Topology and its Applications 30 (1988) 63-88
North-Holland

MEASURED LAMINATION SPACES FOR SURFACES, FROM THE TOPOLOGICAL VIEWPOINT

A.E. HATCHER
Department of Mathematics, Cornell University, Ithaca, NY 14853-7901, USA

Received 17 February 1987
Revised 29 June 1987

This is an expository paper, developing the basic structure of Thurston's space of measured laminations in a surface.

AMS (MOS) Subj. Class.: 57-02, 57M99
Thurston measured lamination

A significant part of Thurston's fundamental theory of measured laminations in surfaces is of a purely topological nature, not involving hyperbolic geometry. The aim of this paper is to derive much of this topological theory topologically, in as simple and direct a way as possible. Few if any of the results here are new, though some of the proofs may be. While the topological approach in dimension two has the virtue of economy, in one higher dimension it becomes a necessity, as the measured lamination space of a 3-manifold is, it seems, exclusively a topological object; see [6].

Here is a quick outline of the contents of the paper. Starting with the classical notion of curve systems on a compact surface M, we construct, via standard train tracks associated to a decomposition of M into pairs of pants, a polyhedron $\mathrm{ML}(M)$ whose "integer" points are exactly the isotopy classes of curve systems in M. (In the language of Dehn's remarkably foresightful 1938 paper [4], these integer points are the "arithmetic field" of M.) The global structure of $\mathrm{ML}(M)$, as well as that of its projectivization $\mathrm{PL}(M)$, is then determined by a simple inductive argument. Next, the noninteger points of $\mathrm{ML}(M)$ are interpreted as measured laminations in M, and the piecewise linear structure of $\mathrm{ML}(M)$ is shown to be independent of the choice of standard tracks by studying the equivalent "functional" topology on $\mathrm{ML}(M)$ obtained by viewing measured laminations as length functionals on closed curves in M. This yields Thurston's original definition of $\mathrm{PL}(M)$ as the completion of the projective curve systems in M. As an application, we deduce a weak form of Thurston's structure theorem for diffeomorphisms of M in case $\partial M \neq \emptyset$. Finally, we derive the basic relation between measured laminations in M and actions of $\pi_1 M$ on $\mathbb{R}$-trees (by isometries): A measured lamination has a dual $\mathbb{R}$-tree with

$\pi_1 M$-action, and a $\pi_1 M$-action on an $\mathbb{R}$-tree can be pulled back (nonuniquely) to a measured lamination in M.

The exposition here has inevitably a certain amount in common with earlier accounts, such as [12], [5], [7], [2] and [10]. The reader should consult these for alternative viewpoints, and especially for the important application of PL(M) to compactifying the Teichmüller space of M and for the full strength of Thurston's structure theorem for diffeomorphisms of M.

1. Curve systems

The complex of curve systems

Let M be a compact surface, possibly with boundary. By a *curve system* in M we mean a finite collection of disjointly embedded curves which are either:
- circles not bounding disks in M and not isotopic to components of ∂M, or
- arcs with endpoints of ∂M, not isotopic (rel endpoints) to arcs in ∂M.

Let $\mathscr{CS}(M)$ be the set of isotopy classes of curve systems in M. Identifying a nonempty curve system with any positive number of parallel copies of itself yields the set $\mathscr{PS}(M)$ of projective isotopy classes of curve systems.

Every curve system can be expressed uniquely in the form $n_0C_0+\cdots+n_kC_k$ where the C_i's are connected, nonisotopic curve systems and n_iC_i denotes n_i parallel copies of C_i, $n_i>0$. We can form a simplicial complex PS(M) whose k-simplices correspond bijectively with isotopy classes of such $(k+1)$-tuples $[C_0,\ldots,C_k]$. The various faces of such a k-simplex are obtained by deleting C_i's. $\mathscr{PS}(M)$ is then identified with the set of points of PS(M) having rational, or equivalently integral, barycentric (projective) coordinates, namely, $n_0C_0+\cdots+n_kC_k$ has coordinates $n_0[C_0]+\cdots+n_k[C_k]$. For example, $2C_0+3C_1$ is the point on the edge $[C_0, C_1]$ three-fifths of the way from C_0 to C_1.

Example 1. M is a pair of pants, i.e., S^2 minus three disks. Curve systems can contain only arcs in this case, and there are only six isotopy classes of such arcs: one joining each pair of boundary circles, and one joining each boundary circle to itself and separating the other two boundary circles. Thus PS(M) has six vertices, and these span four triangles as in Fig. 1, making PS(M) itself a 2-simplex. (We think of S^2 here as the one-point compactification of the plane of the page, with the three boundary circles of M shown.)

Example 2. M is a once-punctured torus. Here circles and arcs are classified up to isotopy by their homology classes in $H_1(M, \partial M)$, which can be thought of more geometrically as slopes in $\mathbb{Q}\cup\{\infty\}$. One can work out that PS(M) has the form shown in Fig. 2(a). The edges $[C_0, C_1]$ coming in radially from the rational points $[C_0]$ of the perimeter of the square arise from a circle $C_0\subset M$ and an arc $C_1\subset M$ of the same slope. The rest of the interior of the square is filled up with 2-simplices

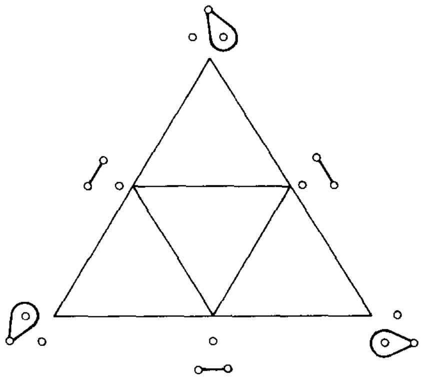


Fig. 1.

$[C_0, C_1, C_2]$, where C_0, C_1, C_2 are arcs of three different slopes. Thus PS(M) consists of the whole square minus the irrational points of its boundary.

Example 3. M is the closed nonorientable surface of Euler characteristic -1. View M as a once-punctured torus with antipodal boundary points identified to form a 1-sided circle $C \subset M$. This identification can also be applied unambiguously (up to isotopy) to curve systems on the punctured torus, giving rise to a copy of Fig. 2(a) in PS(M). The rest of PS(M) consists of edges joining all the rational boundary points in Fig. 2(a) to the vertex $[C]$, making PS(M) a dense subset of a pyramid (topologically, a 2-sphere). The circle C plays a distinguished role here: it is, up to isotopy, the only circle in M whose complement is orientable.

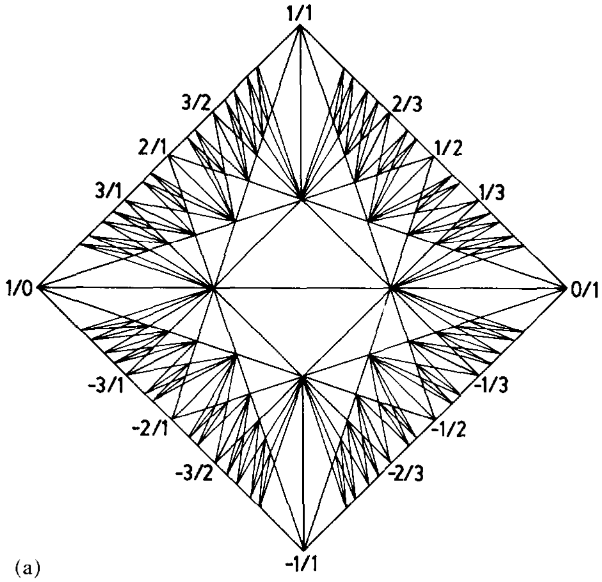
Example 4. M is a once-punctured Klein bottle. this is similar to Example 2, but in some ways simpler. The complex PS(M) is as shown in Fig. 2(b), and completely fills up a square.

Example 5. M is the twice-punctured projective plane. Viewing this as a square with corners deleted (the punctures) and antipodal boundary points identified, one has the picture in Fig. 3.

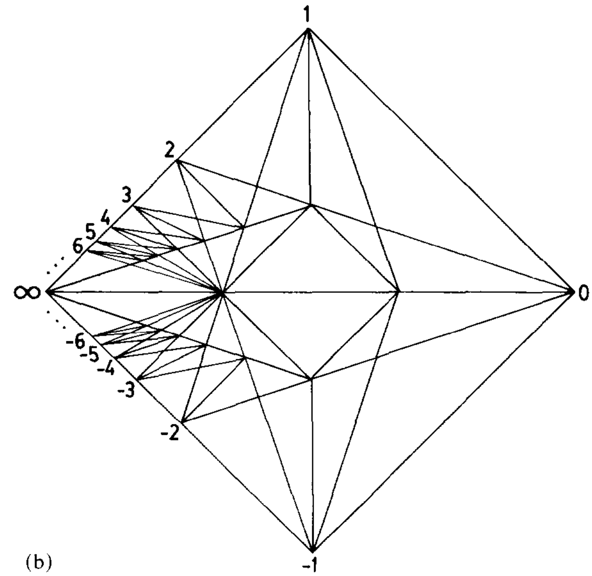
Usually PS(M) is noncompact, being an infinite simplicial complex. But as in Examples 2, 3, and 4, PS(M) turns out to have a natural compactification PL(M) which is a finite polyhedron. The rest of this chapter will be aimed at studying this PL(M) from the viewpoint of curve systems.

Train tracks

A *train track* $\tau \subset M$ is a closed subset locally diffeomorphic to the model in Fig. 4(a). That is, τ is a compact submanifold meeting ∂M transversely, except at finitely many branching points in the interior of M where two arcs of τ merge into one, all three arcs having a common tangent direction. We will usually assume train tracks are *good*, having none of the complementary regions listed in Fig. 5.



(a)



(b)

Fig. 2.

A *measure* on a track τ is an assignment of weights $\alpha_i \geqslant 0$ to the components of the non-branchpoint locus of τ, satisfying the branch equation $\alpha_i = \alpha_j + \alpha_k$ at each branch point as in Fig. 4(a). Such measures $(\alpha_1, \ldots, \alpha_n)$ form a cone $C(\tau)$ in $\mathbb{R}^n$, the intersection of the hyperplanes $\alpha_i = \alpha_j + \alpha_k$ with the quadrant $[0, \infty)^n$. A measure $\alpha = (\alpha_1, \ldots, \alpha_n) \in C(\tau)$ with each α_i an integer determines a curve system $S_\alpha \subset M$ by taking α_i parallel copies of the ith nonsingular arc of τ and matching these arcs

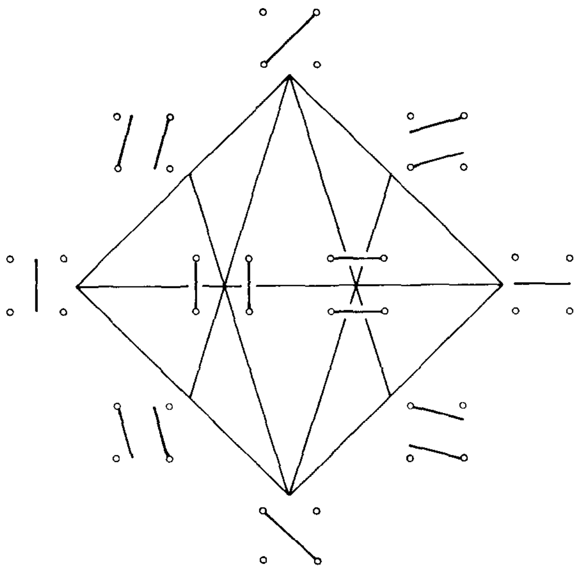
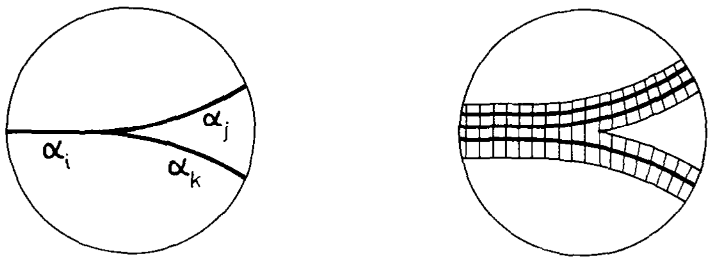


Fig. 3.



(a) (b)

Fig. 4.

together at branch points as in Fig. 4(b), using the branch equations $\alpha_i = \alpha_j + \alpha_k$. We say S_α is *carried* by τ.

Lemma 1.1. *If τ is good and S_α is carried by τ, then $S_\alpha \in \mathscr{CS}(M)$.*

Proof. Let $N(\tau)$ be a fibered neighborhood of τ as in Fig. 4(b), with fibers transverse to S_α. Extend the tangent linefield of S_α to a linefield on $N(\tau)$ transverse to fibers and tangent to $\partial N(\tau) - \partial M$. This linefield can be extended to a linefield on M having isolated singularities, and transverse to ∂M. For such linefields there is an index theory similar to the well-known theory for vector fields. Each isolated singularity is assigned an index, measuring how many times the linefield rotates as one goes around small circles enclosing the singularity; here, 180° counts as one rotation since the lines are not oriented. So for vector fields, the linefield index is twice the vectorfield index. For example, if one takes the foliation of the complex plane by the lines with $\mathrm{Im}(z)$ constant and applies the transformation $z \mapsto z^{2/k}$ (or $\log(z)$ for $k = 0$), one obtains a foliation whose tangent linefield has index $2 - k$ at the singularity at the origin.

The key fact about linefield index is that the sum of the indices at all the singularities is a topological invariant, namely, twice the Euler characteristic. It is not hard to check that the excluded regions in Fig. 5 are exactly those having nonnegative total index. In particular, we note that $\chi(M)$ must be negative if M contains a nonempty good track.

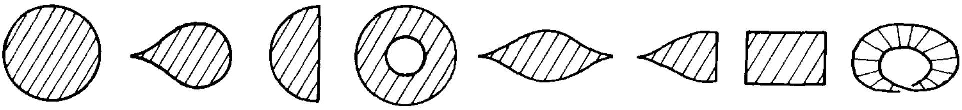


Fig. 5.

Returning now to the linefield we constructed on M, suppose S_α contains a circle bounding a disk in M. Then the total index in this disk is positive. The singularities of the linefield in this disk, however, have a negative sum since they determine the indices of the complementary regions of τ in this disk, So S_α can have no such circle component. Similarly, S_α contains no circle or arc parallel to ∂M. □

Another useful consequence of linefield index theory is the fact that a subtrack of a good track is also good.

Standard tracks

We wish to construct a finite set of "standard" tracks which carry all curve systems on M. These standard tracks are not canonically defined, but depend on choosing disjoint circles $S_1, \ldots, S_k$ which split M into pairs of pants $P_1, \ldots, P_n$ ($n = -\chi(M) > 0$). Some S_i's will be one-sided if M is nonorientable; let $\tilde{S}_i$ denote the boundary of a Möbius band neighborhood of a one-sided S_i.

Consider an arbitrary curve system $S \in \mathscr{CS}(M)$. Isotope S to minimize the number of intersection points with the splitting circles S_i. Then S is the disjoint union of systems S' and S'', where S' consists of circles parallel to the S_i's, and S'' meets each P_j in a curve system in $\mathscr{CS}(P_j)$. Each system $S'' \cap P_j$ is carried by one of the four basic tracks in Fig. 6 (compare Example 1). To reassemble S'' across a two-sided S_i, we must allow for possible twisting along S_i. A track which achieves this is obtained by inserting one of the two "connectors" in Fig. 7(a), which contain S_i as a subset. For one-sided S_i no nontrivial twisting along S_i is possible, so the connector in Fig. 7(b) suffices to carry S'' if S'' meets S_i in this case.

In this way we form $2^t 4^n$ tracks in M, t being the number of two-sided S_i's. All of these tracks, together with their various subtracks, we call *standard* tracks. (Note that the four basic tracks in Fig. 6 have a total of 15 subtracks, one for each edge and vertex of Fig. 1.) These standard tracks suffice to carry S'' and also the

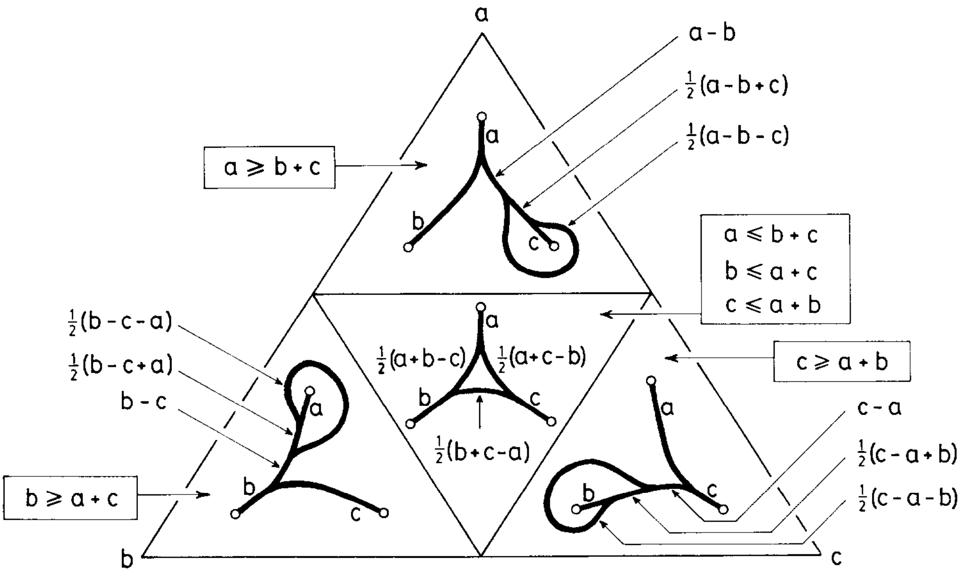


Fig. 6.

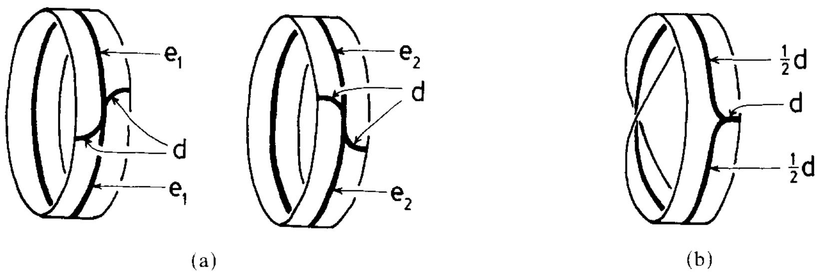


Fig. 7.

components of S' parallel to two-sided S_i's. To carry circles of S' parallel to one-sided S_i's, we enlarge our definition of *standard* to include tracks $\tau \cup S_i$ where τ is a standard track disoint from the one-sided splitting circle S_i. Thus we obtain finitely many standard tracks carrying all curve systems on M.

Observe that the maximal standard tracks have complementary regions all of linefield index -1, so these tracks are good. By an earlier remark, this implies that all standard tracks are good.

The polyhedra ML(M) *and* PL(M)

Each standard track τ has its cone $C(\tau)$ of measures, whose faces are identifiable with the cones $C(\tau')$ associated to the subtracks $\tau' \subset \tau$. Making all such identifications, there results a polyhedron ML(M). Projectivizing this construction by deleting $0 \in$ ML(M) (corresponding to the empty curve system) and factoring out by scalar multiplication, one obtains a finite polyhedron PL(M), with a decomposition into convex linear cells, the projectivized cones $C(\tau)$. Note that the simplices of PS(M) embed linearly into cells of PL(M), namely, $[C_0, \ldots, C_k]$ embeds in the projectivization of $C(\tau)$ for τ a standard track carrying $C_0 \cup \cdots \cup C_k$.

For example, if M is a pair of pants then ML(M) can be identified with the first octant of $\mathbb{R}^3$, with coordinates (a, b, c) in Fig. 6, and PL(M) = PS(M), the subdivided 2-simplex in Fig. 1.

Let ML$(M)_{\mathbb{Z}}$ denote the points of ML(M) having integer coordinates.

Proposition 1.2. *The map* ML$(M)_{\mathbb{Z}} \to \mathscr{CS}(M)$ *is a bijection.*

It follows that the rational (equivalently, integral) points of PL(M) correspond bijectively with $\mathscr{PS}(M)$. These rational points are dense in PL(M) since the branch equations which define the cones $C(\tau)$ have integer coefficients, implying that rational points are dense in $C(\tau)$.

Proof. It remains to show injectivity. We will do this using *intersection numbers.* If γ is any loop in M, not necessarily embedded, let $i_\gamma : \mathscr{CS}(M) \to [0, \infty)$ be the function which assigns to a curve system S the minimum number of intersection points of S with loops homotopic to γ.

Lemma 1.3. *For γ transverse to S, $i_\gamma(S) = |S \cap \gamma|$ iff no arc of $\gamma - S$ can be homotoped rel its endpoints into S.*

Proof. Suppose there is a homotopy $F: S^1 \times I \to M$ of $\gamma = F|S^1 \times \{0\}$ which decreases the number of intersection points with S. We may assume F is transverse to S, so $F^{-1}(S)$ is a one-dimensional submanifold of $S^1 \times I$ containing, by hypothesis, at least one arc with both endpoints on $S^1 \times \{0\}$. An innermost such arc cuts off from $S^1 \times I$ a half-disk, and restricting F to this half-disk gives a homotopy of an arc of $\gamma - S$ rel endpoints into S. $\square$

From this lemma it follows incidentally that if γ is an embedded loop, then $i_\gamma(S)$ is also the minimum number of intersection points of S with embedded loops isotopic to γ.

Injectivity in Proposition 1.2 follows immediately from the following lemma.

Lemma 1.4. *There exist finitely many embedded loops γ_m in M such that any two curve systems corresponding to distinct points of* ML$(M)_{\mathbb{Z}}$ *are distinguished by their intersection numbers with one of the γ_m's.*

Proof. For a start, we take for γ_m's all the boundary curves of the pairs of pants P_j, i.e., the components of ∂M, the two-sided S_i's, and the $\tilde{S}_i$'s associated to one-sided S_i's. Intersection numbers with these γ_m's are given by the three boundary weights a, b, c of each of the basic tracks in Fig. 6, by Lemma 1.3.

Next consider a two-sided S_i. We shall choose two more curves γ_m such that intersection numbers with these two curves detect the "twist parameters" e_1 and e_2 in Fig. 7(a). There are two subcases, according to whether the two "sides" of S_i belong to the same or different P_j's. One of the two choices for γ_m is shown in Fig. 8(a). We are interested in how the twist parameter e_i affects the intersection number with this γ_m. We note that the intersection number of a curve system S carried by a standard track with γ_m can be computed using only the subsurface $M_0 \subset M$ shown in Fig. 8(a), namely it is the intersection number of $S \cap M_0$ with γ_m in M_0 (Exercise).

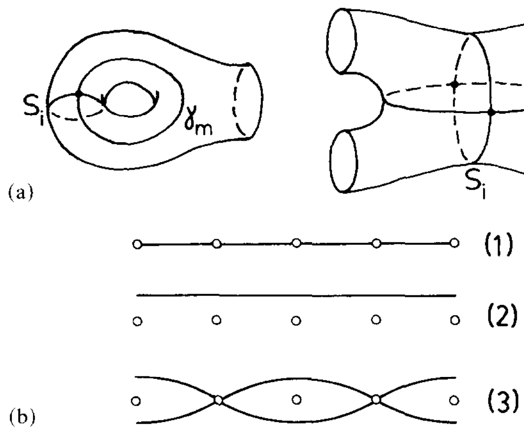


Fig. 8.

To see the effect of varying e_i on this intersection number, it is convenient to lift to a covering space $\mathbb{R}^2 - \mathbb{Z}^2$ of M_0. In case M_0 is the punctured torus, this cover is the universal cover of the torus with the preimages of the puncture deleted. If M_0 is a 4-punctured sphere, there is the well-known 2-sheeted branched covering of a sphere by a torus, branched at the four punctures, and $\mathbb{R}^2 - \mathbb{Z}^2$ is again the universal cover of the torus with preimages of the four punctures deleted; the group of deck transformations in this case is generated by 180° rotations about the points of $\mathbb{Z}^2$. We choose coordinates in $\mathbb{R}^2 - \mathbb{Z}^2$ so that S_i lifts to a slope ∞ line and γ_m lifts to a slope 0 line. For the punctured torus case, curve systems can be isotoped to consist of parallel copies of:

(1) arcs which lift to lines of rational slope passing through points of $\mathbb{Z}^2$, and
(2) circles which lift to lines of rational slope disjoint from $\mathbb{Z}^2$.

For the 4-punctured sphere case, there are also:

(3) arcs whose lifts are obtained from the lines in (1) by pushing off alternate points in $\mathbb{Z}^2$; see Fig. 8(b), illustrating the slope 0 case.

Looking in this $\mathbb{R}^2 - \mathbb{Z}^2$ cover, we can see that increasing e_1 increases the slope of the lifted curves, while increasing e_2 decreases slopes. Intersection number with γ_m therefore increases with e_i. So γ_m detects the amounts of twisting, with at most a two-fold ambiguity, the direction of twisting. To remedy this small deficiency,

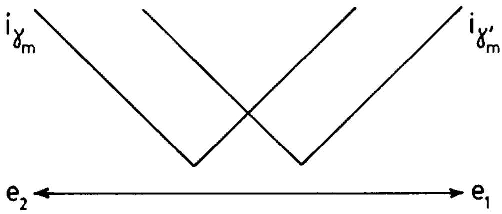


Fig. 9.

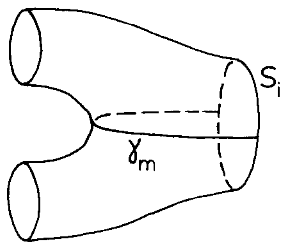


Fig. 10.

take, in addition to γ_m, the loop γ'_m obtained from γ_m by Dehn twist along S_i. See Fig. 9 for sketches of the graphs of these intersection numbers as functions of e_i. (One graph is a translate of the other.)

Finally, consider a one-sided S_i. Here we are not concerned with twisting along S_i, but with detecting the presence of parallel copies of S_i in curve systems which are disjoint from $\tilde{S}_i$. This can be done by intersection number with one loop γ_m as shown in Fig. 10. □

The Global Structure of ML(M) *and* PL(M)

Let b be the number of boundary components of M and let $\chi = \chi(M)$.

Proposition 1.5. ML(M) *is piecewise linearly homeomorphic to* $\mathbb{R}^{-3\chi-b} \times [0,\infty)^b$, *preserving scalar multiplication, with projection to the* $[0,\infty)$ *factors given by the weights at the b boundary circles of M. Consequently* PL(M) *is piecewise linearly homeomorphic to the join of a sphere* $S^{-3\chi-b-1}$ *with a simplex* Δ^{b-1}.

Proof. This will be by induction on k, the number of circles S_i in the splitting of M into pairs of pants. We have already observed that the result holds when M is a pair of pants. This yields the initial step of the induction, $k=0$, since ML for a disjoint union of two surfaces can clearly be identified with the product of the two ML's.

For the induction step, consider first the case of splitting M along a two-sided S_i to form the surface M'. Referring to Fig. 8(a), we see that in passing from M' to M, two boundary weights d_1 and d_2 are set equal, and new weights e_1 and e_2

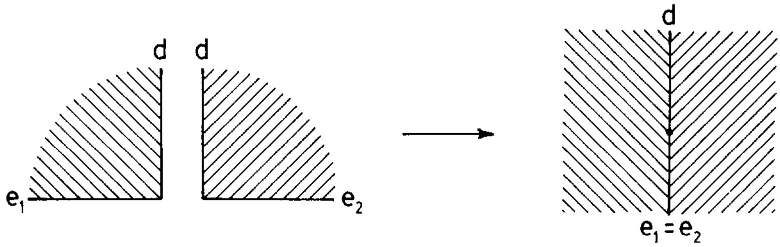


Fig. 11.

measuring the twisting around S_i are introduced. Thus the coordinates $(d_1, d_2) \in [0,\infty)^2$ are replaced by coordinates (d, e_1) or (d, e_2). These two new quadrants $[0,\infty)^2$ have their boundaries identified as in Fig. 11 to form, piecewise linearly, a full plane $\mathbb{R}^2$, since if either $d = 0$ or $e_1 = e_2 = 0$, the two alternative subtracks in Fig. 7(a) coincide. Thus in $\mathbb{R}^{-3\chi-b} \times [0,\infty)^b$, a subproduct $[0,\infty)^2$ in the second factor shifts to an $\mathbb{R}^2$ in the first factor, completing the induction step in this case.

If S_i is one-sided, we form M from M' by adjoining the Möbius band in Fig. 7(b). The boundary weight $d \in [0,\infty)$ for M' is carried along to the new track for M, becoming $\frac{1}{2}d$ around the new branch. In addition, if $d = 0$ we are allowed to enlarge our track by adding a copy of the loop S_i, with arbitrary weight $e \geq 0$. If both d and e are zero we have a common subtrack, so the new factor $[0,\infty) = \{e \geq 0\}$

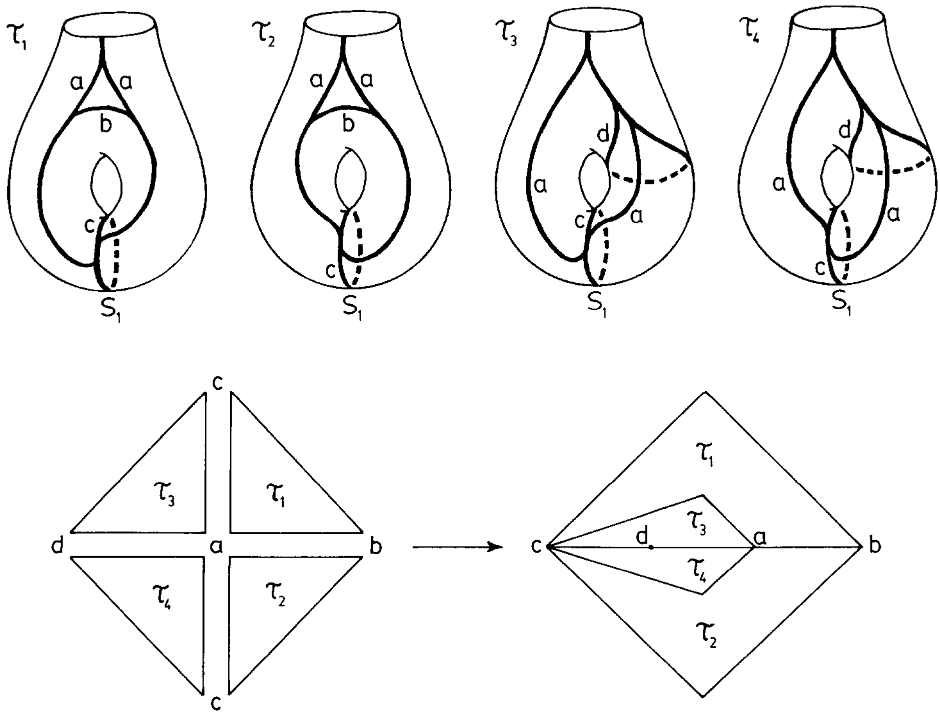


Fig. 12.

intersects the old $[0,\infty)=\{d\geqslant 0\}$ only in the origin, and the union of these two $[0,\infty)$'s is an $\mathbb{R}$. So in $\mathbb{R}^{-3\chi-b}\times[0,\infty)b$, a factor $[0,\infty)$ shifts to an $\mathbb{R}$ factor when we pass from M' to M. This finishes the induction step. $\square$

As an example, consider again the once-punctured torus, which is split into one pair of pants P_1 by a circle S_1. *A priori*, the four basic tracks on P_1 and the two connectors near S_1 would give eight tracks on the punctured torus. But since we must have equal weights at the two boundary circles of P_1 which are identified to form S_1, only two of the basic tracks on P_1 are really needed. Thus the four standard tracks in Fig. 12 cover $\mathrm{ML}(M)$. Each of these standard tracks has an octant of $\mathbb{R}^3$ for its cone of weights, as indicated. Projectively, this gives four 2-simplices which fit together piecewise linearly to form a square, as shown. This is $\mathrm{PL}(M)\approx S^1 * \Delta^0$. Figure 2 can be superimposed on this square to show how $\mathrm{PS}(M)$ embeds in $\mathrm{PL}(M)$.

2. Measured laminations

The idea here is to provide an interpretation for the nonintegral points of $\mathrm{ML}(M)$, and likewise for the irrational points of $\mathrm{PL}(M)$, as topological objects in M. At the same time we shall eliminate the dependence of these spaces on a choice of decomposition of M into pairs of pants, which was part of their original definition.

Construction of measured laminations

Given a track τ and a positive measure $\alpha\in C(\tau)$ (i.e., with all weights $\alpha_i>0$), one can construct a foliation N_α of the fibered neighborhood $N(\tau)$, as follows. Decompose $N(\tau)$ as a union of rectangles lying over the nonsingular arcs of τ, the fibers of $N(\tau)$ giving the vertical direction in these rectangles. Imagine the ith rectangle as a strip of paper of height α_i. The branch equations $\alpha_i=\alpha_j+\alpha_k$ allow these paper rectangles to be matched together at their vertical ends to form a copy of $n(\tau)$, with the foliation N_α determined by the horizontal lines in the rectangles. N_α has singularities at the cusp points of $\partial N(\tau)$. To eliminate these singularities, we slit N_α open along its finitely many singular leaves, starting at the cusps. In some cases, for example if all the weights α_i are rational, the singular leaves of N_α are compact, and the slitting open process is finite, yielding a thickening L_α of a curve system on M, with the product foliation (or twisted product, in the case of Möbius band components of L_α). In general, though, there may be noncompact singular leaves of N_α, and some care must be taken to damp down the magnitude of the slitting fast enough so that the process converges (details left to the reader). The result of this slitting is the *measured lamination* L_α, which lies in $N(\tau)$ transverse to the fibers.

An alternative viewpoint which avoids the subtleties of slitting along noncompact leaves is to consider equivalence classes of N_α's under the equivalence relation generated by isotopy and by slitting only along compact arcs in leaves, starting at cusps as before.

Let $\mathcal{ML}(M)$ denote the set of equivalence classes of measured N_α's carried by good tracks in M, and let $\mathcal{PL}(M)$ be the projectivization of $\mathcal{ML}(M)$, taking non-empty N_α's and factoring out by scalar multiplication of α.

Remarks. (1) It is not hard to check that two N_α's slit open to isotopic L_α's if and only if they are equivalent. So elements of $\mathcal{ML}(M)$ can be viewed as isotopy classes of measured laminations.

(2) Instead of slitting N_α open to a lamination we could collapse $M - N_\alpha$ onto a suitable spine, so as to form a foliation F_α of M having isolated singularities with negative index. Here too to recover $\mathcal{ML}(M)$ one needs an equivalence relation, generated by collapsing leaves joining singularities. This was Thurston's original approach.

(3) Slitting along compact arcs corresponds to an analogous "splitting" operation on τ. So elements of $\mathcal{ML}(M)$ can also be identified with suitable equivalence classes of measured tracks τ; see [7].

In each equivalence class in $\mathcal{ML}(M)$ there is a unique (up to isotopy) representative N'_α obtained by slitting completely all the compact singular leaves, since slitting a compact subarc of a noncompact singular leaf can be realized by isotopy. Certain components of N'_α are foliated by parallel compact leaves, forming a thickening of some curve system in M. We claim: In the other components of N'_α there are no compact leaves, and the noncompact singular leaves (along which one would slit to form L_α) are dense. To see this, consider a short vertical arc in N'_α with one endpoint on a fixed nonsingular leaf. This arc can be translated horizontally along leaves (the idea of "holonomy") to any prescribed distance without encountering cusps of N'_α if the vertical segment is short enough. If the given leaf is compact, a suitably short vertical arc therefore eventually returns to its initial position exactly, and the leaf has a neighborhood of compact leaves. So compact leaves in N'_α are open; they are also obviously closed, and so form certain components of N'_α. If translation of a fixed vertical arc could be continued indefinitely along a noncompact leaf without being obstructed by a cusp of N'_α, it would return infinitely often to the same fiber of N'_α with a vertical translation, and so give this fiber infinite measure, which is impossible. Thus translation of every vertical segment along a non-compact leaf must eventually meet a cusp, making the singular leaves dense in the noncompact-leaf components of N'_α.

Two consequences of this are:

- The fully slit open lamination L_α meets fibers of $N(\tau)$ only in intervals and Cantor sets.
- Only the compact-leaf components of N'_α can meet ∂M, since obviously a noncompact singular leaf along which one would slit to form L_α cannot meet ∂M, the only endpoint of this leaf being a cusp point in the interior of M.

Length functions

Given N_α as above and a loop γ in M, consider loops homotopic to γ which meet N_α in finitely many arcs which are either vertical (in fibers of $N(\tau)$) or horizontal (in leaves of N_α); such loops we call PVH—piecewise vertical or horizontal. PVH loops have a *length*, the total length of all their vertical segments. Define $\ell_\gamma(N_\alpha)$ to be the infimum of the lengths of all PVH loops homotopic to γ. (We will show that this infimum is always realized, assuming τ is good.) The notion of "vertical" depends on the vertical fiber structure of N_α, which is not really part of the data of N_α. To avoid this, we could instead use PTH paths: piecewise either transverse to N_α or horizontal. Length can be defined just as well for PTH paths, and leads to the same $\ell_\gamma(N_\alpha)$.

Note that $\ell_\gamma(N_\alpha)$ is linear with respect to scalar multiplication of α. Also, $\ell_\gamma(N_\alpha)$ is clearly constant on equivalence classes in $\mathcal{ML}(M)$, and so defines a map $\ell_\gamma : \mathcal{ML}(M) \to [0, \infty)$. If the measure α is integral, it is easy to see that $\ell_\gamma(N_\alpha) = i_\gamma(S_\alpha)$ where S_α is the curve system associated to $\alpha \in C(\tau)$. So ℓ_γ extends i_γ from $\mathcal{CS}(M)$ to $\mathcal{ML}(M)$.

If α has some coordinates $\alpha_i = 0$, this corresponds to passing to a subtrack of τ on which α becomes positive, in other words to a face of $C(\tau)$. So we can regard ℓ_γ as a function $C(\tau) \to [0, \infty)$.

Proposition 2.1. *For τ a good track and γ a loop in M, the function $\ell_\gamma : C(\tau) \to [0, \infty)$ is piecewise linear.*

Remark. If τ is allowed to have complementary digons (disks with two cusps), ℓ_γ may not be continuous at $\partial C(\tau)$. For example, for the track in Fig. 13, if the weights b and c are equal and $a > 0$, then the indicated curve γ can be homotoped (rel its boundary) so that in the part of the surface shown, its length is zero. On the other hand when $a = 0$ this cannot be done, and one suddenly has a contribution of $b + c$ to ℓ_γ.

Proof of Proposition 2.1. We begin by putting γ into a minimal position with respect to τ. We consider γ's which are divided into finitely many smoothly immersed *segments* which lie either inside τ, or outside τ with endpoints on τ. For example,

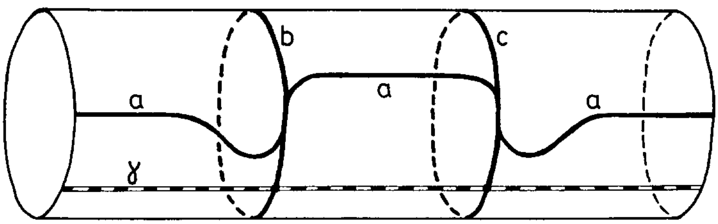


Fig. 13.

if γ is transverse to τ, the intersection points with τ determine a segmentation of γ. For a segmented γ, any backtracking along τ can easily be eliminated (by homotopy of γ) without adding new segments, so we may assume that at the endpoints of its segments, γ either leaves or enters τ, or γ stays in τ but reverses direction and switches branches at a cusp of τ. Possibly γ is a single segment consisting of a smoothly immersed circle in τ or $M-\tau$ (a "segment" without endpoints).

We choose γ within its homotopy class to have the minimum number of segments outside τ, and among γ's with this number of segments outside τ, the minimum number of segments inside τ.

The track τ is obtained from N_α by collapsing fibers of N_α. Let γ_1 be a PVH loop which projects to γ under this collapse (and hence is homotopic to γ). Keeping γ_1 PVH, we can by a finite number of vertical deformations pull it *taut* in N_α, so that it has no configuration as in Fig. 14 which can be shortened by pushing vertically a horizontal piece of γ_1.

Fig. 14.

The claim is that in this taut position, γ_1 achieves the minimum length within its homotopy class. If not, there is a shorter PVH loop γ' homotopic to γ_1. We may assume the measure α is rational and positive, since both γ' and the taut γ_1 can be chosen to vary continuously (by vertical deformation) with small perturbations of α. Rescaling, we may then assume α is integral.

Let A be the union of the fibered rectangles in M such that collapsing the fibers of A to points converts the foliated submanifold L_α to N_α. When we reverse this collapsing process, γ_1 can be "expanded" to a loop γ_2 by inserting certain fibers of A. We may assume the vertical parts of γ_2 in L_α are complete fibers. The length of γ_2 is the number of intersections, counted with multiplicity, of γ_2 with C, the "core" leaves of L_α, of which L_α is a thickening.

Since γ_2 is not shortest in its homotopy class, it intersects C more often than necessary. By Lemma 1.3 this implies there is a homotopy of a piece of γ_2 which eliminates two intersections of γ_2 with C. Thus there is a map $f: D^2 \to M$ taking one arc $\partial_+ D^2$ of ∂D^2 to γ_2, the remainder of ∂D^2 being $\partial_- D^2 = f^{-1}(C)$. Since $f(\partial_+ D^2)$ meets $\operatorname{int}(L_\alpha)$ only in two vertical arcs containing its two endpoints, we may push the path $f(\partial_- D^2)$ vertically away from C to make $f(\partial D^2)$ disjoint from $\operatorname{int}(L_\alpha)$. Then we may assume that $f(D^2) \subset M - \operatorname{int}(L_\alpha)$; see Fig. 15 for an example. The path $f(\partial_+ D^2)$ must now lie in $\partial L_\alpha \cup A$, otherwise homotoping it via f to $f(\partial_- D^2)$ would yield a homotopy of γ to a loop with fewer segments outside τ. We may assume that $f(\partial_- D^2)$ is a monotone path in its leaf and that f is transverse to ∂L_α and to $\partial_v A$, the vertical part of ∂A. Circles of $f^{-1}(\partial_v A)$ in $\operatorname{int}(D^2)$ can be eliminated by redefining f in $\operatorname{int}(D^2)$.

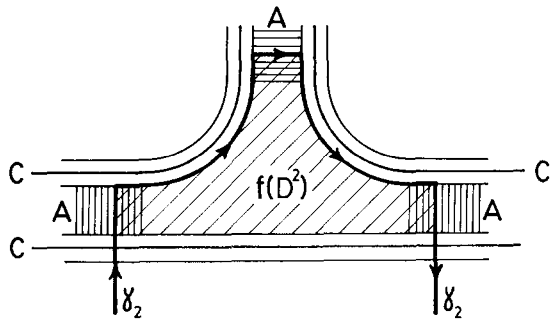
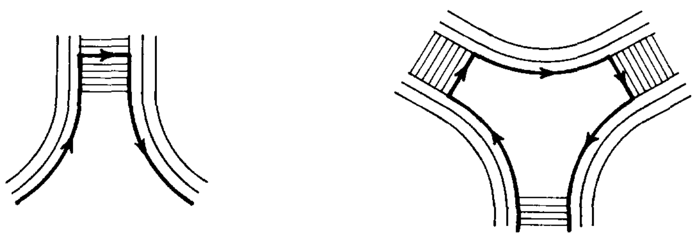


Fig. 15.

Consider now an arc $\boldsymbol{a}$ of $f^{-1}(\partial_v A)$ in $\operatorname{int}(D^2)$.

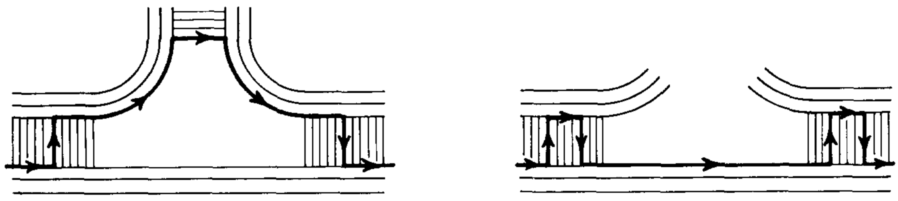
(i) Suppose both endpoints of $\boldsymbol{a}$ lie on $\partial_- D^2$ and $\boldsymbol{a}$ is an edgemost such arc in D^2. If both endpoints of $\boldsymbol{a}$ map to the same endpoint of the component $\boldsymbol{c}$ of $\partial_v A$ containing $f(\boldsymbol{a})$, then when the path $f(\partial_- D^2)$ passes through this endpoint of $\boldsymbol{c}$ at the two endpoints of $\boldsymbol{a}$, it does so in opposite directions (by transversality), contradicting monotonicity of $f(\partial_- D^2)$ in its leaf. If, on the other hand, the two endpoints of $\boldsymbol{a}$ map to different endpoints of $\boldsymbol{c}$, we may take f to be a homeomorphism of $\boldsymbol{a}$ onto $\boldsymbol{c}$, and f restricted to the disk in D^2 cut off by $\boldsymbol{a}$ gives a homotopy in $M - \operatorname{int}(L_\alpha \cup A)$ of $\boldsymbol{c}$ to an arc in $\partial(M - \operatorname{int}(L_\alpha))$, fixing $\partial \boldsymbol{c}$. So $\boldsymbol{c}$ cuts off a disk from $M - \operatorname{int}(L_\alpha)$ which corresponds to a complementary monogon of τ—the second excluded region in Fig. 5. Thus there can be no arcs $\boldsymbol{a}$ with both endpoints on $\partial_- D^2$.

(ii) Suppose that $\boldsymbol{a}$ has both endpoints on $\partial_+ D^2$ and that $\boldsymbol{a}$ is an edgemost such arc in D^2. Then the arc of $\partial_+ D^2$ cut off by $\boldsymbol{a}$ maps by f either into A (Fig. 16(a)) or into the closure of a complementary region of $L_\alpha \cup A$ (Fig. 16(b)). In the former case this arc would have to arise from backtracking of γ in τ, contrary to our choice of γ. In the latter case, this arc of $\partial_+ D^2$ mapping to $\partial(L_\alpha \cup A)$ must correspond to a part of γ which cycles around (perhaps more than once) an n-cusped boundary component of a complementary region R of $M - \tau$. This cycle is null-homotopic in R, via f on the disk in D^2 cut off by $\boldsymbol{a}$, so R is a disk, and $n \geqslant 3$ since τ is good. In this case γ would not be minimal. Hence there can be no arcs $\boldsymbol{a}$ with both endpoints on $\partial_+ D^2$.



(a) (b)

Fig. 16.



(a) (b)

Fig. 17.

(iii) The remaining arcs a have one endpoint on each of $\partial_+ D^2$ and $\partial_- D^2$. These arcs divide D^2 into regions mapping either into A or into components of $M - \text{int}(L_\alpha \cup A)$. An arc of $\partial_+ D^2$ in the boundary of the latter type of region in D^2 must map to the leaf of L_α containing $f(\partial_- D^2)$, for otherwise, as in (ii), the region in D^2 would map to an n-cusped disk component of $M - \tau$ and γ would again be nonminimal, as shown for example in Fig. 17(a). With a suitable choice of the "expansion" γ_2 of γ_1 (eliminating detours as in Fig. 17(b)) we may now assume f maps all of $\partial_+ D^2$ into the leaf containing $f(\partial_- D^2)$. But this contradicts the tautness of γ_1. The claim that γ_1 minimizes length within its homotopy class is therefore proved.

As mentioned before, a taut γ_1 can clearly be chosen to vary continuously with α, so $\ell_\gamma(N_\alpha)$ is at least a continuous function of α. To show piecewise linearity, look at a piece of γ_1 projecting to a segment of γ in τ. This piece of γ_1 moves monotonically through a string of rectangles of N_α formed by the fibers through cusp points, as shown in Fig. 18. The heights of these rectangles are given by weights α_i. The net vertical distance between any two horizontal edges in this chain of rectangles is a linear function of the α_i's with $\mathbb{Z}$ coefficients. This can be seen by induction on the number of rectangles between the two horizontal edges; in the induction step one either adds or subtracts an α_i, since two adjacent rectangles always have a horizontal edge on the same level.

Where two horizontal edges are at the same height therefore defines a hyperplane in $C(\tau)$ (if it is not all of $C(\tau)$). On the complementary components of the union of all such hyperplanes, the length of γ_1 is a $\mathbb{Z}$-linear function of α, since each stretch of γ_1 going monotonically (in the weak sense) up or down has length given by the vertical distance between two horizontal edges of rectangles. (This holds true even in the special case that γ is a smoothly immersed circle in τ and γ_1 is globally

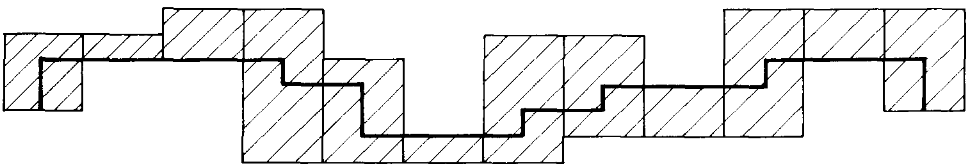


Fig. 18.

monotone up or down, since the length of γ_1 in this case is the vertical distance from a horizontal edge to itself as measured all the way around γ_1.) $\square$

We remark that we have proved slightly more than stated in Proposition 2.1, since we have shown that where ℓ_γ is linear, it has integer coefficients. This will be useful in proving Lemma 2.2 below.

A version of Proposition 2.1 for embedded loops γ (which follows from 2.1 and Lemma 2.4 below) is sketched in [10].

Intrinsic structure via length functions

Call a collection of n loops $\gamma \subset M$ *injective* if the map $\mathscr{CS}(M) \to [0, \infty)^n$ having the associated intersection number functions i_γ as coordinates is injective. Lemma 1.4 says that injective collections exist.

Lemma 2.2. *For an injective collection of loops γ, the map $\ell: \mathrm{ML}(M) \to [0, \infty)^n$ whose coordinates are the associated length functions ℓ_γ is a piecewise linear homeomorphism onto its image.*

Proof. We can subdivide $\mathrm{ML}(M)$ into finitely many polyhedral cones on each of which ℓ is $\mathbb{Z}$-linear, in particular $\mathbb{Q}$-linear. If ℓ were noninjective on one cone, it would be noninjective on rational points, by linear algebra. (The kernel of a linear transformation with $\mathbb{Q}$ coefficients has the same dimension over $\mathbb{Q}$ as over $\mathbb{R}$.) Clearing denominators, ℓ would then be non-injective on integer points, contrary to hypothesis. Similarly, if ℓ took two points in different cones to the same point, this would already happen for a pair of integer points. So ℓ is injective on all of $\mathrm{ML}(M)$. Since ℓ is linear on the finitely many cones covering $\mathrm{ML}(M)$, it must be a piecewise linear homeomorphism onto its image. $\square$

Note that the image of ℓ in Lemma 2.2 is independent of the choice of decomposition of M into pairs of pants used to define $\mathrm{ML}(M)$ originally. Namely, it is the closure of the set of rays from the origin passing through ℓ-images of the integer points $\mathrm{ML}(M)_{\mathbb{Z}} = \mathscr{CS}(M)$. In other words, $\ell(\mathrm{ML}(M)) = \ell(\mathscr{ML}(M))$, since length functions $\ell_\gamma : C(\tau) \to [0, \infty)$ are continuous for good tracks τ.

An immediate consequence is that $\mathrm{ML}(M)$ has a well-defined piecewise linear structure (over $\mathbb{Q}$), independent of the decomposition of M into pairs of pants. Since all the piecewise linear maps involved here preserve scalar multiplication, it follows that $\mathrm{PL}(M)$ has a well-defined, intrinsic piecewise projective-$\mathbb{Q}$-linear structure.

To get a description of $\mathrm{ML}(M)$ completely independent of any choices, we could identify it with its image under the embedding $\ell: \mathrm{ML}(M) \to [0, \infty)^\infty$ whose coordinates are the length function ℓ_γ associated to *all* the homotopy classes of loops γ in M. Alternatively, we could take all embedded loops γ. (However, the piecewise linear structure is perhaps somewhat more obscure in the infinite dimensional space $[0, \infty)^\infty$.)

This yields a description of the space $\mathrm{PL}(M)$ in more elementary terms: It is the closure of the image of the projectivization of the map $i:\mathscr{CS}(M)\to[0,\infty)^\infty$ whose coordinates are the intersection number functions i_γ.

Normal form for measured laminations

Let a decomposition of M into pairs of pants be given, defining the polyhedron $\mathrm{ML}(M)$.

Proposition 2.3. *The map* $\mathrm{ML}(M)\to\mathscr{ML}(M)$ *is a bijection, hence* $\ell:\mathscr{ML}(M)\to[0,\infty)^\infty$ *is injective.*

Proof. The map $\mathrm{ML}(M)\to\mathscr{ML}(M)$ is injective since its composition with ℓ is, by Lemma 2.2. Surjectivity will use the following lemma.

Lemma 2.4. *Given* $N_\alpha\in\mathscr{ML}(M)$ *and a curve system* γ *consisting of circles, then* N_α *can be slit and isotoped until each circle of* γ *either is a leaf of* N_α *or meets* N_α *in fibers of* $N(\tau)$ *whose total length realizes the minimum length in the homotopy class of that circle.*

Proof. This involves a variation on the proof of Proposition 2.1 to assure that γ remains embedded. We consider PVH embedded γ's which are divided into finitely many segments which lie either outside $N(\tau)$, with endpoints on $\partial N(\tau)$, or inside $N(\tau)$, proceeding monotonically (in the weak sense) through fibers, and transverse to fibers through cusp points of $\partial N(\tau)$. Define the complexity of such a γ to be the lexicographically ordered triple (#segments outside $N(\tau)$, #segments inside $N(\tau)$, #points of intersection with cusp fibers). Choose γ of minimal complexity within its isotopy class. This implies that γ passes monotonically through the rectangles of $N(\tau)$ bounded by the cusp fibers, except for arcs as in Fig. 19(a).

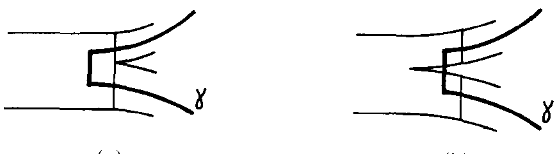


Fig. 19.

Now pull γ taut, by vertical *homotopy*. Again this taut γ achieves the minimum length within its homotopy class. The proof of this is virtually identical with the one given before, and will be left from the reader to verify in detail. The one minor difference (aside from the fact that we are now minimizing complexity over the isotopy class of γ, rather than the homotopy class) is that an edgemost arc $\boldsymbol{a}$ with both endpoints on ∂_+D^2, cutting off a disk from D^2 which maps to A, gives rise to the backtracking in Fig. 19(a), which is not ruled out. The presence of such arcs $\boldsymbol{a}$

makes little difference to the rest of the argument however, which eventually eliminates them anyway.

Consider now the process of pulling γ taut. Slitting as in Fig. 19(b) to eliminate nonmonotonicity of γ with respect to fibers of $N(\tau)$ will not affect the process, so we may assume γ is monotonic. Eliminating configurations as in Fig. 14 may introduce self-intersections of γ, but these are of a special kind: they can be removed by small perturbations of γ.

Having γ in taut position, we first slit N_α along its compact singular leaves, which introduces some segments of γ outside $N(\tau)$. Achieving the desired conclusion for the compact-leaf components of the resulting N'_α is very easy, so we need only worry about the other components. In these, the singular leaves are dense, as observed earlier in this section, so we may slit N'_α until each vertical piece of γ has an endpoint on $\partial N'_\alpha$. Then with a little more slitting, the remaining horizontal pieces can be eliminated. Finally, γ can be perturbed back to an embedding. This is homotopic, hence isotopic, to the original γ. $\square$

Proof of Proposition 2.3 (continued). Take N_α satisfying the conclusion of Lemma 2.4 with γ the circles determining the given decomposition of M into pairs of pants P_j. After further slitting, N_α decomposes as the disjoint union of N'_α and N''_α where N'_α consists of all circle leaves of N_α parallel to circles of γ. Choose a track τ carrying N_α which is the disjoint union of tracks τ' and τ'' carrying N'_α and N''_α, where $\tau' \subset \gamma$ and τ'' is transverse to γ. We may take τ to be a splitting of a good track carrying the original N_α, so τ is good, hence also τ' and τ''.

The essential point is to see now that all leaves of $N''_\alpha \cap P_j$ are compact. Fixing j, we may assume the previous slitting of N_α eliminated any compact singular leaves of $N''_\alpha \cap P_j$. As observed earlier, all leaves of $N''_\alpha \cap P_j$ meeting ∂P_j must then be compact, forming certain foliated rectangle components of $N''_\alpha \cap P_j$. Let $\dot{N}_j$ be the union of the remaining components of $N''_\alpha \cap P_j$, carried by a subtrack τ_j of $\tau'' \cap P_j$ contained in $\mathrm{int}(P_j)$. This τ_j must be good in P_j, since the only possible complementary region in Fig. 5 is an annulus with one boundary circle in τ_j and the other boundary circle in ∂P_j, but this would force N_j, and hence N''_α, to have circle leaves parallel to ∂P_j, contrary to the definition of N''_α. If $N_j \neq \emptyset$, then $C(\tau_j) \neq \emptyset$, so $C(\tau_j)$ would contain rational points corresponding to curve systems in P_j disjoint from ∂P_j. We conclude that $N_j = \emptyset$, verifying that all leaves of $N'_\alpha \cap P_j$ are compact arcs.

Next we pinch N''_α (the reverse of the slitting operation), without disturbing N'_α, so that for all j, no complementary region of $N''_\alpha \cap P_j$ in P_j is one of the types in Fig. 5. The only bad regions which might occur are half-digons (sixth type in Fig. 5) and rectangles, since half-disks (third type) cannot occur, otherwise some component of γ would not achieve the minimum length in its homotopy class. Pinch according to the rule: half-digons first, then rectangles. Pinching rectangles might create half-digons, which should then be pinched before any more rectangles are pinched, to be sure that no digons are created. Since the total number of complementary regions in the P_j's decreases with each pinching, this is a finite process.

After this pinching, each $N''_\alpha \cap P_j$ is carried by a standard track in P_j, and so clearly N_α itself is carried by a standard track. □

Application to diffeomorphisms of nonclosed surfaces

Using the topological machinery developed so far, we can easily prove the following proposition.

Proposition 2.5. *Suppose ∂M is nonempty. Then a diffeomorphism $f: M \to M$ can be isotoped so that either*

(1) *f has finite order, or*

(2) *f leaves invariant a curve system consisting of circles in M, or*

(3) *f takes a measured lamination in* $\operatorname{int}(M)$ *to a scalar multiple of itself.*

Proof. A diffeomorphism f induces a homeomorphism of $\mathcal{ML}(M)$, given by a permutation of the coordinate length functions ℓ_γ, and hence also a homeomorphism of $\mathcal{PL}(M)$. This is a ball if $\partial M \neq \emptyset$, so by the Brouwer fixed point theorem, there must be a (nonempty) measured lamination $L_\alpha \subset M$ with $f(L_\alpha) = L_{\lambda\alpha}$ for some $\lambda > 0$. If L_α is disjoint from ∂M, we are in case (3), so we may assume L_α meets ∂M. Leaves of L_α meeting ∂M are compact. These yield a curve system consisting of arcs in M which is invariant under f. Let M' be M split open along these arcs, and let M'' be M' minus those components which are either disks, or annuli with one boundary circle in ∂M. If $M'' \neq \emptyset$ then the circles of $\partial M''$ which are not circles of ∂M give rise to a system of circles in M invariant under f, and we are in case (2). If $M'' = \emptyset$, f can clearly be isotoped to be of finite order. □

In case (2), M and f can be split open along the invariant circle system to obtain a simpler situation which can be analyzed inductively. Only a small amount of information is lost in this splitting process: Dehn twists along the invariant circles. If we are in case (3) but not case (2), then the invariant lamination L_α has no compact leaves, and each complementary region of L_α must be a disk with some number of cusps on its boundary, and possibly one puncture (component of ∂M) in its interior. (The full strength of Thurston's theorem on surface diffeomorphisms—which holds also when M is closed—is that in this last case, f also preserves another measured lamination L_β transverse to L_α, with $f(L_\alpha) = L_{\lambda\alpha}$ and $f(L_\beta) = L_{\beta/\lambda}$ for some $\lambda \neq 1$.)

Proposition 2.6. *Laminations without compact leaves are dense in the subspace $\mathcal{ML}_0(M) \approx \mathbb{R}^{-3\chi - b}$ of $\mathcal{ML}(M)$ consisting of laminations disjoint from ∂M if and only if M is orientable.*

Proof. The laminations in $\mathcal{ML}_0(M)$ containing a given circle $C \subset M$ as a leaf are obtained by assigning a weight in $(0, \infty)$ to C and then adjoining a lamination in $\mathcal{ML}_0(M - C)$. If M is orientable, C is two-sided, so $M - C$ has two more boundary circles than M, making $\mathcal{ML}_0(M-C)$ have dimension two less than $\mathcal{ML}_0(M)$. Thus, laminations containing C as a leaf are in the image of a piecewise linear map $(0,\infty)\times\mathcal{ML}_0(M-C)\to\mathcal{ML}_0(M)$, this image having codimension at least one (equal to one, in fact). Since C ranges over a countable set of possibilities, the rest of $\mathcal{ML}_0(M)$, consisting of laminations without compact leaves, is dense in $\mathcal{ML}_0(M)$.

When M is non-orientable and C is one-sided, $\mathcal{ML}_0(M-C)$ has dimension only one less than $\mathcal{ML}_0(M)$, so this argument fails. In fact, taking C as a one-sided S_i in a pair of pants decomposition of M, the proof of Proposition 1.5 shows that laminations containing C correspond to the open set in $\mathcal{ML}_0(M)\approx\mathbb{R}^{-3\chi-b}$ where the coordinate of $\mathbb{R}^{-3\chi-b}$ corresponding to C is positive. □

3. Trees

Axiomatics

By a *tree* (or more precisely, an $\mathbb{R}$-tree) we mean a metric space (T, d) such that:

(1) Any two points $x, y\in T$ are the endpoints of a unique *segment* $[x, y]$, i.e., a subset isometric to a closed interval in $\mathbb{R}$.

(2) $[x,y]\cap[x,z]=[x,w]$ for some w.

(3) $[x,y]\cup[y,z]=[x,z]$ if $[x,y]\cap[y,z]=\{y\}$.

This is the definition in the foundational paper [1]. Other basic references are [3] and [8]. (In fact, axiom (2) follows from (1) since d takes values in $\mathbb{R}$, rather than in a more general ordered abelian group.)

The uniqueness of segments $[x, y]$ with given endpoints tends to be hard to verify in practice. This problem is avoided with the following proposition.

Proposition 3.1. *A metric space (T, d) is a tree if there is a function assigning to each unordered pair $x, y\in T$ a segment $[x, y]$ in T, such that* (2) *and* (3) *hold.*

Proof. Choose a basepoint $x_0\in T$. According to Theorem 3.17 of [1], the following two axioms characterize trees:

(4) Segments $[x_0, x]$ exist for all $x\in T$.

(5) Letting $x\wedge y=\frac{1}{2}(d(x,x_0)+d(y,x_0)-d(x,y))$, then $x\wedge z\geq\min(x\wedge y, y\wedge z)$ for all $x, y, z\in T$.

Under the hypotheses of the Proposition, $x\wedge y$ is $d(x_0, w)$, where $[x_0,x]\cap[x_0,y]=[x_0,w]$. Then (5) follows by simply listing the possible configurations for the segments joining x_0 to x, y, and z, shown in Fig. 20. □

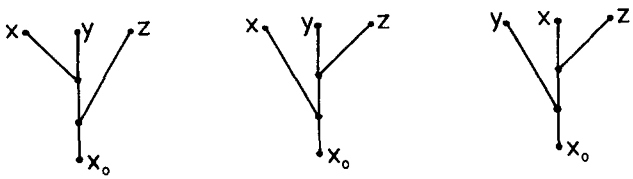


Fig. 20.

The dual tree to a measured lamination

Given a measured lamination N_α carried by a good track τ, let $\tilde{N}_\alpha$ be its preimage in the universal cover $\tilde{M}$ of M. For $x, y \in \tilde{M}$ define $\tilde{d}(x, y)$ to be the infimum of the lengths of all paths in $\tilde{M}$ joining x and y, where as in section 2 we take paths which are piecewise either transverse to leaves of $\tilde{N}_\alpha$ or in leaves of $\tilde{N}_\alpha$, and we measure lengths via the lifted measure from N_α. Then $\tilde{d}$ clearly defines a pseudometric on $\tilde{M}$ (symmetric and satisfying the triangle inequality, but not positive definite—points in the same leaf being at $\tilde{d}$-distance zero, for example). Let (T, d) be the associated metric space, obtained by identifying points of $\tilde{M}$ of $\tilde{d}$-distance zero apart.

A basic fact is that, given $x, y \in \tilde{M}$, there exists a path joining x and y whose length is $\tilde{d}(x, y)$. This is proved by the argument for the existence of minimum length loops in the proof of Proposition 2.1, the only difference being that now we use paths with fixed endpoints rather than loops. (Details left to the reader.) A consequence of this fact is that points of T are either nonsingular leaves of $\tilde{N}_\alpha$ or connected unions of singular leaves with adjoining complementary regions of $\tilde{N}_\alpha$. The latter type of points of T, countable in number, can be thought of as the "vertices" of T, though they can be dense in T. (This happens for example when M is the punctured torus and N_α corresponds to an irrational point of $\partial\mathrm{PL}(M)$, such an N_α is obtained from a foliation of the torus by lines of irrational slope, by splitting open along the leaf containing the puncture.)

Proposition 3.2. *(T, d) is a tree.*

For α rational this is easily seen: T is evidently a 1-dimensional simplicial complex, and it is simply-connected because loops in T can be lifted to loops in $\tilde{M}$. (Note that goodness of τ is not needed in this case.)

Proof. First a preliminary observation: Nonsingular leaves $\tilde{\lambda}$ of $\tilde{N}_\alpha$ are properly embedded submanifolds of $\tilde{M}$. This follows from the stronger assertion that $\tilde{\lambda}$ meets each fiber of $N(\tilde{\tau})$ at most once. For if $\tilde{\lambda}$ met a fiber of $N(\tilde{\tau})$ twice, the arc of $\tilde{\lambda}$ between the two intersection points would enclose a complementary region of $N(\tilde{\tau})$ of linefield index greater than zero, hence $N(\tau)$ also would have such a complementary region. (Lift to $\tilde{M}$ a linefield on M transverse to fibers of $N(\tau)$ and ∂M, and having all singularities of negative index.)

Now to prove the proposition we use the criteria of Proposition 3.1. For the segment $[x, y]$ we take the image in T of a shortest path in $\tilde{M}$. To verify that this segment is well-defined, and at the same time check axiom (2), consider two shortest paths in $\tilde{M}$ with the same initial point. The intersection of their images in T is closed, being the intersection of two closed sets. The intersection is also connected. For otherwise we would have an embedded circle in T. This would contain points corresponding to nonsingular leaves of $\tilde{N}_\alpha$, since there are only countably many singular leaves. Projecting to this circle in T there would be a loop in $\tilde{M}$ meeting

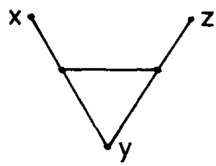


Fig. 21.

a nonsingular leaf of $\tilde{N}_\alpha$ in one point, transversely. This is impossible since nonsingular leaves are properly embedded in $\tilde{M}$.

For (3), let segments $[x, y]$ and $[y, z]$ be given with $[x, y]\cap[y, z]=\{y\}$. If $[x, z]$ differed from $[x, y]\cup[y, z]$ we would have the configuration of Fig. 21 and again T would contain an embedded circle. $\square$

Pullback laminations

Let an action of $\pi_1 M$ on a tree T be given, the elements of $\pi_1 M$ acting as isometries of T. Choose a triangulation of M, and lift this to a triangulation of the universal cover $\tilde{M}$ invariant under the action of $\pi_1 M$ by deck transformations. We construct an equivariant map $\tilde{f}: \tilde{M}\to T$ inductively over skeleta $\tilde{M}^{(i)}$, as follows. On $\tilde{M}^{(0)}$, let $\tilde{f}$ be an arbitrary equivariant map. (Choose lifts $\tilde{v}$ of the vertices v of M, let $\tilde{f}$ be arbitrary on these $\tilde{v}$'s, then extend over $\tilde{M}^{(0)}$ by equivariance.) Extend $\tilde{f}$ over 1-simplices $[v_0, v_1]$ by sending $[v_0, v_1]$ linearly to the unique segment joining $\tilde{f}(v_0)$ to $\tilde{f}(v_1)$ in T. This is automatically equivariant. For convenience, modify this choice of $\tilde{f}$ to be constant near vertices. To extend $\tilde{f}$ over a 2-simplex $[v_0, v_1, v_2]$ consider the generic case shown in Fig. 22, where $Y=\tilde{f}(\partial[v_0, v_1, v_2])$ consists of three edges meeting at a central vertex. In $[v_0, v_1, v_2]$ insert three bands of parallel arcs as shown, of thicknesses given by the lengths of the corresponding edges of Y. Then let $\tilde{f}$ map $[v_0, v_1, v_2]$ to Y by the obvious collapse. In degenerate cases some of the three edges of Y may have length zero, and the corresponding bands in $[v_0, v_1, v_2]$ are deleted. This prescription is automatically equivariant. Down in M we have therefore constructed a foliation N_α of a neighborhood $N(\tau)$ of a track τ which in each 2-simplex has the standard form shown in Fig. 22 (or a subtract of this, in degenerate cases).

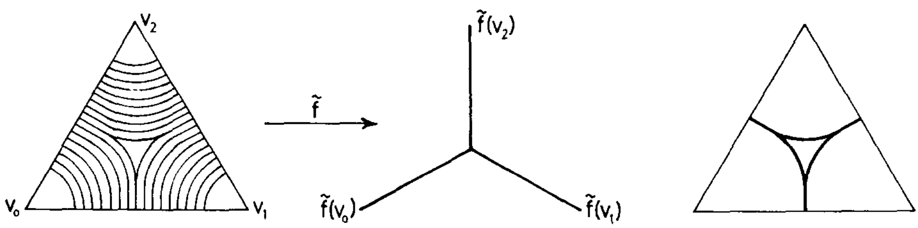
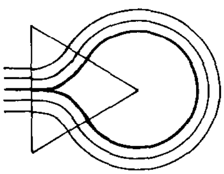


Fig. 22.

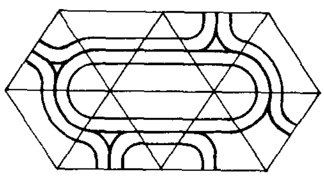
Proposition 3.3. *By-rechoosing $\tilde{f}$ on $\tilde{M}^{(0)}$ (and possibly retriangulating M in case M is nonorientable) we can arrange that the track τ carrying N_α is good.*

Proof. As Fig. 23(a) shows, complementary monogons (the second picture in Fig. 5) cannot occur since $\tilde{f}$ is monotone on edges. A smooth disk complementary region gives rise to a family of parallel trivial circle leaves as in Fig. 23(b). We can rechoose $\tilde{f}$ on the vertices inside this disk so as to eliminate a maximal such family of parallel circles, producing N_α carried by a subtrack of τ. A complementary smooth half-disk or an annulus with one boundary circle on ∂M is treated similarly. For other product complementary regions, $\tilde{f}$ can be redefined on vertices so as to transfer a family of parallel leaves on N_α across such a product. Using the fact that the closure of a product region of $M-\tau$ cannot be all of M (otherwise M would have Euler characteristic zero), a maximal family can be transferred so as to give a new N_α carried by a subtrack of τ. Repeating these steps as often as possible (a finite process since at each step we pass to a subtrack), we eventually eliminate all complementary regions except Möbius bands—finishing the proof if M is orientable.

A complementary Möbius band has a core circle C in $M^{(1)}$. Retriangulate, as in Fig. 24, by inserting a triangulated neighborhood of C. Replace τ by $\tau \cup C$, making the complementary Möbius band into a complementary annulus. Now as in an earlier step, transfer a maximal family of leaves across this annulus onto a neighborhood of C, thereby passing to a subtrack of $\tau \cup C$. Repeat this until the complementary region bordering C has negative index, which will eventually occur since $\chi(M)<0$. $\square$



(a)



(b)

Fig. 2?.

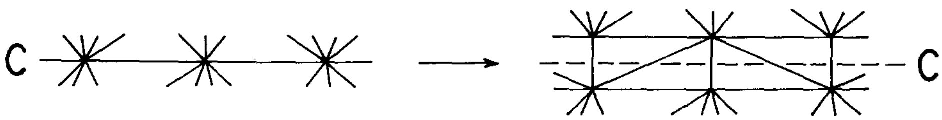


Fig. 24.

The pullback lamination given by Proposition 3.3 is not in general uniquely determined by T and the $\pi_1 M$ action. The relations between dual trees and pullback laminations are the following:

(1) A lamination is a pullback of its dual tree.

(2) Given T with $\pi_1 M$ action, there is an equivariant quotient map from the dual tree of any pullback lamination to T which preserves lengths of paths.

References

[1] R. Alperin and H. Bass, Length functions of group actions on Λ-trees, in: S.M. Gersten and J.R. Stallings, Eds., Combinatorial Group Theory and Topology, Ann. Math. Studies 111 (1987) 265-378.
[2] A. Casson, Automorphisms of surfaces after Nielsen and Thurston, notes by S. Bleiler, University of Texas at Austin, 1983.
[3] M. Culler and J. Morgan, Group action on $\mathbb{R}$-trees, Proc. London Math. Soc. 55 (1987) 571-604.
[4] M. Dehn, Die Gruppe der Abbildungsklassen, Acta Math. 69 (1938) 135-206.
[5] A. Fathi, F. Laudenbach and V. Poénaru, Travaux de Thurston sur les surfaces, Astérisque 66-67 (1979).
[6] A. Hatcher, Measured lamination spaces for 3-manifolds, in preparation.
[7] J. Harer and R. Penner, Combinatorics of train tracks, to appear.
[8] J. Morgan and P. Shalen, Valuations, trees, and degenerations of hyperbolic structures, Ann. Math. 120 (1984) 401-476.
[9] A. Papadopoulos, Réseaux ferroviaires, difféomorphismes pseudo-anosov, etc., Orsay preprint 1983.
[10] A. Papadopoulos, Réseaux ferroviaires et courbes simple sur les surfaces, C.R. Acad. Sci. Paris 297 (1983) 565-568.
[11] M. Scharlemann, The complex of curves on non-orientable surfaces, J. London Math. Soc. 25 (2) (1982) 171-184.
[12] W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces, preprint.