. Assume that $u \notin F$ is a systole of M_g. The copies C_i, $i = 1, \dots, 4$, induce a partition of u into k, say, parts $u_1, \dots, u_k$ (the indices respect the natural order of the parts). Each part u_j relates two sides of a copy C_i. If these sides are not neighbours, then $L(u_j) > s$. Assume that u_j and u_{j+1} both relates two sides (of the corresponding copy C_i) which are neighbours. Then $L(u_j) + L(u_{j+1}) > s$. Since $L(u) \le 2s$ it follows that $k \le 3$. If $k = 3$, then the three parts u_j must all relate two sides (of the corresponding copy C_i) which are neighbours. But this is impossible for a simple closed geodesic u in M_g. So we have $k = 2$. But then each part u_j relates two sides (of the corresponding copy C_i) which are not neighbours hence $L(u) > 2s$. This shows that u cannot exist.

(iii) Let $u, v \in F$ such that u and v intersect. Let $F_0 = F \setminus \{u, v\}$. Note that F_0 fills up, but every strict subset of F_0 does not fill up. It follows by Lemma 9 that $\partial Min(F_0) = \emptyset$ and that $Min(F_0)$ is homeomorphic to R^{2g-1}; moreover, $j_{F_0}(M) = 1$ for all $M \in Min(F_0)$. We now show that $M_g \in Min(F_0)$.

For every $w \in F$ there exists an orientation reversing involution ψ_w of M_g which fixes w pointwise. In F, w has two "neighbours" (elements of F which intersect w), therefore, ψ_w fixes them setwise. By induction we obtain that $\psi(w') = w'$ for every $w' \in F$. Let $M \in Min(F_0)$. Since the involutions ψ_w act as elements of the mapping class group, it follows (by the uniqueness in the definition of the map ϕ_F) that M must be invariant with respect to them. Therefore, M contains (like M_g) four copies C_i, $i = 1, \dots, 4$, which are right-angled $(2g + 2)$-gons (with disjoint interior). In other words, M is determined by a right-angled $(2g + 2)$-gon. The space of the right-angled $(2g + 2)$-gons can be parametrized by $2g - 1$ parameters, more precisely, this space is homeomorphic to R^{2g-1} and has empty boundary. Since $Min(F_0)$ is also homeomorphic to R^{2g-1} with empty boundary, the two spaces may be canonically identified and it follows that $M_g \in Min(F_0)$.

(iv) By Theorem 37 in Section 6, M_g is a critical point of *syst*.

It follows by the same argument as in (iii) that $Min(F) = Min(F_0)$. This implies that $j_F(M) = 3$ for all $M \in Min(F)$. Therefore, M_g is a non–degenerate critical point of *syst* of index $4g - 5$. q.e.d.

Definition. The *virtual cohomological dimension vcd* of $\Gamma(g, n)$ is defined as the cohomological dimension *cd* of a torsion free subgroup Γ of $\Gamma(g, n)$ of finite index.

If G is a discrete group, then $cd(G)$ is defined as the supremum of those m such that $H^m(G, A) \neq 0$ for some G-module A (H^m being a cohomology group).

Remark. (i) If G is a discrete group with torsion, then $cd(G) = \infty$. On the other hand, $vcd(\Gamma(g, n))$ is finite and does not depend on the choice of the torsion free subgroup Γ of $\Gamma(g, n)$ of finite index (see for example [30], [5]). The exact value of $vcd(\Gamma(g, n))$ is known for every pair (g, n) by Harer [6] (see also [7] and [10]). In particular, we have

$$vcd(\Gamma(g, 0)) = 4g - 5. \tag{6}$$

(ii) If $syst$ is a topological Morse function of $T(g, 0)$, then the maximal index for a critical point of $syst$ is at least $4g-5$ by Theorem 36. On the other hand, it follows from (6) that there cannot exist a Morse function on $T(g, 0)$, invariant with respect to $\Gamma(g, 0)$, such that the difference between the maximal and the minimal indice of the critical points is smaller than $4g - 5$ (see for example [30]).

(iii) Until now, no contractible simplicial complex on $T(g, 0)$, invariant with respect to $\Gamma(g, 0)$, has been found with the best dimension $4g - 5$ (see for example [7]). It is quite possible that $syst$ gives rise to such a complex (for $g = 2$, this will be proved in Theorem 44).

6. Classification of the critical points of *syst* in some cases

Recall first that if M is a surface with $M \notin Pos(S(M))$, then $S(M)$ must fill up. In closed surfaces or in surfaces with only one cusp, systoles can intersect at most once (see [24]), therefore, if the set of systoles fills up, all systoles are non-separating in these cases.

The following result is taken from [24].

Theorem 37. *Let M be a (g, n)-surface such that $Isom(M)$ is isomorphic to a (non–trivial) quotient of a Fuchsian triangle group. Let $S(M)$ have a subset which fills up and on which $Isom(M)$ acts transitively. Then M is a critical point for* syst.

Remark. The theorem holds without the hypothesis that $S(M)$ has a subset which fills up and on which $Isom(M)$ acts transitively. Namely, assume that there exists $\xi \in Tan(M)$ such that $\xi(S(M)) \geq 0$. Let

$$\xi_0 = \sum_\gamma \gamma(\xi),$$

where γ runs over $Isom(M)$. Then ξ_0 is invariant with respect to $Isom(M)$ and $\xi_0(S(M)) \neq 0$. This implies that the earthquake path induced by ξ_0 is also invariant with respect to $Isom(M)$. But this is impossible since the points with the same automorphism group as M are isolated in the Teichmüller space of M. I thank Christophe Bavard for helping me with this argument.

Theorem 38. *In $T(1,1)$, syst has exactly two non–isometric critical points M_0 and M_1 which moreover are non–degenerate and of index 0 and 1, respectively. M_0 has three systoles and corresponds to the modular torus. M_1 has two systoles which intersect orthogonally; $Isom(M_1)$ is a quotient of the Fuchsian triangle group $(2,4,\infty)$.*

Proof. By [24] M_0 is the surface of $T(1,1)$ with the longest systole and is therefore a critical point of *syst*; moreover, M_0 is non-degenerate by Theorem 25.

All other (non–isometric to M_0) surfaces of $T(1,1)$ have at most two systoles. For any further $M \in T(1,1)$ with $S(M) \notin Pos(S(M))$, $S(M)$ has two elements which, moreover, must intersect orthogonally.

$M = M_1$ is determined by this condition; moreover, by Theorem 37, M_1 is critical. Non-degeneracy of M_1 is clear since $j_{S(M_1)}(M) = 1$ for all $M \in Min(S(M_1))$. q.e.d.

Remark. For the Euclidean tori (of fixed volume), *syst* has the analogous critical points as in the case of $(1,1)$-surfaces, namely, the hexagonal torus corresponds to M_0 and the square torus corresponds to M_1 of Theorem 38.

In the introduction I mentioned the function $-\log\det\Delta(M)$ and the conjecture that it is a Morse function. The conjecture is true for Euclidean tori, and the hexagonal torus and the square torus are the unique critical points for this function (see [23]).

Corollary 39. *For $T(1,1)$, syst is a topological Morse function and the mass formula of Theorem 34 gives*

$$\frac{1}{6} - \frac{1}{4} = -\frac{1}{12} = \chi(\Gamma(1,1))\,.$$

Proof. That *syst* is a Morse function, follows by the proof of Theorem 38 and by Theorem 30. Note further that $|Isom(M_0)| = 6$ and $|Isom(M_1)| = 4$ where M_0 and M_1 are defined as in Theorem 38.

q.e.d.

Corollary 40. *(i) For $T(0,4)$,* syst *has exactly two non–isometric critical points M_0 and M_1 which moreover are non-degenerate and of indices* 0 *and* 1*, respectively. M_0 has three systoles and corresponds to $\mathsf{H}/\Gamma(3)$, where H is the upper halfplane and $\Gamma(3)$ is the principal congruence subgroup of level 3 of the modular group. M_1 has two systoles which intersect orthogonally; $Isom(M_1)$ is a quotient of the Fuchsian triangle group $(2,4,\infty)$.*

(ii) In the case of $(0,4)$-surfaces syst *is a topological Morse function, and the mass formula of Theorem 34 gives*

$$2-3=-1=\chi(\Gamma(0,4))\,.$$

Proof. There is a well–known isomorphism between $T(1,1)$ and $T(0,4)$, induced by an isomorphism between the simple closed geodesics of a $(1,1)$-surface and the simple closed geodesics of a $(0,4)$-surface. This, together with Theorem 38, proves (i). Concerning (ii) note that $|Isom(M_0)|=12$ and $|Isom(M_1)|=8$. q.e.d.

Theorem 41. *Let f be a (topological) Morse function on $T(1,1)$ which is invariant with respect to $\Gamma(1,1)$. Then M_0 and M_1 (defined as in Theorem 38) are critical points of f. Moreover, M_0 is a minimum or a maximum for f while M_1 is extremal or not extremal, depending on f.*

Proof. (i) Since f is a Morse function, f induces a complex from which we can calculate $\chi(\Gamma(1,1))$ as in Corollary 39. Let $M\in T(1,1)$, M not isometric to M_0 nor to M_1. Then it is well known that $|Isom(M)|=2$. Since $\chi(\Gamma(1,1))=-1/12$, it follows that M_0 and M_1 must be critical points of f. Moreover, in the formula of Theorem 34, the parity of the index of M_0 must be even which implies that M_0 is a maximum or a minimum for f.

(ii) If $f=syst$, then the index of M_1 is 1. Let now f be the following function.

$$f:T(1,1)\longrightarrow \mathsf{R}_+,\quad M\mapsto \min\{L_M(u)+L_M(v):u,v\},$$

where u,v are (any) simple closed geodesics of $\mathcal{M}\in T(1,1)$ which intersect exactly once. Let us look for the critical points of f. Obviously, M_0 is a (local) maximum of f while M_1 is the global minimum of f. There is a third critical point M_f of f defined as follows. Let

$S(M_0) = \{u_1, u_2, u_3\}$. Let $0 < \alpha_i \leq \pi/2$ be the intersection angle between u_i and u_{i+1}, $i = 1, 2, 3$, (taking the indices modulo 3). Then M_f is given by

$$\cos \alpha_1 = 2 \cos \alpha_2 = 2 \cos \alpha_3 \,.$$

It is easy to verify that M_f is a critical point of f of index 1 and that f has no further critical points (or points corresponding to the boundary points of *syst*). The proof of Theorem 30 applies to f and clearly M_0, M_1 and M_f are non-degenerate. Therefore, f is a topological Morse function of $T(1,1)$, invariant with respect to $\Gamma(1,1)$. The existence of this f (and of *syst*) shows that the parity of the index of M_1 as a critical point is not constant. q.e.d.

Remark. The phenomenon described in Theorem 41 is surprising. There are surfaces which are critical points for all Morse functions (which are invariant with respect to the corresponding mapping class group). But some of them are *stable* in the sense that the parity of their index does not change while other are *unstable* (the parity of the index changes). What lies behind this stability property?

Theorem 42. *In $T(1,2)$,* syst *has exactly five mutually nonisometric critical points M_i, $i = 0, \dots, 4$, which moreover are non–degenerate, namely,*

- M_0 with five systoles and index 0;

- M_1 with four systoles a_1, a_2, b_1, b_2 such that the a_i are mutually disjoint and the b_i are mutually disjoint, $i = 1, 2$; the index of M_1 is 1;

- M_2 with four systoles such that two of them intersect all other systoles; the index of M_2 is 1;

- M_3 with three systoles which all mutually intersect; the index of M_3 is 2.

- M_4 with three systoles a, b, c such that a and c are disjoint; the index of M_4 is 2.

Proof. (i) Every $M \in T(1,2)$ is hyperelliptic and each non-separating simple closed geodesic u can be described by two numbers in the set $\{1, 2, 3, 4\}$ corresponding to the two fixed points on u of the hyperelliptic involution. Two systoles in M can intersect at most once (compare [24]); therefore, if the set of systoles fills up, then the systoles are non-separating and different systoles cannot have two common fixed points of the hyperelliptic involution.

(ii) M_0 is the unique point in $T(1,2)$ with five systoles and has moreover the maximal length of the systole; see [24]. By Theorem 25, M_0 is non-degenerate.

(iii) If $S(M)$ fills up and has four systoles a,b,c,d, then there are two combinatorial possibilities, namely $\{a,b,c,d\} = \{12,34,13,24\}$ which corresponds to M_1 and $\{a,b,c,d\} = \{12,13,23,34\}$ which corresponds to M_2 (we use the notation defined in (i)). Since we assume that $M \notin Pos(S(M))$ it follows that in the first case, the angles in the four intersection points must all be equal which determine M_1 uniquely. In the second case, the angles in the intersection points of d with b and c must be equal determine M_2 uniquely.

$Isom(M_1)$ is a quotient of the $(2,4,\infty)$ triangle group and therefore critical by Theorem 37. That M_2 is critical can be verified by a calculation.

(iv) A set of two systoles cannot fill up, so we remain with the possibility that the set of systoles fills up and has three elements a,b,c. There are two combinatorial possibilities, namely $\{a,b,c\} = \{12,13,14\}$ which corresponds to M_3 and $\{a,b,c\} = \{12,23,34\}$ which corresponds to M_4. In the first case, the angles in the intersection point (the three systoles intersect in the same point) must all be equal which determine M_3 uniquely. In the second case, the angles in the intersection points of a with b and of b with c must be orthogonal which determine M_4 uniquely.

That M_3 and M_4 are critical can be easily verified.

(v) Note that by (i), $j_{S(M_0)}(M) = 1$ for all $M \in Min(S(M_0))$. Since $S(M_0)$ contains $S(M_i)$, $i = 1, \dots, 4$, it follows that $j_{S(M_i)}(M) = 1$ for all $M \in Min(S(M_i))$, $i = 1, \dots, 4$. This proves that all five critical points are non-degenerate and that their index is as claimed. q.e.d.

Corollary 43. *For $T(1,2)$,* syst *is a topological Morse function and the mass formula of Theorem 34 gives*

$$\frac{1}{2} - \frac{1}{4} - 1 + \frac{1}{3} + \frac{1}{2} = \frac{1}{12} = \chi(\Gamma(1,2)).$$

Proof. Clear by the proof of Theorem 42 and by Theorem 30. (Note that we have $|Isom(M_0)| = |Isom(M_4)| = 4$, $|Isom(M_1)| = 8$, $|Isom(M_3)| = 6$, and $|Isom(M_2)| = 2$.) q.e.d.

Theorem 44. *In $T(2,0)$,* syst *has exactly four mutually nonisometric critical points M_i, $i = 0, \dots, 3$, which moreover are non-degenerate, namely,*

- M_0 *with* 12 *systoles and index* 0*;*
- M_1 *with* 9 *systoles and index* 1*;*
- M_2 *with* 6 *systoles and index* 3*;*
- M_3 *with* 5 *systoles and index* 2.

Proof. Surfaces of genus 2 are hyperelliptic, and the systoles can be characterized by two numbers in the set $\{1,2,3,4,5,6\}$ corresponding to the two fixed points of the hyperelliptic involution on the systole.

(i) By [24] $T(2,0)$ has a unique surface M_0 with maximal length of the systoles. It has 12 systoles and $|Isom(M_0)| = 48$. M_0 is non–degenerate by Theorem 25.

(ii) $T(2,0)$ has a one-parameter family of surfaces with a set of 9 systoles (which can be chosen as the set $\{12,13,23,56,46,45,14,25,36\}$) which contains a point M_1 characterized by the following geometric property. M_1 has a separating simple closed geodesic z which intersects the systoles 14, 25, 36 orthogonally in both intersection points and is disjoint to the other systoles. The automorphism group of M_1 has order 12 (the two subsurfaces separated by z have the same automorphism group of order 6 as M_0 in the case of $(1,1)$-surfaces; moreover, there is an automorphism interchanging these two subsurfaces).

One easily verifies that M_1 is critical for *syst*.

That M_1 is non-degenerate will be proved in Theorem 47 below.

(iii) $T(2,0)$ contains a surface M_2, isometric to the surface also denoted by M_2 in Theorem 36. Therefore, M_2 is critical and non-degenerate of index 3 and has an automorphism group of order 24 (which is a quotient of the $(2,4,6)$ triangle group).

(iv) $T(2,0)$ has a surface M_3 with an automorphism group of order 10 which is a quotient of the $(2,5,10)$ triangle group. M_3 has a set of 5 systoles $\{12,23,34,45,15\}$ and the angles in all intersection points are equal. By Theorem 37, M_3 is a critical point of *syst*.

It follows from [27] that $j_{S(M_3)}(M) = 1$ for all $M \in Min(S(M_3))$. Therefore, M_3 is non-degenerate of index 2.

(v) A set of three systoles cannot fill up. If $S(M)$ fills up and has four elements $\{a,b,c,d\}$ such that $M \notin Pos(S(M))$, then $S(M)$ must be of the form $\{12,23,34,45\}$ and all angles in the intersection points are orthogonal. In other words, $\{a,b,c,d\}$ must be a subset of $S(M_2)$, but then the surface is automatically M_2 by the proof of Theorem 36.

If $S(M)$ fills up and has five elements such that $M \notin Pos(S(M))$, then, by a similar argument, $M = M_3$. Thus by [25], when $M \in T(2,0)$ with $M \notin Pos(S(M))$ such that $|S(M)| \geq 6$, M is isometric to one of

M_i, $i = 0, 1, 2$. q.e.d.

Corollary 45. *For* $T(2,0)$, syst *is a topological Morse function and the mass formula of Theorem 34 gives*

$$\frac{1}{48} - \frac{1}{12} - \frac{1}{24} + \frac{1}{10} = -\frac{1}{240} = \chi(\Gamma(2,0))\,.$$

Corollary 46. *(i) In* $T(0,6)$, syst *has exactly four mutually non-isometric critical points* M_i, $i = 0, \dots, 3$, *which moreover are non-degenerate, namely,*

- M_0 *with* 12 *systoles and index* 0, *with the automorphism group of the octahedron;*

- M_1 *with* 9 *systoles and index* 1, *with the automorphism group of a regular prism;*

- M_2 *with* 6 *systoles and index* 3, *with an isometry of order 6 (*$Isom(M_2)$ *is a quotient of the* $(2, 6, \infty)$ *triangle group);*

- M_3 *with* 5 *systoles and index* 2, *with an isometry of order 5 (*$Isom(M_3)$ *is a quotient of the* $(5, \infty, \infty)$ *triangle group).*

(ii) For $T(0,6)$, syst *is a topological Morse function, and the mass formula of Theorem 34 gives*

$$\frac{720}{24} - \frac{720}{6} - \frac{720}{12} + \frac{720}{5} = -6 = \chi(\Gamma(0,6))\,.$$

Proof. Let $M \in T(0,6)$ such that $S(M)$ fills up. Let $u \in S(M)$ and assume that u separates M into two $(0,4)$-subsurfaces. But as one verifies by a calculation, this is impossible since u is intersected by another systole (instead of a calculation one may also use the surface M_1). Therefore, every $u \in S(M)$ separates M into a $(0,5)$-subsurface and into a $(0,3)$-subsurface. The latter contains two different cusps A_1 and A_2. Therefore, systoles in M can be characterized by these two cusps. In other words, if A_i, $i = 1, \dots, 6$, are the cusps of M, then the elements of $S(M)$ can be characterized by (i,j) with $1 \leq i < j \leq 6$. This is the analogous characterization as for systoles in $(2,0)$-surfaces which are non-separating. Therefore, the well known isometry between $T(0,6)$ and $T(2,0)$ maps critical points of *syst* to critical points of *syst* and vice versa.

The corollary then follows from Theorem 44 (note that surfaces in $T(0,6)$ do not have a hyperelliptic involution which fixes all simple closed geodesics: therefore $|Isom(M_i)|$, $M_i \in T(0,6)$, is only half of $|Isom(M_i)|$, $M_i \in T(2,0)$, $i = 0, \dots, 3$). q.e.d.

Theorem 47. *Let $M_1 \in T(2,0)$ be defined as in the proof of Theorem 44. Let $F = S(M_1)$. Then:*
(a) M_1 is a non-degenerate critical point of syst *of index 1,*
(b) M_1 is not F-regular.

Proof. (i) Let

$$F = \{(12), (13), (23), (56), (46), (45), (14), (25), (36)\}$$

as in the proof of Theorem 44. Let

$$F \supset G = \{(12), (13), (23), (45), (46), (56)\}$$

and let $F \supset H = \{(14), (25), (36)\}$. Let z be the (separating) simple closed geodesic of M_1 which is disjoint to all elements of G. Let $F \supset F_1 = G \cup \{(14)\}$, let $F \supset F_2 = G \cup \{(25)\}$, let $F \supset F_3 = G \cup \{(36)\}$.

Let $M \in Min(F_1)$. Then in M, z must intersect orthogonally the element $(14) \in F_1$ (note that the two intersection angles of z with (14) are always equal since surfaces in $T(2,0)$ are hyperelliptic). Since $\xi_z(F_1) = 0$ and since on the other hand, the length functions of the elements of F_1 parametrize the surface up to the twist along z, it follows that $\dim W_{F_1}(M) = 5$ for all $M \in Min(F_1)$ (recall that $\dim T(2,0) = 6$). The same argument shows that $\dim W_{F_i}(M) = 5$ for all $M \in Min(F_i)$, $i = 2, 3$.

Assume now that M_1 is F-regular. Then there exists an F-regular neighbourhood U of M_1 in $Min(F)$ such that

$$Min(F_1) \cap U = Min(F_2) \cap U = Min(F_3) \cap U,$$

since $\dim W_F(M) = \dim W_{F_i}(M) = 5$ for all $M \in U$, $i = 1, 2, 3$. Therefore, in $M \in U$, z intersects orthogonally all elements of H. But the subspace of $T(2,0)$ with this property has only dimension 3, a contradiction. Therefore, M_1 is not F-regular which proves (b).

(ii) For every $M \in T(2,0)$ let $\alpha(M)$ be one of the two equal directed angles from z to the geodesic $(14) \in F$, $0 < \alpha(M) < \pi$. For every α $(0 < \alpha < \pi)$ put

$$T(\alpha) = \{M \in T(2,0) : \alpha(M) = \alpha\}.$$

Since $\alpha(M)$ can be used as a parameter in $T(2,0)$, it is clear that $T(\alpha) \subset T(2,0)$ is a differentiable submanifold of dimension 5.

Let $\mathcal{F} \subset T(2,0)$ be the one-parameter family containing the surfaces where all elements of F have the same length. Then for every $0 < \alpha < \pi$, there is a unique element $M(\alpha)$ in $\mathcal{F} \cap T(\alpha)$.

Let U be an open neighbourhood of M_1 in $T(2,0)$. Assume that $syst$ restricted to $U \cap T(\alpha)$ has its unique maximum in $M(\alpha)$. As in Lemma 24 we then construct a differentiable tangent vector field $\tilde{\theta}$ on $U \cap T(\alpha) \setminus \{M(\alpha)\}$ such that $syst$ is strictly growing along the integral curves induced by $\tilde{\theta}$ which moreover all end in $M(\alpha)$. Let $\epsilon > 0$. If this holds for every $\alpha \in [\pi/2 - \epsilon, \pi/2 + \epsilon]$, then we obtain the desired homeomorphic parametrization of U by putting, for $M \in U \cap T(\alpha)$,

$$M \longrightarrow \left(\sqrt{syst(M(\alpha)) - syst(M_1)}\; x, \sqrt{syst(M(\alpha)) - syst(M)}\; y\right),$$

where $x = \pm 1$ (depending on the sign of $\pi/2 - \alpha$), and $y \in \mathsf{R}^5$ is a vector of norm 1, induced by the flow of $\tilde{\theta}$ on $T(\alpha) \cap U$. This proves (a).

(iii) It remains to show that there exists an $\epsilon > 0$ such that for all $\alpha \in [\pi/2 - \epsilon, \pi/2 + \epsilon]$, $M(\alpha)$ is the unique maximum of $syst$ restricted to $T(\alpha) \cap U$.

For $\alpha = \pi/2$ we can choose U so that $U \cap T(\pi/2) \subset Min(F_1)$ (compare (i)). It follows that M_1 is the unique maximum of $syst$ restricted to $U \cap T(\pi/2)$. Fix now α so that $|\alpha - \pi/2| = \epsilon' > 0$ is small. In $M(\alpha)$ execute a twist deformation along z such that we obtain $N_0 \in T(\pi/2)$. For ϵ' small enough, N_0 is in $Min(F_1)$. Let $M' \in U \cap T(\alpha)$. Execute a twist deformation along z in M' such that we obtain $N_1 \in T(\pi/2)$. Again, we may assume that $N_1 \in Min(F_1)$. Therefore, there exists $u \in F_1$ such that $L_{N_1}(u) < L_{N_0}(u)$. If $u \in G$ we are done. Thus, we may assume that $u = (14)$. Let $4t' = L_{N_1}(u)$ and $4t = L_{N_0}(u)$. We then have

$$\sinh(L_{M'}(u)/4) = \frac{\sinh t'}{\sin\alpha}, \quad \sinh(L_{M(\alpha)}(u)/4) = \frac{\sinh t}{\sin\alpha}.$$

Since $t' < t$, u is smaller in M' than in $M(\alpha)$ and we are done. q.e.d.

Theorem 48. *In $T(0,5)$, there are, up to isometry, exactly three different critical points of* syst, *which moreover all are non-degenerate, namely*

M_0 with 6 systoles and index 0:
M_1 with 4 systoles and index 1;
M_2 with 5 systoles and index 2.

Proof. (i) Denote by A_i the cusps of $M \in T(0,5)$, $i = 1, \dots, 5$. Every simple closed geodesic u of M can be characterized by the two cusps which lie in the $(0,3)$-subsurface $Y(u)$ of M of which u is a boundary geodesic. We note $u = (ij)$ if these two cusps are A_i and A_j, $1 \leq i < j \leq 5$. We then call A_i, A_j the associated cusps of u. Clearly, two systoles of M cannot have both the same two associated cusps.

(ii) By [25] there exists $M_0 \in T(0,5)$ with 6 systoles which is a local maximum for *syst*. Therefore, M_0 is critical and non-degenerate by Theorem 25.

Let $M_1 \in T(0,5)$ have the isometry group of a Euclidean pyramid with a square as basis. Then M_1 is critical for *syst* by Theorem 37 ($Isom(M_1)$ is a quotient of the $(4, \infty, \infty)$ triangle group). A calculation yields that $S(M_1)$ can be described as

$$S(M_1) = \{(12), (23), (34), (14)\}.$$

Since $S(M_1)$ is admissible in the sense of Section 4, it follows by Corollary 32 that M_1 is a non-degenerate critical point for *syst* of index 1.

Let $M_2 \in T(0,5)$ such that M_2 has an isometry of order 5. It follows by Theorem 37 that M_2 is critical for *syst* ($Isom(M_2)$ is a quotient of the $(2, 5, \infty)$ triangle group); moreover, $F_2 := S(M_2)$ can be described as

$$F_2 = \{(12), (23), (34), (45), (15)\}.$$

Let $M \in Min(F_2)$. There exists an orientation reversing involution $\phi \in \Gamma(0,5)$ with $\phi(u) = u$ for all $u \in F_2$. It follows that ϕ is an isometry of M and that $\xi_u(v) = 0$ for all $u, v \in F_2$. Let $u, v \in F_2$ be disjoint. Then $\xi_u, \xi_v \in Tan(M)$ are linearly independent and $\xi_u(F_2) = \xi_v(F_2) = 0$ which shows that $j_{F_2}(M) \geq 3$. Since $j_{F_2}(M_2) = 3$ as it is easy to see, it follows that M_2 is a non-degenerate critical point of *syst* of index 2.

(iii) It remains to show that for all other $M \in T(0,5)$ we have $M \in Pos(S(M))$. So assume that $S(M)$ fills up. Then $S(M)$ has at least three elements. If every cusp of M is associated cusp for at most two different elements of $S(M)$, then clearly $S(M) \subset S(M_1)$ or $S(M) \subset S(M_2)$. In the latter case we must have equality since $S(M)$ fills up. If $S(M) \subset S(M_1)$, then ϕ is an isometry of M and it is easy to see that $S(M) = S(M_1)$.

By the existence of M_1 it is not possible that $S(M)$ has four elements which have the same associated cusp.

We thus can assume that $S(M)$ has three elements which have the same associated cusp. A calculation yields that $S(M)$ must be a subset of $S(M_0)$ and it is then easy to see that $S(M) = S(M_0)$. q.e.d.

Corollary 49. *For* $T(0,5)$, syst *is a topological Morse. The mass formula of Theorem 34 gives*

$$\frac{120}{6} - \frac{120}{4} + \frac{120}{10} = 2 = \chi(\Gamma(0,5))\,.$$

Corollary 50. *Among all* $(0,5)$*-surfaces,* M_0 *(defined as in Theorem 48) has the longest possible systole.*

Theorem 51. *In the above treated cases the following critical points of* syst *correspond to arithmetic Fuchsian groups.*

(a) Both critical points in the case $(g,n) = (1,1)$ *as well as in the case* $(g,n) = (0,4)$*,*

(b) the surfaces M_1 *and* M_3 *in the case* $(g,n) = (1,2)$*,*

(c) all four critical points in the case $(g,n) = (2,0)$*,*

(d) the surfaces M_0 *and* M_2 *in the case* $(g,n) = (0,6)$*,*

(e) the surfaces M_0 *and* M_1 *in the case* $(g,n) = (0,5)$*.*

Proof. (i) If a surface M has cusps and a corresponding Fuchsian group is arithmetic, then this group has only elements with traces which are square roots of rational integers (compare [32]). By a calculation, the systoles of the surfaces M_0, M_2, M_4 in the case $(g,n) = (1,2)$ as well as the surfaces M_1, M_3 in the case $(g,n) = (0,6)$ and the surface M_2 in the case $(g,n) = (0,5)$ are not of this form and hence the corresponding Fuchsian groups are not arithmetic.

(ii) As already mentioned, M_0 and M_1 in the case $(g,n) = (1,1)$ (as well as in the case $(g,n) = (0,4)$) correspond to Fuchsian groups which are subgroups of the modular group and the $(2,4,\infty)$ triangle group, respectively. M_1 in the case $(g,n) = (1,2)$ also corresponds to a Fuchsian group which is a subgroup of the $(2,4,\infty)$ triangle group. In the case $(g,n) = (2,0)$, the surfaces M_0, M_2, M_3 have a corresponding Fuchsian group which is subgroup of the $(2,3,8)$, $(2,4,6)$, $(2,5,10)$ triangle group, respectively. If $(g,n) = (0,6)$ then M_0 and M_2 have a corresponding Fuchsian group which is a subgroup of the $(2,3,\infty)$ and the $(2,6,\infty)$ triangle group, respectively. If $(g,n) = (0,5)$, then M_0, M_1 correspond to a Fuchsian group which is a subgroup of the $(2,\infty,\infty)$,

$(4, \infty, \infty)$ triangle group, respectively. All these triangle groups (and their subgroups) are arithmetic by Takeuchi [33].

(iii) In order to show that all traces of the elements of a Fuchsian group are of the necessary form (described in [32]) to give an arithmetic group, it is sufficient to control this for a system of generating elements (of the Fuchsian group) and some relations. A calculation proves that M_3 in the case $(g, n) = (1, 2)$ and M_1 in the case $(g, n) = (2, 0)$ are arithmetic. q.e.d.

References

[1] A. Ash, *On eutactic forms,* Canad. J. Math. **29** (1977) 1040–1054.

[2] ———, *On the existence of eutactic forms,* Bull. London Math. Soc. **12** (1980) 192–196.

[3] C. Bavard, *Systole et invariant d'Hermite,* J. Reine Angew. Math. **482** (1997) 93–120.

[4] ———, *Familles hyperboliques de réseaux symplectiques,* Preprint.

[5] K. S. Brown, *Cohomology of groups.* Springer, Berlin, 1982.

[6] J. L. Harer, *The virtual cohomological dimension of the mapping class group of an orientable surface,* Invent. Math. **84** (1986) 157–176.

[7] ———, *The cohomology of moduli space of curves,* Theory of moduli, (Sernesi ed.), Lecture Notes in Math., Springer, Berlin, **1337** (1988) 138–221.

[8] J. L. Harer & D. Zagier, *The Euler characteristic of the moduli space of curves,* Invent. Math. **85** (1986) 457–485.

[9] K. Ireland & M. Rosen, *A classical introduction to modern number theory,* Springer, Berlin, 1982.

[10] N. V. Ivanov., *Complexes of curves and the Teichmüller modular group,* Uspekhi Mat. Nauk **42** (1987) 43–91; English transl. Russian Math. Surveys **42** (1987) 55–107.

[11] S. P. Kerckhoff, *The Nielsen realization problem,* Ann. of Math. **117** (1983) 235–265.

[12] ———, *Earthquakes are analytic,* Comment. Math. Helv. **60** (1985) 17–30.

[13] ———, *Lines of minima in Teichmüller space,* Duke Math. J. **65** (1992) 187–213.

[14] M. Kontsevich, *Intersection theory on the moduli space of curves and the matrix Airy function.* Comm. Math. Phys. **147** (1992) 1–23.

[15] R. S. Kulkarni, *Symmetries of surfaces,* Topology **26** (1987) 195–203.

[16] R.S. Kulkarni & C. Maclachlan, *Cyclic p-groups of symmetries of surfaces,* Glasgow Math. J. **33** (1991) 213–221.

[17] J. Martinet, *Les réseaux parfaits des espaces euclidiens,* Masson, Paris 1996.

[18] M. Morse, *Topologically non-degenerate functions on a compact n-manifold M,* J. Anal. Math. **7** (1959) 189–208.

[19] D. Mumford, *A remark on Mahler's compactness theorem,* Proc. Amer. Math. Soc. **28** (1971) 289–294.

[20] B. Osgood, R. Phillip &; P. Sarnak, *Extremals of determinants of Laplacians,* J. Funct. Anal. **80** (1988) 148–211.

[21] ———, *Moduli space, heights and isospectral sets of plane domains,* Ann. of Math. **129** (1989) 293–362.

[22] R. C. Penner, *Perturbative series and the moduli space of Riemann surfaces.* J. Differential Geom. **27** (1988) 35–53.

[23] P. Sarnak, *Extremal geometries,* Extremal Riemann surfaces, (Quine/Sarnak eds.), Contemp. Math. Amer. Math. Soc. **201** (1997) 1–7.

[24] P. Schmutz, *Riemann surfaces with shortest geodesic of maximal length,* Geom. Funct. Anal. **3** (1993, 564–631.

[25] ———, *Systoles of Riemann surfaces,* Manuscripta Math. **85** (1994) 429–447.

[26] P. Schmutz Schaller, *Geometry of Riemann surfaces based on closed geodesics,* Bull. Amer. Math. Soc. **35** (1998) 193–214.

[27] ———, *A cell decomposition of Teichmüller space based on geodesic length functions,* Preprint.

[28] ———, *A systolic geometric cell decomposition for the space of once-holed Riemann surfaces of genus 2,* To appear in Topology.

[29] ———, *Riemann surfaces with longest systole and an improved Voronoï algorithm,* To appear in Arch. Math. (Basel).

[30] J.P. Serre, *Cohomologie des groupes discrets,* Prospects in Mathematics, Ann. of Math. Stud. **70** (1971) 77–169.

[31] C.L. Siegel, *On the theory of indefinite quadratic forms,* Ann. of Math. **45** (1944) 577–622; Gesammelte Abhandlungen II, Springer, Berlin, 1966, 421–466.

[32] K. Takeuchi, *A characterization of arithmetic Fuchsian groups,* J. Math. Soc. Japan **27** (1975) 600–612.

[33] ———, *Arithmetic triangle groups,* J. Math. Soc. Japan **29** (1977) 91–106.

[34] W. P. Thurston, *Earthquakes in two-dimensional hyperbolic geometry,* Low-dimensional topology and Kleinian groups, (Epstein ed.), Cambridge Univ. Press, 1986, 91–112.

[35] ———, *A spine for Teichmüller space,* Preprint, 1986.

[36] A. J. Tromba, *Teichmüller theory in Riemannian geometry,* Birkhäuser, Boston, 1992.

[37] G. Voronoï, *Sur quelques propriétés des formes quadratiques positives parfaites,* J. Reine Angew. Math. **133** (1908) 97–178.

[38] M. Wolf, *The Teichmüller theory of harmonic maps,* J. Differential Geom. **29** (1989) 449–479.

[39] S.A. Wolpert, *Geodesic length functions and the Nielsen problem,* J. Differential Geom. **25** (1987) 275–296.

Université de Genève