William P. Thurston

The Geometry and Topology of Three-Manifolds

Electronic version 1.1 - March 2002

http://www.msri.org/publications/books/gt3m/

This is an electronic edition of the 1980 notes distributed by Princeton University. The text was typed in TEX by Sheila Newbery, who also scanned the figures. Typos have been corrected (and probably others introduced), but otherwise no attempt has been made to update the contents. Genevieve Walsh compiled the index.

Numbers on the right margin correspond to the original edition's page numbers.

Thurston's *Three-Dimensional Geometry and Topology*, Vol. 1 (Princeton University Press, 1997) is a considerable expansion of the first few chapters of these notes. Later chapters have not yet appeared in book form.

Please send corrections to Silvio Levy at levy@msri.org.

CHAPTER 9

Algebraic convergence

9.1. Limits of discrete groups

It is important for us to develop an understanding of the geometry of deformations of a given discrete group. A qualitative understanding can be attained most concretely by considering limits of sequences of groups. The situation is complicated by the fact that there is more than one reasonable sense in which a group can be the limit of a sequence of discrete groups.

Definition 9.1.1. A sequence $\{\Gamma_i\}$ of closed subgroups of a Lie group G *converges geometrically* to a group Γ if

(i) each $\gamma \in \Gamma$ is the limit of a sequence $\{\gamma_i\}$, with $\gamma_i \in \Gamma_i$, and
(ii) the limit of every convergent sequence $\{\gamma_{i_j}\}$, with $\gamma_{i_j} \in \Gamma_{i_j}$, is in Γ.

Note that the geometric limit Γ is automatically closed. The definition means that Γ_i's look more and more like Γ, at least through a microscope with limited resolution. We shall be mainly interested in the case that the Γ_i's and Γ are discrete. The *geometric topology* on closed subgroups of G is the topology of geometric convergence.

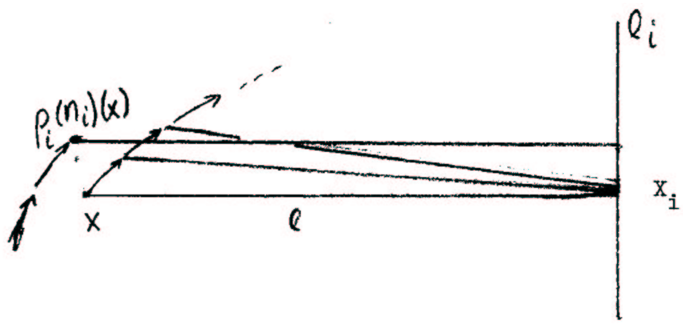
The notion of geometric convergence of a sequence of discrete groups is closely related to geometric convergence of a sequence of complete hyperbolic manifolds of bounded volume, as discussed in 5.11. A hyperbolic three-manifold M determines a subgroup of $\mathrm{PSL}(2,\mathbb{C})$ well-defined *up to conjugacy.* A specific representative of this conjugacy class of discrete groups corresponds to a choice of a base frame: a base 9.2
point p in M together with an orthogonal frame for the tangent space of M at p. This gives a specific way to identify $\tilde{M}$ with H^3. Let $O(\mathcal{H}_{[\epsilon,\infty)})$ consist of all base frames contained in $M_{[\epsilon,\infty)}$, where M ranges over $\mathcal{H}$ (the space of hyperbolic three-manifolds with finite volume). $O(\mathcal{H}_{[\epsilon,\infty)})$ has a topology defined by geometric convergence of groups. The topology on $\mathcal{H}$ is the quotient topology by the equivalence relation of conjugacy of subgroups of $\mathrm{PSL}(2,\mathbb{C})$. This quotient topology is not well-behaved for groups which are not geometrically finite.

Definition 9.1.2. Let Γ be an abstract group, and $\rho_i : \Gamma \to G$ be a sequence of representations of Γ into G. The sequence $\{\rho_i\}$ *converges algebraically* if for every $\gamma \in \Gamma$, $\{\rho_i(\gamma)\}$ converges. The limit $\rho : \Gamma \to G$ is called the algebraic limit of $\{\rho_i\}$.

DEFINITION 9.1.3. Let Γ be a countable group, $\{\rho_i\}$ a sequence of representations of Γ in G with $\rho_i(\Gamma)$ discrete. $\{\rho_i\}$ *converges strongly* to a representation ρ if ρ is the algebraic limit of $\{\rho_i\}$ *and* $\rho\Gamma$ is the geometric limit of $\{\rho_i\Gamma\}$.

EXAMPLE 9.1.4 (Basic example). There is often a tremendous difference between algebraic limits and geometric limits, growing from the following phenomenon in a sequence of cyclic groups.

Pick a point x in H^3, a "horizontal" geodesic ray l starting at x, and a "vertical" plane through x containing the geodesic ray. Define a sequence of representations 9.3
$\rho_i : Z \to \mathrm{PSL}(2,\mathbb{C})$ as follows. Let x_i be



the point on l at distance i from x, and let l_i be the "vertical" geodesic through x_i: perpendicular to l and in the chosen plane. Now define ρ_i on the generator 1 by letting $\rho_i(1)$ be a screw motion around l_i with fine pitched thread so that $\rho_i(1)$ takes x to a point at approximately a horizontal distance of 1 from x and some high power $\rho_i(n_i)$ takes x to a point in the vertical plane a distance of 1 from x. The sequence $\{\rho_i\}$ converges algebraically to a parabolic representation $\rho : \mathbb{Z} \to \mathrm{PSL}(2,\mathbb{C})$, while $\{\rho_i\mathbb{Z}\}$ converges geometrically to a parabolic subgroup of rank 2, generated by $\rho(\mathbb{Z})$ plus an additional generator which moves x a distance of 1 in the vertical plane. 9.4

This example can be described in matrix form as follows. We make use of one-complex parameter subgroups of $\mathrm{PSL}(2,\mathbb{C})$ of the form

$$\begin{bmatrix} \exp w & a\sinh w \\ 0 & \exp -w \end{bmatrix},$$

with $w \in \mathbb{C}$. Define ρ_n by

$$\rho_n(1) = \begin{bmatrix} \exp w_n & n\sinh w_n \\ 0 & \exp -w_n \end{bmatrix}$$

where $w_n = 1/n^2 + \pi i/n$.

Thus $\{\rho_n(1)\}$ converges to

$$\begin{bmatrix} 1 & \pi i \\ 0 & 1 \end{bmatrix}$$

while $\{\rho_n(n)\}$ converges to

$$\begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

This example can be easily modified without changing the algebraic limit so that $\{\rho_i(\mathbb{Z})\}$ has no geometric limit, or so that its geometric limit is a one-complex-parameter parabolic subgroup, or so that the geometric limit is isomorphic to $\mathbb{Z} \times \mathbb{R}$.

9.5

This example can also be combined with more general groups: here is a simple case. Let Γ be a Fuchsian group, with M_Γ a punctured torus. Thus Γ is a free group on generators a and b, such that $[a, b]$ is parabolic. Let $\rho : \Gamma \to \mathrm{PSL}(2, \mathbb{C})$ be the identity representation. It is easy to see that $\mathrm{Tr}\, \rho'[a, b]$ ranges over a neighborhood of 2 as ρ' ranges over a neighborhood of ρ. Any nearby representation determines a nearby hyperbolic structure for $M_{[\epsilon,\infty)}$, which can be thickened to be locally convex except near $M_{(0,\epsilon]}$. Consider representations ρ_n with an eigenvalue for

$$\rho_n[a, b] \sim 1 + C/n^2 + \pi i/n.$$

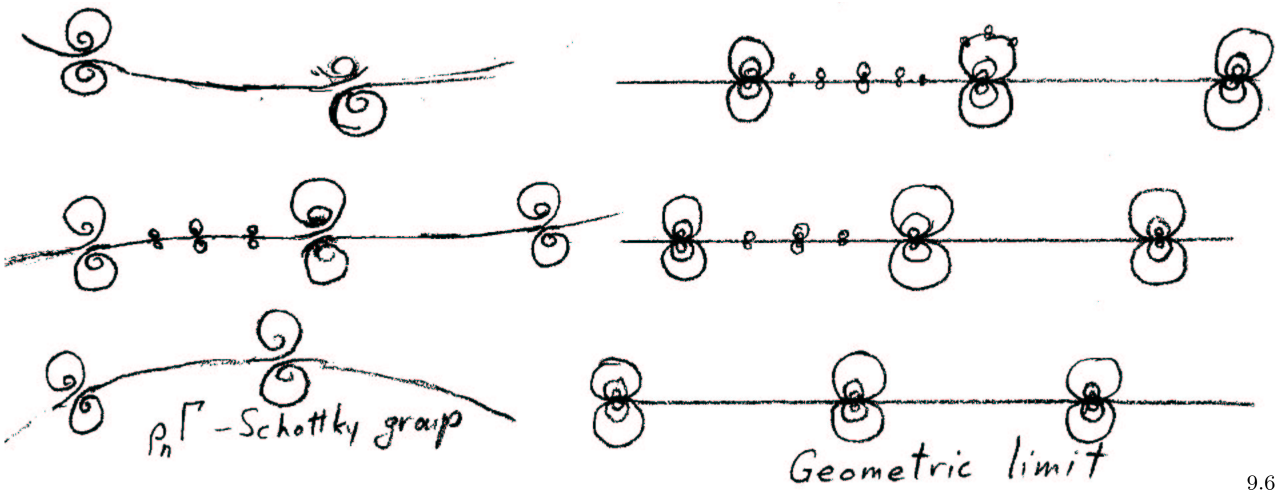
$\rho_n[a, b]$ translates along its axis a distance of approximately $2\,\mathrm{Re}(C)/n^2$, while rotating an angle of approximately

$$\frac{2\pi}{n} + \frac{2\,\mathrm{Im}(C)}{n^2}.$$

Thus the n-th power translates by a distance of approximately $2\,\mathrm{Re}(C)/n$, and rotates approximately

$$2\pi + \frac{2\,\mathrm{Im}(C)}{n}.$$

The axis moves out toward infinity as $n \to \infty$. For C sufficiently large, the image of ρ_n will be a geometrically finite group (a Schottky group); a compact convex manifold with $\pi_1 = \rho_n(\Gamma)$ can be constructed by piecing together a neighborhood of $M_{[\epsilon,\infty)}$ with (the convex hull of a helix)/$\mathbb{Z}$. The algebraic limit of $\{\rho_n\}$ is ρ, while the geometric limit is the group generated by $\rho(\Gamma) = \Gamma$ together with an extra parabolic generator commuting with $[a, b]$.



9.6

Troels Jørgensen was the first to analyze and understand this phenomenon. He showed that it is possible to iterate this construction and produce examples as above where the algebraic limit is the fundamental group of a punctured torus, but the geometric limit is not even finitely generated. See § .

Here are some basic properties of convergence of sequences of discrete groups.

Proposition 9.1.5. *If* $\{\rho_i\}$ *converges algebraically to* ρ *and* $\{\rho_i \Gamma\}$ *converges geometrically to* Γ'*, then* $\Gamma' \supset \rho\Gamma$*.*

Proof. Obvious. □

Proposition 9.1.6. *For any Lie group* G*, the space of closed subgroups of* G *(with the geometric topology) is compact.*

Proof. Let $\{\Gamma_i\}$ be any sequence of closed subgroups. First consider the case that there is a lower bound to the "size" $d(e, \gamma)$ of elements of $\gamma \in \Gamma_i$. Then there is an upper bound to the number of elements of Γ_i in the ball of radius γ about e, for every γ. The Tychonoff product theorem easily implies the existence of a subsequence converging geometrically to a discrete group.

Now let S be a maximal subspace of $T_e(G)$, the tangent space of G at the identity element e, with the property that for any $\epsilon > 0$ there is a Γ_i whose ϵ-small elements fill out all directions in S, within an angle of ϵ. It is easy to see that S is closed under Lie brackets. Furthermore, a subsequence $\{\Gamma_{i_j}\}$ whose small elements fill out S has 9.7 the property that all small elements are in directions near S. It follows, just as in the previous case, that there is a subsequence converging to a closed subgroup whose tangent space at e is S.

Corollary 9.1.7. *The set of complete hyperbolic manifolds* N *together with base frames in* $N_{[\epsilon,\infty)}$ *is compact in the geometric topology.*

□

COROLLARY 9.1.8. *Let Γ be any countable group and $\{\rho_i\}$ a sequence of discrete representations of Γ in $\mathrm{PSL}(2,\mathbb{C})$ converging algebraically to a representation ρ. If $\rho\Gamma$ does not have an abelian subgroup of finite index then $\{\rho_i\}$ has a subsequence converging geometrically to a discrete group $\Gamma' \supset \circ\Gamma$. In particular, $\rho\Gamma$ is discrete.*

PROOF. By 9.1.7, there is a subsequence converging geometrically to *some* closed group Γ'. By 5.10.1, the identity component of Γ' must be abelian; since $\rho\Gamma \subset \Gamma'$, the identity component is trivial. □

Note that if the ρ_i are all faithful, then their algebraic limit is also faithful, since there is a lower bound to $d(\rho_i\gamma x, x)$. These basic facts were first proved in ????

Here is a simple example negating the converse of 9.1.8. Consider any discrete group $\Gamma \subset \mathrm{PSL}(2,\mathbb{C})$ which admits an automorphism ϕ of infinite order: for instance, Γ might be a fundamental group of a surface. The sequence of representations ϕ^i has no algebraically convergent subsequence, yet $\{\phi^i\Gamma\}$ converges geometrically to Γ. 9.8

There are some simple statements about the behavior of limit sets when passing to a limit. First, if Γ is the geometric limit of a sequence $\{\Gamma_i\}$, then each point $x \in L_\Gamma$ is the limit of a sequence $x_i \in L_{\Gamma_i}$. In fact, fixed points x (eigenvectors) of non-trivial elements of $\gamma \in \Gamma$ are dense in L_Γ; for high i, Γ_i must have an element near γ, with a fixed point near x. A similar statement follows for the algebraic limit ρ of a sequence of representations ρ_i. Thus, the limit set cannot suddenly increase in the limit. It may suddenly decrease, however. For instance, let $\Gamma \subset \mathrm{PSL}(2,\mathbb{C})$ be any finitely generated group. Γ is residually finite (see §), or in other words, it has a sequence $\{\Gamma_i\}$ of subgroups of finite index converging geometrically to the trivial group (e). $L_{\Gamma_i} = L_\Gamma$ is constant, but $L_{(e)}$ is empty. It is plausible that every finitely generated discrete group $\Gamma \subset \mathrm{PSL}(2,\mathbb{C})$ be a geometric limit of groups with compact quotient.

We have already seen (in 9.1.4) examples where the limit set suddenly decreases in an algebraic limit.

Let Γ be the fundamental group of a surface S with finite area and $\{\rho_i\}$ a sequence of faithful quasi-Fuchsian representations of Γ, preserving parabolicity. Suppose $\{\rho_i\}$ converges algebraically to a representation ρ as a group without any additional parabolic elements. Let N denote $N_{\rho(\Gamma)}$, N_i denote $N_{\rho_i(\Gamma)}$, etc. 9.9

THEOREM 9.2. *N is geometrically tame, and $\{\rho_i\}$ converges strongly to ρ.*

PROOF. If the set of uncrumpled maps of S into N homotopic to the standard map is compact, then using a finite cover of $\mathcal{GL}(S)$ carried by nearly straight train tracks, one sees that for any discrete representation ρ' near ρ, every geodesic lamination γ of S is realizable in N' near its realizations in N. (Logically, one can think of uncrumpled surfaces as equivariant uncrumpled maps of M^2 into H^3, with the compact-open topology, so that "nearness" makes sense.) Choose any subsequence of the ρ_i's so that the bending loci for the two boundary components of M_i converge in $\mathcal{GL}(S)$. Then the two boundary components must converge to locally convex disjoint embeddings of S in N (unless the limit is Fuchsian). These two surfaces are homotopic, hence they bound a convex submanifold M of N, so $\rho(\Gamma)$ is geometrically finite.

Since $M_{[\epsilon,\infty)}$ is compact, strong convergence of $\{\rho_i\}$ follows form 8.3.3: no unexpected identifications of N can be created by a small perturbation of ρ which preserves parabolicity.

If the set of uncrumpled maps of S homotopic to the standard map is not compact, then it follows immediately from the definition that N has at least one geometrically infinite tame end. We must show that both ends are geometrically tame. The possible phenomenon to be wary of is that the bending loci β_i^+ and β_i^- of the two boundary components of M_i might converge, for instance, to a single point λ in $\mathcal{GL}(S)$. (This would be conceivable if the "simplest" homotopy of one of the two boundary compo- 9.10
nents to a reference surface which persisted in the limit first carried it to the vicinity of the other boundary component.) To help in understanding the picture, we will first find a restriction for the way in which a hyperbolic manifold with a geometrically tame end can be a covering space.

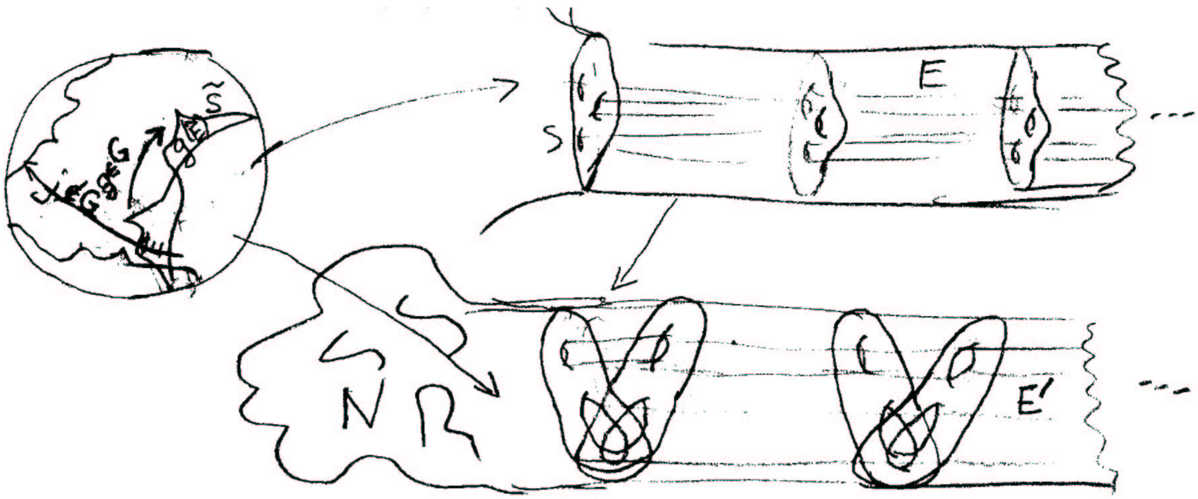
Definition 9.2.1. Let N be a hyperbolic manifold, P a union of horoball neighborhoods of its cusps, E' an end of $N - P$. E' is *almost geometrically tame* if some finite-sheeted cover of E' is (up to a compact set) a geometrically tame end. (Later we shall prove that if E is almost geometrically tame it is geometrically tame.)

Theorem 9.2.2. *Let N be a hyperbolic manifold, and $\tilde{N}$ a covering space of N such that $\tilde{N} - \tilde{P}$ has a geometrically infinite tame end E bounded by a surface $S_{[\epsilon,\infty)}$. Then either N has finite volume and some finite cover of N fibers over S^1 with fiber S, or the image of E in $N - P$, up to a compact set, is an almost geometrically tame end of N.*

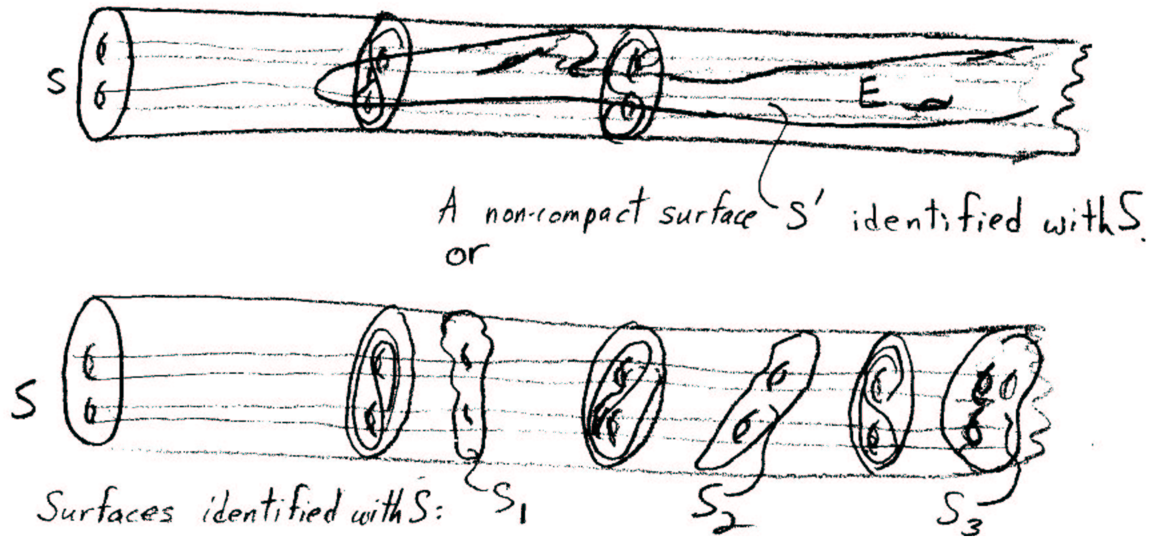
Proof. Consider first the case that all points of E identified with $S_{[\epsilon,\infty)}$ in the projection to N lie in a compact subset of E. Then the local degree of the projection of E to N is finite in a neighborhood of the image of S. Since the local degree is constant except at the image of S, it is everywhere finite.

Let $G \subset \pi_1 N$ be the set of covering transformations of H^3 over N consisting of 9.11
elements g such that $g\tilde{E} \cap \tilde{E}$ is all of $\tilde{E}$ except for a bounded neighborhood of $\tilde{S}$. G

is obviously a group, and it contains $\pi_1 S$ with finite index. Thus the image of E, up to compact sets, is an almost geometrically tame end of N.

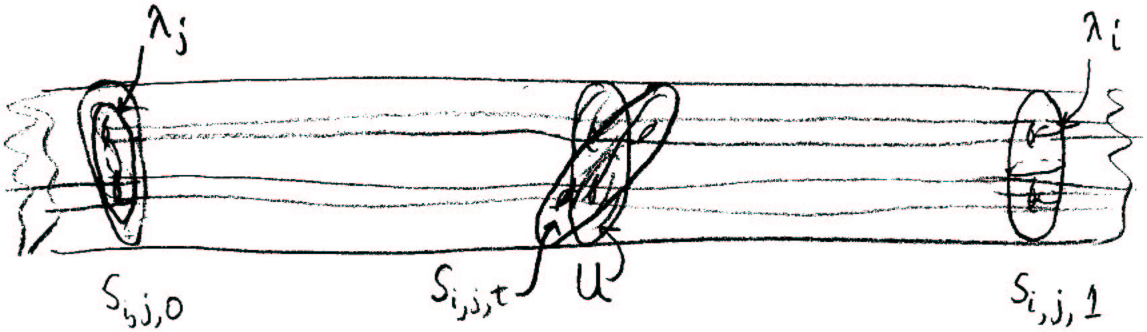


The other case is that $S_{[\epsilon,\infty)}$ is identified with a non-compact subset of E by projection to N. Consider the set I of all uncrumpled surfaces in E whose images intersect the image of $S_{[\epsilon,\infty)}$. Any short closed geodesic on an uncrumpled surface of E is homotopic to a short geodesic of E (not a cusp), since E contains no cusps other than the cusps of S. Therefore, by the proof of 8.8.5, the set of images of I in N is precompact (has a compact closure). If I itself is not compact, then N has a finite cover which fibers over S^1, by the proof of 8.10.9. If I is compact, then (since uncrumpled surfaces cut E into compact pieces), infinitely many components of the set of points identified with $S_{[\epsilon,\infty)}$ are compact and disjoint from S. 9.12



These components consist of immersions of k-sheeted covering spaces of S injective on π_1, which must be homologous to $\pm k\,[S]$. Pick two disjoint immersions with the same sign, homologous say to $-k\,[S]$ and $-l\,[S]$. Appropriate multiples of these cycles are homologous by a compactly supported three-chain which maps to a three-cycle in $N - P$, hence N has finite volume. Theorem 9.2.2 now follows from 8.10.9. $\square$

We continue the proof of Theorem 9.2. We may, without loss of generality, pass to a subsequence of representations ρ_i such that the sequences of bending loci $\{\beta_i^+\}$ and $\{\beta_i^-\}$ converge, in $\mathcal{PL}_0(S)$, to laminations β^+ and β^-. If β^+, say, is realizable for the limit representation ρ, then any uncrumpled surface whose wrinkling locus contains β^+ is embedded and locally convex—hence it gives a geometrically finite end of N. The only missing case for which we must prove geometric tameness is that neither 9.13
β^+ nor β^- is realizable. Let $\lambda_i^\epsilon \in \mathcal{PL}_0(S)$ (where $\epsilon = +, -$) be a sequence of geodesic laminations with finitely many leaves and with transverse measures approximating β_i^ϵ closely enough that the realization of λ_i^ϵ in N_i is near the realization of β_i^ϵ. Also suppose that $\lim \lambda_i^\epsilon = \beta^\epsilon$ in $\mathcal{PL}_0(S)$. The laminations λ_i^ϵ are all realized in N. They must tend toward ∞ in N, since their limit is not realized. We will show that they tend toward ∞ in the ϵ-direction. Imagine the contrary—for definiteness, suppose that the realizations of $\{\lambda_i^+\}$ in N go to ∞ in the $-$ direction. The realization of each λ_i^+ in N_j must be near the realization in N, for high enough j. Connect λ_j^+ to λ_i^+ by a short path $\lambda_{i,j,t}$ in $\mathcal{PL}_0(S)$. A family of uncrumpled surfaces $S_{i,j,t}$ realizing the $\lambda_{i,j,t}$ is not continuous, but has the property that for t near t_0, $S_{i,j,t}$ and S_{i,j,t_0} have points away from their cusps which are close in N. Therefore, for every uncrumpled surface U between $S_{i,j,0}$ and $S_{i,j,1}$ (in a homological sense), there is some t such that $S_{i,j,t} \cap U \cap (N - P)$ is non-void.



9.14

Let γ be any lamination realized in N, and U_j be a sequence of uncrumpled surfaces realizing γ in N_j, and converging to a surface in N. There is a sequence $S_{i(j),j,t(j)}$ of uncrumpled surfaces in N_j intersecting U_j whose wrinkling loci tend toward β^+.

Without loss of generality we may pass to a geometrically convergent subsequence, with geometric limit Q. Q is covered by N. It cannot have finite volume (from the analysis in Chapter 5, for instance), so by 8.14.2, it has an almost geometrically tame end E which is the image of the $-$ end E_- of N. Each element α of $\pi_1 E$ has a finite power $\alpha^k \in \pi_1 E_-$. Then a sequence $\{\alpha_i\}$ approximating α in $\pi_1(N_i)$ has the property that the α_i^k have bounded length in the generators of $\pi_1 S$, this implies that the α_i have bounded length, so α is in fact in $\pi_1 E_-$, and $E_- = E$ (up to compact sets). Using this, we may pass to a subsequence of $S_{i(j),j,t}$'s which converge to an uncrumpled surface R in E. R is incompressible, so it is in the standard homotopy class. It realizes β^+, which is absurd.

there is no 8.14.2

We may conclude that N has two geometrically tame ends, each of which is mapped homeomorphically to the geometric limit Q. (This holds whether or not they are geometrically infinite.) This implies the local degree of $N \to Q$ is finite one or two (in case the two ends are identified in Q). But any covering transformation α of N over Q has a power (its square) in $\pi_1 N$, which implies, as before, that $\alpha \in \pi_1 N$, so that $N = Q$. This concludes the proof of 9.2. □

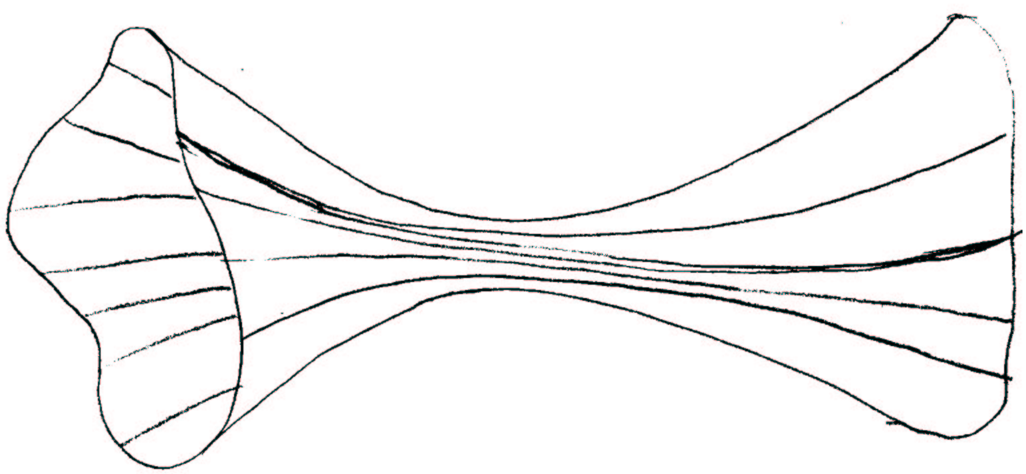
9.3. The ending of an end

In the interest of avoiding circumlocution, as well as developing our image of a geometrically tame end, we will analyze the possibilities for non-realizable laminations in a geometrically tame end.

We will need an estimate for the area of a cylinder in a hyperbolic three-manifold. Given any map $f : S^1 \times [0,1] \to N$, where N is a convex hyperbolic manifold, we may straighten each line $\theta \times [0,1]$ to a geodesic, obtaining a ruled cylinder with the same boundary.

THEOREM 9.3.1. *The area of a ruled cylinder (as above) is less than the length of its boundary.*

PROOF. The cylinder can be C^0-approximated by a union of small quadrilaterals each subdivided into two triangles. The area of a triangle is less than the minimum of the lengths of its sides (see p. 6.5). □



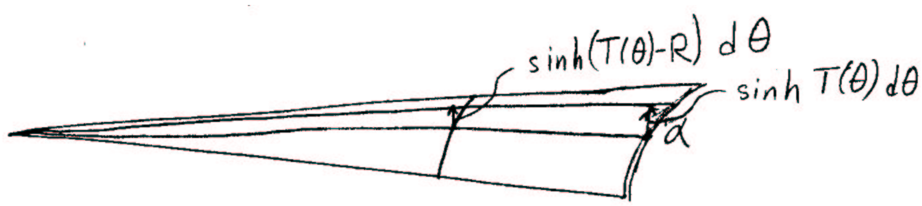
9.16

If the two boundary components of the cylinder C are far apart, then most of the area is concentrated near its boundary. Let γ_1 and γ_2 denote the two components of ∂C.

THEOREM 9.3.2. *Area* $(C - \mathcal{N}_r \gamma_1) \leq e^{-r}\, l(\gamma_1) + l(\gamma_2)$ *where* $r \geq 0$ *and* l *denotes length.*

This is derived by integrating the area of a triangle in polar coordinates from any vertex:

$$A = \int \int_0^{T(\theta)} \sinh t \, dt \, d\theta = \int (\cosh T(\theta) - 1)\, d\theta$$



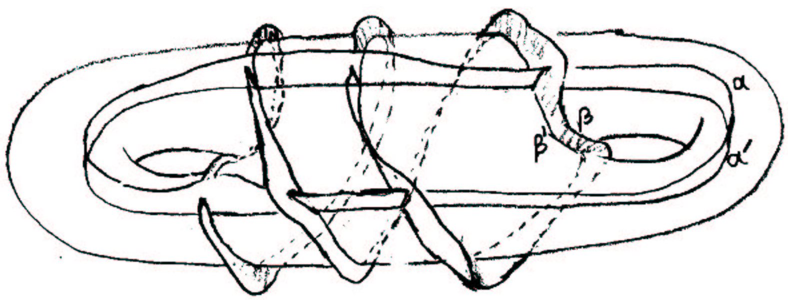
The area outside a neighborhood of radius r of its far edge α is

$$\int \cosh\,(T(\theta) - r) - 1 \, d\theta < e^{-r} \int \sinh T(\theta)\, d\theta < e^{-r}\, l(\alpha).$$

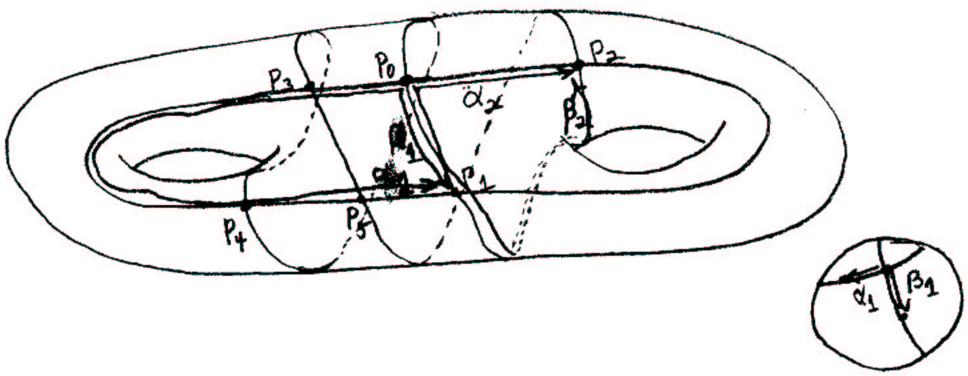
This easily implies 9.3.2

Let E be a geometrically tame end, cut off by a surface $S_{[\epsilon,\infty)}$ in $N-P$, as usual. A curve α in E homotopic to a simple closed curve α' on S gives rise to a ruled cylinder $C_\alpha : S^1 \times [0,1] \to N$.

Now consider two curves α and β homotopic to simple closed curves α' and β' 9.17
on S. One would expect that if α' and β' are forced to intersect, then either α must intersect C_β or β must intersect C_α, as in 8.11.1



We will make this more precise by attaching an invariant to each intersection. Let us assume, for simplicity, that α' and β' are geodesics with respect to some hyperbolic structure on S. Choose one of the intersection points, p_0, of α' and β' as a base point for N. For each other intersection point p_i, let α_i and β_i be paths on α' and β' from p_0 to p_i. Then $\alpha_i * \beta_i^{-1}$ is a closed loop, which is non-trivial in $\pi_1(S)$ when $i \neq 0$ since two geodesics in $\tilde{S}$ have at most one intersection.



There is some ambiguity, since there is more than one path from α_0 to α_i on 9.18
α'; in fact, α_i is well-defined up to a power of α'. Let $\langle g\rangle$ denote the cyclic group generated by an element g. Then $\alpha_i \cdot \beta_i^{-1}$ gives a well-defined element of the double coset space $\langle\alpha'\rangle\backslash\pi_1(S)/\langle\beta'\rangle$. [The double coset $H_1 g H_2 \in H_1\backslash G/H_2$ of an element $g \in G$ is the set of all elements $h_1 g h_2$, where $h_i \in H_i$.] The double cosets associated to two different intersections p_i and p_j are distinct: if $\langle\alpha'\rangle\alpha_i\beta_i^{-1}\langle\beta'\rangle = \langle\alpha'\rangle a_j\beta_j^{-1}\langle\beta'\rangle$, then there is some loop $\alpha_j^{-1}\alpha'^k\alpha_i\beta_i^{-1}\beta'^l\beta_j$ made up of a path on α' and a path on

β' which is homotopically trivial—a contradiction. In the same way, a double coset $D_{x,y}$ is attached to each intersection of the cylinders C_α and C_β. Formally, these intersection points should be parametrized by the domain: thus, an intersection point means a pair $(x, y) \in (S^1 \times I) \times (S' \times I)$ such that $C_\alpha x = C_\beta y$.

Let $i(\gamma, \delta)$ denote the number of intersections of any two simple geodesics γ and δ on S. Let $D(\gamma, \delta)$ be the set of double cosets attached to intersection points of γ and δ (including p_0). Thus $i(\gamma, \delta) = |D(\gamma, \delta)|$. $D(\alpha, C_\beta)$ and $D(C_\alpha, \beta)$ are defined similarly.

Proposition 9.3.3. $|\alpha \cap C_\beta| + |C_\alpha \cap \beta| \geq i(\alpha', \beta')$. *In fact*

$$D(a, C_\beta) \cup D(C_\alpha, \beta) \supset D(\alpha', \beta').$$

9.19

Proof. First consider cylinders C'_α and C'_β which are contained in E, and which are nicely collared near S. Make C'_α and C'_β transverse to each other, so that the double locus $L \subset (S^1 \times I) \times (S^1 \times I)$ is a one-manifold, with boundary mapped to $\alpha \cup \beta \cup \alpha' \cup \beta'$. The invariant $D_{(x,y)}$ is locally constant on L, so each invariant occurring for $\alpha' \cap \beta'$ occurs for the entire length of interval in L, which must end on α or β. In fact, each element of $D(\alpha', \beta')$ occurs as an invariant of an odd number of points $\alpha \cup \beta$.

Now consider a homotopy h_t of C'_β to C_β, fixing $\beta \cup \beta'$. The homotopy can be perturbed slightly to make it transverse to α, although this may necessitate a slight movement of C_β to a cylinder C''_β. Any invariant which occurs an odd number of times for $a \cap C'_\beta$ occurs also an odd number of times for $\alpha \cap C''_\beta$. This implies that the invariant must also occur for $a \cap C_\beta$. □

Remark. By choosing orientations, we could of course associate signs to intersection points, thereby obtaining an algebraic invariant $\mathcal{D}(\alpha', \beta') \in \mathbb{Z}^{\langle\alpha'\rangle\backslash\pi_1 S/\langle\beta'\rangle}$. Then 9.3.3 would become an equation,

$$\mathcal{D}(\alpha', \beta') = \mathcal{D}(\alpha, C_\beta) + \mathcal{D}(C_\alpha, \beta).$$

Since $\pi_1(S)$ is a discrete group, there is a restriction on how closely intersection points can be clustered, hence a restriction on $|D(\alpha, c_\beta)|$ in terms of the length of α times the area of C_β.

9.20

Proposition 9.3.4. *There is a constant K such that for every curve α in E with distance R from S homotopic to a simple closed curve α' on S and every curve β in E not intersecting C_α and homotopic to a simple curve β' on S,*

$$i(\alpha', \beta') \leq K\big[l(\alpha) + \big(l(\alpha) + 1\big)\big(l(\beta) + e^{-R} + l(\beta')\big)\big].$$

Proof. Consider intersection points $(x, y) \in S^1 \times (S^1 \times I)$ of α and C_β. Whenever two of them, (x, y) and (x', y'), are close in the product of the metrics induced from N, there is a short loop in N which is non-trivial if $D_{(x,y)} \neq D_{(x',y')}$.

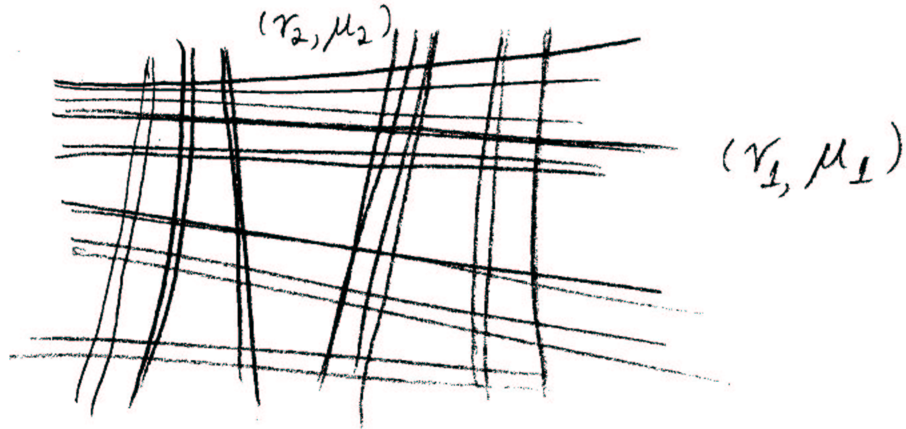
Case (i). α is a short loop. Then there can be no short non-trivial loop on C_β near an intersection point with α. The disks of radius ϵ on C_β about intersection points with α have area greater than some constant, except in special cases when they are near ∂C_β. If necessary, extend the edges of C_β slightly, without substantially changing the area. The disks of radius ϵ must be disjoint, so this case follows from 9.3.2 and 9.3.3.

Case (ii). α is not short. Let $E \subset C_\beta$ consist of points through which there is a short loop homotopic to β. If (x, y) and (x', y') are intersection points with $D_{x,y} \neq D_{x',y'}$ and with y, y' in E, then x and x' cannot be close together—otherwise two distinct conjugates of β would be represented by short loops through the same point. The number of such intersections is thus estimated by some constant times $l(\alpha)$. 9.21

Three intersections of α with $C_\beta - E$ cannot occur close together. $S^1 \times (C_\beta - E)$ contains the balls of radius ϵ, with multiplicity at most 2, and each ball has a definite volume. This yields 9.3.4. □

Let us generalize 9.3.4 to a statement about measured geodesic laminations. Such a lamination (γ, μ) on a hyperbolic surface S has a well-defined "average length" $l_S(\gamma, \mu)$. This can be defined as the total mass of the measure which is locally the product of the transverse measure μ with one-dimensional Lebesgue measure on the leaves of γ. Similarly, a realization of γ in a homotopy class $f : S \to N$ has a length $l_f(\gamma, \mu)$. The length $l_S(\gamma, \mu)$ is a continuous function on $\mathcal{ML}_0(S)$, and $l_f(\gamma)$ is a continuous function where defined. If γ is realized a distance of R from an uncrumpled surface S, then $l_f(\gamma, \mu) \leq (1/\cosh R) l_S(\gamma, \mu)$. This implies that if f preserves non-parabolicity, l_f extends continuously over all of $\mathcal{ML}_0$ so that its zero set is the set of non-realizable laminations.

The intersection number $i\big((\gamma_1, \mu_1), (\gamma_2, \mu_2)\big)$ of two measured geodesic laminations is defined similarly, as the total mass of the measure $\mu_1 \times \mu_2$ which is locally the product of μ_1 and μ_2. (This measure $\mu_1 \times \mu_2$ is interpreted to be zero on any common leaves of γ_1 and γ_2.)



9.22

Given a geodesic lamination γ realized in E, let d_γ be the miniaml distance of an uncrumpled surface through γ from $S_{[\epsilon,\infty)}$.

THEOREM 9.3.5. *There is a constant K such that for any two measured geodesic laminations (γ_1, μ_1) and $(\gamma_2, \mu_2) \in \mathcal{ML}_0(S)$ realized in E,*

$$i\big((\gamma_1, \mu_1), (\gamma_2, \mu_2)\big) \leq K \cdot e^{-2R} l_S(\gamma_1, \mu_1) \cdot l_S(\gamma_2, \mu_2)$$

where $R = \inf(d_{\gamma_1}, d_{\gamma_2})$.

PROOF. First consider the case that γ_1 and γ_2 are simple closed geodesics which are not short. Apply the proof of 9.3.4 first to intersections of γ_1 with C_{γ_2}, then to intersections of C_{γ_1} with γ_2. Note that $l_S(\gamma_i)$ is estimated from below by $e^R l(\gamma_i)$, so the terms involving $l(\gamma_i)$ can be replaced by $Ce^{-R} l(\gamma_i)$. Since γ_1 and γ_2 are not short, one obtains

$$i(\gamma_1, \gamma_2) \leq K \cdot e^{-2R}\, l_S(\gamma_1)\, l_S(\gamma_2),$$

for some constant K. Since both sides of the inequality are homogeneous of degree one in γ_1 and γ_2, it extends by continuity to all of $\mathcal{ML}_0(S)$. □

Consider any sequence $\{(\gamma_i, \mu_i)\}$ of measured geodesic laminations in $\mathcal{ML}_0(S)$ whose realizations go to ∞ in E. If (λ_1, μ_1) and (λ_2, μ_2) are any two limit points of this sequence, 9.3.5 implies that $i(\lambda_1, \lambda_2) = 0$: in other words, the leaves do not cross. The union $\lambda_1 \cup \lambda_2$ is still a lamination. 9.23

DEFINITION 9.3.6. The *ending* lamination $\epsilon(E) \in \mathcal{GL}(S)$ is the union of all limit points λ_i, as above.

Clearly, $\epsilon(E)$ is compactly supported and it admits a measure with full support. The set $\Delta(E) \subset \mathcal{PL}_0(S)$ of all such measures on $\epsilon(E)$ is closed under convex combinations, hence its intersection with a local coordinate system (see p. 8.59) is convex.

In fact, a maximal train track carrying $\epsilon(E)$ defines a single coordinate system containing $\Delta(E)$.

The idea that the realization of a lamination depends continuously on the lamination can be generalized to the ending lamination $\epsilon(E)$, which can be regarded as being realized at ∞.

PROPOSITION 9.3.7. *For every compact subset K of E, there is a neighborhood U of $\Delta(E)$ in $\mathcal{PL}_0(S)$ such that every lamination in $U - \Delta(E)$ is realized in $E - K$.*

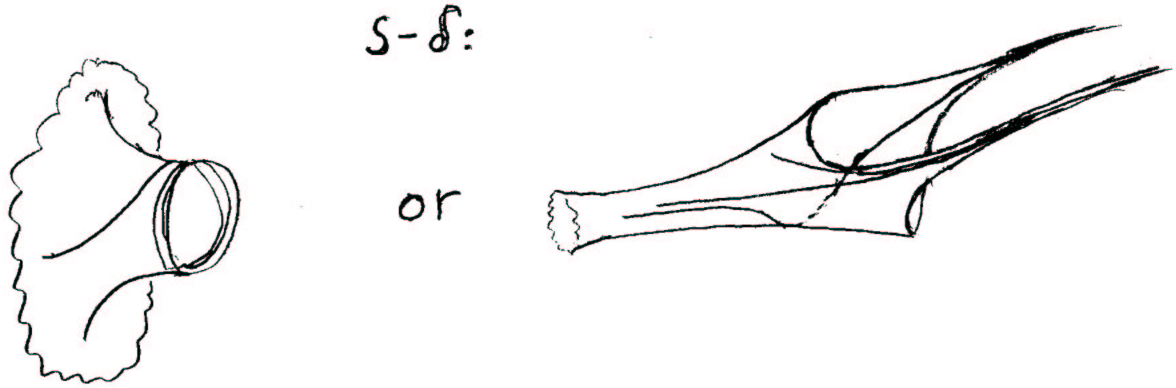
PROOF. It is convenient to pass to the covering of N corresponding to $\pi_1 S$. Let S' be an uncrumpled surface such that K is "below" S' (in a homological sense). Let $\{V_i\}$ be a neighborhood basis for $\Delta(E)$ such that $V_i - \Delta(E)$ is path-connected, and let $\lambda_i \in V_i - \Delta(E)$ be a sequence whose realizations go to ∞ in E. If there is any point $\pi_i \in V_i - \Delta(E)$ which is a non-realizable lamination or whose realization is not "above" S', connect λ_i to π_i by a path in V_i. There must be some element of this path whose realization intersects $S'_{[\epsilon,\infty)}$ (since the realizations cannot go to ∞ while in E.) Even if certain non-peripheral elements of S are parabolic, excess pinching of non-peripheral curves on uncrumpled surfaces intersecting S' can be avoided if S' is far from S, since there are no extra cusps in E. Therefore, only finitely many such π_i's can occur, or else there would be a limiting uncrumpled surface through S realizing the unrealizable. □

PROPOSITION 9.3.8. *Every leaf of $\epsilon(E)$ is dense in $\epsilon(E)$, and every non-trivial simple curve in the complement of $\epsilon(E)$ is peripheral.*

PROOF. The second statement follows easily from 8.10.8, suitably modified if there are extra cusps. The first statement then follows from the next result:

PROPOSITION 9.3.9. *If γ is a geodesic lamination of compact support which admits a nowhere zero transverse measure, then either every leaf of γ is dense, or there is a non-peripheral non-trivial simple closed curve in $S - \gamma$.*

PROOF. Suppose $\delta \subset \gamma$ is the closure of any leaf. Then δ is also an open subset of γ: all leaves of γ near δ are trapped forever in a neighborhood of δ. This is seen by considering the surface $S - \delta$.



9.25

An arc transverse to these leaves would have positive measure, which would imply that a transverse arc intersecting these leaves infinitely often would have infinite measure. (In general, a closed union of leaves $\delta \subset \gamma$ in a general geodesic lamination has only a finite set of leaves of γ intersecting a small neighborhood.)

If $\delta \neq \gamma$, then δ has two components, which are separated by some homotopically non-trivial curve in $S - \gamma$. □

□

Corollary 9.3.10. *For any homotopy class of injective maps $f : S \to N$ from a hyperbolic surface of finite area to a complete hyperbolic manifold, if f preserves parabolicity and non-parabolicity, there are $n = 0$, 1 or 2 non-realizable laminations ϵ_i $[1 \leq i \leq n]$ such that a general lamination γ on S is non-realizable if and only if the union of its non-isolated leaves is an ϵ_i.*

9.4. Taming the topology of an end

We will develop further our image of a geometrically tame end, once again to avoid circumlocution.

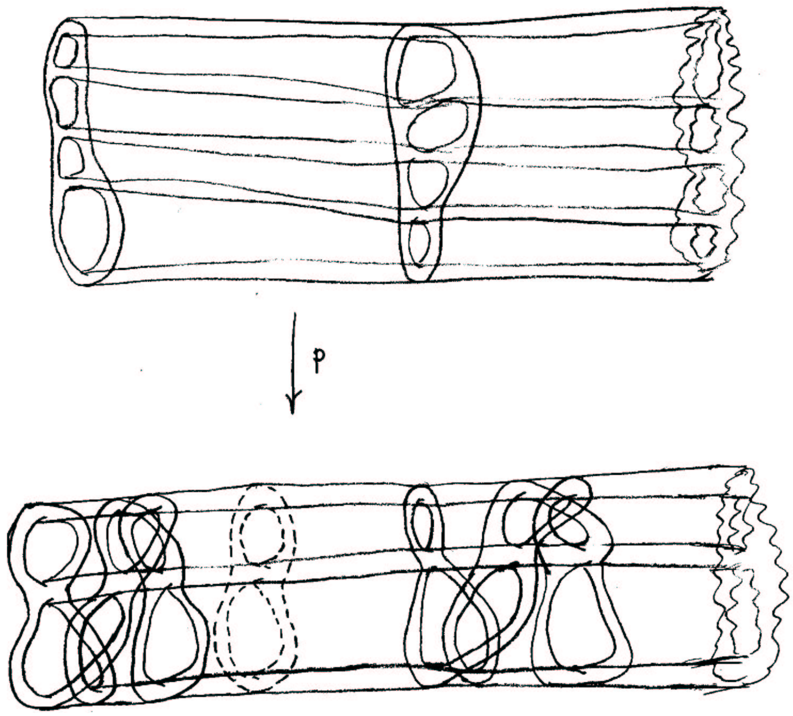
Theorem 9.4.1. *A geometrically tame end $E \subset N - P$ is topologically tame. In other words, E is homeomorphic to the product $S_{[\epsilon,\infty)} \times [0, \infty)$.*

Theorem 9.4.1 will be proved in §§9.4 and 9.5.

Corollary 9.4.2. *Almost geometrically tame ends are geometrically tame.*

Proof that 9.4.1 implies 9.4.2. Let E' be an almost geometrically tame end, finitely covered (up to compact sets) by a geometrically tame end $E = S_{[\epsilon,\infty)} \times [0, \epsilon)$, 9.26
with projection $p : E \to E'$. Let $f : E' \to [0, \epsilon)$ be a proper map. The first step is to find an incompressible surface $S' \subset E'$ which cuts it off (except for compact sets).

Choose t_0 high enough that $p : E \to E'$ is defined on $S_{[\epsilon,\infty)} \times [t_0, \infty)$, and choose $t_1 > t_0$ so that $p\big(S_{[\epsilon,\infty)} \times [t_1, \infty)\big)$ does not intersect $p(S_{[\epsilon,\infty)} \times t_0)$.



9.27

Let $r \in [0, \infty)$ be any regular value for f greater than the supremum of $f \circ p$ on $S_{[\epsilon,\infty)} \times [0, t_1)$. Perform surgery (that is, cut along circles and add pairs of disks) to $f^{-1}(r)$, to obtain a not necessarily connected surface S' in the same homology class which is incompressible in

$$E' - p\big(S_{[\epsilon,\infty)} \times [0, t_0)\big).$$

The fundamental group of S' is still generated by loops on the level set $f = r$. S' is covered by a surface $\tilde{S}'$ in E. $\tilde{S}'$ must be incompressible in $E-$ otherwise there would be a non-trivial disk D mapped into $S_{[\epsilon,\infty)} \times [t_1, \infty)$ with boundary on $\tilde{S}$; $p \circ D$ would be contained in

$$E' - p(S_{[\epsilon,\infty)} \times [0, t_0])$$

so S' would not be incompressible (by the loop theorem). One deduces that $\tilde{S}'$ is homotopic to $S_{[\epsilon,\infty)}$ and S' is incompressible in $N - P$.

If E is geometrically finite, there is essentially nothing to prove—E corresponds to a component of $\partial\tilde{M}$, which gives a convex embedded surface in E'. If E is geometrically infinite, then pass to a finite sheeted cover E'' of E which is a regular cover of E'. The ending lamination $\epsilon(E'')$ is invariant under all diffeomorphisms (up to compact sets) of E''. Therefore it projects to a non-realizable geodesic lamination $\epsilon(E')$ on S'. □

Proof of 9.4.1. We have made use of one-parameter families of uncrumpled surfaces in the last two sections. Unfortunately, these surfaces do not vary continuously. To prove 9.4.1, we will show, in §9.5, how to interpolate with more general surfaces, to obtain a (continuous) proper map $F\colon S_{[\epsilon,\infty)} \times [0,\infty) \to E$. The theorem will follow fairly easily once F is constructed: 9.28

Proposition 9.4.3. *Suppose there is a proper map $F : S_{[\epsilon,\infty)} \times [0,\infty) \to E$ with $F(S_{[\epsilon,\infty)} \times 0)$ standard and with $F\big(\partial S_{[\epsilon,\infty)} \times [0,\infty)\big) \subset \partial(N - P)$. Then E is homeomorphic to $S_{[\epsilon,\infty)} \times [0,\infty)$.*

Proof of 9.4.3. This is similar to 9.4.2. Let $f : E \to [0,\infty)$ be a proper map. For any compact set $K \subset E$, we can find a $t_1 > 0$ so that $F\big(S_{[\epsilon,\infty)} \times [t_1,\infty)\big)$ is disjoint from K. Let r be a regular value for f greater than the supremum of $f \circ F$ on $S_{[\epsilon,\infty)} \times [0,t_1]$. Let $S' = f^{-1}(r)$ and $S'' = (f \circ F)^{-1}(r)$. $F : S'' \to S'$ is a map of degree one, so it is surjective on π_1 (or else it would factor through a non-trivial covering space on S', hence have higher degree). Perform surgery on S' to make it incompressible in the complement of K, without changing the homology class. Now S' must be incompressible in E; otherwise there would be some element α of $\pi_1 S'$ which is null-homotopic in E. But α comes from an element β on S'' which is null-homotopic in $S_{[\epsilon,\infty)} \times [t_1,\infty)$, so its image α is null-homotopic in the complement of K. It follows that S' is homotopic to $S_{[\epsilon,\infty)}$, and that the compact region of E cut off by S' is homeomorphic to $S_{[\epsilon,\infty)} \times I$. By constructing a sequence of such disjoint surfaces going outside of every compact set, we obtain a homeomorphism with $S_{[\epsilon,\infty)} \times [0,\infty)$. □

□

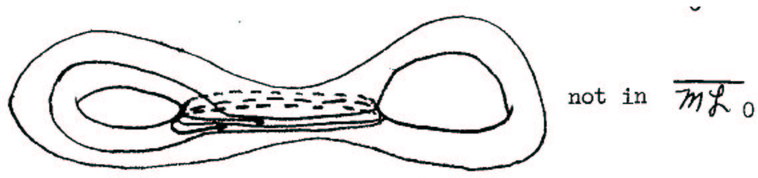
9.5. Interpolating negatively curved surfaces

Now we turn to the task of constructing a continuous family of surfaces moving out 9.29
to a geometrically infinite tame end. The existence of this family, besides completing the proof of 9.4.1, will show that a geometrically tame end has uniform geometry, and it will lead us to a better understanding of $\mathcal{ML}_0(S)$.

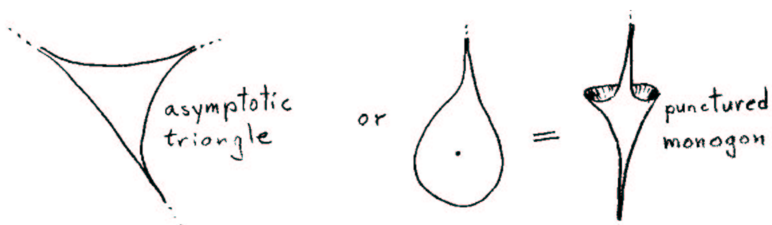
We will work with surfaces which are totally geodesic near their cusps, on esthetic grounds. Our basic parameter will be a family of compactly supported geodesic laminations in $\mathcal{ML}_0(S)$. The first step is to understand when a family of uncrumpled surfaces realizing these laminations is continuous and when discontinuous.

DEFINITION 9.5.1. For a lamination $\gamma \in \mathcal{ML}_0(S)$, let $\mathcal{T}_\gamma$ be the limit set in $\mathcal{GL}(S)$ of a neighborhood system for γ in $\mathcal{ML}_0(S)$. ($\mathcal{T}_\gamma$ is the "qualitative tangent space" of $\mathcal{ML}_0(S)$ at γ).

Let $\overline{\mathcal{ML}}_0(S)$ denote the closure of the image of $\mathcal{ML}_0(S)$ in $\mathcal{GL}(S)$. Clearly $\overline{\mathcal{ML}}_0(S)$ consists of laminations with compact support, but not every lamination with compact support is in $\overline{\mathcal{ML}}_0(S)$:



Every element of $\overline{\mathcal{ML}}_0$ is in $\mathcal{T}_\gamma$ for some $\gamma \in \mathcal{ML}_0$. Let us say that an element $\gamma \in \overline{\mathcal{ML}}_0$ is *essentially complete* if γ is a maximal element of $\overline{\mathcal{ML}}_0$. If $\gamma \in \mathcal{ML}_0$, then γ is essentially complete if and only if $\mathcal{T}_\gamma = \gamma$. A lamination γ is maximal among all compactly supported laminations if and only if each region of $S - \gamma$ is an asymptotic 9.30
triangle or a neighborhood of a cusp of S with one cusp on its boundary—a punctured monogon.

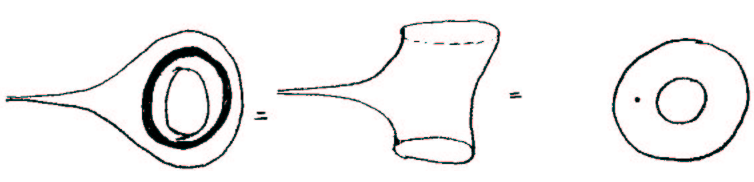


(These are the only possible regions with area π which are simply connected or whose fundamental group is peripheral.) Clearly, if $S - \gamma$ consists of such regions, then γ is essentially complete. There is one special case when essentially complete laminations are not of this form; we shall analyze this case first.

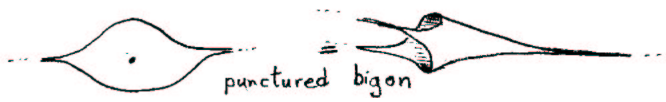
PROPOSITION 9.5.2. *Let $T - p$ denote the punctured torus. An element*

$$\gamma \in \overline{\mathcal{ML}}_0(T - p)$$

is essentially complete if and only if $(T - p) - \gamma$ is a punctured bigon.

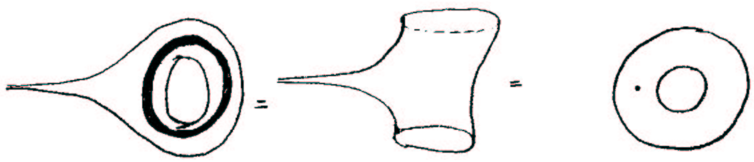


If $\gamma \in \mathcal{ML}_0(T - p)$, then either γ has a single leaf (which is closed), or every leaf of γ is non-compact and dense, in which case γ is essentially complete. If γ has a single closed leaf, then $\mathcal{T}_\gamma$ consists of γ and two other laminations:



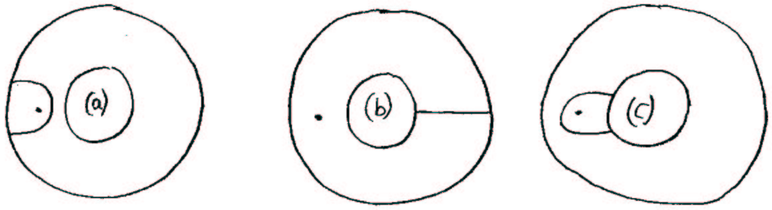
9.31

Proof. Let $g \in \mathcal{ML}_0(T - p)$ be a compactly supported measured lamination. First, note that the complement of a simple closed geodesic on $T - p$ is a punctured annulus,

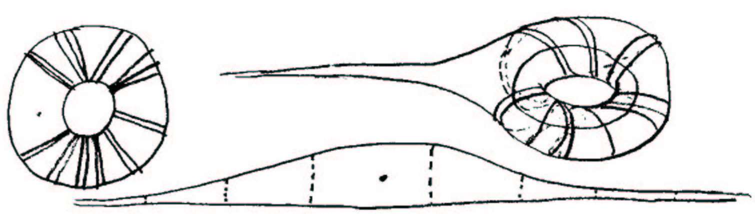


which admits no simple closed geodesics and consequently no geodesic laminations in its interior. Hence if γ contains a closed leaf, then γ consists only of this leaf, and otherwise (by 9.3.9) every leaf is dense.

Now let α be any simple closed geodesic on $T - p$, and consider γ cut apart by α. No end of a leaf of γ can remain forever in a punctured annulus, or else its limit set would be a geodesic lamination. Thus α cuts leaves of γ into arcs, and these arcs have only three possible homotopy classes:

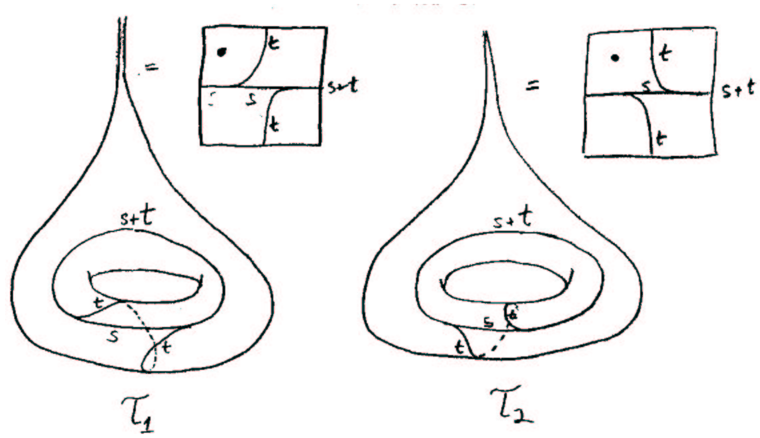


If the measure of the set of arcs of type (a) is m_a, etc., then (since the two boundary components match up) we have $2m_a + m_b = 2m_c + m_b$. But cases (a) and (c) 9.32
are incompatible with each other, so it must be that $m_a = m_c = 0$. Note that γ is orientable: it admits a continuous tangent vector field. By inspection we see a complementary region which is a punctured bigon.



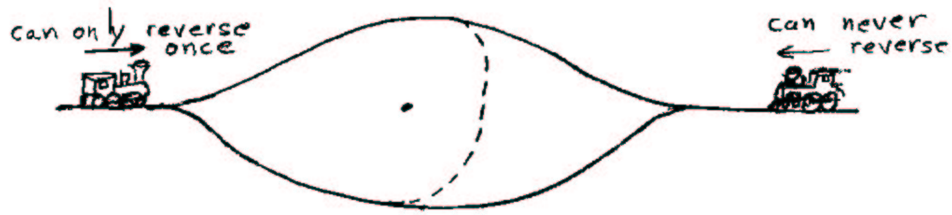
Since the area of a punctured bigon is 2π, which is the same as the area of $T - p$, this is the only complementary region.

It is now clear that a compactly supported measured lamination on $T - p$ with every leaf dense is essentially complete—there is nowhere to add new leaves under a small perturbation. If γ has a single closed leaf, then consider the families of measures on train tracks:

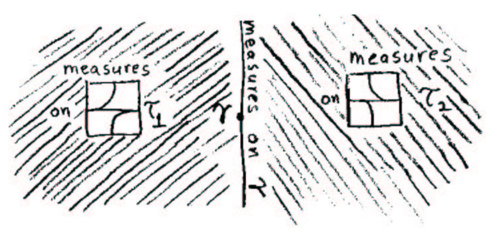


9.33

These train tracks cannot be enlarged to train tracks carrying measures. This can be deduced from the preceding argument, or seen as follows. At most one new branch could be added (by area considerations), and it would have to cut the punctured bigon into a punctured monogon and a triangle.



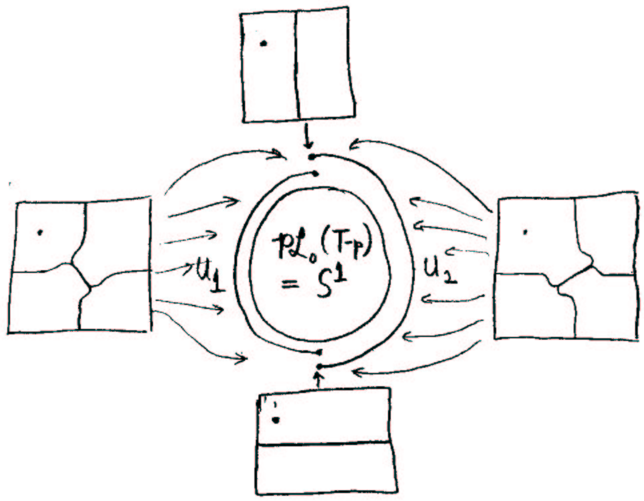
The train track is then orientable in the complement of the new branch, so a train can traverse this branch at most once. This is incompatible with the existence of a positive measure. Therefore $\mathcal{ML}_0(T - p)$ is two-dimensional, so τ_1 and τ_2 carry a neighborhood of γ.



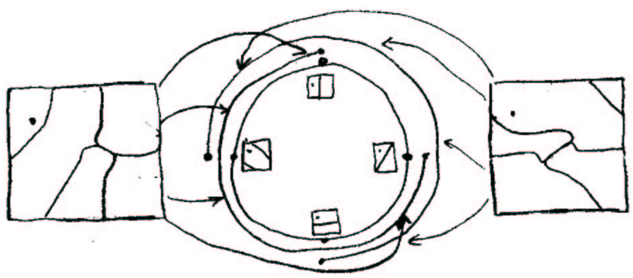
It follows that τ_γ is as shown. □

PROPOSITION 9.5.3. *$\mathcal{PL}_0(T - p)$ is a circle.*

PROOF. The only closed one-manifold is S^1. That $\mathcal{PL}_0(T-p)$ is one-dimensional follows from the proof of 9.5.2. Perhaps it is instructive in any case to give a covering of $\mathcal{PL}_0(T-p)$ by train track neighborhoods: 9.34



or, to get open overlaps,

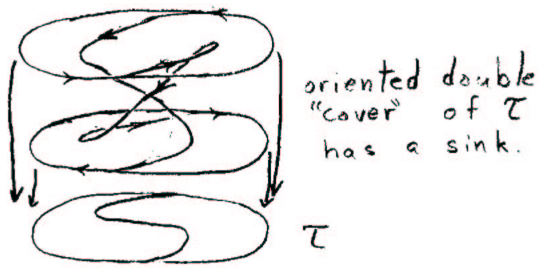


□

PROPOSITION 9.5.4. *On any hyperbolic surface S which is not a punctured torus, an element $\gamma \in \overline{\mathcal{ML}_0}(S)$ is essentially complete if and only if $S-\gamma$ is a union of triangles and punctured monogons.*

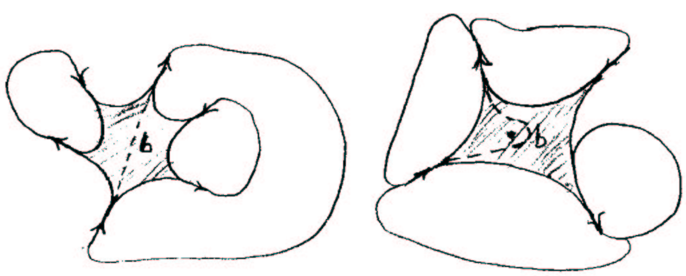
PROOF. Let γ be an arbitrary lamination in $\mathcal{ML}_0(S)$, and let τ be any train track approximation close enough that the regions of $S-\tau$ correspond to those of $S-\gamma$. If some of these regions are not punctured monogons or triangles, we will add extra 9.35
branches in a way compatible with a measure.

First consider the case that each region of $S-\gamma$ is either simply connected or a simple neighborhood of a cusp of S with fundamental group $\mathbb{Z}$. Then τ is connected. Because of the existence of an invariant measure, a train can get from any part of τ to any other. (The set of points accessible by a given oriented train is a "sink," which can only be a connected component.) If τ is not orientable, then every oriented train can get to any position with any orientation. (Otherwise, the oriented double "cover" of τ would have a non-trivial sink.)

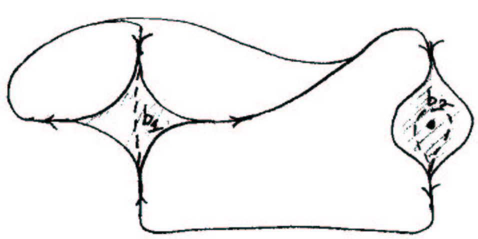


In this case, add an arbitrary branch b to τ, cutting a non-atomic region (of area $> \pi$). Clearly there is some cyclic train path through b, so $\tau \cup b$ admits a positive measure.

If τ is oriented, then each region of $S - \tau$ has an even number of cusps on its boundary. The area of S must be 4π or greater (since the only complete oriented surfaces of finite area having $\chi = -1$ are the thrice punctured sphere, for which $\mathcal{ML}_0$ 9.36
is empty, and the punctured torus). If there is a polygon with more than four sides, it can be subdivided using a branch which preserves orientation, hence admits a cyclic train path. The case of a punctured polygon with more than two sides is similar.

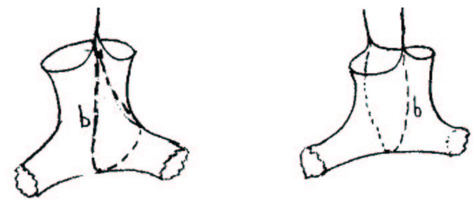


Otherwise, $S - \gamma$ has at least two components. Add one branch b_1 which reverses positively oriented trains, in one region, and another branch b_2 which reverses negatively oriented trains in another.



There is a cyclic train path through b_1 and b_2 in $\tau \cup b_1 \cup b_2$, hence an invariant measure.

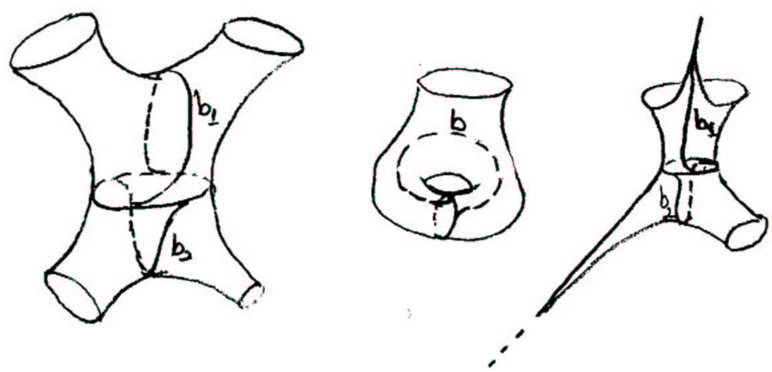
Now consider the case when $S - \tau$ has more complexly connected regions. If 9.37
a boundary component of such a region R has one or more vertices, then a train pointing away from R can return to at least one vertex pinting toward R. If R is not an annulus, hook a new branch around a non-trivial homotopy class of arcs in R with ends on such a pair of vertices.



If R is an annulus and each boundary component has at least one vertex, then add one or two branches running across R which admit a cyclic train path.

If R is not topologically a thrice punctured disk or annulus, we can add an interior closed curve to R.

Any boundary component of R which is a geodesic α has another region R' (which may equal R) on the other side. In this case, we can add one or more branches in R and R' tangent to α in opposite directions on opposite sides, and hooking in ways 9.38
similar to those previously mentioned.



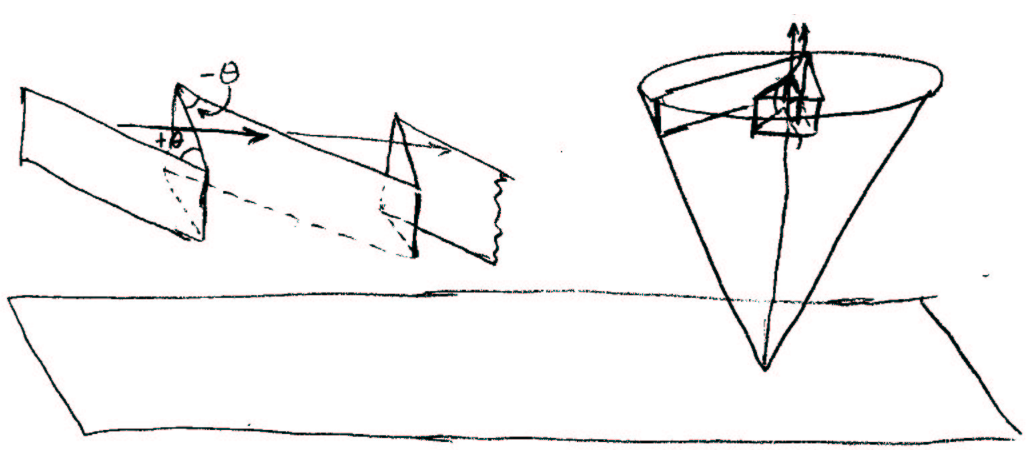
From the existence of these extensions of the original train track, it follows that an element $\gamma \in \mathcal{ML}_0$ is essentially complete if and only if $S - \gamma$ consists of triangles and punctured monogons. Furthermore, every $\gamma \in \overline{\mathcal{ML}_0}$ can be approximated by essentially complete elements $\gamma' \in \mathcal{ML}_0$. In fact, an open dense set has the property that the ϵ-train track approximation τ_ϵ has only triangles and punctured monogons as complementary regions, so generically *every* τ_ϵ has this property. The characterization of essential completeness then holds for $\overline{\mathcal{ML}_0}$ as well. □

Here is some useful geometric information about uncrumpled surfaces.

PROPOSITION 9.5.5. (i) *The sum of the dihedral angles along all edges of the wrinkling locus $w(S)$ tending toward a cusp of an uncrumpled surface S is 0.* (*The sum is taken in the group $S^1 = \mathbb{R} \bmod 2\pi$.*)

(ii) *The sum of the dihedral angles along all edges of $w(S)$ tending toward any side of a closed geodesic γ of $w(S)$ is $\pm\alpha$, where α is the angle of rotation of parallel translation around γ.* (*The sign depends on the sense of the spiralling of nearby geodesics toward γ.*) 9.39

PROOF. Consider the upper half-space model, with either the cusp or the end of $\tilde{\gamma}$ toward which the geodesics in $w(S)$ are spiralling at ∞. Above some level (in case (a)) or inside some cone (in case (b)), S consists of vertical planes bent along vertical lines. The proposition merely says that the total angle of bending in some fundamental domain is the sum of the parts.

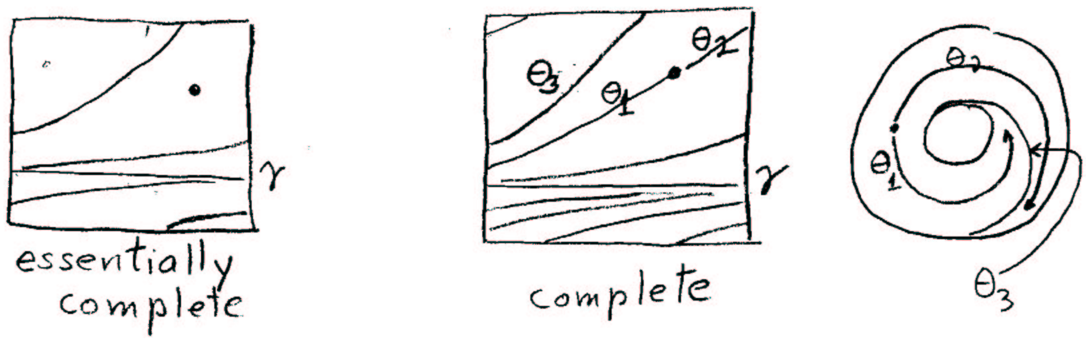


□

COROLLARY 9.5.6. *An uncrumpled surface realizing an essentially complete lamination in $\overline{\mathcal{ML}_0}$ in a given homotopy class is unique. Such an uncrumpled surface is totally geodesic near its cusps.*

PROOF. If the surface S is not a punctured torus, then it has a unique completion obtained by adding a single geodesic tending toward each cusp. By 9.5.5, an uncrumpled surface cannot be bent along any of these added geodesics, so we obtain 9.40
9.5.6.

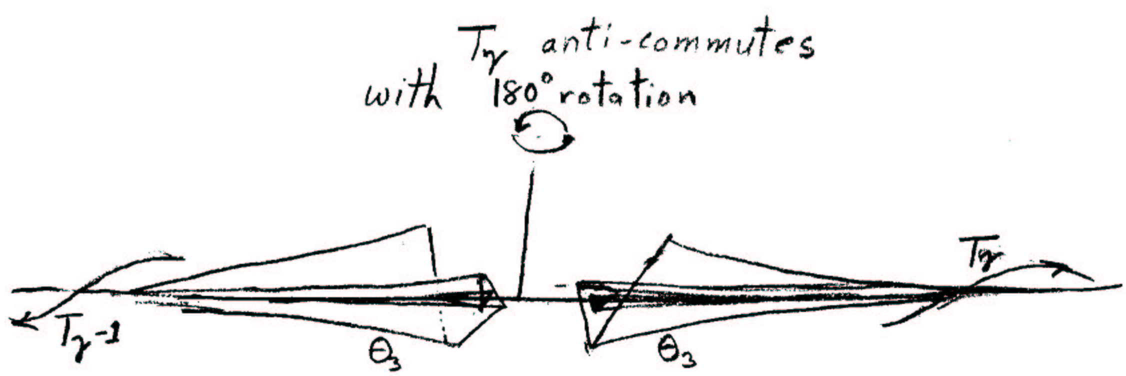
If S is the punctured torus $T - p$, then we consider first the case of a lamination γ which is an essential completion of a single closed geodesic. Complete γ by adding two closed geodesics going from the vertices of the punctured bigon to the puncture.



If the dihedral angles along the infinite geodesics are θ_1, θ_2 and θ_3, as shown, then by 9.5.5 we have

$$\theta_1 + \theta_2 = 0, \quad \theta_1 + \theta_3 = \alpha, \quad \theta_2 + \theta_3 = \alpha,$$

where α is some angle. (The signs are the same for the last two equations because any hyperbolic transformation anti-commutes with a 180° rotation around any perpendicular line.)



9.41

Thus $\theta_1 = \theta_2 = 0$, so an uncrumpled surface is totally geodesic in the punctured bigon. Since simple closed curves are dense in $\mathcal{ML}_0$, every element $g \in \mathcal{ML}_0$ realizable in a given homotopy class has a realization by an uncrumpled surface which is totally geodesic on a punctured bigon. If γ is essentially complete, this means its realizing surface is unique. □

PROPOSITION 9.5.7. *If γ is an essentially complete geodesic lamination, realized by an uncrumpled surface U, then any uncrumpled surface U' realizing a lamination γ' near γ is near U.*

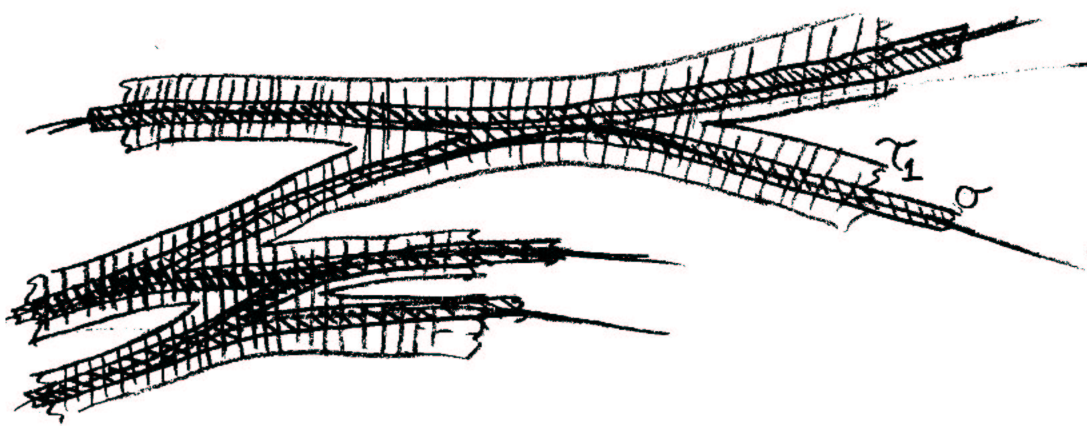
Proof. You can see this from train track approximations. This also follows from the uniqueness of the realization of γ on an uncrumpled surface, since uncrumpled surfaces realizing laminations converging to γ must converge to a surface realizing γ. □

Consider now a typical path $\gamma_t \in \mathcal{ML}_0$. The path γ_t is likely to consist mostly of essentially complete laminations, so that a family of uncrumpled surfaces U_t realizing γ_t would be usually (with respect to t) continuous. At a countable set of values of t, γ_t is likely to be essentially incomplete, perhaps having a single complementary quadrilateral. Then the left and right hand limits U_{t-} and U_{t+} would probably exist, 9.42
and give uncrumpled surfaces realizing the two essential completions of γ_t. In fact, we will show that any path γ_t can be perturbed slightly to give a "generic" path in which the only essentially incomplete laminations are ones with precisely two distinct completions. In order to speak of generic paths, we need more than the topological structure of $\mathcal{ML}_0$.

Proposition 9.5.8. $\mathcal{ML}$ *and* $\mathcal{ML}_0$ *have canonical PL (piecewise linear) structures.*

Proof. We must check that changes of the natural coordinates coming from maximal train tracks (pp. 8.59-8.60) are piecewise linear. We will give the proof for $\mathcal{ML}_0$; the proof for $\mathcal{ML}$ is obtained by appropriate modifications.

Let γ be any measured geodesic lamination in $\mathcal{ML}_0(S)$. Let τ_1 and τ_2 be maximal compactly supported train tracks carrying γ, defining coordinate systems ϕ_1 and ϕ_2 from neighborhoods of γ to convex subsets of R^n (consisting of measures on τ_1 and τ_2). A close enough train track approximation σ of γ is carried by τ_1 and τ_2.



9.43

The set of measures on σ go linearly to measures on τ_1 and τ_2. If σ is a maximal compact train track supporting a measure, we are done—the change of coordinates $\phi_2 \circ \phi_2^{-1}$ is linear near γ. (In particular, note that if γ is essentially complete, change of coordinates is *always* linear at γ). Otherwise, we can find a finite set of enlargements of σ, $\sigma_1, \ldots, \sigma_k$, so that every element of a neighborhood of γ is closely approximated by one of the σ_i. Since every element of a neighborhood of γ is carried by τ_1 and τ_2, it follows that (if the approximations are good enough) each of the σ_i is carried by τ_1 and τ_2. Each σ_i defines a convex polyhedron which is mapped linearly by ϕ_1 and ϕ_2, so $\phi_2 \circ \phi_1^{-1}$ must be PL in a neighborhood of γ. □

REMARK 9.5.9. It is immediate that change of coordinates involves only rational coefficients. In fact, with more care $\mathcal{ML}$ and $\mathcal{ML}_0$ can be given a piecewise integral linear structure. To do this, we can make use of the set $\mathcal{D}$ of integer-valued measures supported on finite collections of simple closed curves (in the case of $\mathcal{ML}_0$); $\mathcal{D}$ is analogous to the integral lattice in $\mathbb{R}^n$. $\mathrm{GL}_n\,\mathbb{Z}$ consists of linear transformations of $\mathbb{R}^n$ which preserve the integral lattice. The set V_τ of measures supported on a given train track τ is the subset of some linear subspace $V \subset \mathbb{R}^n$ which satisfies a finite number of linear inequalities $\mu(b_i) > 0$. Thus V_τ is the convex hull of a finite number of lines, each passing through an integral point. The integral points in U are closed under integral linear combinations (when such a combination is in U), so they determine an 9.44
integral linear structure which is preserved whenever U is mapped linearly to another coordinate system.

Note in particular that the natural transformations of $\mathcal{ML}_0$ are volume-preserving.

The structure on $\mathcal{PL}$ and $\mathcal{PL}_0$ is a piecewise integral *projective* structure. We will use the abbreviations PIL and PIP for piecewise integral linear and piecewise integral projective.

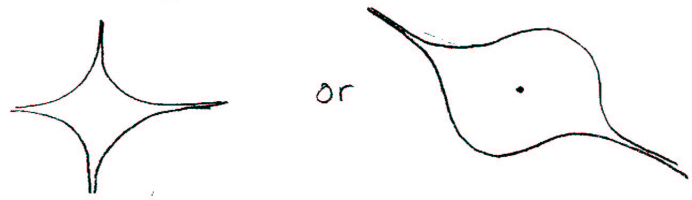
DEFINITION 9.5.10. The *rational depth* of an element $\gamma \in \mathcal{ML}_0$ is the dimension of the space of rational linear functions vanishing on γ, with respect to any natural local coordinate system. From 9.5.8 and 9.5.9, it is clear that the rational depth is independent of coordinates.

PROPOSITION 9.5.11. *If γ has rational depth 0, then γ is essentially complete.*

PROOF. For any $\gamma \in \mathcal{ML}_0$ which is not essentially complete we must construct a rational linear function vanishing on γ. Let τ be some train track approximation of γ which can be enlarged and still admit a positive measure. It is clear that the set of measures on τ spans a proper rational subspace in any natural coordinate system coming from a train track which carries τ. (Note that measures on τ consist of positive linear combinations of integral measures, and that every lamination carried by τ is approximable by one *not* carried by τ.) □ 9.45

PROPOSITION 9.5.12. *If $\gamma \in \mathcal{ML}_0$ has rational depth 1, then either γ is essentially complete or γ has precisely two essential completions. In this case either*

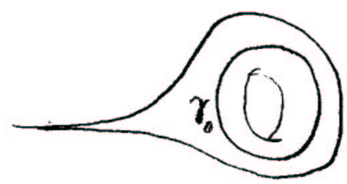
A. *γ has no closed leaves, and all complementary regions have area π or 2π. There is only one region with area 2π unless γ is oriented and* $\text{area}(S) = 4\pi$ *in which case there are two. Such a region is either a quadrilateral or a punctured bigon.*



or

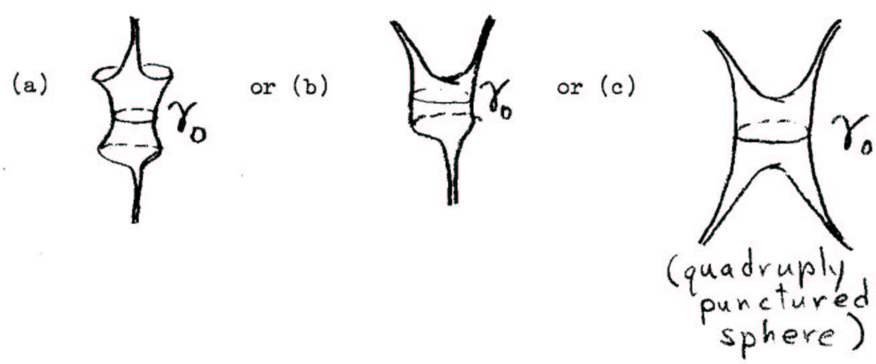
B. *γ has precisely one closed leaf γ_0. Each region touching γ_0 has area 2π. Either*

1. *S is a punctured torus*

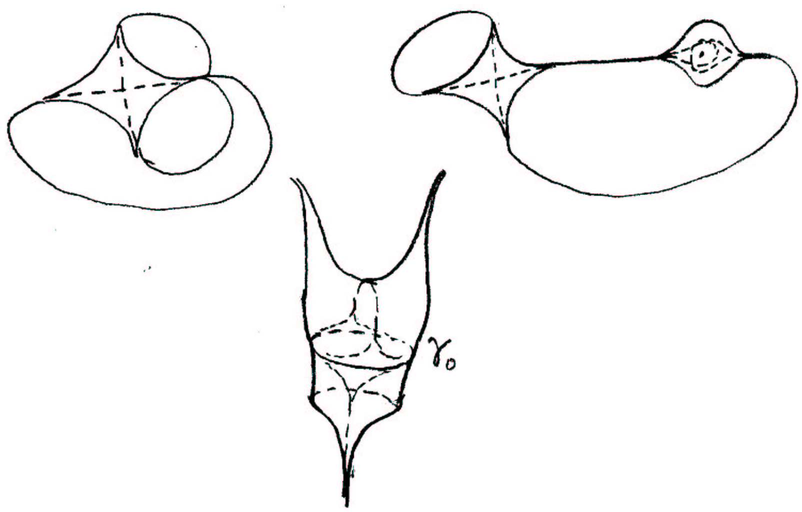


or

2. *γ_0 touches two regions, each a one-pointed crown or a devils cap.*



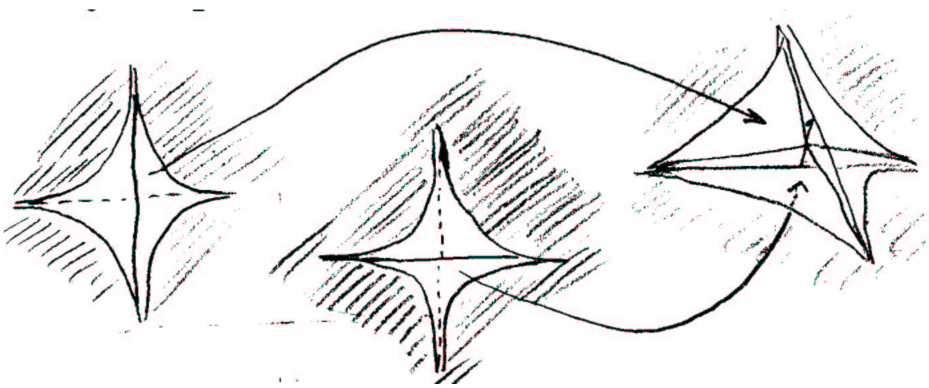
Proof. Suppose γ has rational depth 1 and is not essentially complete. Let τ be a close train track approximation of γ. There is some finite set $\tau_1, \dots, \tau_k$ of essentially complete enlargements of τ which closely approximate every γ' in a neighborhood of γ. Let σ carry all the τ_i's and let V_σ be its coordinate system. The set of γ corresponding to measures carried by a given proper subtrack of a τ_i is a proper rational subspace of V_σ. Since γ is in a unique proper rational subspace, V_τ, the set of measures V_{τ_i} carried on any τ_i must consist of one side of V_τ. (If V_{τ_i} intersected both sides, by convexity γ would come from a measure positive on all branches of τ_i). Since this works for any degree of approximation of nearby laminations, γ has precisely two essential completions. A review of the proof of 9.5.4 gives the list of possibilities for $\gamma \in \mathcal{ML}_0$ with precisely two essential completions. The ambiguity in the essential completions comes from the manner of dividing a quadrilateral or other region, and the direction of spiralling around a geodesic.



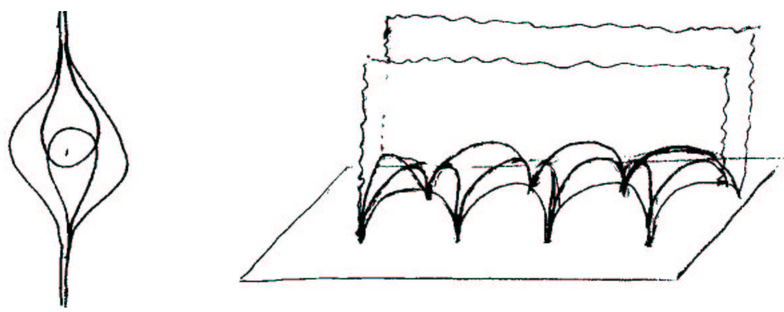
$\square$ 9.47

REMARK. There are good examples of $\gamma \in \mathcal{ML}_0$ which have large rational depth but are essentially complete. The construction will occur naturally in another context.

We return to the construction of continuous families of surfaces in a hyperbolic three-manifold. To each essentially incomplete $\gamma \in \mathcal{ML}_0$ of rational depth 1, we associate a one-parameter family of surfaces U_s, with U_0 and U_1 being the two uncrumpled surfaces realizing γ. U_s is constant where U_0 and U_1 agree, including the union of all triangles and punctured monogons in the complement of γ. The two images of any quadrilateral in $S - \gamma$ form an ideal tetrahedron. Draw the common perpendicular p to the two edges not in $U_0 \cap U_1$, triangulate the quadrilateral with 4 triangles by adding a vertex in the middle, and let this vertex run linearly along p, from U_0 to U_1. This extends to a homotopy of S straight on the triangles.

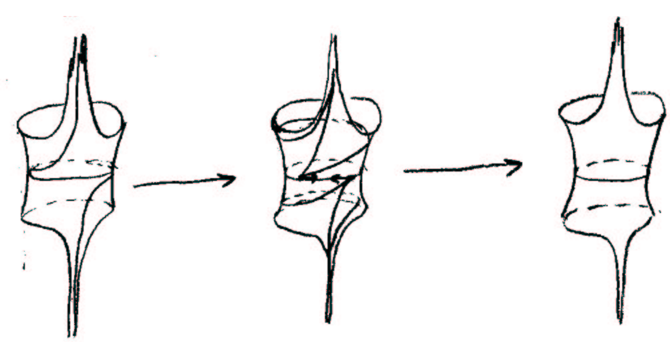


The two images of any punctured bigon in $S-\gamma$ form a solid torus, with the generating curve parabolic. The union of the two essential completions in this punctured bigon gives a triangulation except in a neighborhood of the puncture, with two new 9.48
vertices at intersection points of added leaves.



Draw the common perpendiculars to edges of the realizations corresponding to these intersection points, and homotope U_0 to U_1 by moving the added vertices linearly along the common perpendiculars.

When γ has a closed leaf γ_0, the two essential completions of γ have added leaves spiralling around γ_0 in opposite directions. U_0 can be homotoped to U_1 through surfaces with added vertices on γ_0.



9.49

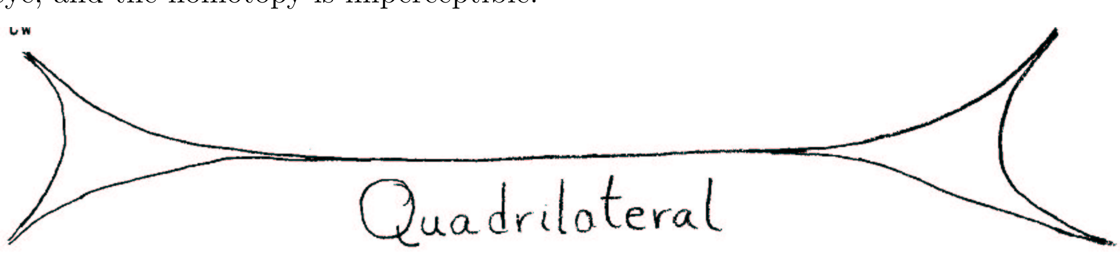
Note that all the surfaces U_s constructed above have the property that any point on U_s is in the convex hull of a small circle about it on U_s. In particular, it has curvature ≤ -1; curvature -1 everywhere except singular vertices, where negative curvature is concentrated.

Theorem 9.5.13. *Given any complete hyperbolic three-manifold N with geometrically tame end E cut off by a hyperbolic surface $S_{[\epsilon,\infty)}$, there is a proper homotopy $F : S_{[\epsilon,\infty)} \times [0,\infty) \to N$ of S to ∞ in E.*

Proof. Let V_τ be the natural coordinate system for a neighborhood of $\epsilon(E)$ in $\mathcal{ML}_0(S)$, and choose a sequence $\gamma_i \in V_\tau$ limiting on $\epsilon(E)$. Perturb the γ_i slightly so that the path γ_t $[0 \leq t \leq \infty]$ which is linear on each segment $t \in [i,\, i+1]$ consists of elements of rational depth 0 or 1. Let U_t be the unique uncrumpled surface realizing γ_t when γ_t is essentially complete. When t is not essentially complete, the left and right hand limits U_{t+} and U_{t-} exist. It should now be clear that F exists, since one can cover the closed set $\{U_{t\pm}\}$ by a locally finite cover consisting of surfaces homotopic by small homotopies, and fill in larger gaps between U_{t+} and U_{t-} by the homotopies constructed above. Since all interpolated surfaces have curvature ≤ -1, and they all realize a γ_t, they must move out to ∞. An explicit homotopy can actually be defined, using a new parameter r which is obtained by "blowing up" all parameter values of t with rational depth 1 into small intervals. Explicitly, these parameter values can be enumerated in some order $\{t_j\}$, and an interval of length 2^{-j} inserted 9.50
in the r-parameter in place of t_j. Thus, a parameter value t corresponds to the point or interval

$$r(t) = \left[t + \sum_{\{j \mid t_j < t\}} 2^{-} j,\ t + \sum_{\{j \mid t_j \leq t\}} 2^{-j} \right].$$

Now insert homotopies as constructed above in *each* blown up interval. It is not so obvious that the family of surfaces is still continuous when an infinite family of homotopies is inserted. Usually, however, these homotopies move a very small distance—for instance, γ_t may have a quadrilateral in $S - \gamma_t$, but for all but a locally small number of t's, this quadrilateral looks like two asymptotic triangles to the naked eye, and the homotopy is imperceptible.



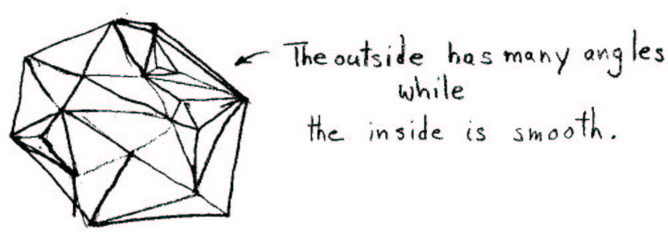
Formally, the proof of continuity is a straightforward generalization of the proof of 9.5.7. The remark which is needed is that if S is a surface of curvature ≤ -1 with a (pathwise) isometric map to a hyperbolic surface homotopic to a homeomorphism, then S is actually hyperbolic and the map is isometric—indeed, the area of S is not greater than the area of the hyperbolic surface. □

REMARKS. 1. There is actually a canonical line of hyperbolic structures on S joining those of U_{t+} and U_{t-}, but it is not so obvious how to map them into E nicely. 9.51

2. An alternative approach to the construction of F is to make use of a sequence of triangulations of S. Any two triangulations with the same number of vertices can be joined by a sequence of elementary moves, as shown.

Although such an approach involves more familiar methods, the author brutally chose to develop extra structure.

3. There should be a good analytic method of constructing F by using harmonic mappings of hyperbolic surfaces. Realizations of geodesic laminations of a surface are analogous to harmonic mappings coming from points at ∞ in Teichmüller space. The harmonic mappings corresponding to a family of hyperbolic structures on S moving along a Teichmüller geodesic to $\epsilon(E)$ ought to move nicely out to ∞ in E. A rigorous proof might involve good estimates of the energy of a map, analogous to §9.3.

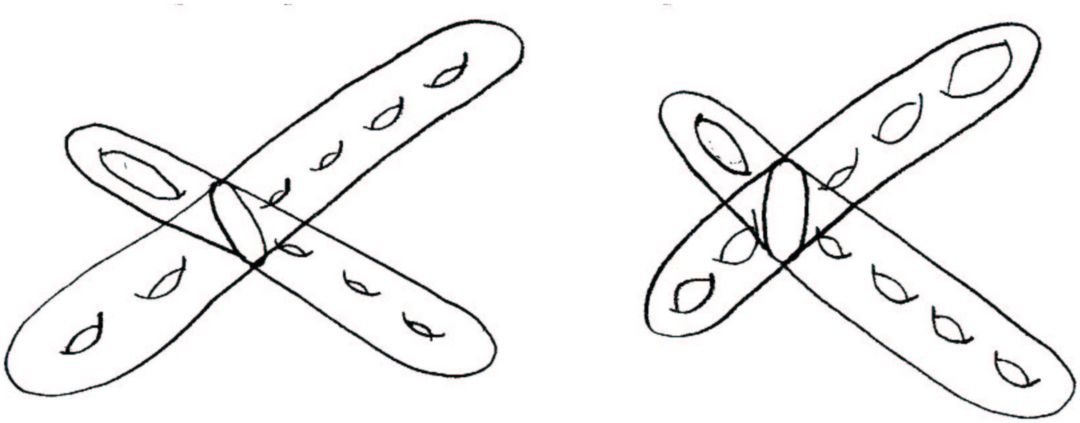


9.52

9.6. Strong convergence from algebraic convergence

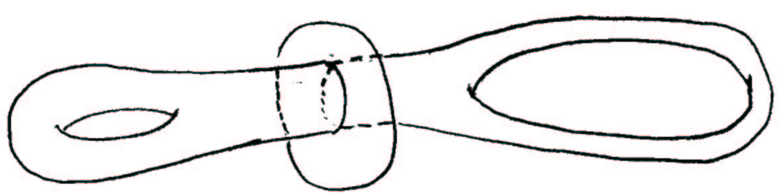
We will take another step in our study of algebraic limits. Consider the space of discrete faithful representations ρ of a fixed torsion free group Γ in $\mathrm{PSL}_2(\mathbb{C})$. The set $\Pi_\rho \subset \Gamma$ of parabolics—i.e., elements $\gamma \in \Gamma$ such that $\rho(\gamma)$ is parabolic—is an important part of the picture; we shall assume that $\Pi_\rho = \Pi$ is constant. When a sequence ρ_i converges algebraically to a representation ρ where $\Pi = \Pi_{\rho_i}$ is constant by $\Pi_\rho \supset \Pi$ is strictly bigger, then elements $\gamma \in \Pi_\rho - \Pi$ are called *accidental parabolics*. The incidence of accidental parabolics can create many interesting phenomena, which we will study later.

One complication is that the quotient manifolds $N_{\rho_i\Gamma}$ need not be homeomorphic; and even when they are, the homotopy equivalence given by the isomorphism of fundamental groups need not be homotopic to a homeomorphism. For instance, consider three-manifolds obtained by gluing several surfaces with boundary, of varying genus, in a neighborhood of their boundary. If every component has negative Euler characteristic, the result can easily be given a complete hyperbolic structure. The homotopy type depends only on the identifications of the boundary components of the original surfaces, but the homeomorphism type depends on the order of arrangement around each image boundary curve.



9.53

As another example, consider a thickened surface of genus 2 union a torus as shown.



It is also easy to give this a complete hyperbolic structure. The fundamental group has a presentation

$$\langle a_1, b_1, a_2, b_2, c : [a_1, b_1] = [a_2, b_2], \big[[a_1, b_1] = c, \big] = 1\rangle.$$

This group has an automorphism

$$a_1 \mapsto a_1,\ b_1 \mapsto b_1,\ c \mapsto c,\ a_2 \mapsto c a_2 c^{-1},\ b_2 \mapsto c b_2 c^{-1}$$

which wraps the surface of genus two around the torus. No non-trivial power of this automorphism is homotopic to a homeomorphism. From an algebraic standpoint there are infinitely many distinct candidates for the peripheral subgroups.

One more potential complication is that even when a given homotopy equivalence is homotopic to a homeomorphism, and even when the parabolic elements correspond,

there might not be a homeomorphism which preserves cusps. This is easy to picture for a closed surface group Γ: when Π is the set of conjugates of powers of a collection of simple closed curves on the surface, there is not enough information in Π to say which curves must correspond to cusps on which side of S. Another example is when Γ is a free group, and Π corresponds to a collection of simple closed curves on the boundary of a handlebody with fundamental group Γ. The homotopy class of a 9.54
simple closed curve is a very weak invariant here.

Rather than entangle ourselves in cusps and handlebodies, we shall confine ourselves to the case of real interest, when the quotient spaces admit cusp-preserving homeomorphisms.

We shall consider, then, geometrically tame hyperbolic manifolds which have a common model, (N_0, P_0). N_0 should be a compact manifold with boundary, and P_0 (to be interpreted as the "parabolic locus") should be a disjoint union of regular neighborhoods of tori and annuli on ∂N_0, with fundamental groups injecting into $\pi_1 N_0$. Each component of $\partial N_0 - P_0$ should be incompressible.

THEOREM 9.6.1. *Let* (N_0, P_0) *be as above. Suppose that* $\rho_i : \pi_1 N \to \mathrm{PSL}(2, \mathbb{C})$ *is a sequence of discrete, faithful representations whose quotient manifolds* N_i *are geometrically tame and admit homeomorphisms* (*in the correct homotopy class*) *to* N_0 *taking horoball neighborhoods of cusps to* P_0. *If* $\{\rho_i\}$ *converges algebraically to a representation* ρ, *then the limit manifold* N *is geometrically tame, and admits a homeomorphism* (*in the correct homotopy class*) *to* N_0 *which takes horoball neighborhoods of cusps to* P_0.

We shall prove this first with an additional hypothesis:

9.6.1a. Suppose also that no non-trivial non-peripheral simple curve of a component of $\partial N_0 - P_0$ *is homotopic* (*in* N_0) *to* P_0. 9.55

REMARKS. The proof of 9.6.1 (without the added hypothesis) will be given in §9.8. There is no §9.8.

The main case is really when all N_i are geometrically finite. One of the main corollaries, from 8.12.4, is that $\rho(\pi_1 N_0)$ satisfies the property of Ahlfors: its limit set has measure 0 or measure 1.

PROOF OF 9.6.1a. It will suffice to prove that every sequence $\{\rho_i\}$ converging algebraically to ρ has a subsequence converging strongly to ρ. Thus, we will pass to subsequences whenever it is convenient.

Let $S_1, \ldots, S_k$ be the components of $\partial N_0 - P_0$, each equipped with a complete hyperbolic metric of finite area. (In other words, their boundary components are made into punctures.) For each i, let P_i denote a union of horoball neighborhoods of cusps of N_i, and let $E_{i,1}, \ldots, E_{i,k}$ denote the ends of $N_i - P_i$ corresponding to $S_1, \ldots, S_k$.

Some of the $E_{i,j}$ may be geometrically finite, others geometrically infinite. We can pass (for peace of mind) to a subsequence so that for each i, the $E_{i,j}$ are all geometrically finite or all geometrically infinite. We pass to a further subsequence so the sequences of bending or ending laminations $\{\beta_{i,j}\}_i$ or $\{\epsilon_{i,j}\}_i$ converge in $\mathcal{GL}_{(S_j)}$. Let χ_j be the limit.

If χ_j is realizable in N, then all nearby laminations have realizations for all representations near ρ, and the $E_{i,j}$ must have been geometrically finite. An uncrumpled surface U realizing χ_j is in the convex hull M of N and approximable by boundary 9.56
components of the convex hulls M_i Since the limit set cannot suddenly increase in the algebraic limit (p. 9.8), U must be a boundary component.

If χ_j is not realizable in N, then it must be the ending lamination for some geometrically infinite tame end E of the covering space of N corresponding to $\pi_1 S_j$, since we have hypothesized away the possibility that it represents a cusp. In view of 9.2.2 and 9.4.2, the image E_j of E in $N - P$ is a geometrically tame end of $N - P$, and $\pi_1 E = \pi_1 S_j$ has finite index in $\pi_1 E_j$.

In either case, we obtain embeddings in $N - P$ of oriented surfaces S'_j finitely covered by $S_{j[\epsilon,\infty)}$. We may assume (after an isotopy) that these embeddings are disjoint, and each surface cuts off (at least) one piece of $N-P$ which is homeomorphic to the product $S'_j \times [0,\infty)$. Since (N,P) is homotopy equivalent to (N_0, P_0), the image of the cycle $\sum \big[S_{j[\epsilon,\infty)}, \partial S_{j[\epsilon,\infty)}\big]$ in (N,P) bounds a chain C with compact support. Except in a special case to be treated later, the S'_j are pairwise non-homotopic and the fundamental group of each S'_j maps isomorphically to a unique side in $N - P$. C has degree 0 "outside" each S'_j and degree some constant l elsewhere. Let N' be the region of $N - P$ where C has degree l. We see that N is geometrically tame, and homotopy equivalent to N'.

The Euler characteristic is a homotopy invariant, so $\chi(N) = \chi(N') = \chi(N_0)$. This imples $\chi(\partial N') = \chi(\partial N_0)$ (by the formula $\chi(\partial M^3) = 2\chi(M^3)$) so in fact the 9.57
finite sheeted covering $S_{j[\epsilon,\infty)} \to S'_j$ has only one sheet—it is a homeomorphism.

Let Q be the geometric limit of any subsequence of the N_i. N is a covering space of Q. Every boundary component of the convex hull M of N is the geometric limit of boundary components of the M_i; consequently, M covers the convex hull of Q. This covering can have only finitely many sheets, since $M - P$ is made of a compact part together with geometrically infinite tame ends. Any element $\alpha \in \pi_1 Q$ has some finite power $\alpha^k \in \pi_1 N$ $[k \geq 1]$. In any torsion-free subgroup of $\mathrm{PSL}(2,\mathbb{C})$, an element has at most one k-th root (by consideration of axes). If we write α as the limit of elements $\rho_i(g_i)$, $g_i \in \pi_1 N_0$, by this remark, g_i must be eventually constant so α is actually in the algebraic limit $\pi_1 N$. $Q = N$, and ρ_i converges strongly to ρ.

A cusp-preserving homeomorphism from N to some N_i, hence to N_0, can be constructed by using an approximate isometry of N' with a submanifold of $N_i - P_i$, for high enough i. The image of N' is homotopy equivalent to N_i, so the fundamental group of each boundary component of N' must map surjectively, as well as injectively, to the fundamental group of the neighboring component of $(N_i, P_i) - N'$. This implies that the map of N' into N_i extends to a homeomorphism from N to N_i.

There is a special case remaining. If any pair of the surfaces S'_i constructed in $N - P$ is homotopic, perform all such homotopies. Unless $N - P$ is homotopy equivalent to a product, the argument continues as before—there is no reason the cover of S'_i must be a connected component of $\partial N_0 - P_0$.

When $N - P$ is homotopy equivalent to the oriented surface S'_1 in it, then by a standard argument $N_0 - P_0$ must be homeomorphic to $S'_1 \times I$. This is the case essentially dealt with in 9.2. The difficulty is to control both ends of $N - P$—but the argument of 9.2 shows that the ending or bending laminations of the two ends of $N_i - P_i$ cannot converge to the same lamination, otherwise the limit of some intermediate surface would realize χ_i. This concludes the proof of 9.6.1a. □

9.7. Realizations of geodesic laminations for surface groups with extra cusps, with a digression on stereographic coordinates

In order to analyze geometric convergence, and algebraic convergence in more general cases, we need to clarify our understanding of realizations of geodesic laminations for a discrete faithful representation ρ of a surface group $\pi_1(S)$ when certain non-peripheral elements of $\pi_1(S)$ are parabolic. Let $N = N_{\rho\pi_1 S}$ be the quotient three-manifold. Equip S with a complete hyperbolic structure with finite area. As in §8.11, we may embed S in N, cutting it in two pieces the "top" N_+ and the "bottom" N_-. Let γ_+ and γ_- be the (possibly empty) cusp loci for N_+ and N_-, and denote by $S_{1+}, \dots, S_{j+}$ and $S_{1-}, \dots, S_{k-}$ the components of $S - \gamma_+$ and $S - \gamma_-$ (endowed with complete hyperbolic structures with finite area). Let $E_{1+}, \dots, E_{j+}$ and $E_{1-}, \dots, E_{k-}$ denote the ends of $N - P$, where P is the union of horoball neighborhoods of all cusps.

A compactly supported lamination on S_{i+} or S_{i-} defines a lamination on S. In particular, $\epsilon(E_{i\pm})$ may be thought of as a lamination on S for each geometrically infinite tame end of $E_{i\pm}$.

Proposition 9.7.1. *A lamination $\gamma \in \mathcal{GL}_0(S)$ is realizable in N if and only if γ contains no component of γ_+, no component of γ_-, and no $\epsilon(E_{i+})$ or $\epsilon(E_{i-})$.*

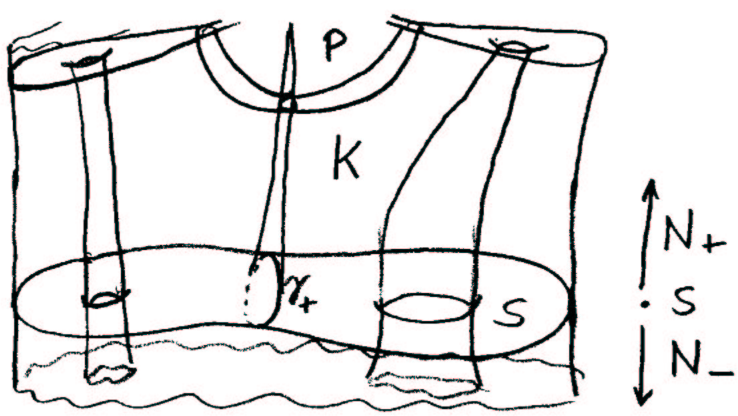
Proof. If γ contains any unrealizable lamination, it is unrealizable, so the necessity of the condition is immediate.

Let $\gamma \in \mathcal{ML}_0(S)$ be any unrealizable compactly supported measured lamination. If γ is not connected, at least one of its components is unrealizable, so we need only consider the case that γ is connected. If γ has zero intersection number with any components of γ_+ or γ_-, we may cut S along this component, obtaining a simpler surface S'. Unless γ is the component of γ_+ or γ_- in question, S' supports γ, so we pass to the covering space of N corresponding to $\pi_1 S'$. The new boundary components of S' are parabolic, so we have made an inductive reduction of this case.

We may now suppose that γ has positive intersection number with each component of γ_+ and γ_-. Let $\{\beta_i\}$ be a sequence of measures, supported on simple closed curves non-parabolic in N which converges to γ. Let $\{U_i\}$ be a sequence of uncrumpled surfaces realizing the β_i. If U_i penetrates far into a component of P 9.60
corresponding to an element α in γ_+ or γ_-, then it has a large ball mapped into P; by area considerations, this ball on U_i must have a short closed loop, which can only be in the homotopy class of α. Then the ratio

$$l_S(\beta_i)/i(\beta_i, \alpha) \geq l_{U_i}(\beta_i)/i(\beta_i, \alpha)$$

is large. Therefore (since $i(\gamma, \alpha)$ is positive and $l_S(\gamma)$ is finite) the U_i, away from their cusps, remain in a bounded neighborhood of $N - P$ in N. If γ_+ (say) is non-empty, one can now find a compact subset K of N so that any U_i intersecting N_+ must intersect K.



By the proof of 8.8.5, if infinitely many U_i intersected K, there would be a convergent subsequence, contradicting the non-realizability of γ. The only remaining possibility is that we have reached, by induction, the case that either N_+ or N_- has no extra cusps, and γ is an ending lamination. 9.61

A general lamination $\gamma \in \mathcal{GL}(S)$ is obtained from a possibly empty lamination which admits a compactly supported measure by the addition of finitely many non-compact leaves. (Let $\delta \subset \gamma$ be the maximal lamination supporting a positive transverse measure. If l is any leaf in $\gamma - \delta$, each end must come close to δ or go to ∞

in S, otherwise one could enlarge δ. By area considerations, such leaves are finite in number.) From §8.10, γ is realizable if and only if δ is. □

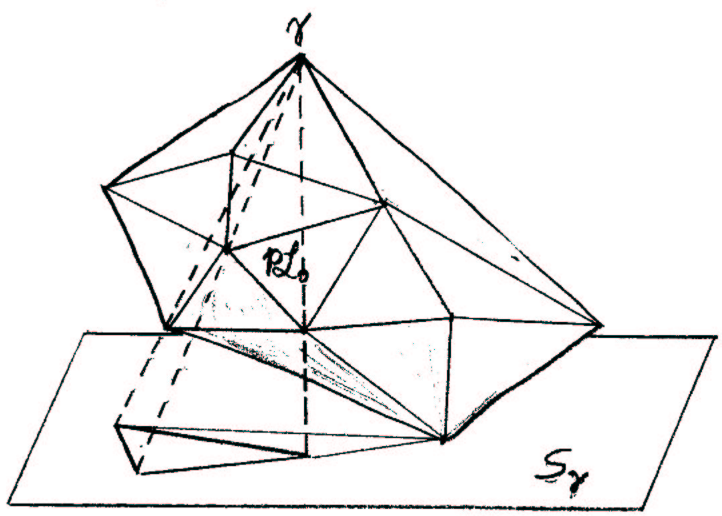
The picture of unrealizable laminations in $\mathcal{PL}_0(S)$ is the following. Let Δ_+ consist of all projective classes of transverse measures (allowing degenerate non-trivial cases) on $\chi_+ = \gamma_+ \cup U_i \epsilon(E_{i+})$. Δ_+ is convex in a coordinate system V_τ coming from any train track τ carrying χ_+.

To see a larger, complete picture, we must find a larger natural coordinate system. This requires a little stretching of our train tracks and imaginations. In fact, it is possible to find coordinate systems which are quite large. For any $\gamma \in \mathcal{PL}_0$, let $\Delta_\gamma \subset \mathcal{PL}_0$ denote the set of projective classes of measures on γ.

Proposition 9.7.2. *Let γ be essentially complete. There is a sequence of train tracks τ_i, where τ_i is carried by τ_{i+1}, such that the union of natural coordinate systems $S_\gamma = U_i V_{\tau_i}$ contains all of $\mathcal{PL}_0 - \Delta_\gamma$.*

The proof will be given presently.

Since τ_i is carried by τ_{i+1}, the inclusion $V_{\tau_i} \subset V_{\tau+1}$ is a projective map (in $\mathcal{ML}_0$, 9.62
the inclusion is linear). Thus S_γ comes naturally equipped with a projective structure. We have not made this analysis, but the typical case is that $\gamma = \Delta_\gamma$. We think of S_γ as a stereographic coordinate system, based on projection from γ. (You may imagine $\mathcal{PL}_0$ as a convex polyhedron in $\mathbb{R}^n$, so that changes of stereographic coordinates are piecewise projective, although this finite-dimensional picture cannot be strictly correct, since there is no fixed subdivision sufficient to make all coordinate changes.)



Corollary 9.7.3. *$\mathcal{PL}_0(S)$ is homeomorphic to a sphere.*

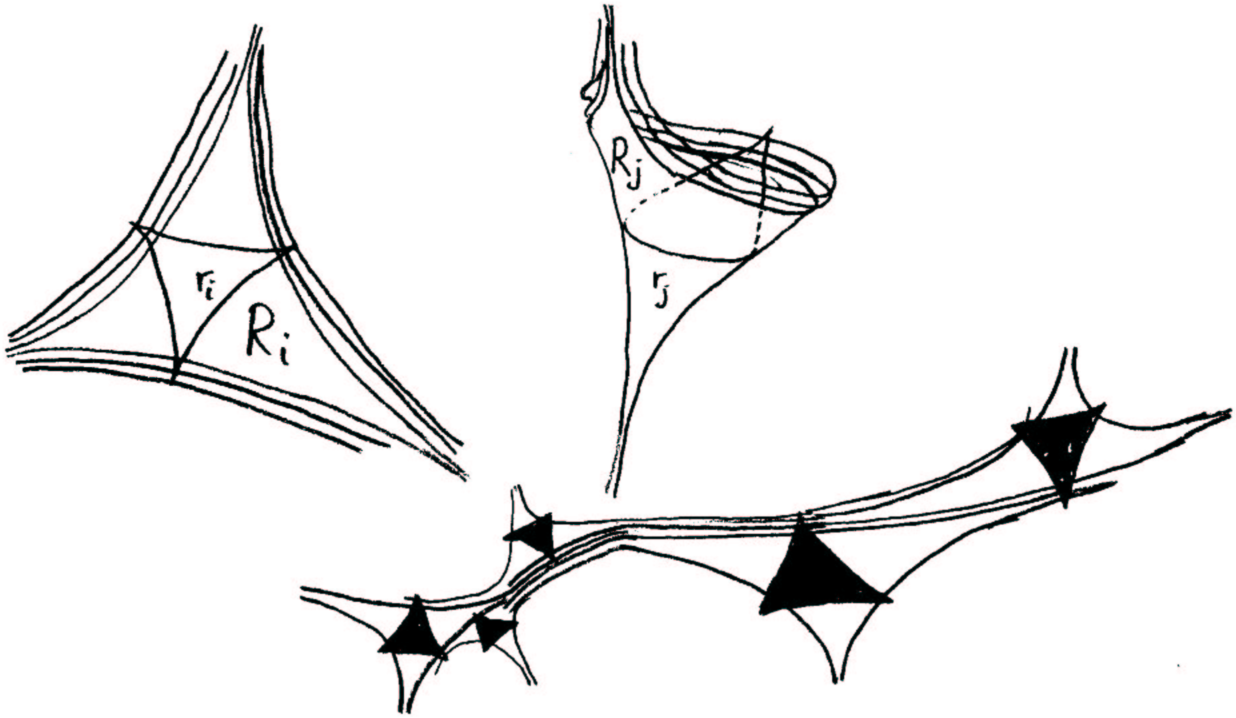
PROOF THAT 9.7.2 IMPLIES 9.7.3. Let $\gamma \in \mathcal{PL}_0(S)$ be any essentially complete lamination. Let τ be any train track carrying γ. Then $\mathcal{PL}_0(S)$ is the union of two coordinate systems $V_\tau \cup S_\tau$, which are mapped to convex sets in Euclidean space. 9.63
If $\Delta_\gamma \neq \gamma$, nonetheless the complement of Δ_γ in V_τ is homeomorphic to $V_\tau - \gamma$, so $\mathcal{PL}_0(S)$ is homeomorphic to the one-point compactification of S_γ. □

COROLLARY 9.7.4. *When $\mathcal{PL}_0(S)$ has dimension greater than* 1*, it does not have a projective structure.* (In other words, the *pieces* in changes of coordinates have not been eliminated.)

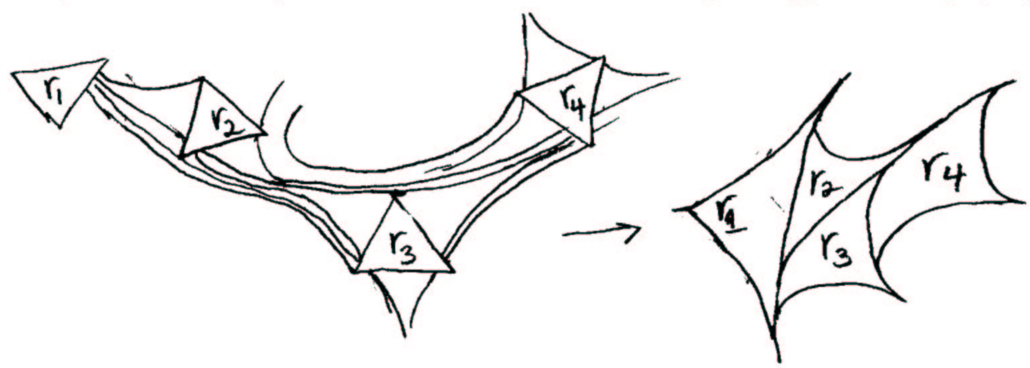
PROOF THAT 9.7.3 IMPLIES 9.7.4. The only projective structure on S^n, when $n > 1$, is the standard one, since S^n is simply connected. The binary relation of antipodality is natural in this structure. What would be the antipodal lamination for a simple closed curve α? It is easy to construct a diffeomorphism fixing α but moving any other given lamination. (If $i(\gamma, \alpha) \neq 0$, the Dehn twist around α will do.) □

REMARK. When $\mathcal{PL}_0(S)$ is one-dimensional (that is, when S is the punctured torus or the quadruply punctured sphere), the PIP structure *does* come from a projective structure, equivalent to $\mathbb{R}P^1$. The natural transformations of $\mathcal{PL}_0(S)$ are necessarily integral—in $\mathrm{PSL}_2(\mathbb{Z})$.

PROOF OF 9.7.2. Don't blink. Let γ be essentially complete. For each region R_i of $S - \gamma$, consider a smaller region r_i of the same shape but with finite points, rotated so its points alternate with cusps of R_i and pierce very slightly through the sides of R_i, ending on a leaf of γ. 9.64



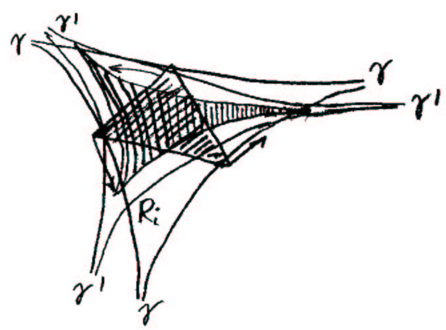
By 9.5.4, 9.5.2 and 9.3.9, both ends of each leaf of γ are dense in γ, so the regions r_i separate leaves of γ into arcs. Each region of $S - \gamma - U_i r_i$ must be a rectangle with two edges on ∂r_i and two on γ, since r_i covers the "interesting" part of R_i. (Or, prove this by area, χ). Collapse all rectangles, identifying the r_i edges with each other, and obtain a surface S' homotopy-equivalent to S, made of $U_i r_i$, where ∂r_i projects to a train track τ. (Equivalently, one may think of $S - U_i r_i$ as made of very wide corridors, with the horizontal direction given approximately by γ).



9.65

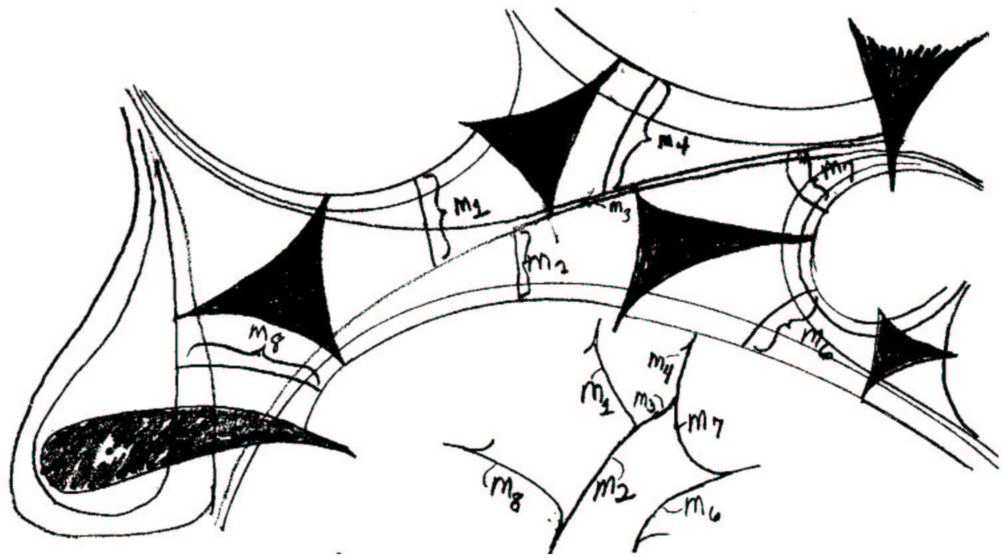
If we take shrinking sequences of regions $r_{i,j}$ in this manner, we obtain a sequence of train tracks τ_j which obviously have the property that τ_j carries τ_k when $j > k$. Let $\gamma' \in \mathcal{PL}_0(S) - \Delta_\gamma$ be any lamination not topologically equivalent to γ. From the density in γ of ends of leaves of γ, it follows that whenever leaves of γ and γ' cross, they cross at an angle. There is a lower bound to this angle. It also follows that $\gamma \cup \gamma'$ cuts S into pieces which are compact except for cusps of S.

When R_i is an asymptotic triangle, for instance, it contains exactly one region of $S - \gamma - \gamma'$ which is a hexagon, and all other regions of $S - \gamma - \gamma'$ are rectangles. For sufficiently high j, the r_{ij} can be isotoped, without changing the leaves of γ which they touch, into the complement of γ'. It follows that γ' projects nicely to τ_j.



$\square$

Stereographic coordinates give a method of computing and understanding intersection number. The transverse measure for γ projects to a "tangential" measure ν_γ on each of the train tracks τ_i: i.e., $\nu_\gamma(b)$ is the γ-transverse length of the sides of the rectangle projecting to b.



It is clear that for any $\alpha \in \mathcal{ML}_0$ which is determined by a measure μ_α on τ_i

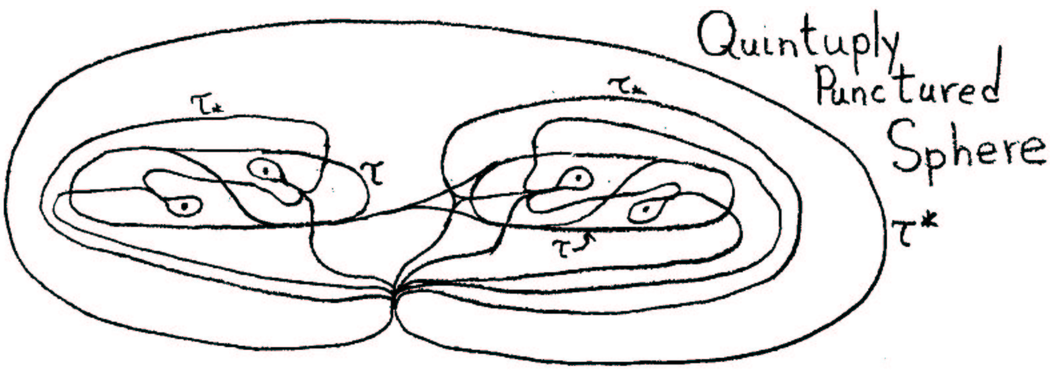
9.7.5. $$i(\alpha,\gamma) = \sum_b \mu_\alpha(b) \cdot \nu_\gamma(b).$$

Thus, in the coordinate system V_{τ_i} in $\mathcal{ML}_0$, intersection with γ is a linear function.

To make this observation more useful, we can reverse the process of finding a family of "transverse" train tracks τ_i depending on a lamination γ. Suppose we are given 9.67
an essentially complete train track τ, and a non-negative function (or "tangential" measure) ν on the branches of b, subject only to the triangle inequalities

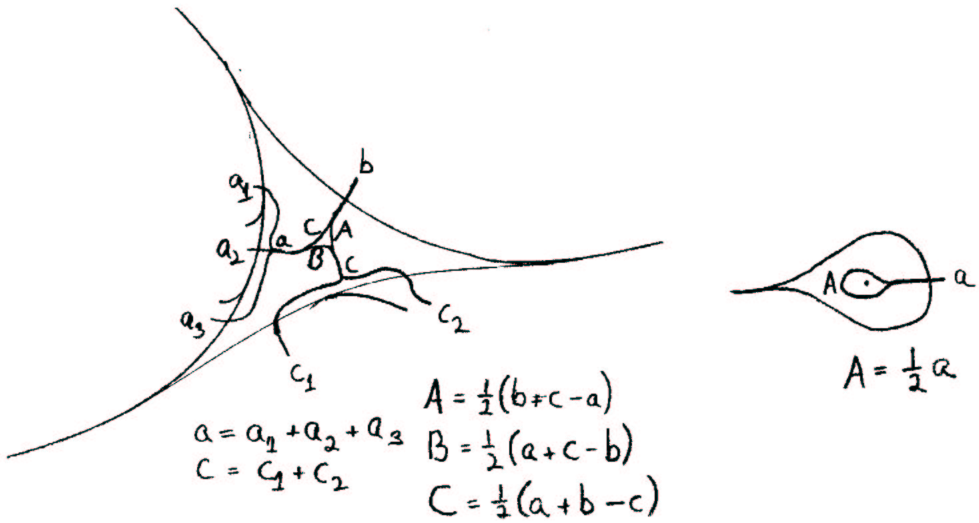
$$a + b - c \geq 0 \quad a + c - b \geq 0 \quad b + c - a \geq 0$$

whenever a, b and c are the total ν-lengths of the sides of any triangle in $S - \tau$. We shall construct a "train track" τ^* dual to τ, where we permit regions of $S - \tau^*$ to be bigons as well as ordinary types of admissible regions—let us call τ^* a *bigon track*.

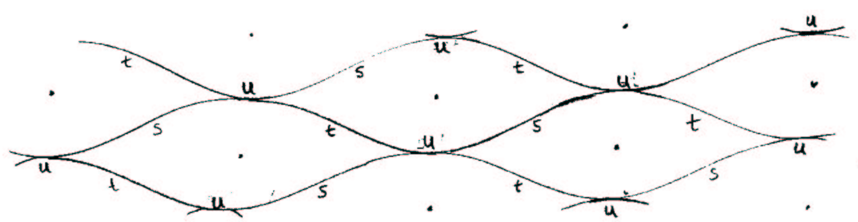


τ^* is constructed by shrinking each region R_i of $S-\tau$ and rotating to obtain a region $R_i^* \subset R_i$ whose points alternate with points of R_i. These points are joined using one more branch b^* crossing each branch b of τ; branches b_1^* and b_2^* are confluent at a vertex of R^* whenever b_1 and b_2 lie on the same side of R. Note that there is a bigon in $S-\tau^*$ for each switch in τ.

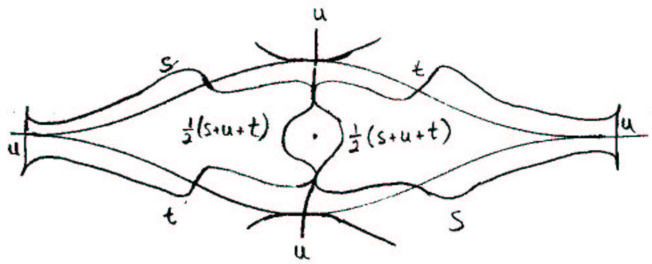
The tangential measure ν for τ determines a transverse measure defined on the branches of τ^* of the form b^*. This extends uniquely to a transverse for τ^* when S is not a punctured torus. 9.68



When S is the punctured torus, then τ must look like this, up to the homeomorphism (drawn on the abelian cover of $T-p$):



Note that each side of the punctured bigon is incident to each branch of τ. Therefore, the tangential measure ν has an extension to a transverse measure ν^* for τ^*, which is unique if we impose the condition that the two sides of R^* have equal transverse measure.

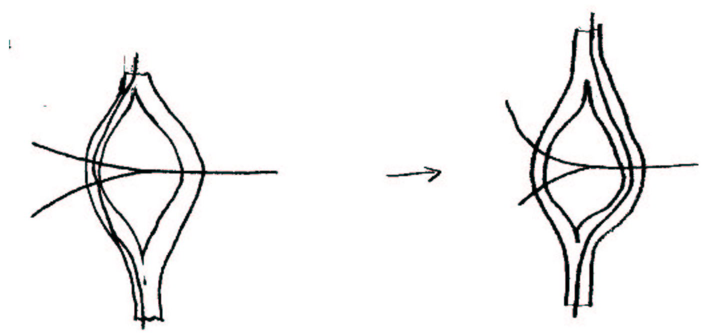


9.69

A transverse measure on a bigon track determines a measured geodesic lamination, by the reasoning of 8.9.4. When τ is an essentially complete train track, an open subset of $\mathcal{ML}_0$ is determined by a function μ on the branches of τ subject to a condition for each switch that

$$\sum_{b\in\mathcal{I}} \mu(b) = \sum_{b\in\mathcal{O}} \mu(b),$$

where $\mathcal{I}$ and $\mathcal{O}$ are the sets of "incoming" and "outgoing" branches. Dually, "tangential" measure ν on the branches of τ determines an element of $\mathcal{ML}_0$ (via ν^*), but two functions ν and ν' determine the same element if ν is obtained from ν' by a process of adding a constant to the incoming branches of a switch, and subtracting the same constant from the outgoing branches—or, in other words, if $\nu - \nu'$ annihilates all transverse measures for τ (using the obvious inner product $\nu \cdot \mu = \sum \nu(b)\mu(b)$). In fact, this operation on ν merely has the effect of switching "trains" from one side of a bigon to the other.



(Some care must be taken to obtain ν' from ν by a sequence of elementary "switching" operations without going through negative numbers. We leave this as an exercise to 9.70
the reader.)

Given an essentially complete train track τ, we now have two canonical coordinate systems V_τ and V_τ^* in $\mathcal{ML}_0$ or $\mathcal{PL}_0$. If $\gamma \in V_\tau$ and $\gamma^* \in V_\tau^*$ are defined by measures μ_γ and ν_{γ^*} on τ, then $i(\gamma, \gamma^*)$ is given by the inner product

$$i(\gamma, \gamma^*) = \sum_{b\in\tau} \mu_\gamma(b)\nu_{\gamma^*}(b).$$

To see this, consider the universal cover of S. By an Euler characteristic or area argument, no path on $\tilde{\tau}$ can intersect a path on $\tilde{\tau}^*$ more than once. This implies the formula when γ and γ' are simple geodesics, hence, by continuity, for all measured geodesic laminations.

Proposition 9.7.4. *Formula* 9.7.3 *holds for all* $\gamma \in V_\tau$ *and* $\gamma^* \in V_\tau^*$. *Intersection number is a bilinear function on* $V_\tau \times V_\tau^*$ *(in* $\mathcal{ML}_0$*).* □

This can be interpreted as a more intrinsic justification for the linear structure on the coordinate systems V_τ—the linear structure can be reconstructed from the embedding of V_τ in the dual space of the vector space with basis $\gamma^* \in V_\tau^*$.

Corollary 9.7.5. *If* $\gamma, \gamma' \in \mathcal{ML}_0$ *are not topologically conjugate and if at least one of them is essentially complete, then there are neighborhoods* U *and* U' *of* γ *and* γ' *with linear structures in which intersection number is bilinear.*

9.71

Proof. Apply 9.7.4 to one of the train tracks τ_i constructed in 9.7.2. □

Remark. More generally, the only requirement for obtaining this local bilinearity near γ and γ' is that the complementary regions of $\gamma \cup \gamma'$ are "atomic" and that $S - \gamma$ have no closed non-peripheral curves. To find an appropriate τ, simply burrow out regions of r_i, "transverse" to γ with points going between strands of γ', so the regions r_i cut all leaves of γ into arcs. Then collapse to a train track carrying γ' and "transverse" to γ, as in 9.7.2.

What is the image of $\mathbb{R}^n$ of stereographic coordinates S_γ for $\mathcal{ML}_0(S)$? To understand this, consider a system of train tracks

$$\tau_1 \longrightarrow \tau_2 \longrightarrow \cdots \longrightarrow \tau_k \longrightarrow \cdots$$

defining S_γ. A "transverse" measure for τ_i pushes forward to a "transverse" measure for τ_j, for $j > i$. If we drop the restriction that the measure on τ_i is non-negative, still it often pushes forward to a positive measure on τ_j. The image of S_γ is the set of 9.72

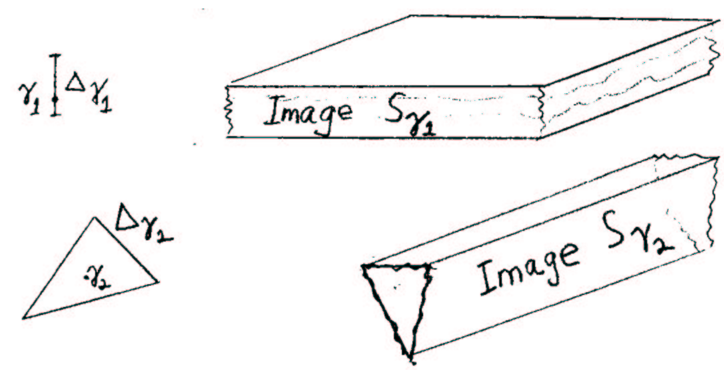
such arbitrary "transverse" measures on τ_1 which eventually become positive when pushed far enough forward.

For $\gamma' \in \Delta_\gamma$, let $\nu_{\gamma'}$ be a "tangential" measure on τ_1 defining γ'.

PROPOSITION 9.7.6. *The image of S_γ is the set of all "transverse," not necessarily positive, measures μ on τ_1 such that for all $\gamma' \in \Delta_\gamma$, $\nu_{\gamma'} \cdot \mu > 0$.*

(Note that the functions $\nu_{\gamma'} \cdot \mu$ and $\nu_{\gamma''} \cdot \mu$ are distinct for $\gamma' \neq \gamma''$.)

In particular, note that if $\Delta_\gamma = \gamma$, the image of stereographic coordinates for $\mathcal{ML}_0$ is a half-space, or for $\mathcal{PL}_0$ the image is $\mathbb{R}^n$. If Δ_γ is a k-simplex, then the image of S_γ for $\mathcal{PL}_0$ is of the form $\operatorname{int}(\Delta^k) \times \mathbb{R}^{n-k}$. (This image is defined only up to projective equivalence, until a normalization is made.)



PROOF. The condition that $\nu_{\gamma'} \cdot \mu > 0$ is clearly necessary: intersection number $i(\gamma', \gamma'')$ for $\gamma' \in \Delta_\gamma$, $\gamma'' \in S_\gamma$ is bilinear and given by the formula 9.73

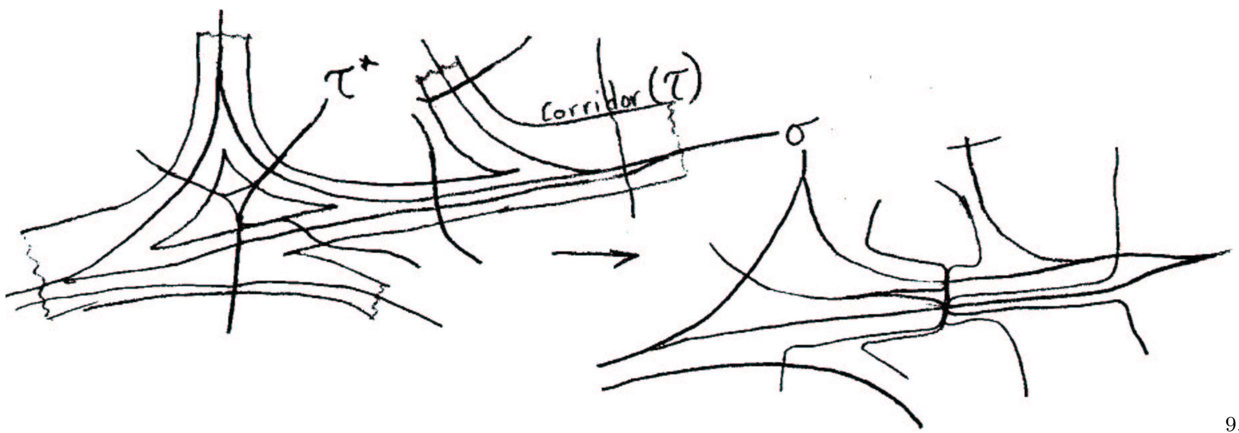
$$i(\gamma', \gamma'') = \nu_{\gamma'} \cdot \mu_{\gamma''}.$$

Consider any transverse measure μ on τ_1 such that μ is always non-positive when pushed forward to τ_i. Let b_i be a branch of τ_i such that the push-forward of μ is non-positive on b_i. This branch b_i, for high i, comes from a very long and thin rectangle ρ_i. There is a standard construction for a transverse measure coming from a limit of the average transverse counting measures of one of the sides of ρ_i. To make this more concrete, one can map ρ_i in a natural way to τ_j^* for $j \leq i$.

(In general, whenever an essentially complete train track τ carries a train track σ, then σ^* carries τ^*

$$\sigma \to \tau$$
$$\sigma^* \leftarrow \tau^*.$$

To see this, embed σ in a narrow corridor around τ, so that branches of τ^* do not pass through switches of σ. Now σ^* is obtained by squeezing all intersections of branches of τ^* with a single branch of σ to a single point, and then eliminating any bigons contained in a single region of $S - \sigma$.)



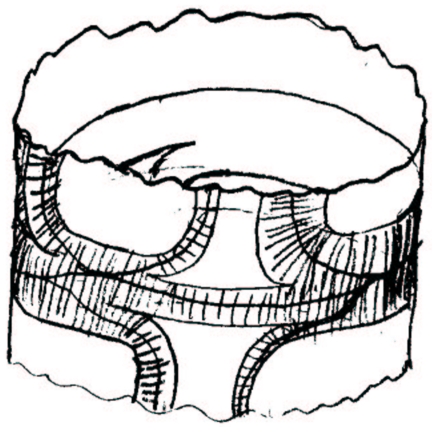
9.74

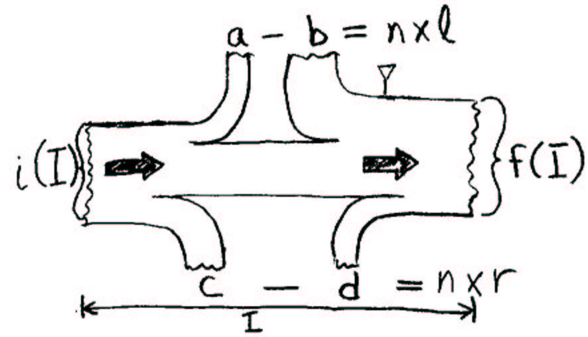
On τ_1^*, ρ_i is a finite but very long path. The average number of times ρ_i tranverses a branch of τ_1^* gives a function ν_i which almost satisfies the switch condition, but not quite. Passing to a limit point of $\{\nu_i\}$ one obtains a "transverse" measure ν for τ_1^*, whose lamination topologically equals γ, since it comes from a transverse measure on τ_i^*, for all i. Clearly $\nu \cdot \mu \leq 0$, since ν_i comes frm a function supported on a single branch b_i^* of τ_i^*, and $\mu(b_i) < 0$. □

For $\gamma \in \mathcal{ML}_0$ let $Z_\gamma \subset \mathcal{ML}_0$ consist of γ' such that $i(\gamma, \gamma') = 0$. Let C_γ consist of laminations γ' not intersecting γ, i.e., such that support of γ' is disjoint from the support of γ. An arbitrary element of Z_γ is an element of C_γ, together with some measure on γ. The same symbols will be used to denote the images of these sets in $\mathcal{PL}_0(S)$.

Proposition 9.7.6. *The intersection of Z_γ with any of the canonical coordinate systems X containing γ is convex.* (In $\mathcal{ML}_0$ or $\mathcal{PL}_0$.)

Proof. It suffices to give the proof in $\mathcal{ML}_0$. First consider the case that γ is a simple closed curve and $X = V_\tau$, for some train track τ carrying γ. Pass to the cylindrical covering space C of S with fundamental group generated by γ. The path of γ on C is embedded in the train track $\tilde{\tau}$ covering τ. From a "transverse" measure m on $\tilde{\tau}$, construct corridors on C with a metric giving them the proper widths. 9.75





For any subinterval I of γ, let $\mathrm{nxr}(I)$ and $\mathrm{nxl}(I)$ be (respectively) the net right hand exiting and the net left hand exiting in the corresponding to I; in computing this, we weight entrances negatively. (We have chosen some orientation for γ). Let $i(I)$ be the initial width of I, and $f(I)$ be the final width.

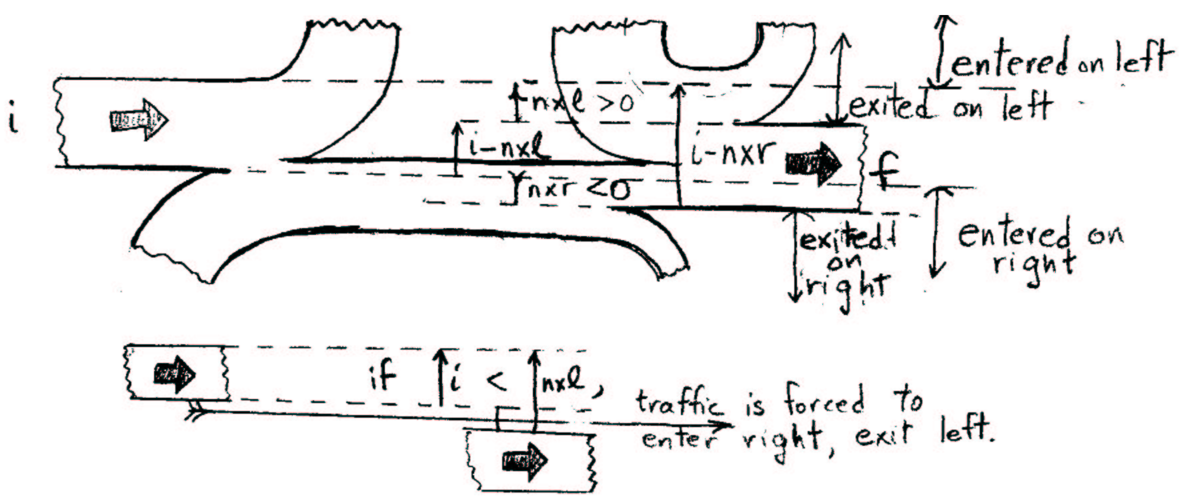
If the measure m comes from an element γ', then $\gamma' \in Z_\gamma$ if and only if there is no "traffic" entering the corridor of γ on one side and exiting on the other. This implies the inequalities

$$i(I) \geq \mathrm{nxl}(I)$$

and

$$i(I) \geq \mathrm{nxr}(I)$$

for all subintervals I.



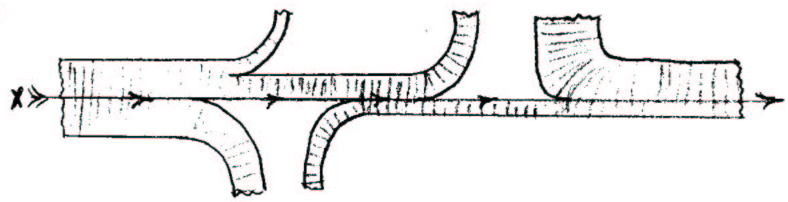
9.76

It also implies the equation

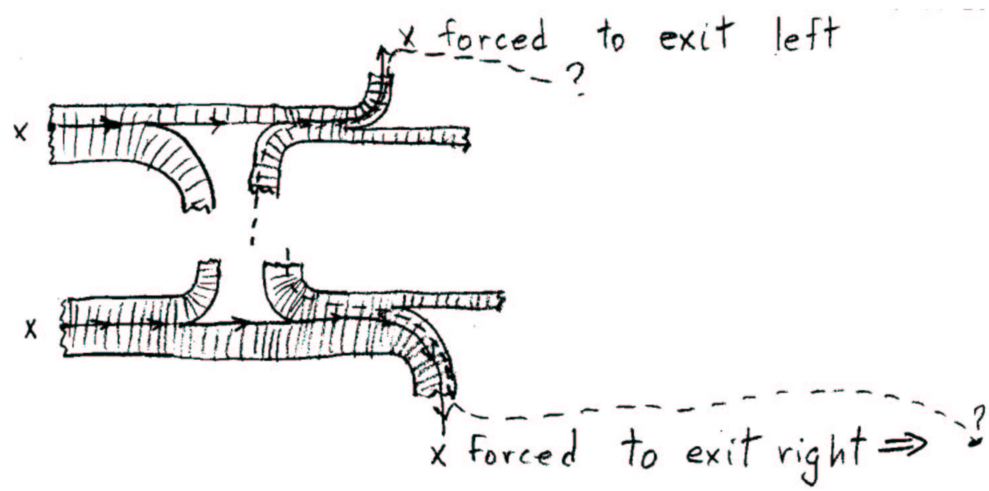
$$\mathrm{nxl}(\gamma) = 0,$$

so that any traffic travelling once around the corridor returns to its inital position. (Otherwise, this traffic would spiral around to the left or right, and be inexorably forced off on the side opposite to its entrance.)

Conversely, if these inequalities hold, then there is some trajectory going clear around the corridor and closing up. To see this, begin with any cross-section of the corridor. Let x be the supremum of points whose trajectories exit on the right. Follow the trajectory of x as far as possible around the corridor, always staying in the corridor whenever there is a choice.



The trajectory can never exit on the left—otherwise some trajectory slightly lower would be forced to enter on the right and exit on the left, or vice versa. Similarly, it can't exit on the right. Therefore it continues around until it closes up.



9.77

Thus when γ is a simple closed curve, $Z_\gamma \cap V_\tau$ is defined by linear inequalities, so it is convex.

Consider now the case $X = V_\tau$ and γ is connected but not a simple geodesic. Then γ is associated with some subsurface $M_\gamma \subset S$ with geodesic boundary defined to be the minimal convex surface containing γ. The set C_γ is the set of laminations not intersecting int (M_γ). It is convex in V_τ, since

$$C_\gamma = \bigcap \{Z_\alpha | \alpha \text{ is a simple closed curve } \subset \text{ int } (M_\gamma)\}.$$

A general element γ' of Z_γ is a measure on $\gamma \cup \gamma''$, so Z_γ consists of convex combinations of Δ_γ and C_γ: hence, it is convex.

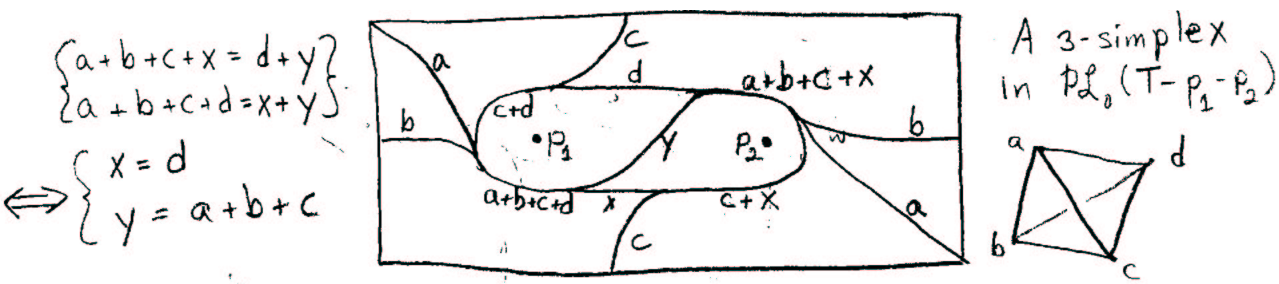
If γ is not connected, then Z_γ is convex since it is the intersection of $\{Z_{\gamma_i}\}$, where the γ_i are the components of γ.

The case X is a stereographic coordinate system follows immediately. When $X = V_\tau^*$, consider any essentially complete $\gamma \in V_\tau$. From 9.7.5 it follows that V_τ^* is linearly embedded in S_γ. (Or more directly, construct a train track (without bigons) carrying τ^*; or, apply the preceding proof to bigon track τ^*.) □

Remark. Note that when γ is a union of simple closed curves, C_γ in $\mathcal{PL}_0(S)$ is homeomorphic to $\mathcal{PL}_0(S-\gamma)$, regarded as a complete surface with finite area—i.e., 9.78
C_γ is a sphere. When γ has no component which is a simple closed curve, C_γ is convex. Topologically, it is the join of $\mathcal{PL}_0(S - \bigcup S_\gamma)$ with the simplex of measures on the boundary components of the S_{γ_i}, where the S_{γ_i} are subsurfaces associated with the components γ_i of γ.

Now we are in a position to form an image of the set of unrealizable laminations for $\rho\pi_1 S$. Let $U_+ \subset \mathcal{PL}_0$ be the union of laminations containing a component of χ_+ and define U_- similarly, so that γ is unrealizable if and only if $\gamma \in U_+ \cup U_-$. U_+ is a union of finitely many convex pieces, and it is contained in a subcomplex of $\mathcal{PL}_0$ of codimension at least one. It may be disjoint from U_-, or it may intersect U_- in an interesting way.

Example. Let S be the twice punctured torus. From a random essentially complete train track,

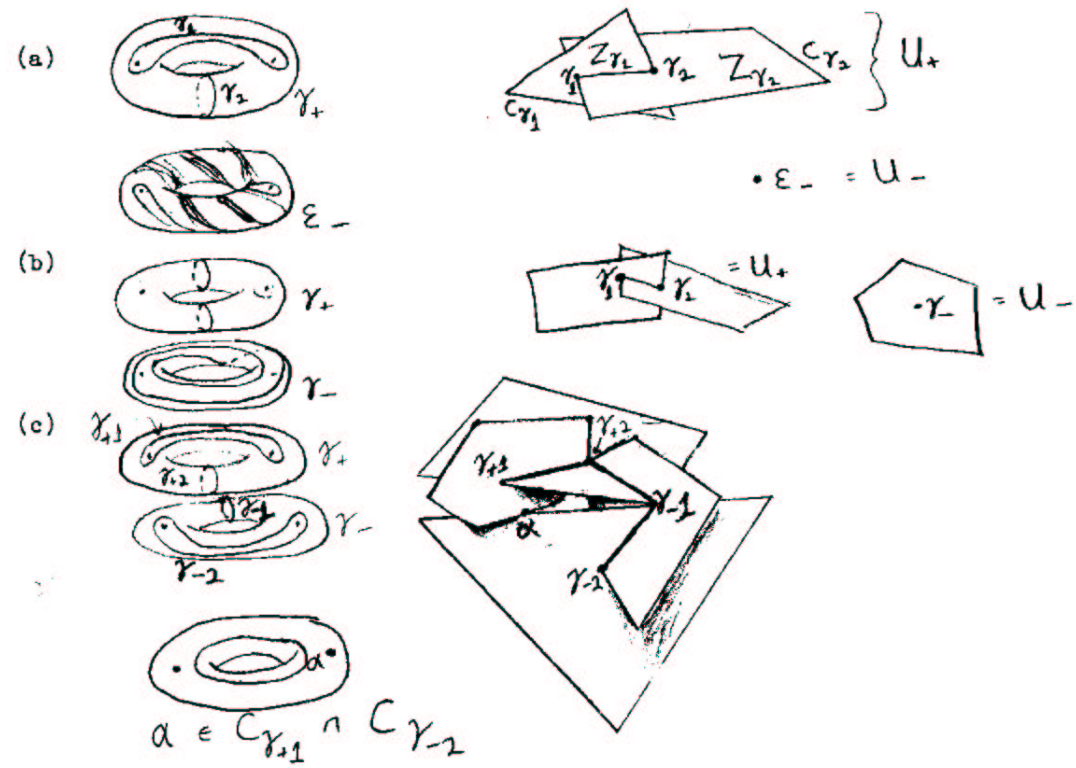


we compute that $\mathcal{ML}_0$ has dimension 4, so $\mathcal{PL}_0$ is homeomorphic to S^3. For any simple closed curve α on S, C_α is $\mathcal{PL}_0(S-\alpha)$,

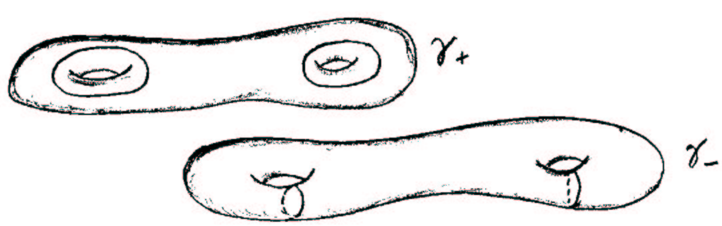
9.79

where $S - \alpha$ is either a punctured torus union a (trivial) thrice punctured sphere, or a 4-times punctured sphere. In either case, C_α is a circle, so Z_α is a disk.

Here are some sketches of what U_+ and $U-$ can look like.

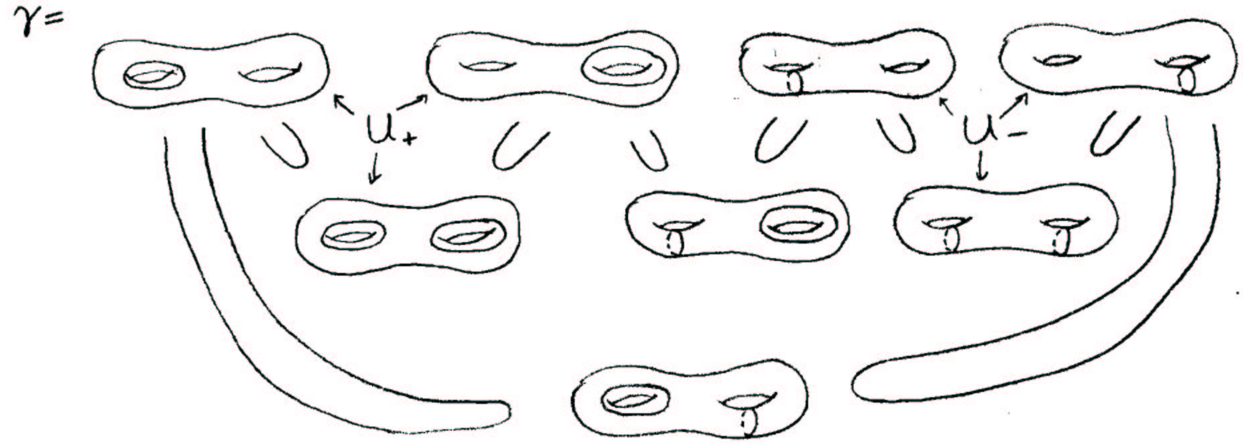


Here is another example, where S is a surface of genus 2, and $U_+(S) \cup U_-(S)$ has the homotopy type of a circle (although its closure is contractible):



9.80

In fact, $U_+ \cup U_-$ is made up of convex sets $Z_\gamma - C_\gamma$, with relations of inclusion as diagrammed:



The closures all contain the element α; hence the closure of the union is starlike:

9.9-1

9.9. Ergodicity of the geodesic flow

We will prove a theorem of Sullivan (1979): There is no §9.8

Theorem 9.9.1. *Let M^n be a complete hyperbolic manifold (of not necessarily finite volume). Then these four conditions are equivalent:*

(a) *The series*

$$\sum_{\gamma \in \pi_1 M^n} \exp\big(-(n-1)\, d(x_0, \gamma x_0)\big)$$

diverges. (Here, $x_0 \in H^n$ is an arbitrary point, γx_0 is the image of x_0 under a covering transformation, and $d(\ ,\ )$ is hyperbolic distance).

(b) *The geodesic flow is not dissipative. (A flow ϕ_t on a measure space (X, μ) is dissipative if there exists a measurable set $A \subset X$ and a $T > 0$ such that $\mu(A \cap \phi_t(A)) = 0$ for $t > T$, and $X = \cup_{t \phi_t(A)}$.)*

(c) *The geodesic flow on $T_1(M)$ is recurrent. (A flow ϕ_t on a measure space (X, μ) is recurrent when for every measure set $A \subset X$ of positive measure and every $T > 0$ there is a $t \geq T$ such that $\mu(A \cap \phi_t(A)) > 0$.)*

(d) *The geodesic flow on $T_1(M)$ is ergodic.*

Note that in the case M has finite volume, recurrence of the geodesic flow is immediate (from the Poincaré recurrence lemma). The ergodicity of the geodesic flow in this case was proved by Eberhard Hopf, in ??. The idea of (c) → (d) goes back to Hopf, and has been developed more generally in the theory of Anosov flows ??. 9.9-2

Corollary 9.9.2. *If the geodesic flow is not ergodic, there is a non-constant bounded superharmonic function on M.*

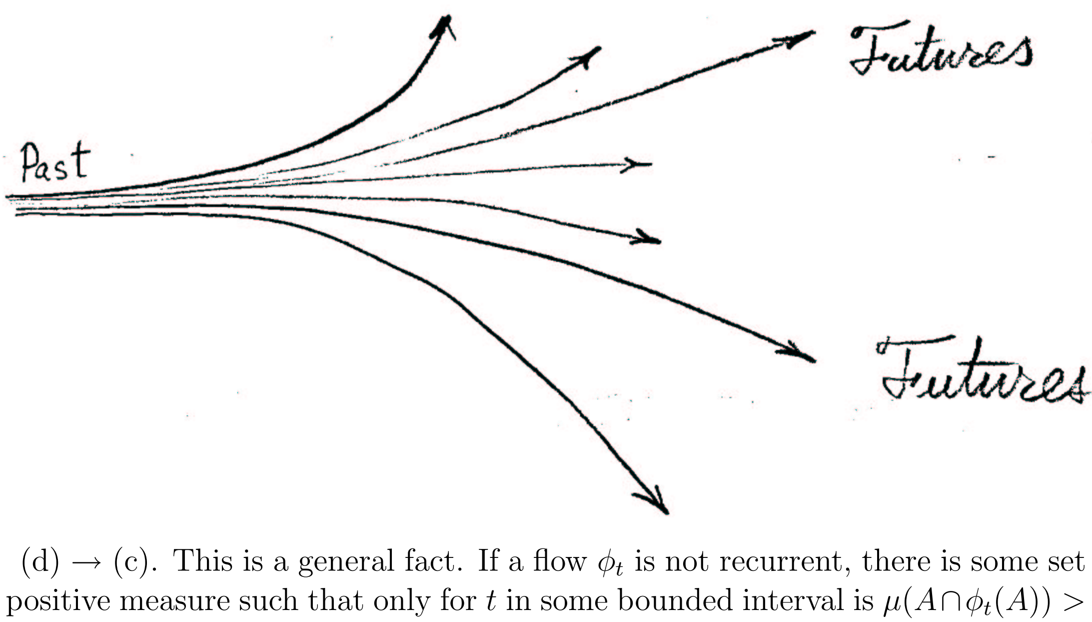
Proof of 9.9.2. Consider the Green's function $g(x) = \int_{d(x,x_0)}^{\infty} \sin h^{1-n} t\, dt$ for hyperbolic space. (This is a harmonic function which blows up at x_0.) By (a), the series $\sum_{\gamma \in \pi_1 M} g \circ \gamma$ converges to a function, invariant by γ, which projects to a Green's function G for M. The function $f = \arctan G$ (where $\arctan \infty = \pi/2$) is a bounded superharmonic function, since arctan is convex. □

Remark. The convergence of the series (a) is actually equivalent to the existence of a Green's function on M, and also equivalent to the existence of a bounded superharmonic function. See (Ahlfors, Sario) for the case $n = 2$, and [ ] for the general case.

Corollary 9.9.3. *If Γ is a geometrically tame Kleinian group, the geodesic flow on $T_1(H^n/\Gamma)$ is ergodic if and only if $L_\Gamma = S^2$.*

Proof of 9.9.3. From 9.9.2 and 8.12.3. □

Proof of 9.9.1. Sullivan's proof of 9.9.1 makes use of the theory of Brownian motion on M^n. This approach is conceptually simple, but takes a certain amount of technical background (or faith). Our proof will be phrased directly in terms of geodesics, but a basic underlying idea is that a geodesic behaves like a random path: its future is "nearly" independent of its past. 9.9-2a



9.9-3

(d) → (c). This is a general fact. If a flow ϕ_t is not recurrent, there is some set A of positive measure such that only for t in some bounded interval is $\mu(A \cap \phi_t(A)) > 0$.

Then for any subset $B \subset A$ of small enough measure, $\cup_t \phi_t(B)$ is an invariant subset which is proper, since its intersection with A is proper.

(c) $\to$ (b). Immediate.

(b) $\to$ (a). Let B be any ball in H^n, and consider its orbit ΓB where $\Gamma = \pi_1 M$. For the series of (a) to diverge means precisely that the total apparent area of ΓG as seen from a point $x_0 \in H^n$, (measured with multiplicity) is infinite.

In general, the underlying space of a flow is decomposed into two measurable parts, $X = D \cup R$, where ϕ_t is dissipative on D (the union of all subsets of X which eventually do not return) and recurrent on R. The reader may check this elementary fact. If the recurrent part of the geodesic flow is non-empty, there is some ball B in M^n such that a set of positive measure of tangent vectors to points of B give rise to geodesics that intersect B infinitely often. This clearly implies that the series of (a) diverges.

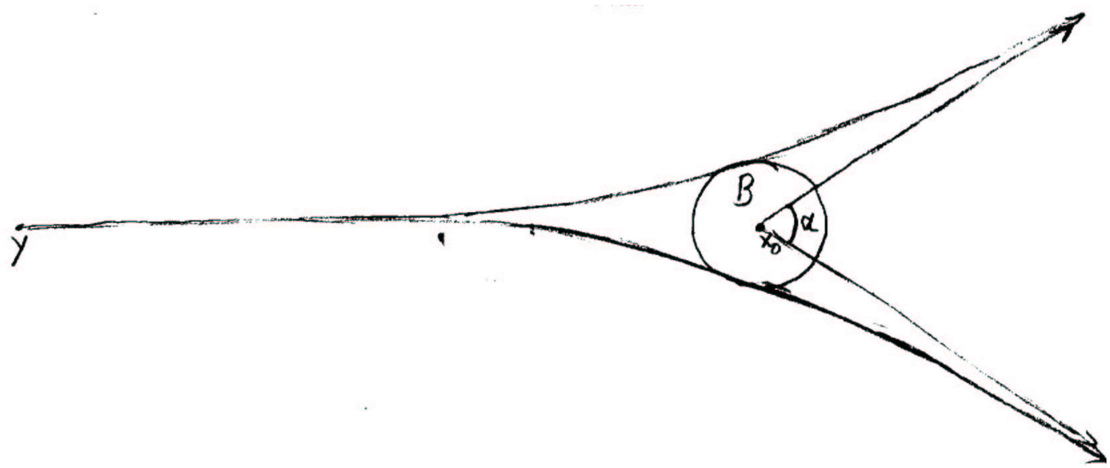
The idea of the reverse implication (a) $\to$ (b) is this: if the geodesic flow is dissipative there are points x_0 such that a positive proportion of the visual sphere is not covered infinitely often by images of some ball. Then for *each* "group" of geodesics that return to B, a definite proportion must eventually escape ΓB, because future and past are nearly independent. The series of (a) can be regrouped as a geometric progression, so it converges. We now make this more precise.

Recall that the term "visual sphere" at x_0 is a synonym to the "set of rays" emanating from x_0. It has a metric and a measure obtained from its identification with the unit sphere in the tangent space at x_0. 9.9-4

Let $x_0 \in M^n$ be any point and $B \subset M^n$ any ball. If a positive proportion of the rays emanating from x_0 pass infinitely often through B, then for a slightly larger ball B', a definite proportion of the rays emanating from *any* point $x \in M^n$ spend an infinite amount of time in B', since the rays through x are parallel to rays through x_0. Consequently, a subset of $T_1(B')$ of positive measure consists of vectors whose geodesics spend an infinite total time in $T_1(B')$; by the Poincaré recurrence lemma, the set of such vectors is a recurrent set for the geodesic flow. (b) holds so (a) $\to$ (b) is valid in this case. To prove (a) $\to$ (b), it remains to consider the case that almost every ray from x_0 eventually escapes B; we will prove that (a) fails, i.e., the series of (a) converges.

Replace B by a slightly smaller ball. Now almost every ray from almost every point $x \in M$ eventually escapes the ball. Equivalently, we have a ball $B \subset H^n$ such that for every point $x \in H^n$, almost no geodesic through x intersects ΓB, or even $\Gamma(N_\epsilon(B))$, more than a finite number of times.

Let x_0 be the center of B and let α be the infimum, for $y \in H^n$, of the diameter of the set of rays from x_0 which are parallel to rays from y which intersect B. This infimum is positive, and very rapidly approached as y moves away from x_0.



9.9-5

Let R be large enough so that for every ball of diameter greater than α in the visual sphere at x_0, at most (say) half of the rays in this ball intersect $\Gamma N_\in(B)$ at a distance greater than R from x_0. R should also be reasonably large in absolute terms and in comparison to the diameter of B.

Let x_0 be the center of B. Choose a subset $\Gamma' \subset \Gamma$ of elements such that: (i) for every $\gamma \in \Gamma$ there is a $\gamma' \in \Gamma'$ with $d(\gamma' x_0, \gamma x_0) < R$. (ii) For any γ_1 and γ_2 in Γ', $d(\gamma_1 x_0, \gamma_2 x_0) \geq R$.

Any subset of Γ maximal with respect to (ii) satisfies (i).

We will show that $\sum_{\gamma' \in \Gamma'} \exp(-(n-1)\, d(x_0, \gamma' x_0))$ converges. Since for any γ' there are a bounded number of elements $\gamma \in \Gamma$ so that $d(\gamma x_0, \gamma' x_0) < R$, this will imply that the series of (a) converges.

Let $<$ be the partial ordering on the elements of Γ' generated by the relation $\gamma_1 < \gamma_2$ when $\gamma_2 B$ eclipses $\gamma_1 B$ (partially or totally) as viewed from x_0; extend $<$ to be transitive.

Let us denote the image of γB in the visual sphere of x_0 by B_γ. Note that when $\gamma' < \gamma$, the ratio $\operatorname{diam}(B_{\gamma'}) / \operatorname{diam}(B_\gamma)$ is fairly small, less than 1/10, say. Therefore $\cup_{\gamma' < \gamma} B_{\gamma'}$ is contained in a ball concentric with B_γ of radius 10/9 that of B_γ.

Choose a maximal independent subset $\Delta_1 \subset \Gamma'$ (this means there is no relation $\delta_1 < \delta_2$ for any $\delta_1, \delta_2 \in \Delta_1$). Do this by successively adjoining any γ whose B_γ has largest size among elements not less than any previously chosen member. Note that area $(\cup_{\delta \in \Delta} B_\delta) / \operatorname{area}(\cup_{\gamma \in \Gamma'} B_\gamma)$ is greater than some definite (a priori) constant: $(9/10)^{n-1}$ in our example. Inductively define $\Gamma'_0 = \Gamma'$, $\gamma'_{i+1} = \Gamma'_i - \Delta_{i+1}$ and define $\Delta_{i+1} \subset \Gamma_i$ similarly to Δ_1. Then $\Gamma' = \cup_{i=1}^{\infty} \Delta_i$. 9.9-6

For any $\gamma \in \Gamma'$, we can compare the set B_γ of rays through x_0 which intersect $\gamma(B)$ to the set C_γ of parallel rays through γX_0.

Any ray of B_γ which re-enters $\Gamma'(B)$ after passing through $\gamma'(B)$, is within ϵ of the parallel ray of C_γ by that time. At most half of the rays of C_γ ever enter $N_\epsilon(\Gamma' B)$.

The distortion between the visual measure of B_γ and that of C_γ is modest, so we can conclude that the set of reentering rays, $B_\gamma \cap \bigcup_{\gamma' < \gamma} B_{\gamma'}$, has measure less than 2/3 the measure of B_γ.

We conclude that, for each i,

$$\begin{aligned} &\operatorname{area}\big(\bigcup_{\gamma\in\Gamma'_{i+1}} B_\gamma\big) - \operatorname{area}\big(\bigcup_{\gamma\in\Gamma'_i} B_\gamma\big) \\ &\geq 1/3 \operatorname{area}\big(\bigcup_{\delta\in\Delta_{i+1}} B_\delta\big) \\ &\geq 1/3\cdot(9/10)^{n-1}\operatorname{area}\big(\bigcup_{\gamma\in\Gamma'_i} B_\gamma\big). \end{aligned}$$

The sequence $\{\operatorname{area}(\bigcup_{\gamma\in\Gamma'_i} B_\gamma)\}$ decreases geometrically. This sequence dominates the terms of the series $\sum_i \operatorname{area} \cup_{\delta\in\Delta_i} B_\delta = \sum_{\gamma\in\Gamma'} \operatorname{area}(B_\gamma)$, so the latter converges, which completes the proof of (a) $\to$ (b). 9.9-7

(b) $\to$ (c). Suppose $R \subset T_1(M^n)$ is any recurrent set of positive measure for the geodesic flow ϕ_t. Let B be a ball such that $R \cap T_1(B)$ has positive measure. Almost every forward geodesic of a vector in R spends an infinite amount of time in B. Let $A \subset T_1(B)$ consist of all vectors whose forward geodesics spend an infinite time in B and let ψ_t, $t \geq 0$, be the measurable flow on A induced from ϕ_t which takes a point leaving A immediately back to its next return to A.

Since ψ_t is measure preserving, almost every point of A is in the image of ψ_t for all t and an inverse flow ψ_{-t} is defined on almost all of A, so the definition of A is unchanged under reversal of time. Every geodesic parallel in either direction to a geodesic in A is also in A; it follows that $A = T_1(B)$. By the Poincaré recurrence lemma, ψ_t is recurrent, hence ϕ_t is also recurrent.

(c) $\to$ (d). It is convenient to prove this in the equivalent form, that if the action of Γ on $S^{n-1}_\infty \times S^{n-1}_\infty$ is recurrent, it is ergodic. "Recurrent" in this context means that for any set $A \subset S^{n-1} \times S^{n-1}$ of positive measure, there are an infinite number of elements $\gamma \in \Gamma$ such that $\mu(\gamma A \cap A) > 0$. Let $I \subset S^{n-1} \times S^{n-1}$ be any measurable set invariant by Γ. Let $-B_1$ and $B_2 \subset S^{n-1}$ be small balls. Let us consider what I must look like near a general point $x = (x_1, x_2) \in B_1 \times B_2$. If γ is a "large" element of Γ such that γx is near x, then the preimage of γ of a product of small ϵ-ball around γx_1 and γx_2 is one of two types: it is a thin neighborhood of one of the factors, $(x_1 \times B_2)$ or $(B_1 \times x_2)$. (γ must be a translation in one direction or the other along an axis from approximately x_1 to approximately x_2.) Since Γ is recurrent, almost every point $x \in B_1 \times B_2$ is the preimage of elements γ of both types, of an infinite number of 9.9-8 points where I has density 0 or 1. Define

$$f(x_1) = \int_{B_2} \chi_I(x_1, x_2)\, dx_2,$$

where χ_I is the characteristic function of I, for $x_1 \in B_1$ (using a probability measure on B_2). By the above, for almost every x_1 there are arbitrarily small intervals around x_1 such that the average of f in that interval is either 0 or 1. Therefore f is a characteristic function, so $I \cap B_1 \times B_2$ is of the form $S \times B_2$ (up to a set of measure zero) for some set $S \subset B_1$.

Similarly, I is of the form $B_1 \times R$, so I is either $\emptyset \times \emptyset$ or $B_1 \times B_2$ (up to a set of measure zero). $\square$