

## THE LARGE-SCALE GEOMETRY OF HILBERT MODULAR GROUPS

BENSON FARB & RICHARD SCHWARTZ

### Abstract

Let  $G$  be the rank-2 semisimple Lie group  $PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ . In this paper we give a canonical isomorphism between the quasi-isometry group and the commensurator group of an irreducible, nonuniform lattice in  $G$ . The most familiar of these lattices are the classical Hilbert modular groups  $PSL_2(\mathcal{O}_d)$ , where  $\mathcal{O}_d$  is the ring of integers in the real quadratic field  $\mathbf{Q}(\sqrt{d})$ . As corollaries to this theorem we obtain the following results:

1. The complete quasi-isometry classification of lattices in  $G$ .
2. Let  $\Gamma$  be any finitely generated group. If  $\Gamma$  is quasi-isometric to an irreducible, nonuniform lattice  $\Lambda$  in  $G$ , then  $\Gamma$  is a finite extension of an irreducible, nonuniform lattice commensurable with  $\Lambda$  in  $G$ .
3. Two irreducible, nonuniform lattices in  $G$  are quasi-isometric iff they are commensurable. In particular, no two distinct classical Hilbert modular groups are quasi-isometric.

### 1. Introduction

This paper introduces a new technique for studying rigidity of lattices in semisimple Lie groups, in particular for studying quasi-isometries without the usual equivariance assumptions as in Mostow Rigidity. The idea is to develop coarse metrical versions of some basic topological principles, for example Alexander duality and Jordan separation, and to apply these principles to pinning down the structure of quasi-isometries of a given lattice. This theory of “coarse topology” is developed and applied in §4 and §5. As a consequence we will prove the first known

---

Received September 20, 1995. Both authors are supported in part by N.S.F. Postdoctoral Fellowships.

quasi-isometric rigidity result for nonuniform lattices in a higher rank semisimple group, where the lack of negative curvature (as in rank-one) makes explicit geometric methods for nonuniform lattices (as in [4]) quite intractable.

Another new idea essential to our work is that of “shadows”, which also allows one to drop the usual group equivariance assumptions. The notion of shadow is given in §3, and is developed and applied in §6 and §7.7.

### 1.1. Background

A quasi-isometry is a map that distorts distances by a (uniformly) bounded amount; this relation captures the large-scale geometry of a space. Quasi-isometries are the maps one considers when studying the geometry of finitely generated groups endowed with the word metric. For many finitely generated groups, the quasi-isometry type of the group actually determines its algebraic structure; for example, a celebrated theorem of Gromov states that any group quasi-isometric to a nilpotent group is virtually nilpotent.

In his plenary address to the I.C.M. in 1983, Gromov laid out a program of studying and classifying infinite discrete groups up to quasi-isometry (see his recent paper/book [1] for an extensive discussion). Much of the motivation for this line of research came from the ideas of Furstenberg and Mostow in their approaches to rigidity (See Mostow’s book [2] for a reference.) In all of the cases above, lattices in semisimple Lie groups are the basic examples considered.

A group acting properly discontinuously by isometries on a (proper) metric space  $X$  is quasi-isometric to  $X$  if and only if the action is cocompact. Hence all uniform (i.e., cocompact) lattices in a Lie group  $G$  are quasi-isometric to  $G$  (with its left-invariant metric); hence to each other. In this way the large-scale geometry of uniform lattices reduces to that of the ambient Lie group; or what is the same thing, the associated symmetric space.

For nonuniform lattices  $\Gamma$  acting on a symmetric space  $X$ , the story is quite different. In this case one chops off every cusp of the quotient  $X/\Gamma$  and looks at the lifts of each cusp to  $X$ , giving a  $\Gamma$ -equivariant union of horoballs in  $X$ . These horoballs are disjoint precisely when  $\Gamma$  is so-called  $Q$ -rank one (e.g. when  $X$  is a hyperbolic or any  $R$ -rank-one symmetric space, or when  $\Gamma$  acts irreducibly on a product of hyperbolic planes). Then  $\Gamma$  acts cocompactly by isometries on  $X$  with these horoballs deleted, giving a quasi-isometry of  $\Gamma$  with this “neutered space”.

Thus the geometry of  $\Gamma$  lies in the geometry of this *packing* of horoballs. Two different nonuniform lattices acting on  $X$  look very much alike since their associated packings and neutered spaces are qualitatively the same. For this reason, the large-scale geometry of nonuniform lattices has not, until recently, been well understood.

In [4], one of us gave the complete quasi-isometry classification of lattices in rank-one lie groups (i.e., the lattices acting on the hyperbolic spaces). As a consequence, in rank-one Lie groups, two nonuniform lattices are quasi-isometric if and only if they are commensurable; equivalently, two nonuniform lattices are quasi-isometric if and only if their associated finite volume orbifold quotients have isometric finite covers.

We have the feeling that the principle “quasi-isometric iff commensurable” should hold for nonuniform lattices in a very wide class of Lie groups, in particular for irreducible lattices in semisimple Lie groups. (The “if” part is trivial.) It is one of the purposes of this paper to verify this conjecture for the simplest higher rank Lie group, namely  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ .

### 1.2. Statement of results

It is natural to consider only irreducible lattices in  $G$ . A lattice in  $G$  is said to be irreducible if it does not contain a finite index subgroup which splits as a product of lattices in  $PSL_2(\mathbf{R})$ . Selberg showed ([6]) that every irreducible lattice in  $G$  is commensurable to one of the classical Hilbert modular groups; hence studying the coarse geometry of irreducible lattices in  $G$  reduces to the study of the coarse geometry of Hilbert modular groups. Prasad [3] proved strong rigidity for these lattices, and more generally for irreducible  $\mathbf{Q}$ -rank one lattices in semisimple Lie groups.

Recall that a classical Hilbert modular group is a group of the form  $PSL_2(\mathcal{O}_d)$ , where  $\mathcal{O}_d$  is the ring of integers in the real quadratic field  $\mathbf{Q}(\sqrt{d})$ . If  $M \in PSL_2(\mathcal{O}_d)$  is a matrix, then let  $\bar{M}$  denote the matrix obtained by replacing each entry of  $M$  by its Galois conjugate. Surprisingly, the representation

$$PSL_2(\mathcal{O}_d) \longrightarrow PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$$

$$M \longmapsto (M, \bar{M})$$

is an irreducible lattice in  $G$ .

Given two irreducible lattices  $\Lambda_1, \Lambda_2 \subset G$ , we say that an element  $g \in G$  *commensurates*  $\Lambda_1$  to  $\Lambda_2$  if the intersection  $\Lambda_2 \cap g\Lambda_1g^{-1}$  has

finite index in  $\Lambda_2$ . Such a group element automatically induces a quasi-isometry between  $\Lambda_1$  and  $\Lambda_2$ . One of our main results states that every quasi-isometry between  $\Lambda_1$  and  $\Lambda_2$  arises in this fashion.

**Main Theorem.** *Suppose  $\Lambda_1$  and  $\Lambda_2$  are nonuniform irreducible lattices in  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ . Then any quasi-isometry between  $\Lambda_1$  and  $\Lambda_2$  is equivalent to a group element which commensurates  $\Lambda_1$  to  $\Lambda_2$ , and hence the natural map*

$$\text{Comm}(\Lambda) \longrightarrow \text{QI}(\Lambda)$$

*is an isomorphism for a nonuniform lattice  $\Lambda$  in  $G$ .*

Roughly, two quasi-isometries are said to be *equivalent* if they differ by a uniformly bounded additive constant (see [4] for a precise definition). We note here that the commensurator group  $\text{Comm}(SL_2(\mathcal{O}_d))$  contains  $PGL_2(\mathbf{Q}(\sqrt{d}))$  as a subgroup of index 2; hence the Main Theorem may be viewed as a concrete computation of the quasi-isometry group  $\text{QI}(SL_2(\mathcal{O}_d))$ .

The Main Theorem basically reconstructs the Lie group  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$  and the lattice  $\Lambda < G$  only from the quasi-isometry type of  $\Lambda$ . Once this reconstruction is available, it is possible to prove the following extremely general result, which characterizes each irreducible, nonuniform lattice in  $G$  among all finitely generated groups by its large-scale geometry.

**Corollary 1.** *Let  $\Gamma$  be any finitely generated group. If  $\Gamma$  is quasi-isometric to an irreducible, nonuniform lattice  $\Lambda$  in  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ , then  $\Gamma$  is a finite extension of an irreducible, nonuniform lattice commensurable with  $\Lambda$  in  $G$ .*

Here are some other corollaries, all of which are rather immediate from the Main Theorem and Corollary 1.

**Corollary 2.** *Two irreducible, nonuniform lattices in  $PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$  are quasi-isometric if and only if they are commensurable.*

In particular, two distinct classical Hilbert modular groups cannot be quasi-isometric.

The Main Theorem also holds for the groups  $PSL_2(\mathcal{O}_d)$  with  $d < 0$ . This is a consequence of [4] (Main Theorem), since these groups are nonuniform lattices in  $PSL_2(\mathbf{C})$ . We can combine our results with those of [4] to make a definitive statement about the large-scale geometry of  $2 \times 2$  matrix groups over quadratic fields:

**Corollary 3a.** *Let  $\mathcal{O}_{d_1}$  and  $\mathcal{O}_{d_2}$  be rings of integers over quadratic fields, real or imaginary. Then  $SL_2(\mathcal{O}_{d_1})$  and  $SL_2(\mathcal{O}_{d_2})$  are quasi-isometric if and only if  $d_1 = d_2$ .*

**Corollary 3b.** *Let  $\Gamma$  be a finitely generated group which is quasi-isometric to  $SL_2(\mathcal{O}_d)$ , where  $d$  is a square-free integer. Then  $\Gamma$  is commensurable with  $SL_2(\mathcal{O}_d)$ .*

Finally, we can classify all lattices in  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$  up to quasi-isometry:

**Corollary 4 [Q.i., classification of lattices].** *There are 3 distinct quasi-isometry classes of lattices in  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ :*

1. *Uniform lattices.*
2. *Nonuniform, reducible lattices.*
3. *Nonuniform, irreducible lattices.*

*Class 1 has a single quasi-isometry class. Class 2 has two quasi-isometry classes: one for products of noncocompact and cocompact fuchsian groups, and one for products of noncocompact fuchsian groups. Class 3 consists of infinitely many distinct quasi-isometry classes, precisely one for each (commensurability class of) nonuniform, irreducible lattice in  $G$ .*

### 1.3. Overview of the proof of the Main Theorem

As is done in [4], we construct a geometric model of the irreducible lattice  $\Gamma$  in  $PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ , called a *neutered space*. The neutered space  $\Omega$  is formed by taking the product of two hyperbolic planes  $X = \mathbf{H}^2 \times \mathbf{H}^2$ , deleting a  $\Gamma$ -equivariant disjoint union of horoballs in  $X$ , and giving the resulting 4-manifold  $\Omega$  with boundary the path metric. The lattice  $\Gamma$  acts properly and cocompactly by isometries on  $\Omega$ , hence  $\Gamma$  is quasi-isometric to  $\Omega$ . The point of using the neutered space in place of the lattice is that it has more geometric and topological structure.

The first main step in the proof is showing that the quasi-isometry between neutered spaces (roughly) takes boundaries to boundaries. In contrast to the rank-one case of [4], we have to contend with directions of zero sectional curvature. As far as we can see, the zero curvature directions make explicit geometric methods quite intractable. Instead, we develop a “large-scale (algebraic) topology”, and exploit the rough analogy between quasi-isometries and homeomorphisms. The arguments in this step are coarse versions of familiar homological arguments from algebraic topology. (Incidentally, this large-scale topology can be used to give new proofs of the corresponding results in [4].)

The components of the boundary of the neutered space  $\Omega$  are horospheres in  $X$ ; these horospheres naturally carry the structure of the

three-dimensional solvable Lie group  $Sol$ . By the first part of the proof, we know that a quasi-isometry between neutered spaces induces a quasi-isometry of  $Sol$ . One key idea is that the structure of the neutered space places *a priori* constraints on the nature of this induced quasi-isometry. The next part of the proof exploits this fact in a detailed analysis of these special quasi-isometries of  $Sol$ . We look at the action of the quasi-isometry on an “internal boundary” of  $Sol$ , which philosophically plays a similar role as the boundary at infinity in proofs of Mostow Rigidity. It is this *internal boundary*, not the Tits boundary of  $X$ , that is our central object of study here. We show that the map on this internal boundary is a product of bilipshitz maps.

The third part of the proof consists of zooming-in at a pair of points of differentiability on the internal boundary, in order to find an isometry to commensurate one lattice to another. While similar in spirit to the geometric limiting arguments given in [4], several interesting difficulties arise.

#### 1.4. Comments

B.F. supplied the insight that the results of [4] should extend to the case of  $\mathcal{Q}$ -rank one lattices, as well as a geometric understanding of these lattices. R.S. supplied the topological techniques used in the proof, the analysis of the induced quasi-isometries of  $Sol$ , and the geometric limit argument.

We view the results of this paper as an interesting special case in the attempt to understand the large-scale geometry of lattices in Lie groups. The essential feature of the lattices here are that they have  $\mathcal{Q}$ -rank one, which puts certain restraints on their large-scale topology. We hope that the techniques developed in this paper may also shed some light on the perplexing case of higher  $\mathcal{Q}$ -rank lattices in semisimple Lie groups.

Thanks go to Michael Freedman, Slava Krushkal, Pierre Pansu, Christophe Pittet, and Peter Teichner for helpful conversations about this paper.

## 2. Background

In this short chapter we provide some necessary background material that we will use throughout the paper.

### The word metric

A finitely generated group  $\Gamma$  may be endowed with the *word metric* by declaring the distance between two elements  $g_1, g_2 \in \Gamma$  to be the minimum value  $k$  such that the group element  $g_1 g_2^{-1}$  can be expressed as the product of  $k$  generators. (The generating set is assumed to be symmetrical, in the sense that  $g$  is a generator iff  $g^{-1}$  is a generator.)

#### Quasi-isometries

Let  $M$  be a metric space. A subset  $N \subset M$  is called a  $K$ -net if every point of  $M$  is within  $K$  of some point of  $N$ . Two metric spaces  $(M_1, d_1)$  and  $(M_2, d_2)$  are said to be  $K$ -quasi-isometric if there are  $K$ -nets  $N_1 \subset M_1$  and  $N_2 \subset M_2$  and maps  $\phi_j : N_j \rightarrow N_{j+1}$  having the following properties:

1.  $d_{j+1}(\phi_j(x), \phi_j(y)) \in [1/K, K]d_j(x, y)$  provided that  $d_j(x, y) \geq K$ .
2.  $d_j \circ d_{j+1}$  does not move any point more than  $K$  units in  $M_j$ .

The indices here are taken mod 2. Two metric spaces are said to be quasi-isometric if there is a  $K$ -quasi-isometry between them for some  $K > 1$ .

What we call a quasi-isometry is called a *coarse* quasi-isometry by some authors. Quasi-isometric equivalence is easily seen to be an equivalence relation on metric spaces.

#### Examples.

1. A group endowed with the word metric for one finite generating set is quasi-isometric to the same group with the word metric for any other finite generating set.
2. A finite extension of a finitely generated group  $\Gamma$  and a finite index subgroup of  $\Gamma$  are both quasi-isometric to  $\Gamma$ .
3. It is a fundamental observation (made by Svarc and Milnor) that if  $M$  is a closed Riemannian manifold, then  $\pi_1(M)$  is quasi-isometric to the universal cover  $\widetilde{M}$ . This observation holds more generally: let  $X$  be a proper geodesic metric space and let  $\Gamma$  be a discrete group of isometries acting properly discontinuously on  $X$ . If  $\Gamma \backslash X$  is compact, then  $\Gamma$  is finitely generated and is quasi-isometric to  $X$ .

Suppose  $q_j : A_j \rightarrow B_j$  are quasi-isometries from  $M$  to itself, defined relative to the nets  $A_j$  and  $B_j$ . Then  $q_1$  and  $q_2$  are said to be *equivalent* if there are constants  $K, K'$  having the following properties:

1. The nets  $A_1, A_2, B_1, B_2$  are  $K$ -nets.
2. If  $x \in A_1$  and  $y \in A_2$  are within  $K$  of each other, then  $q_1(x)$  and  $q_2(y)$  are within  $K'$  of each other.
3. If  $x \in B_1$  and  $y \in B_2$  are within  $K$  of each other, then  $q_1^{-1}(x)$  and  $q_2^{-1}(y)$  are within  $K'$  of each other.

It is routine to verify that the above relation is an equivalence relation, and that, modulo this relation, the quasi-isometries of the metric space  $M$  form a group. We will call this group the *quasi-isometry group* of  $M$ , and denote it by  $QI(M)$ .

A path  $\alpha : [0, p] \rightarrow X$  is a  $K$ -quasi-geodesic if it is a  $K$ -quasi-isometric embedding of  $[0, p]$  into  $X$ .

### 3. The geometry of cusps

In this chapter we describe some of the geometry underlying the Hilbert modular groups. The geometry of horospheres in  $\mathbf{H}^2 \times \mathbf{H}^2$  will play a particularly important role.

#### 3.1. Special horospheres

Let  $G$  denote the six-dimensional Lie group  $PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ , and let  $X$  denote the symmetric space  $X = G/K$  for  $G$ , where  $K = SO(2) \times SO(2)$  is a maximal compact subgroup of  $G$ . Hence  $X$  is a product of hyperbolic planes  $X = \mathbf{H}^2 \times \mathbf{H}^2$ . The full group  $Isom(X)$  of isometries of  $X$  is an extension of  $G$  by the cyclic group of order 2 generated by the isometry of  $X$  which exchanges its  $\mathbf{H}^2$  factors.

In general, a horoball of  $X$  is defined to be the geometric limit of unboundedly large metric balls, provided that this limit exists, and it not all of  $X$ . However, for concreteness, we shall confine ourselves to certain special horospheres. The remainder of this section is devoted to describing these special horospheres.

A horocircle of  $\mathbf{H}^2$  is said to be *based* at a point  $p \in \partial\mathbf{H}^2$  if it is tangent to  $p$ . This is the usual description. Given a horocircle  $\psi$  based  $p$ , we define  $\psi_n$  to be the horocircle with the following properties:

1.  $\psi_n$  is based at  $p$ .
2. The signed distance from  $\psi_n$  to  $\psi$  is  $n$ .

As a convention, we will say that  $n > 0$  means that  $\psi_n$  is contained in the horoball which is bounded by  $\psi$ . We will sometimes call  $\psi$  and  $\psi_n$  *parallel horocircles*.



A *special horosphere* is indexed by a quadruple  $(a, \alpha, b, \beta)$ . Here,  $a$  and  $b$  are points of  $\partial\mathbf{H}^2$ , and  $\alpha$  and  $\beta$  are horocircles based respectively at these points. The special horosphere  $\sigma = (a, \alpha, b, \beta) \subset X$  is defined to be the set

$$\sigma = \{(x, y) \in \mathbf{H}^2 \times \mathbf{H}^2 \mid x \in \alpha_t, y \in \beta_{-t}, t \in \mathbf{R}\}.$$

We equip  $\sigma$  with the path metric induced from  $X$ . In this metric,  $\sigma$  carries the geometry of the three dimensional Lie group *Sol*. In fact, if we use the induced path metric along  $\alpha_t$  and  $\beta_{-t}$ , we can identify  $\sigma$  with  $\mathbf{R}^3 = (x, y, t)$  equipped with the metric:

$$dt^2 + e^s(e^{2t}dx^2 + e^{-2t}dy^2).$$

The number  $s$  depends on the choice of  $\alpha$  and  $\beta$ . Also, there is the usual ambiguity as to which point is identified to the "origin" of *Sol*. We will use  $\sigma$  and *Sol* interchangeably.

It is a well known fact, mentioned in [1], and easy to prove, that the induced path metric on  $\sigma$  is lipschitz equivalent to the metric on  $X$  restricted to  $\sigma$ . This fact, incidentally, is in marked contrast to the situation in rank-one symmetric spaces - e.g. the hyperbolic plane.

Let  $\gamma_a$  and  $\gamma_b$  be geodesics of  $\mathbf{H}^2$  which emanate, respectively, from  $a$  and  $b$ . The set  $(\gamma_a \times \gamma_b) \cap \sigma$  is called a *vertical line*. Under the identification above, this vertical line is identified with a line  $x = \text{const1.}, y = \text{const2.}$  in *Sol*. From either description, it is clear that the space of vertical lines is topologically equivalent to  $\mathbf{R}^2$ .

We will let  $T_n(\sigma)$  denote the  $n$  tubular neighborhood of  $\sigma$  in  $X$ . Given a vertical line  $l \subset \sigma$ , we let  $T_n(l)$  denote the  $n$  tubular neighborhood of  $l$  in  $\sigma$ . The following lemma is easy to verify:

**Lemma 3.1 (Shadow Lemma).** *Suppose  $\sigma_1$  and  $\sigma_2$  are two special horospheres indexed by  $(a_1, \alpha_1, b_1, \beta_1)$  and  $(a_2, \alpha_2, b_2, \beta_2)$  with  $a_1 \neq a_2$  and  $b_1 \neq b_2$ . Let  $V(k) = \sigma_1 \cap T_k(\sigma_2)$ . Then there is a unique vertical line  $l \subset \sigma_1$  such that for any  $k$  with  $V(k)$  nonempty, there exists  $n$  such that :*

$$l \subset V(k) \subset T_n(l).$$

*The number  $n$  also depends on the horospheres.*

The line  $l$  guaranteed by Lemma 3.1 (Shadow Lemma) is called the *shadow* of  $\sigma_2$  on  $\sigma_1$ . The notion of a shadow is probably the central concept of this paper.

### 3.2. The Geometry of *Sol*

In this subsection we recall some properties of the 3-dimensional solvable Lie group *Sol* endowed with its left invariant metric. As a Lie group *Sol* is a split extension of  $\mathbf{R}^2$  by  $\mathbf{R}$  and is topologically just  $\mathbf{R}^3$ . It is useful to think of *Sol* as the 3-manifold  $\mathbf{R}^3$  endowed with the metric

$$dt^2 + e^{2t}dx^2 + e^{-2t}dy^2,$$

where we think of the  $t$ -direction as vertical. *Sol* has two foliations by totally geodesic hyperbolic planes: The foliation by the planes  $y = \text{const.}$  and the foliation by “upside down” hyperbolic planes  $x = \text{const.}$ . Each leaf of each of these foliations is a totally geodesic hyperbolic plane. *Sol* is also foliated by the horizontal planes  $t = \text{const.}$ . One interesting feature of *Sol* is that its sectional curvature is positive along these horizontal planes.

The group  $\text{Isom}(\text{Sol})$  has eight connected components; the identity component being *Sol* itself. The other components come from the isometries  $(x, y, t) \mapsto (\pm x, \pm y, t)$  and  $(x, y, t) \mapsto (\pm y, \pm x, -t)$ .

### 3.3. Neutered spaces

We define a *neutered space* to be the complement of any disjoint collection of special horospheres in  $X$  such that, for any two of the horospheres, if they are indexed by  $(a_1, \alpha_1, b_1, \beta_1)$  and  $(a_2, \alpha_2, b_2, \beta_2)$  respectively, then  $a_1 \neq a_2$  and  $b_1 \neq b_2$ . Such a collection is at most countable. It is understood that the neutered space is equipped with the path metric, induced from the metric on  $X$ .

There is a second metric on a neutered space, given by the restriction of the metric of  $X$ . In contrast to the cases in [4], this metric is (uniformly) lipschitz equivalent to the path metric. This essentially follows from the fact that, on individual horospheres of  $X$ , the two metrics are lipschitz equivalent. For more details, see [1].

### Hilbert modular groups

Hilbert modular groups are among the simplest and most concrete examples of irreducible (nonuniform) lattices in a higher rank semisimple Lie group.

Let  $d > 1$  be a square-free integer, and let  $\mathcal{O}_d$  denote the ring of integers in the real quadratic number field  $\mathbf{Q}(\sqrt{d})$ . Explicitly,  $\mathcal{O}_d$  has the following  $\mathbf{Z}$ -basis:

$$\mathcal{O}_d = \mathbf{Z} + \mathbf{Z}[\sqrt{d}] \quad \text{if } d \equiv 2, 3 \pmod{4},$$

$$\mathcal{O}_d = \mathbf{Z} + \mathbf{Z}\left[\frac{1+\sqrt{d}}{2}\right] \quad \text{if } d \equiv 1 \pmod{4}.$$

The group  $PSL_2(\mathcal{O}_d)$  arises from the group of automorphisms of the  $\mathcal{O}_d$ -module  $\mathcal{O}_d \oplus \mathcal{O}_d$ , and is called a (classical) *Hilbert modular group*.

Denote by  $\bar{\eta}$  the Galois conjugate of  $\eta \in \mathcal{O}_d$ ; so if  $\eta = a + b\sqrt{d}$ , then  $\bar{\eta} = a - b\sqrt{d}$ . Denote by  $\bar{A}$  the matrix in  $PSL_2(\mathcal{O}_d)$  formed from the matrix  $A \in PSL_2(\mathcal{O}_d)$  by replacing each entry of  $A$  by its Galois conjugate. Then the image of the representation

$$\begin{aligned} PSL_2(\mathcal{O}_d) &\longrightarrow PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R}), \\ M &\longmapsto (M, \bar{M}) \end{aligned}$$

is a nonuniform irreducible lattice in  $G$ .  $PSL_2(\mathcal{O}_d)$  has only finitely many orbits of parabolic fixed points; hence the quotient of  $\mathbf{H}^2 \times \mathbf{H}^2$  by this lattice gives a 4-dimensional orbifold with finitely many cusps, called a *Hilbert modular surface*. It is well-known (see, e.g., [7]) that the number of cusps of this orbifold equals the class number  $Q(\sqrt{d})$ .

### 3.5. Canonical models

Given  $PSL_2(\mathcal{O}_d)$ , we know there are finitely many orbits of parabolic points. Let  $\Omega_d$  be the space formed from  $X$  by deleting a  $PSL_2(\mathcal{O}_d)$ -equivariant disjoint union of special horospheres around every parabolic fixed point, and give  $\Omega_d$  the path metric. The result is that  $PSL_2(\mathcal{O}_d)$  acts properly cocompactly by isometries on  $\Omega_d$ , so that, choosing any point  $x \in \Omega_d$ , the natural map

$$\begin{aligned} PSL_2(\mathcal{O}_d) &\rightarrow \Omega_d, \\ \gamma &\mapsto \gamma \cdot x \end{aligned}$$

is a quasi-isometry. We call  $\Omega_d$  the *canonical model* associated to the Hilbert modular group  $PSL_2(\mathcal{O}_d)$ . It is easy to see that  $\Omega_d$  is a neutered space. Actually,  $\Omega_d$  is not uniquely defined, for we may replace our equivariant family by another one. However, we fix one such neutered space, one for each  $d$ , once and for all.

## 4. Detecting boundary components

### 4.1. Overview

Let  $q : \Omega_1 \rightarrow \Omega_2$  be a quasi-isometry between neutered spaces, defined relative to the nets  $N_j \subset \Omega_j$ . We say that  $N_j$  is *adapted* to  $\Omega_j$  if, for every component  $\sigma \subset \partial\Omega_j$ ,  $N_j \cap \sigma$  is a (uniform) net of  $\sigma$ .

We say that  $q$  is *adapted to the pair*  $(\Omega_1, \Omega_2)$  if the following conditions hold:

1.  $q$  is defined on nets  $N_j$  adapted to  $\Omega_j$ .
2. For each horosphere  $\sigma_1 \subset \partial\Omega_1$ , there is a horosphere  $\sigma_2 \subset \partial\Omega_2$  such that  $q(N_1 \cap \sigma_1) \subset \sigma_2$ .
3. Conditions 1 and 2 hold for the inverse map  $q^{-1}$ .

In other words,  $q$  is adapted to  $(\Omega_1, \Omega_2)$  if  $q$  takes horospheres to horospheres. The goal of this chapter is to prove the following:

**Theorem 4.1 (Boundary Detection Theorem).** *Suppose  $q_0 : \Omega_1 \rightarrow \Omega_2$  is an arbitrary quasi-isometry. Then  $q_0$  is equivalent to a quasi-isometry  $q_1 : \Omega_1 \rightarrow \Omega_2$  which is adapted to the pair  $(\Omega_1, \Omega_2)$ .*

This result states that  $q_0$  takes points of  $N_1 \cap \sigma_1$  into a uniformly small tubular neighborhood of some unique horosphere  $\sigma_2 \subset \partial\Omega_2$ .

#### 4.2. Coarse separation

Given any metric space  $X$  and a closed subset  $S \subset X$ , we say that the *inner radius* of  $S$  is the largest value of  $s \in [0, \infty]$  such that a metric ball of radius  $s$  is contained in  $S$ . Let  $T_r(S)$  denote the  $r$ -tubular neighborhood of  $S$ . We say that  $S$  has the *strong separation property* if there is a fixed constant  $r < \infty$  having the following property: For every  $K > 0$ , there are at least two distinct connected components of  $X - T_r(S)$  which have inner radius at least  $K$ . We call  $r$  the *separation constant* of  $S$  in  $X$ .

In Chapter 5, we will prove:

**Theorem 4.2 (Coarse Separation Theorem, Special Case).** *Let  $\Omega \subset X$  be a neutered space, let  $\sigma$  be a horosphere of  $X$ , and let  $q : \sigma \rightarrow \Omega$  be a  $K$ -quasi-isometric embedding. Then  $q(\sigma)$  has the strong separation property in  $X$ , and the separation constant only depends on  $K$ .*

**Corollary 4.3 (Packing Theorem, Special Case).** *Let  $\sigma$  be a horosphere of  $X$ , and  $q : \sigma \rightarrow \Omega$  a  $K$ -quasi-isometric embedding. Then  $q$  is a  $K'$ -quasi-isometry. Here  $K'$  only depends on  $K$ .*

#### 4.3. Proof of the Boundary Detection Theorem

We begin with the following crucial lemma:

**Lemma 4.4 (Avoiding a horosphere).** *Let  $\sigma$  be a horosphere boundary component of the neutered space  $\Omega \subset \mathbf{H}^2 \times \mathbf{H}^2$ , and let  $p, q$  be points in  $\Omega$  which are a distance at least  $C > 0$  from  $\sigma$ . Then there is a path  $\gamma$  from  $p$  to  $q$  in  $\Omega$  lying outside the  $C$ -neighborhood of  $\sigma$ .*

*Proof.* Without loss of generality, we may assume that  $p$  and  $q$  are the same distance  $k$  from  $\sigma$ . Let  $\Psi$  denote the union of other horospheres of  $\Omega$ . From the Lemma 3.1 (Shadow Lemma) and the fact that the horospheres in  $\Psi$  are disjoint, the intersection

$$T_k(\sigma) \cap \Psi$$

consists of a countable disjoint union of sets, each of which is contained in a tubular neighborhood of a vertical line. Hence,

$$T_k(\sigma) \cap \Omega$$

is connected, as desired. q.e.d.

The fact that shadows have codimension at least two inside horospheres is essential to the proof of Lemma 4.5; indeed, the conclusion of the lemma is not true for the hyperbolic plane.

Let  $q = q_0$ . We may assume that  $q$  is defined relative to a net  $N_1$  which is adapted to  $\Omega_1$ . Below, the constants  $K_1, K_2, K_3, \dots$  will only depend on the quasi-isometry constant of  $q$ , and on the neutered spaces  $\Omega_j$ .

**Lemma 4.5.** *Let  $\sigma_1 \subset \partial\Omega_1$  be a horosphere, then there is a constant  $K_1$  and a horosphere  $\sigma_2 \subset \partial\Omega_2$  such that every point of  $\sigma_2$  is within  $K_1$  of some point of  $q(N_1 \cap \sigma_1)$ .*

*Proof.* We will assume that the lemma is false, and obtain a contradiction to the Coarse Separation Theorem. Let  $T_r$  be the  $r$  tubular neighborhood of  $q(\sigma_1)$ .

Let  $x_1, x_2 \subset X$  be any two points which are at least  $m$  away from  $q(\sigma_1)$ . Provided that  $m > r$ , these points will belong to  $X - T_r$ . We will show that, if  $m$  is sufficiently large, these paths belong to the same connected component of  $X - T_r$ , which contradicts the Coarse Separation Theorem.

Let  $x = x_1$ . Suppose that  $x$  does not belong to  $\Omega_2$ . This is to say that  $x$  belongs to some deleted horoball  $\alpha$  with  $\partial\alpha \subset \Omega_2$ . By assumption, there is some point  $y \in \alpha$  which remains at least  $m - r$  from  $T_r$ . If  $m$  is sufficiently large, then clearly there is a path connecting  $x$  to  $y$  which avoids  $T_r$ . (The point is that  $T_r$  only penetrates  $r$  units into the horoball  $\alpha$ .) The same argument, applied to  $x_2$ , says that we can assume, without loss of generality, that  $x_j$  belongs to  $\Omega_2$ .

Since  $q$  is a quasi-isometry, there are points  $p_j \subset \Omega_1$  such that  $z_j = q(p_j)$  is very close to  $y_j$ . Note that any short path from  $y_j$  to

$z_j$  will avoid  $T_r$ , if  $m$  is sufficiently large. Hence, all we have to do is connect  $z_1$  to  $z_2$ .

Note that, if  $m$  is large, then  $p_j$  is very far from  $\sigma$ . From Lemma 4.5 there is a path  $\gamma$  connecting  $p_1$  to  $p_2$ , which also remains very far from  $\sigma$ . The path  $q(\gamma)$  connects  $z_1$  to  $z_2$  and remains very far from  $q(\sigma)$ , as desired. q.e.d.

**Lemma 4.6.** *Let  $\sigma_1, \sigma_2$  be as in Lemma 4.5. Then every point of  $q(N_1 \cap \sigma_1)$  stays within  $K_5$  of  $\sigma_2$ .*

*Proof.* Given a point  $x \in \sigma_2$ , we define  $\phi(x)$  to be a point  $y \in \sigma_1$  such that  $q(y)$  is closest metrically to  $x$ . (If there is more than one, choose randomly.) We assert that  $\phi$  is a  $K_3$ -quasi-isometry. To do this, we will show that  $\phi$  is bi-lipschitz when defined relative to a sufficiently sparse net of  $\sigma_2$ .

Choose two points  $x_1, x_2 \in \sigma_2$ . From Lemma 4.5 are points  $p_1, p_2 \subset \sigma_1$  such that  $q(p_j) = y_j$ , and  $y_j$  is within  $K_1$  of  $x_j$ . If  $x_1$  and  $x_2$  are sufficiently far apart, then the distance from  $y_1$  to  $y_2$  is within a factor of 2 of the distance from  $x_1$  and  $x_2$ . Furthermore, once  $x_1$  and  $x_2$  are sufficiently far apart, then the map  $q^{-1}$  (choosing any inverse images) is  $K_2$  bi-lipschitz on the pair  $\{x_1, x_2\}$ . Hence  $\phi$  is  $2K_2$  bi-lipschitz on  $\{x_1, x_2\}$ .

Thus,  $\phi$  is a  $K_3$ -quasi-isometric embedding from  $\sigma_2$  into  $\sigma_1$ . It follows from the Packing Theorem that  $\phi$  is a  $K_4$ -quasi-isometry. But this immediately implies that every point of  $N_1 \cap \sigma_1$  gets mapped to within some  $K_5$  of  $\sigma_2$ . q.e.d.

For each point  $x \in N_1 \cap \sigma_1$ , define  $q_1(x)$  to be the point of  $\sigma_2$  which is closest (within  $K_5$ ) of  $q(x)$ . The same reasoning as in Lemma 4.6 yields that  $q_1$  is a  $K_6$ -quasi-isometry. By construction,  $q_1$  is adapted to the pair  $(\Omega_1, \Omega_2)$ . This completes the proof of the Boundary Detection Theorem.

## 5. Coarse topology

The purpose of this chapter is to give a proof of the Coarse Separation Theorem and the Packing Theorem, which were used in Chapter 4. Before stating these theorems, we will develop some topological notions which allow us to state and prove a quite general theorem. After introducing each general notion, we will prove a lemma showing that the relevant special case is indeed an example.

**5.1. Unpinched spheres**

Let  $(M, d)$  be a smooth Riemannian  $(n + 1)$ -manifold. Let  $S^n$  denote the standard unit sphere of  $\mathbf{R}^{n+1}$ . Fix, once and for all, some small constant  $\delta_0 = \pi/100$ . (The precise value of  $\delta_0$  is rather arbitrary.) We say that a piecewise smoothly embedded sphere  $\Sigma \subset M$  is *r-unpinched* if there is a homeomorphism  $h : \Sigma \rightarrow S^n$ , where  $S^n$  is the unit sphere, such that the Euclidean distance from  $h(x)$  to  $h(y)$  is less than  $\delta_0$  provided that  $d(x, y) < r$ . Note that  $d$  is *not* the path metric on  $\Sigma$ , but the ambient metric in  $M$ .

We say that  $M$  has an *expanding sphere* provided that  $M$  contains an  $r$ -unpinched sphere for every  $r > 0$ . We remark that having an expanding sphere is an extremely weak condition.

**Lemma 5.1.** *Let  $\sigma$  be a special horosphere in  $X = \mathbf{H}^2 \times \mathbf{H}^2$ , equipped with the invariant Riemannian metric  $d$ . Then  $(\sigma, d)$  has an expanding sphere*

*Proof.* Let  $Q$  denote the unit cube in  $\mathbf{R}^3$ . Let  $\Sigma_n \subset Sol$  be the boundary of the “box”:

$$\sum_n = \{(x, y, t) : |t| \leq n, |x| \leq \exp(\exp(n)), |y| \leq \exp(\exp(n))\}.$$

(The nesting of exponential functions is deliberate.) Let  $h_n : \Sigma_n \rightarrow Q$  be the obvious piecewise affine map from  $\Sigma_n$  to  $Q$ .

Given any point  $p = (x_0, y_0, t_0) \in \Sigma_n$ , and a point  $(x, y, t) \in T_r(p)$ , we have the following:

1.  $|x - x_0| \leq r \exp(n)$ .
2.  $|y - y_0| \leq r \exp(n)$ .
3.  $|t - t_0| \leq r$ .

From these equations, it is easy to verify that (for fixed  $r$ ),  $h_n(T_r \cap \Sigma_n)$  has diameter tending to zero as  $n \rightarrow \infty$ . If we compose  $h_n$  with any piecewise smooth map from  $Q$  to the unit sphere, we see that  $\Sigma_n$  is  $r$  unpinched for sufficiently large  $n$ . q.e.d.

**5.2. Uniform contractibility**

Following Gromov [1], we say that  $(M, d)$  is *uniformly contractible* if there is a function  $\alpha : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  having the following property: If a continuous map of a finite simplicial complex  $\Delta \rightarrow M$  is contained in an  $r$ -ball, then it is contractible in an  $\alpha(r)$ -ball. The function  $\alpha$  is not

supposed to depend on the dimension of  $\Delta$ . Since  $Sol$  is diffeomorphic to  $\mathbf{R}^3$ , and has a transitive group of isometries, it is clearly uniformly contractible.

### 5.3. Coarse separation theorem

For the notion of strong separation, see Section 4. The following theorem is a large-scale version of the Jordan Separation Theorem.

**Theorem 5.2 (Coarse Separation Theorem).** *Suppose that  $J$  and  $Y$  are Riemannian manifolds, diffeomorphic to Euclidean spaces, with  $\dim(J) = \dim(Y) - 1$ . Suppose that  $q : J \rightarrow Y$  is a  $K$ -quasi-isometric embedding. Then  $q$  has the strong separation property provided that*

1.  $J$  has an expanding sphere, and is uniformly contractible.
2.  $Y$  is uniformly contractible.

*The separation constant  $r$  of  $q(J)$  only depends on the triple  $(K, J, Y)$ .*

We remark that the strong separation property is a bit strange, in that it does not rule out the possibility that the complement  $Y - T_r(J)$  consists of infinitely many connected components, some having arbitrarily large inner diameter. A more precise statement, that  $Y - T_r(J)$  consists of two infinite inner diameter connected components, is in fact true. However, the proof of this crisper statement would be too long of a digression. For details, see [5].

Taking the special case  $Y = J \times \mathbf{R}$ , we immediately have the following:

**Corollary 5.3 (Packing Theorem).** *Suppose that  $J$  is diffeomorphic to Euclidean space, is uniformly contractible, and has an expanding sphere. Suppose that  $q : J \rightarrow J$  is a  $K$ -quasi-isometric embedding. Then  $q$  is a  $K'$ -quasi-isometry, and  $K'$  only depends on the pair  $(K, J)$ .*

The Packing Theorem has a bit of history. It was proven for the Euclidean plane by Furstenberg, and general Euclidean space and other symmetric spaces by Block-Weinberger. Our theorem proves much more. For example, it works for solvable groups homeomorphic to  $\mathbf{R}^n$ . Details of this will appear in [5].

As another corollary of the Coarse Separation Theorem, we derive the following:

**Theorem 5.4 (Coarse Separation Theorem, Special Case).** *Let  $\Omega \subset X$  be a neutered space, let  $\sigma$  be a horosphere of  $X$ , and let*



$q : \sigma \rightarrow \Omega$  be a  $K$ -quasi-isometric embedding. Then  $q(\sigma)$  has the strong separation property in  $X$ .

*Proof.* Since the metric on  $\Omega$  is uniformly lipschitz equivalent to the restriction of the ambient metric on  $X$ , the inclusion  $i : \Omega \rightarrow X$  is a quasi-isometric embedding. Hence the composition  $\bar{q} = i \circ q : \sigma \rightarrow X$  is a quasi-isometric embedding. Now apply the Coarse Separation Theorem. q.e.d.

The special case of the Packing Theorem used in Chapter 4 follows immediately from the three preceding results.

The remainder of this chapter is devoted to proving the Coarse Separation Theorem.

#### 5.4. Continuous extension

Let  $T$  be a locally finite simplicial triangulation of  $J$ . Subdividing if necessary, we can guarantee that there is a uniform bound to the size of simplices of  $T$ . Let  $T^{(k)}$  be the  $k$ -skeleton of  $T$ . By altering  $q$  trivially, we can assume that  $q = q_0$  is defined on the zero skeleton  $T^{(0)}$ . Inductively, we extend  $q_k$  from  $T^{(k)}$  to  $T^{(k+1)}$ . The fact that the simplices of  $T$  have uniformly bounded diameter, and that  $Y$  is uniformly contractible, implies that the extension  $q_{k+1}$  is still a quasi-isometry. By induction, then, the final map  $q_n$  is continuous and defined on all of  $J$ , and is still a quasi-isometry. Furthermore, the quasi-isometry constant of  $q_n$  only depends on  $T$ ,  $Y$ , and the quasi-isometry constant of  $q$ .

If  $r$  is sufficiently large, then the  $r$ -tubular neighborhood of  $q(J)$  will contain  $q_n(J)$ . Hence, it is sufficient to reformulate the Coarse Separation Theorem as follows:

**Reformulation of Coarse Separation Theorem.** Provided that the quasi-isometry  $q$  is continuous and defined on all of  $J$ , the complement  $Y - q(J)$  contains at least two distinct connected components of arbitrarily large inner radius.

#### 5.5. Images of Unpinched Spheres

Let  $\Sigma$  be an unpinched sphere. By choosing a triangulation of  $\Sigma$ , we consider  $q(\Sigma)$  as a singular homology class in  $H_n(T)$ , where  $T \subset Y$  is any subset containing  $q(\Sigma)$ .

**Lemma 5.5 (Homology Lemma).** *Let  $\Sigma$  be an unpinched sphere, let  $k$  be a fixed constant, and let  $T \supset q(\Sigma)$  be any compact subset of the  $k$ -tubular neighborhood of  $q(\Sigma)$ . Then there is a second constant  $r$  such that  $q(\Sigma)$  is an infinite element of  $H_n(T)$  provided that  $\Sigma$  is  $r$ -unpinched. The constant  $r$  only depends on the quasi-isometry constant of  $q$ .*

The point of using the subset  $T$  instead of the  $k$ -tubular neighborhood of  $q(\Sigma)$  is that we will need deal with topologically taut sets below (to apply Alexander Duality).

*Proof.* If  $q(\Sigma)$  is a torsion element of  $H_n(T)$ , then there is some  $(n+1)$ -chain  $\beta$ , with underlying simplicial complex  $b = |\beta|$ , having the following properties:

1.  $\beta(b) \subset T$ .
2.  $\partial b$  consists of  $d > 0$  copies of  $\Sigma$ . Call these copies  $c_1, \dots, c_d$ .
3.  $\beta|_{c_j} = q|_{\Sigma}$ .

Below, we will construct a continuous map  $\eta : b \rightarrow \Sigma$  such that  $\eta|_{\partial b}$  has positive degree. This is a contradiction; it states that  $\Sigma$  is a torsion element in  $H_n(J-x)$ , where  $x$  is any point contained in the ball bounded by  $\Sigma$ .

Recall that there is a homeomorphism  $h : \Sigma \rightarrow S^n$ , which takes  $J$ -metric  $r$ -balls into sets having  $\delta_0 = \pi/100$  Euclidean diameter. We will actually construct a map  $\eta : b \rightarrow S^n$ , then pull this map back to  $\Sigma$ , via the map  $h$ .

If necessary, choose a simplicial subdivision  $b'$  of  $b$  with the following property: If two vertices  $v_1, v_2$  belong to the same simplex, then the points  $\beta(v_1)$  and  $\beta(v_2)$  are at most one unit apart in  $Y$ . We first construct the map  $\eta$  from the vertices of  $b'$  into  $S^n$ . Here is the formula:

1. Choose a vertex  $v \in b'$ .
2. Let  $p_v \in q(\Sigma)$  be a point which minimizes the distance from  $\beta(v)$  to  $q(\Sigma)$ . (If this point is not unique, choose randomly amongst the candidates.)
3. Let  $v'$  be an inverse image, under  $q$ , of  $p_v$ . (If this point is not unique, then choose randomly amongst the candidates.)
4. Let  $\eta(v) = h(v')$ .

For two vertices  $v_1$  and  $v_2$  belonging to the same simplex, the points  $\beta(v_j)$  are at most 1 apart. Hence, the points  $q(v'_j)$  are at most  $2k+1$  apart, since  $\beta \subset T$ . Since  $q$  is a  $K$ -quasi-isometry, the two points  $v'_1$  and  $v'_2$  belong to the same  $K(2k+1)$  metric ball in  $J$ . Choosing  $r > K(2k+1)$ , we have that  $\eta(v_1)$  and  $\eta(v_2)$  have distance at most  $\delta_0 = \pi/100$  on the unit sphere  $S^n$ . This property allows one to extend

$\eta$ , skeleta by skeleta, using spherical totally geodesic interpolations. The resulting map is defined and continuous on all of  $b$ .

Let  $c_j$  be a boundary component of  $b$ . Recall that  $c_j$  is actually a copy of  $\Sigma$ . We will show that the map  $\eta|_{c_j}$  has degree one. This implies that  $h^{-1} \circ \eta$  also has degree one. This in turn implies that  $\eta|_{\partial b}$  has positive degree, which is a contradiction.

Let  $x \subset c_j = \Sigma$ , be a vertex of  $b'$ . By construction,  $\eta(x) = h(x')$ , where  $x' \subset \Sigma$  is some point of  $\Sigma$  which is mapped, via  $q$ , closest metrically to  $q(x)$ . Since  $x \subset \Sigma$ , we have that  $q(x) = q(x')$ . Since  $q$  is a  $K$  quasi-isometry, we know that  $x$  and  $x'$  are at most  $K$  units away in  $J$ . Since  $\Sigma$  is certainly  $K$ -unpinched, we have that  $\eta(x)$  lies at most  $\pi/100$  from  $h(x)$ . Therefore, for general  $y \in c_j$ , the points  $\eta(y)$  and  $h(y)$  lie at most  $\pi/50$  from each other. This easily implies that  $\eta|_{c_j}$  has degree one. q.e.d.

**5.6. Alexander duality**

Let  $Z_q(M)$  denote singular  $q$ -cycles of  $M$ . Let  $H_q(M)$  be the singular homology classes of  $M$ . (The homology coefficients are in  $\mathbf{Z}$ , as usual.) Suppose that  $p + q = n - 1$ . Given two cycles  $z_p \in Z_p(\mathbf{R}^n)$  and  $z_q \in Z_q(\mathbf{R}^n)$ , they have an *algebraic linking number*, called  $link(z_p, z_q)$ . This is defined in the usual way: Pick a generic coboundary  $b_{p+1}$  such that  $\partial b_{p+1} = z_p$ , and count the algebraic intersection number  $b_{p+1} \cap z_q$ .

Suppose that  $A \subset \mathbf{R}^n$  is a compact subset.  $A$  is said to be *taut* if there is a sequence of open subsets  $U_j \supset A$  such that  $U_j \supset U_{j+1}$ , and  $U_j$  deformation retracts to  $A$ , and  $\cap U_j = A$ .

**Lemma 5.6 (Alexander Duality).** *Suppose that  $p + q = n - 1$ , that  $A \subset \mathbf{R}^n$  is compact and taut, and that  $\xi \in H_p(A)$  is not a torsion element. Then there is some  $\xi_q \in H_q(\mathbf{R}^n - A)$  such that  $link(\xi_p, \xi_q) \neq 0$ .*

*Proof.* Here is a sketch. For more details, see any book on algebraic topology. The standard Alexander Duality theorem gives an isomorphism  $f : H_q(\mathbf{R}^n - A) \rightarrow H^p(A; \mathbf{Z})$ . Furthermore, the Universal Coefficient Theorem yields a surjection  $\psi : H^p(A; \mathbf{Z}) \rightarrow Hom(H_p(A), \mathbf{Z})$ . The composition  $\psi \circ f$  is exactly the map  $\alpha \rightarrow link(\alpha, *)$ . If  $\xi_p$  is not torsion in  $H_p(A)$ , then there is some element  $\xi_p^* \in Hom(H_p(A), \mathbf{Z})$  such that  $\xi_p^*(\xi_p) \neq 0$ . Finally, there is some  $\xi_q \in H_q$  such that  $\psi \circ f(\xi_q) = \xi_p^*$ . q.e.d.

**Lemma 5.7 (Linking Lemma).** *For any constant  $k$  the following is true: There is a closed curve  $\gamma \subset Y$ , and smoothly embedded sphere  $\Sigma \subset J$  such that:*

1. Every point of  $\gamma$  remains at least  $k$  from  $q(\Sigma)$ .
2.  $\text{link}(q(\Sigma), \gamma) \neq 0$ .

*Proof.* Let  $T_s$  be the  $s$  tubular neighborhood of  $q(\Sigma)$ . In the Hausdorff topology,  $T_{k+1}$  can be arbitrarily closely approximated by a smooth manifold with boundary. Let  $A$  be such an approximation, such that  $T_k \subset A \subset T_{k+2}$ .  $A$  is clearly taut and compact. It follows from Lemma 5.5 that  $q(\Sigma)$  is not torsion in  $H_n(A)$ . Thus by Lemma 5.6, there is some element of  $H_1(Y - A)$  which links  $q(\Sigma)$  algebraically, but avoids  $A$ . This element is represented by a cycle consisting of finitely many closed curves. One of these curves must also link  $q(\Sigma)$ . Perturbing  $\Sigma$  slightly, we can assume that it is smoothly embedded. q.e.d.

### 5.7. Traversing arcs

Let  $\gamma$  and  $\Sigma$  be as in the conclusion of the Linking Lemma. We define  $I_\gamma \subset J$  to be the set of points  $x \in J$  such that  $q(x) \in \gamma$ . In other words  $I_\gamma$  is the inverse image of the intersection  $\gamma \cap q(J)$ .

Let  $B$  be the ball bounding  $\Sigma$ . By construction of  $\gamma$ ,  $I_\gamma$  does not intersect  $\Sigma$ . Say that a point of  $I_\gamma$  is *interior* if it belongs to  $B$ , and *exterior* otherwise. Say that the interval  $\overline{ab} \subset \gamma$  is *traversing* if:

1.  $a, b \in q(J)$ ,
2.  $q^{-1}(a)$  consists entirely of exterior (interior) points,
3.  $q^{-1}(b)$  consists entirely of interior (exterior) points.

**Lemma 5.8.** *Let  $K$  be the quasi-isometry constant of  $q$ . For any constant  $r$ , there is a constant  $r'$  having the following property: Suppose  $\gamma$  avoids  $q(\Sigma)$  by at least  $r'$ . Given a traversing arc  $\overline{ab} \subset \gamma$ , there is a point  $x \in \overline{ab}$  which remains at least  $r$  from  $q(J)$ , and the constant  $r'$  only depends on the quadruple  $(r, K, J, Y)$ .*

*Proof.* The constants  $r_1, r_2, \dots$  have the desired dependence of constants. Choose a sequence of points  $a = x_0, x_1, \dots, x_n = b$  such that the distance in  $Y$  from  $x_i$  to  $x_{i+1}$  is between  $1/2$  and  $1$ . Define  $\phi(x_j) \subset J$  to be any point of  $J$  whose image under  $q$  is as close as possible to  $x_j$ . Assuming that  $\overline{ab}$  remains within  $r$  from  $q(J)$ , the distance between  $q(\phi(x_j))$  and  $x_j$  is at most  $r$ . Hence, the distance from  $\phi(x_j)$  and  $\phi(x_{j+1})$  is at most  $r_1$ . Since  $J$  is uniformly contractible, we have that, for some  $j$ ,  $\phi(x_j)$  is within  $r_2$  of  $\Sigma$ . But then  $q(\phi(x_j))$  is within  $r_3$  of  $q(\Sigma)$ . This implies that  $x_j$  is within  $r_4$  of  $q(\Sigma)$ . Taking  $r' = r_4$  gives the result. q.e.d.

**5.8. The expanding sphere**

We will now prove that  $Y - q(J)$  has the desired separation properties. If this is false, then there is a constant  $m$  having the following property: If  $x, y \in Y$  are at least  $m$  away from  $q(J)$ , then there is a path  $\alpha(x, y) \subset Y - q(J)$  joining  $x$  to  $y$ .

From the previous section, we have, for any value of  $k$ , a piecewise smoothly embedded sphere  $\Sigma \subset J$ , and a curve  $\gamma \subset X$ , such that:

1.  $\text{link}(q(\Sigma), \gamma) \neq 0$ ,
2.  $\gamma$  avoids  $q(\Sigma)$  by at least  $k$ .

Let  $K$  be the quasi-isometry constant of  $q$ . Choose  $k$  so large that, according to Lemma 5.8, there is a point  $x$  on any traversing arc of  $\gamma$  which avoids  $q(J)$  by at least  $m$ .

**Lemma 5.9 (Curve Modification).** *We can find a curve  $\delta$  having the following virtues:*

1.  $\text{link}(q(\Sigma), \delta) \neq 0$ .
2.  $I_\delta$  has only either interior points, or exterior points.

*Proof.* Choose a maximal sequence of traversing arcs  $\gamma_1, \dots, \gamma_n$ . Let  $x_j \subset \gamma_j$  be a point guaranteed by Lemma 5.8. So,  $x_j$  avoids  $q(J)$  by at least  $m$ . Hence there are paths  $\alpha_j = \alpha(x_j, x_{j+1}) \subset Y - q(J)$  connecting  $x_j$  to  $x_{j+1}$ , where indices are taken mod  $n$ . Also, let  $\beta_j$  denote the subarc of  $\gamma$  connecting  $x_j$  to  $x_{j+1}$ .

The closed loops  $\delta_j = \alpha_j \cup -\beta_j$  are 1-cycles. We assert that some  $\delta_j$  links  $q(\Sigma)$  algebraically. If not, then the cycle  $\alpha_1 \cup \dots \cup \alpha_n$  would link  $q(\Sigma)$  but would not intersect  $q(J)$ . So, some  $\delta_k$  links  $q(\Sigma)$  nontrivially. By the maximality assumption on the original arcs, the arc  $\delta = \delta_k$  has the desired intersection properties. q.e.d.

Let  $B$  be the ball bounding  $\Sigma$ . If  $I_\delta$  has no interior points, then  $q(B)$  does not intersect  $\delta$ . This contradicts the fact that  $\delta$  and  $q(\Sigma)$  are linked. Assume on the other hand that  $I_\delta$  has no exterior points. Then  $\Sigma$  is cobordant to an arbitrarily large sphere  $\Sigma' \supset \Sigma$ . By assumption,  $q(\Sigma')$  and  $\delta$  are linked. However, this implies that  $q(\Sigma')$  must intersect each spanning surface of  $\delta$ . In particular, points of  $q(\Sigma')$  must remain fairly close to  $\delta$ , no matter how large  $\Sigma'$  is. This is a contradiction.

## 6. Peripheral quasi-isometries

### 6.1. Overview

Let  $\Omega_1$  and  $\Omega_2$  be neutered spaces associated to two Hilbert modular groups, and let  $q : \Omega_1 \rightarrow \Omega_2$  be an adapted quasi-isometry.

By Chapter 4,  $q$  induces a quasi-isometry from each horosphere  $\sigma$  of  $\Omega_1$  to some horosphere  $\sigma'$  of  $\Omega_2$ . We call this induced quasi-isometry a *peripheral quasi-isometry*. Each horosphere naturally carries the geometry of the 3-dimensional Lie group  $Sol$ ; hence we think of the quasi-isometry  $q$  as inducing a peripheral quasi-isometry  $\phi : Sol \rightarrow Sol$  on the horosphere  $\sigma \subset \Omega_1$ .

In this Chapter, we will show that  $\phi$  takes vertical lines in  $Sol$  into uniformly thin tubular neighborhoods of vertical lines. This fact will allow us to show that the map induced by  $\phi$  on the space of vertical lines in  $Sol$  is bi-lipschitz, from which all of our rigidity results will follow.

We do not know if an arbitrary (i.e., not necessarily induced by a quasi-isometry of a neutered space) quasi-isometry of  $Sol$  must take vertical lines (nearly) to vertical lines. One major difficulty is that there are quasi-geodesics which stray arbitrarily far from geodesics. It is thus necessary to use the special nature of peripheral quasi-isometries in order to show that vertical lines map to vertical lines. Here is an overview of the argument:

**Step 1.** By considering the shadows on  $\sigma$  of other nearby horospheres, we find a “lamination” of vertical lines on  $\sigma$  which are taken (nearly) to vertical lines under  $\phi$ . Call this lamination of vertical lines  $\mathcal{L}$ .

**Step 2.** Making a bounded modification of  $\phi$ , we produce a quasi-isometry  $\psi : Sol \rightarrow Sol$  equivalent to  $\phi$  which takes vertical lines in  $\mathcal{L}$  to other vertical lines. Furthermore, we can assume that  $\psi$  is a uniformly bilipschitz homeomorphism when restricted to lines of  $\mathcal{L}$ .

**Step 3.** Let  $\pi : Sol \rightarrow \mathbf{R}$  be vertical projection  $\pi(x, y, t) = t$ , and let  $i : \mathbf{R} \rightarrow l$  be inclusion of  $\mathbf{R}$  into a line  $l \in \mathcal{L}$ , having the property that  $\pi \circ i = Id$ . On each line  $l \in \mathcal{L}$ , the map  $f_l = \pi \circ \psi \circ i$  can (a priori) be either orientation reversing or orientation preserving. We prove that  $f_l$  is orientation preserving for all  $l \in \mathcal{L}$ . This shows that there is an entire network of lines which is mapped to vertical lines in the topologically correct manner.

**Step 4.** Using Step 3, we show that, for an arbitrary vertical line  $l \subset Sol$ , the map  $f_l$  is coarsely increasing; that is,  $f_l(y) > f_l(x) + 1$

provided that  $y > x + K$ : The constant  $K$  is independent of the line  $l$ .

**Step 5.** By looking at projections of the image  $\psi(l)$  of a vertical line  $l$  onto hyperbolic planes in  $Sol$ , and using the fact that  $\psi(l)$  must keep travelling upward at a definite rate (i.e.,  $f_l$  is coarsely increasing), we show that  $\psi(l)$  must stay close to 2 hyperbolic planes, hence close to some vertical line.

**Step 6.** By Step 5,  $\psi$  induces a map  $\psi_*$  on the space of vertical lines in  $Sol$ , which we identify with the height-zero plane. Again by Step 5,  $\psi$  preserves the 2 foliations by hyperbolic planes, so  $\psi_*$  is a product  $\psi_* = f_1 \times f_2$ . In this step we show that  $f_1$  and  $f_2$  are bilipshitz.

**Step 7.** We use elementary hyperbolic geometry to show that  $\psi$  acts a uniformly bounded amount away from the identity in the vertical direction; that is,  $\psi$  is equivalent to  $f_1 \times f_2 \times Id$ , where each  $f_j$  is bilipshitz.

**6.2. The lamination of shadows**

Let  $E_0$  be the height-zero plane in the horosphere  $\sigma$  considered as the 3-dimensional Lie group  $Sol$ . The intersection of the Hilbert modular group  $PSL_2(\mathcal{O}_d)$  with the stabilizer subgroup of the horosphere, which itself contains  $Sol$  as a finite index subgroup, gives a cocompact lattice in  $Sol$ ; its intersection with  $E_0$  is a lattice in the plane, which we denote by  $L_d$ . To describe  $L_d$ , note that  $E_0$  is the product of horocircles - one from each copy of  $\mathbf{H}^2$ . These horocircles are stabilized by the image of the subgroup  $L < PSL_2(\mathcal{O}_d)$  under the representation  $L \mapsto (L, \bar{L})$ , where  $L$  is the subgroup of upper-triangular matrices with 1's on the diagonal and an arbitrary element of  $\mathcal{O}_d$  in the upper-right corner. From this it is clear that the lattice  $L_d$  may be described as :

$$L_d = \{(v, \bar{v}) : v \in \mathcal{O}_d\}.$$

For each horosphere  $\sigma$  of  $\Omega_1$ , let  $\tau_1, \tau_2, \dots$  be the horospheres of  $\Omega_1$  which are metrically closest to  $\sigma$ . By equivariance, such horospheres form a countable set, indexed by (a quotient of) the cusp group associated to  $\sigma$ . We define  $\mathcal{L}$  to be the set of shadows of  $\tau_j$  on  $\sigma$ . We call  $\mathcal{L}$  the *lamination of shadows* on  $\sigma$ .

**Lemma 6.1 (Lamination lemma).** *Let  $\phi : \sigma \rightarrow \sigma'$  be a peripheral quasi-isometry between horospheres induced by a quasi-isometry  $q : \Omega_1 \rightarrow \Omega_2$  between two neutered spaces. Then there is a constant  $K$ , independent of  $\sigma$  and  $\sigma'$ , having the following property: Every line of  $\mathcal{L}$  in  $\sigma$  is taken to within  $K$  of a (unique) vertical line in  $\sigma'$ .*

*Proof.* The constants  $K_1, K_2, K_3$  below have the desired independence of  $\sigma$  and  $\sigma'$ . Let  $l \subset \mathcal{L}$ . Let  $d$  be the distance from each  $\tau_i$  to  $\sigma$ . Then every point of  $l$  is exactly  $d$  away from  $\tau_j$  for some  $j$ . Set  $\tau = \tau_j$ . Since  $q$  is a quasi-isometry,  $\sigma' = q(\sigma)$  and  $\tau' = q(\tau)$  are at most  $K_1$  apart, and also disjoint. Furthermore, every point of  $q(l)$  is at most  $K_2$  away from  $\tau'$ . Hence  $q(l)$  is contained in some uniform tubular neighborhood  $T_{K_3}(l')$ , where  $l'$  is the shadow of  $\tau'$  on  $\sigma'$ . q.e.d.

**Remark.** It is precisely the property that the lamination  $\mathcal{L}$  is quasi-preserved that makes use of the fact that the quasi-isometry  $\phi : Sol \rightarrow Sol$  is peripheral (i.e., induced by  $q$ ), as opposed to an arbitrary quasi-isometry of  $Sol$ .

The set of lines  $\mathcal{L}$  is invariant under the subgroup of  $PSL_2(\mathcal{O}_d)$  stabilizing  $\sigma$ . Often, we will identify  $\mathcal{L}$  with its intersection with the height zero Euclidean plane  $E_0$ . This intersection is exactly the Euclidean lattice  $L_d$  described in above.

For cosmetic purposes, we redefine  $\phi$  slightly, in light of Lemma 6.1 (Lamination lemma). Let  $l \subset \mathcal{L}$  be a line. Let  $l'$  be the line of  $Sol$  whose tubular neighborhood contains  $\phi(l)$ . Finally, let  $N \subset \sigma$  be the net relative to which  $\phi$  is defined.

Since every point in  $Sol$  is uniformly close to some point on a line of  $\mathcal{L}$ , we may alter  $\phi$  so that  $N$  is a subset of the union of lines in  $\mathcal{L}$ . For each point  $x \in N$ , let  $x'$  denote the point of  $l'$  which is metrically closest to  $\phi(x)$ . If the points of  $N$  are sufficiently sparse, the map  $x \rightarrow x'$  will be bi-lipschitz and order preserving. Extending the linear map  $x \rightarrow x'$  in the obvious way, we obtain a bi-lipschitz map from  $l$  to  $l'$  which is uniformly close to the original map  $\phi|_l$ . Taking the union of these maps over all lines  $l \in \mathcal{L}$ , we obtain a new quasi-isometry  $\psi : Sol \rightarrow Sol$  which is a bi-lipschitz homeomorphism when restricted to any line of  $\mathcal{L}$ . Furthermore,  $\psi$  is uniformly equivalent to  $\phi$ . In what follows we will work with  $\psi$  rather than  $\phi$ .

### 6.3. Geometry of the lamination

In this section we describe some of the geometry of  $\mathcal{L}$ . As we mentioned above, each line of  $\mathcal{L}$  intersects the height-0 plane  $E_0$  in  $Sol$  in a single point. We will use this intersection point to label the line in question: we write  $l = \langle x, y \rangle$  if  $l \cap E_0 = (x, y, 0)$ .

Given any two vertical lines  $l_1, l_2$  in  $Sol$ , let  $d(l_1, l_2)$  be the metric distance in  $Sol$  between these lines. There are two relevant formulas for the metric distance between these lines, depending on how close they are. Let  $\langle x_j, y_j \rangle$  be the coordinates of  $l_j \cap E_0$ . Then we have:



Formula 1: If  $d(l_1, l_2) \geq 2$  then  $d(l_1, l_2) \sim \log(|x_2 - x_1||y_2 - y_1|)$ .

Formula 2: If  $d(l_1, l_2) \leq 2$  then  $d(l_1, l_2) \sim \sqrt{|x_2 - x_1||y_2 - y_1|}$ .

Here  $\sim$  means within a uniform multiplicative factor.

Formula 2 follows from the fact that any Riemannian manifold is infinitesimally Euclidean. Note that if  $l_1$  and  $l_2$  lie in a hyperbolic plane (equivalently, if the points  $l_i \cap E_0$  have a common coordinate), then  $l_1$  and  $l_2$  are asymptotic to each other, so  $d(l_1, l_2) = 0$ , agreeing with Formula 2.

Note that Formula 1 is invariant under isometries of  $Sol$ . Applying an isometry, we may assume that  $l_1 = \langle 0, 0 \rangle$  and that  $l_2 = \langle n, n \rangle$ . By symmetry, the points  $(0, 0, 0)$  and  $(n, n, 0)$  minimize the distance between the two lines. It is standard that the distance between these points is  $\sim 2 \log(n)$  for large  $n$ .

We say that a line  $l_2 \in \mathcal{L}$  is *adjacent* to a line  $l_1 \in \mathcal{L}$  if no other line of  $\mathcal{L}$  is closer to  $l_1$  than  $l_2$ . By compactness and equivariance, the minimum distance between two adjacent lines of  $\mathcal{L}$  exists and is positive. Indeed, the lines  $\langle 0, 0 \rangle$  and  $\langle 1, 1 \rangle$  are adjacent; their distance lies somewhere between 1 and  $\sqrt{2}$ .

Suppose  $\delta = a + b\sqrt{d}$  is the fundamental solution to Pell's equation:  $a^2 - db^2 = 1$ . Then for each integer  $n$ , the line through  $\langle \delta^n, \bar{\delta}^n \rangle$  is adjacent to  $\langle 0, 0 \rangle$  as well. All such points lie along a hyperbola, and there is a hyperbolic linear automorphism of the lattice  $L_d$  which permutes these points.

**Definition (Pell square).** A *Pell square* is a choice of 4 lines of  $\mathcal{L}$  which are cyclically adjacent. We also think of a Pell square as four finite subsegments of these lines, together with the four paths of minimal length between adjacent segments; that is, a labelled octagon so that (see Figure 1):

1. Each even edge of the octagon is a finite subsegment of a line in  $\mathcal{L}$ .
2. Two successive even edges lie on adjacent lines.
3. The odd edges realize the minimum distance  $d_0$  between adjacent lines.

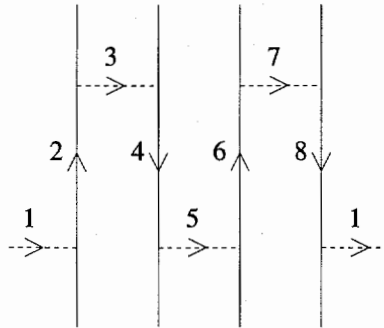


FIGURE 1. A schematic representation of a Pell square.

The reason for the terminology “Pell square” is that one can find such a square containing a given vertical segment by (essentially) solving Pell’s equation 4 times (see the proof of Lemma 6.2 below).

A Pell square involving the line  $l$  determines a unique subsegment  $l' \subset l$ ; the endpoints of  $l'$  are the points of intersection of  $l$  with the arcs of minimal length connecting  $l$  to the adjacent lines of the Pell square. Conversely, given an arbitrary subsegment  $l' \subset l$ , there exists (up to an additive constant) a Pell square containing  $l'$ :

**Lemma 6.2 (Pell lemma).** *Up to a uniformly small additive constant, the following is true: Given two adjacent lines  $l_1, l_2 \in \mathcal{L}$  and some large  $n > 0$ , there exists a Pell square involving subsegments of  $l_1$  and  $l_2$  having length at least  $n$ .*

*Proof.* Applying an isometry, we can assume that  $l_1 = \langle 0, 0 \rangle$ ,  $l_2 = \langle 1, 1 \rangle$ , and  $l_1$  has the origin as a lower endpoint. Let  $\delta$  be the fundamental solution to Pell’s equation. Let  $P_n$  be the Pell square made from the lines  $\langle 0, 0 \rangle, \langle 1, 1 \rangle, \langle \delta^n, \delta^n \rangle, \langle 1 + \delta^n, 1 + \delta^n \rangle$ . The length of Edge 2 in  $P_n$  (in particular) is a linear function in  $n$ . Indeed, all even edges have the same length, by symmetry. q.e.d.

It will also be useful to have the notion of a half-Pell square.

**Definition (Half-Pell square).** A half-Pell square is the same as a Pell square except that:

1. Edge 2 and Edge 4 do not have to belong to  $\mathcal{L}$ . Edge 6 and Edge 8 do have to belong to  $\mathcal{L}$ .

2. Successive even edges lie on vertical lines which are between 1 and  $2d$  apart in  $Sol$ . (Here  $d$  is the integer in  $L_d$ .)
3. Odd edges have length at most  $2d$ .

Note that every Pell square is a half-Pell square.

**Lemma 6.3 (Half-Pell lemma).** *Up to a small additive constant the following is true: given a subsegment  $l'$  of an arbitrary vertical line  $l \subset Sol$ , the segment  $l'$  is part of a half-Pell square.*

*Proof.* We normalize as in the proof of Lemma 6.2 (Pell lemma). Choose any line  $\langle x, y \rangle$  in  $\mathcal{L}$  which has distance between 1 and  $2d$  from the origin. This is possible from the density of  $L_d$  and Formula 2 above. Note that we have translated  $\mathcal{L}$  so that its intersection with  $E_0$  no longer corresponds to  $L_d$ . Let  $H_n$  be the half-Pell square given by  $\langle 0, 0 \rangle, \langle \delta^n, \bar{\delta}^n \rangle, \langle x, y \rangle, \langle x + \delta^n, y + \bar{\delta}^n \rangle$ . Once again, the length of Edge 2 grows linearly in  $n$ , so there is a correct choice of  $n$  up to an additive constant. Note that  $\langle 0, 0 \rangle$  is not necessarily in  $\mathcal{L}$ , and hence neither is  $\langle \delta^n, \bar{\delta}^n \rangle$ . q.e.d.

Recall that we are taking the metric on  $Sol$  to be  $dt^2 + e^t dx^2 + e^{-t} dy^2$ . Let  $\pi_{xt}$  and  $\pi_{yt}$  denote the projections of  $Sol$  onto the  $xt$ -plane and  $yt$ -plane. Note that any plane parallel to the  $xt$ -plane in  $Sol$  has the same intrinsic (hyperbolic) metric as the  $xt$ -plane; similarly for planes parallel to the  $yt$ -plane. The following two facts will be extremely useful throughout the rest of this chapter; their proofs are elementary.

**The divergence principle.** Let  $l_1 = \langle 0, 0 \rangle, l_2 = \langle 1, 1 \rangle$ , or more generally any 2 lines with the property that the rectangle determined by their coordinates is approximately square. Suppose the minimal length path from  $l_1$  to  $l_2$  intersects both lines close to the vertical height  $h$ . Let  $u_i, i = 1, 2$ , be the point on  $l_i$  at height  $h + n$  (resp.  $h - n$ ). Then the points  $\pi_{xt}(u_1)$  and  $\pi_{xt}(u_2)$  are contained in a horocircle in the  $xt$ -plane, and the intrinsic horocircle distance between these two points is  $\sim e^n$  (resp.  $\sim e^{-n}$ ); similarly, the horocircle distance between  $\pi_{yt}(u_1)$  and  $\pi_{yt}(u_2)$  in the  $yt$ -plane is  $\sim e^{-n}$  (resp.  $\sim e^n$ ).

Given two vertical lines in  $Sol$ , the minimal length path between the two lines intersects the lines at some height, which we call the *moment* for the two lines. For example, if  $\delta$  is the fundamental solution of Pell's equation, then the lines  $\langle 0, 0 \rangle$  and  $\langle \delta^n, \bar{\delta}^n \rangle$  have moment  $Kn$  for some constant  $K$  independent of  $n$ .

**Moments are preserved.** There exists a constant  $D > 0$ , depending only on the quasi-isometry  $\psi : Sol \rightarrow Sol$ , with the following

property: If  $l_1, l_2 \in \mathcal{L}$ , and the minimal length path from  $l_1$  to  $l_2$  goes from  $u_1 \in l_1$  to  $u_2 \in l_2$ , then the minimal length path from  $\psi(l_1)$  to  $\psi(l_2)$  intersects  $\psi(l_i)$  within a distance  $D$  from  $\psi(u_i), i = 1, 2$ .

**6.4. Orientation is preserved**

For each line  $l \subset \mathcal{L}$ , consider the bi-lipschitz function  $f_l = \pi \circ \psi \circ i$ , as defined in Step 3 of Section 6. This function is either increasing or decreasing since bilipschitz functions on  $\mathbf{R}$  are increasing or decreasing. By applying an isometry of  $Sol$ , we can assume that  $f_l$  is increasing for at least one line  $l \subset \mathcal{L}$ . In this section, we will show that  $f_m$  is increasing for any other line  $m \subset \mathcal{L}$ . The proof makes use of Pell squares.

**Lemma 6.4 (Orientation lemma).** *The functions  $f_m$  are increasing for all lines  $m \subset \mathcal{L}$ .*

*Proof.* Assume that  $f_m$  is decreasing. By “continuity” we can assume that  $l$  and  $m$  are adjacent lines.

Let  $\psi : Sol \rightarrow Sol$  be a  $K$ -quasi-isometry, and choose  $n \gg K$ . Let  $P_n$  be the Pell square, guaranteed by the Pell lemma (Lemma 6.2), which uses subsegments  $a \subset l$  and  $b \subset m$  of length  $n$  as edges 2 and 4; let  $c$  and  $d$  denote the vertical segments which are edges 6 and 8 of  $P_n$ . Let  $a_i$  and  $a_f$  denote the initial and final points of the subsegment  $a$ ; similarly for  $b, c, d$ .

Let  $P'_n$  denote the image of  $P_n$  under  $\psi$ ; similarly we denote the image of a subsegment, say  $a$ , of  $P_n$  under  $\psi$  by  $a'$ . Recall that  $\psi$  maps each segment of  $P_n$  to a subsegment of some line in  $\mathcal{L}$ . Although  $P'_n$  is not necessarily a closed curve, it may be made so by adding 4 uniformly short arcs.

By assumption the orientation of  $a$  is preserved, and the orientation of  $b$  is reversed under  $\psi$ . Figure 2 shows the 4 possibilities, depending on whether or not the edges  $c$  and  $d$  get reversed.

CASE 1 ( $c^+, d^-$ ) : In this case, all of the edges  $a', b', c', d'$  travel vertically upwards, so that  $d(a'_i, d'_f) \geq 4n/K$ . But  $d(a'_i, d'_f) \leq K \cdot d(a_i, d_f)$ , which is a (uniformly) small number. This gives a contradiction for  $n$  large enough.

CASE 2 ( $c^+, d^+$ ) : Since  $P'_n$  is a closed curve, there must be some point  $x$  of the segment  $d'$  which lies at the same vertical level as the point  $a'_f$  (see Figure 2). Now, since moments are preserved, the moment for the lines containing  $a'$  and  $d'$  occurs near  $d'_f$  and  $a'_i$ . Up to a small additive constant, the points  $x$  and  $a'_f$  on these lines lie an upward vertical distance of at least  $n/K$  from  $d'_f$  and  $a'_i$ , respectively. By the divergence principle,  $\pi_{xt}(x)$  and  $\pi_{xt}(a'_f)$  have horocircle distance  $\sim e^{n/K}$

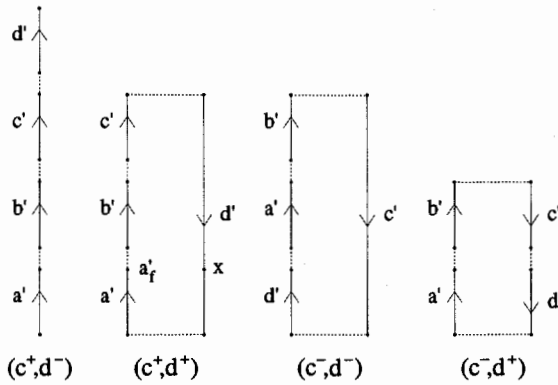


FIGURE 2. This figure shows the four possibilities for the images of the lines making up the Pell square  $P_n$ .

in the  $xt$ -plane, and  $\pi_{yt}(x)$  and  $\pi_{yt}(a'_f)$  have a horocircle distance of  $\sim e^{-n/K}$  in the  $yt$ -plane.

Now consider the path  $\gamma$  from  $a'_f$  through  $b'_f$  and ending at  $x$ . By adding a path of length  $e^{-n/K}$  to  $\gamma$ , we may assume that  $a'_f$  and  $x$  lie in some plane  $P$  which is parallel to the  $xt$ -plane. The path  $\gamma$  is a path of length at most  $4Kn$ , beginning and ending in the plane  $P$ , and  $\gamma$  also stays above the plane  $t = h$  in  $Sol$ . Projecting  $\gamma$  onto the (upside-down hyperbolic) plane  $P$  is length decreasing; following this by downward vertical projection onto the horocircle  $t = h$  in  $P$  is also length decreasing. These projections give a path from  $a'_f$  to  $x$  of length at most  $4Kn$  lying in this horocircle. But  $a'_f$  and  $x$  lie a distance of at least  $e^{n/K}$  from each other as measured in the horocircle metric. This gives a contradiction for large  $n$ ; hence Case 2 cannot occur.

The cases  $(c^-, d^-)$  and  $(c^-, d^+)$  are similar to the previous case. Since none of the four cases can occur, it follows that  $f_m$  must be increasing. q.e.d.

**6.5. Parameter increase for vertical lines**

From the previous sections we know that a peripheral quasi-isometry  $\psi : Sol \rightarrow Sol$  preserves the lamination  $\mathcal{L}$ , and that in fact, after perhaps composing with an isometry of  $Sol$ , the orientation of each line in  $\mathcal{L}$  is preserved. For arbitrary vertical lines, the situation is not so clear. Even though the lines in  $\mathcal{L}$  are taken to vertical lines, little *a priori* can be

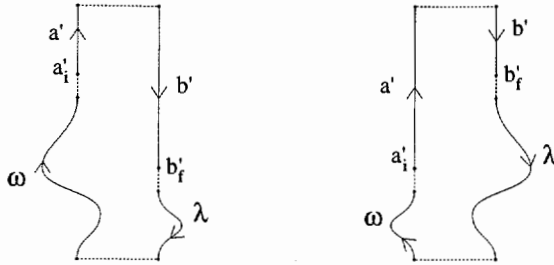


FIGURE 3. This figure illustrates the two possible cases in the proof of Lemma 6.5 (Increasing parameter lemma) : when  $b'_f$  does and does not lie below  $a'_i$ .

said about arbitrary vertical lines. The difficulty is that an arbitrary vertical line possibly comes close to infinitely many *distinct* lines of  $\mathcal{L}$ . In this section, as the first step to understanding the images of arbitrary vertical lines, we prove:

**Lemma 6.5 (Increasing parameter lemma).** *There is a constant  $K$  such that for any vertical line  $l$ , the following holds:  $f_l(y) > f_l(x) + 1$  provided that  $y > x + K$ .*

*Proof.* Long quasigeodesics cannot stay in thin horizontal slabs. Therefore, assuming the lemma false, there are for some very large value of  $n$ , two points  $x, y$  such that  $y \geq x$  and  $f_l(y) = f_l(x) - n$ .

Searching over all triples  $(l, x, y)$ , we may assume that  $y - x$  is (up to one unit from being) as small as possible. (Note that we must have  $y - x > n/K$ , since  $\psi$  is a  $K$ -quasi-isometry.)

Let  $\lambda \subset l$  be the segment connecting  $x$  to  $y$ . By Lemma 6.3 (Half-Pell Lemma), there is a half-Pell square with edges  $\lambda, \omega, a, b$  where both  $a$  and  $b$  are contained in lines which belong to  $\mathcal{L}$ . Note that the images  $a'$  and  $b'$  of  $a$  and  $b$  under  $\psi$  lie on vertical lines. By adding 4 uniformly short arcs to the image of this half-Pell square, we may assume that the image is a closed curve.

Now the point  $b'_f$  has a either lower or heigher height than the point  $a'_i$  (see Figure 3).

First suppose that the point  $b'_f$  sits at a lower vertical height than  $a'_i$ . In this case, the path  $\omega'$  attains a much higher vertical height than the reverse of the path  $\lambda'$ , even though these two paths start at the same height (up to a small additive constant). Hence some subpath of

$\omega$  which is much shorter than  $\lambda$  has image whose vertical parameter is reversed an amount of  $n$ . This contradicts the minimality of  $\lambda$ ; so it must be the case that  $b'_f$  has a height at least as high as  $a'_i$ .

Let  $x \in a'$  be the point of  $a'$  which is at the same vertical height as  $b'_f$ . Let  $E$  be the horizontal plane containing  $x$  and  $b'_f$ .

First note that neither  $\omega'$  nor  $\lambda'$  ever rise above  $E$ , for this would contradict the minimality of  $\lambda$ . Hence the path  $\gamma$  from  $x$  through  $a'_i$  through  $\lambda'_f$  to  $b'_f$  is a path of length at most  $4Kn$  which stays below  $E$ .

The rest of the proof is exactly as in Case 2 of Lemma 6 q.e.d.

**6.6. Images of vertical lines**

In this section we show that the image under  $\psi$  of a vertical line stays (uniformly) close to a vertical line.

Every vertical line of  $Sol$  is contained in two totally geodesic hyperbolic planes. For example, the line  $\langle 0, 0 \rangle$  is contained in the  $xt$ -plane and in the  $yt$ -plane. Furthermore, we have the following lemma whose (obvious) proof we omit.

**Lemma 6.6.** *For any constant  $n$  there is a constant  $n'$  having the following property: The  $n'$ -tubular neighborhood about a vertical line contains the intersection of the  $n$ -tubular neighborhoods of the two hyperbolic planes containing the line.*

**Lemma 6.7.** *Let  $l$  be a vertical line of  $Sol$ , let  $p \in l$  be an arbitrary point, and let  $H$  be a hyperbolic plane through  $\psi(p)$ . Then  $\psi(l)$  lies in the  $C$ -tubular neighborhood about  $H$ . Here  $C$  only depends on the quasi-isometry constant of  $\psi$  and on  $L_d$ .*

*Proof.* The constants  $K_1, K_2, \dots$  below have the desired dependence. By redefining  $\psi$  on  $l$ , we may assume that  $\psi(l)$  is a  $K_1$ -lipschitz curve. For simplicity, we assume that  $H$  is parallel to the  $xt$ -plane. This means that  $H$  is isometric to the upper half plane model of hyperbolic space, by an isometry which *reverses* the direction of vertical lines. (In the other case, the isometry would preserve vertical lines.)

Let  $T_n$  denote the  $n$  tubular neighborhood about  $H$ . Assuming that the lemma is false, we can find a subsegment  $\lambda \subset l$  having the following properties:

1.  $\lambda$  has length  $m$ .
2.  $\psi(\lambda)$  lies completely outside  $T_n$ .
3. Let  $b$  be the lower boundary point of  $\lambda$ . Then  $\psi(b) \subset \partial T_n$ .

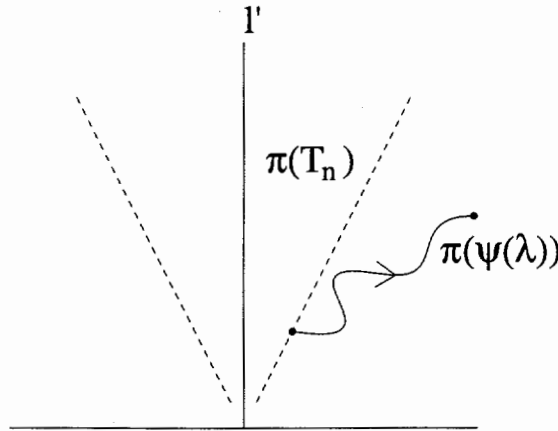


FIGURE 4. The curve  $\pi(\psi(\lambda))$  stays outside an  $n$ -tubular neighborhood of a geodesic in the hyperbolic plane.

Here the pair  $(m, n)$  can be chosen arbitrarily. Let  $t$  be the top endpoint of  $\lambda$ .

Let  $H'$  denote any geodesic hyperbolic plane perpendicular to  $H$ . In other words,  $H'$  is parallel to the  $yt$ -plane. This is to say that  $H'$  is isometric to the upper half plane model of  $\mathbf{H}^2$  by an isometry which preserves the orientation of vertical lines. We will make this identification.

Let  $\pi$  denote projection from  $Sol$  onto  $H'$ . Again, we note that  $\pi$  is distance non-increasing. We have the following information:

1.  $\pi(H) = H \cap H'$  is a vertical line in  $\mathbf{H}^2$ .
2.  $\pi(T_n) = T_n \cap H'$  is the  $n$ -tubular neighborhood about  $l'$  in  $\mathbf{H}^2$ .
3. The curve  $\pi(\psi(\lambda))$  avoids  $\pi(T_n)$ , and has its bottom endpoint on the boundary of this tubular neighborhood.

All of this is shown in Figure 4.

Since  $\psi(\lambda)$  has length at most  $K_1 m$ , the projected curve  $\pi(\psi(\lambda))$  has length at most  $K_1 m$ . Since this curve has its lower endpoint on  $\pi(T_n)$ , the vertical coordinate of  $\pi(\psi(t))$  exceeds that of  $\pi(\psi(b))$  by at most  $K_1 m / \exp(n)$ , from standard hyperbolic geometry.

Since  $\pi$  preserves vertical coordinates, we see that the vertical coordinate of  $\psi(t)$  exceeds that of  $\psi(b)$  by at most  $K_1 m / \exp(n)$ . If both



$m$  and  $n$  are sufficiently large, we get a contradiction to Lemma 6.5 (Increasing parameter lemma).

Consider now the entire parametrized curve  $\bar{l} = \pi(\psi(l))$ . This is a curve in the hyperbolic plane. It has the following property. For any  $p \in \bar{l}$ , and any  $q > p$ , we have that  $\bar{l}(q)$  lies in a thin tubular neighborhood about the vertical line through  $\bar{l}(p)$ . This easily implies that  $\bar{l}$  lies in a tubular neighborhood about some vertical line. q.e.d.

Lemma 6.7 and Lemma 6.6 together imply that the quasi-isometry  $\psi$  takes vertical lines into uniformly thin tubular neighborhoods about vertical lines.

**6.7. The space of vertical lines**

In the previous section, we saw that  $\psi$  takes vertical lines to (tubular neighborhoods about) vertical lines. Since no two vertical lines remain close to each other for the entire length of the line, we see that this vertical line is unique. We identify the space of vertical lines with  $E_0$ . We will call the space of vertical lines the *internal boundary*  $E_0$  of  $Sol$ ; it is this “boundary” that plays the same role as the boundary at infinity in the proof of the rank-one case of Mostow Rigidity. In this section, we study the induced map  $\psi_* : E_0 \rightarrow E_0$ .

Let  $l_1$  and  $l_2$  be two vertical lines. Then  $l_1$  and  $l_2$  lie in a common hyperbolic plane if and only if their tubular neighborhoods intersect in a non-compact set. Also recall that  $\psi$  is orientation preserving when restricted to vertical lines. Combining these two observations, we have that  $\psi$  preserves the two hyperbolic plane foliations of  $Sol$ , and that

$$\psi_* = f_1 \times f_2$$

for some functions  $f_j : \mathbf{R} \rightarrow \mathbf{R}$ . Our goal is to prove that each  $f_j$  is bilipshitz. Incidentally, it is also true that  $\psi$  preserves the horizontal plane foliation of  $Sol$ , though we will not make use of this fact, just yet.

Below, the constants  $K_1, K_2, \dots$  depend only on  $\psi$  and  $L_d$ .

**Lemma 6.8.** *Suppose that  $|x - y| \geq 1$ . Then*

$$|f_j(x) - f_j(y)| \in [1/K_1, K_1]|x - y|.$$

*Proof.* By symmetry, and using inverse maps, it is sufficient to show that  $|f_2(x + \delta) - f_2(x)| < K_1$  provided that  $\delta \leq 1$ . Let  $K_2 = |f_1(1) - f_1(0)|$ . Consider the vertical lines  $l_1 = \langle 0, x \rangle$  and  $l_2 = \langle 1, x + \delta \rangle$ . We have  $d(l_1, l_2) < 2$ , so that  $d(\psi(l_1), \psi(l_2)) < K_2$ . Thus

$$|f_1(0) - f_1(1)||f_2(x) - f_2(x + \delta)| \leq K_3.$$

Hence  $|f_1(x) - f_2(x + \delta)| \leq K_3/K_2$ . q.e.d.

**Lemma 6.9.**  $f_j$  is  $K_4$ -bi-lipschitz

*Proof.* Given the previous lemma, and the existence of inverse maps, it suffices to prove that  $|f_2(x) - f_2(x + \epsilon)| < K_4\epsilon$ , provided that  $\epsilon < 1$ . From the previous lemma, we have

$$f_1(0) - f_1(1/\epsilon) \geq 1/(K_1\epsilon).$$

Let  $l_1 = \langle 0, x \rangle$  and  $l_2 = \langle 1/\epsilon, x + \epsilon \rangle$ . Note that  $d(l_1, l_2) \leq 2$ . Hence

$$d(\psi(l_1), \psi(l_2)) < K_3.$$

The rest of this lemma is just like the previous one. q.e.d.

**6.8. Characterization of peripheral quasi-isometries**

Let  $\phi : Sol \rightarrow Sol$  be a peripheral quasi-isometry. It follows from the work in the previous sections that (after composing with an isometry if necessary)  $\phi$  is uniformly equivalent to a map  $\phi' : Sol \rightarrow Sol$  having the following property:

$$\phi'(x, y, t) = (f_1(x), f_2(y), z(x, y, t)),$$

where the functions  $f_j$  are uniformly bi-lipschitz.

**Lemma 6.10.**  $z(x, y, t)$  is uniformly equivalent to the identity map.

*Proof.* By composing with isometries, it is sufficient to prove that  $z(x, 0, t) = t$ , under the hypothesis that  $f_2(0) = 0$ . In other words, let  $H$  be the hyperbolic  $xt$  plane. Then, we normalize to that  $\phi(H) = H$ , and consider  $z|_H$ .

Let  $\omega = \phi|_H$ . Since  $H$  is totally geodesic,  $\omega$  is a (uniform) quasi-isometry. We consider  $H$  as an isometric copy of the upper half plane. The vertical lines in  $H$  correspond to the vertical lines in the upper half plane model. Let  $S_1 = \partial H = \mathbf{R} \cup \infty$  denote the circle at infinity, and let  $\partial\omega$  denote the canonical extension of  $\omega$  to  $S_1$ . Since  $\omega$  preserves vertical lines, we know that  $\partial\omega(\infty) = \infty$ . Indeed,

$$\partial\omega|_{\mathbf{R}} = f_1.$$

Hence  $\partial\omega$  is uniformly bi-lipschitz.

Consider the new map  $\omega' = \partial\omega \times Id$ . Since  $\partial\omega$  is bi-lipschitz,  $\omega'$  is a quasi-isometry. The two quasi-isometries  $\omega$  and  $\omega'$  have the same boundary extension, and hence are equivalent. The uniformity of the equivalence has the correct dependence of constants. q.e.d.

### 7. Geometric limits

#### 7.1. Some hyperspaces

Let  $X = \mathbf{H}^2 \times \mathbf{H}^2$ . Let  $B_n$  denote the  $n$ -ball about any pre-chosen point  $0 \in X$ . We topologize the closed subsets of  $X$  as follows: Say that a sequence of closed subsets  $S_1, S_2, \dots$  converges to a limit  $S$  provided that, for every pair  $(n, \epsilon)$ , there is an integer  $N$  having the following property: If  $m > N$ , then the  $\epsilon$ -tubular neighborhood of  $S_m \cap B_n$  contains the  $\epsilon$ -tubular neighborhood of  $S \cap B_n$ , and vice versa.

Given a sequence of spaces  $Y_n, Y'_n \subset X$ , and a sequence of  $K$ -quasi-isometries  $q_n : Y_n \rightarrow Y'_n$ , we say that the triple  $(q_n, Y_n, Y'_n)$  converges to the triple  $(q, Y, Y')$  provided the following:

1.  $Y_n \rightarrow Y$  in the Hausdorff topology.
2.  $Y'_n \rightarrow Y'$  in the Hausdorff topology.
3. There exists a constant  $K'$  having the following property: For any fixed  $n$ , there is a constant  $N$  such that  $m > N$  implies that  $q_m$  and  $q$  are  $K'$  equivalent when restricted to  $(Y_m \cup Y) \cap B_n$ .

We will call this topology the *quasi-isometry topology*.

In all that follows, we will use the product upper half-plane model of  $X$ . In this model,  $X$  has coordinates  $(x_1, y_2, x_2, y_2)$  with  $y_j > 0$ ; each factor  $(x_j, y_j)$  is an upper half plane. Let  $\sigma$  be the horosphere in  $X$  given by the equation  $y_1 y_2 = 1$ . We will naturally identify  $\sigma$  with  $Sol$ . So, the coordinates on  $\sigma$  will be  $(x, y, t)$ , as previously. The identification is such that  $(x_1, y_1, x_2, y_2) \in X \cap \sigma$  corresponds to  $(x_1, x_2, \log(y_1) - \log(y_2))$ . In particular,  $(0, 1, 0, 1)$  corresponds to  $(0, 0, 0)$ .

We will say that a map  $\phi : \sigma \rightarrow \sigma$  is *factor preserving* if it can be written in the form

$$\phi(x, y, t) = (f(x), g(y), t).$$

Given such a map, we define  $X\phi = f$ , and  $Y\phi = g$ .

Once and for all, we fix a square-free integer  $d > 0$ , and let  $\Omega = \Omega_d$  denote the neutered space (canonical model) associated to the Hilbert modular group  $SL_2(\mathcal{O}_d)$ . We choose  $\Omega$  so that  $\sigma$  is a peripheral horosphere of  $\Omega$ . Let  $\mathcal{C}$  denote the set of neutered spaces which are isometric to  $\Omega$ , via an isometry which preserves  $\sigma$ .

Let  $\mathcal{S}(K)$  denote the set of triples  $(a, A, A')$  having the following properties:

1.  $A, A' \in \mathcal{C}$ .
2.  $a : A \rightarrow A'$  is a  $K$ -quasi-isometry.
3.  $a$  preserves, and factor preserves,  $\sigma$ .
4.  $a$  fixes the point  $(0, 0, 0) \subset \sigma$ .

Given  $(a, A, A') \in \mathcal{S}(\mathcal{K})$ , let  $Pa$  denote the restriction of  $a$  to  $\sigma$ . From our previous work, we know that

$$Pa = Xa \times Ya \times Id,$$

where  $Xa$  and  $Ya$  are  $C$ -bi-lipshitz maps. The constant  $C$  depends only on  $K$ . Our normalizations guarantee that  $Xa(0) = Ya(0) = 0$ .

Henceforth, we will drop the constants  $C$  and  $K$ . They are implicit in all that is said. We omit the proof of the following lemma, which is completely straightforward, but quite tedious.

**Lemma 7.1.** *The following hold:*

1.  $\mathcal{C}$  is compact in the Hausdorff topology.
2.  $\mathcal{S}$  is compact in the quasi-isometry topology.
3. Suppose that  $(a_n, A_n, A'_n) \rightarrow (b, B, B') \in \mathcal{S}$ . Then  $Xa_n \rightarrow Xb$  and  $Ya_n \rightarrow Yb$  uniformly on compacta in  $\mathbf{R}$ .

## 7.2. Outline

Say that an affine map  $f(x) = mx + b$  is defined over  $\mathbf{Q}(\sqrt{d})$  if both  $m$  and  $b$  belong to this field. Also, define the *conjugate map*  $\bar{f}$  by the equation  $\bar{f}(x) = \bar{m}x + \bar{b}$ . Here  $\bar{m}$  is the Galois conjugate of  $m$ , etc.

The statement of the Main Theorem is that every quasi-isometry of  $\Omega$  is a commensurator. This will follow almost immediately from

**Lemma 7.2 (Rigidity Lemma).** *Suppose  $a : \Omega \rightarrow \Omega$  is a quasi-isometry which preserves, and factor preserves,  $\sigma$ . Then:*

1.  $Xa$  and  $Ya$  are affine maps defined over  $\mathbf{Q}(\sqrt{d})$ .
2.  $Xa$  and  $Ya$  are conjugate maps.

The rest of this chapter is devoted to proving the Rigidity Lemma. Here is an outline of the proof:

**Step 1.** Since  $Xa$  is bi-lipshitz, it is a.e. differentiable. Let  $x_1, x_2$  be two distinct points of differentiability of  $Xa$ . Throughout, we denote the derivative of  $Xa$  by  $Xa'$ . By taking a geometric limit, we will produce two elements  $(c_1, C_1, C'_1), (c_2, C_2, C'_2) \in \mathcal{S}$  such that:

1.  $Xc_1 = Xa'(x_1)$ ; that is,  $Xc_1$  is the linear map "multiply by  $Xa'(x_1)$ ",
2.  $Xc_2 = Xa'(x_2)$ ,
3.  $Yc_1 = Yc_2$ .

**Step 2.** By taking a second geometric limit, we will produce elements  $(e_1, E_1, E'_1), (e_2, E_2, E'_2) \in \mathcal{S}$  such that:

1.  $Xe_1 = Xc_1$ ,
2.  $Xe_2 = Xc_2$ ,
3.  $Ye_1 = Ye_2$ ,
4.  $Ye_1$  and  $Ye_2$  are linear.

**Step 3.** We will isometrically conjugate the triple  $(e_j, E_j, E'_j)$  to a map  $f_j : \Omega \rightarrow \Omega$  such that:

1.  $Xf_j$  is an affine map whose linear part equals  $Xe_j$ ,
2.  $Yf_j$  is an affine map whose linear part equaly  $Ye_j$ .

We will then use the geometry of  $\Omega$  to show that  $Xf_j$  and  $Yf_j$  are conjugate affine maps, defined over  $\mathcal{Q}(\sqrt{d})$ . Since  $Ye_1 = Ye_2$ , the linear parts of  $Xf_1$  and  $Xf_2$  agree. This is to say that  $Xe_1$  and  $Xe_2$  agree. Hence,  $Xa'(x_1) = Xa'(x_2)$ .

**Step 4.** Since  $x_1$  and  $x_2$  were arbitrary points of differentiability,  $Xa'(*)$  is a.e. constant. Since  $Xa$  is bi-lipshitz, this means that  $Xa$  is affine. The same argument, with  $X$  and  $Y$  interchanged, says that  $Ya$  is affine. A repeat of Step 3 now says that these maps are defined over  $\mathcal{Q}(\sqrt{d})$ , and are conjugate to each other.

### 7.3. The zooming principle

Let  $m_n$  denote the linear map of the line which is the multiplication by  $\delta^n$ , for some real number  $\delta > 1$ . Suppose that  $f : \mathbf{R} \rightarrow \mathbf{R}$  has the following properties:

1.  $f(0) = 0$ .

2.  $f'(0)$  exists.

3.  $f$  is bi-lipshitz.

Consider the sequence of maps  $f_n = m_n \circ f \circ m_n^{-1}$ . By the definition of  $f'(0)$ , the sequence of maps  $f_1, f_2, \dots$  converges, uniformly on compacta, to a linear map  $f_\infty$ . Furthermore, the expansion constant of  $f_\infty$  is exactly  $f'(0)$ . In fact, we may write, for shorthand,  $f_\infty = f'(0)$ . We will call this the *zooming principle*, and will rely heavily upon it in what follows.

#### 7.4. First limit

Let  $\delta > 1$  be a unit in  $\mathcal{O}_d$ . Let

$$T_n = \begin{bmatrix} \delta^n & 0 \\ 0 & \delta^{-n} \end{bmatrix}.$$

Let  $Z_n = T_n \times \bar{T}_n$ . Then  $Z_n$  is an isometry of  $X$  which preserves  $\sigma$ . Furthermore,  $PZ_n = m_n \times m_{-n} \times \tau_n$ , where the map  $m_n$  is the multiplication by  $\delta^{2n}$  and the map  $\tau_n$  is the translation by  $n \log(\delta)$ .

Let  $x_1, x_2$  be two (generic) points of differentiability of  $Xa$ . For  $j = 1, 2$ , let  $I_j$  denote the isometry of  $X$  such that

$$PI_j(x, y, t) = (x - x_j, y, t).$$

Let  $J_j$  denote the isometry of  $X$  such that

$$PJ_j(x, y, t) = (x - Xa(x_j), y, t).$$

Define

1.  $B_j = I_j(\Omega)$ ,
2.  $B'_j = J_j(\Omega)$ ,
3.  $b_j = J_j \circ a \circ I_j^{-1}$ .

Note that

1.  $(b_j, B_j, B'_j) \in \mathcal{S}$ ,
2.  $Xb'_j(0) = Xa'(x_j)$ ,
3.  $Yb_1 = Yb_2$ .

For each integer  $n > 0$ , consider the triple

$$(Z_n \circ b_1 \circ Z_n^{-1}, Z_n(B_1), Z_n(B'_1)).$$

For some subsequence  $\xi$ , this collection converges to a triple  $(c_1, C_1, C'_1)$ . By the zooming principle,  $Xc_1$  is linear, and  $Xc_1 = Xb'_1(0)$ .

On some *subsequence of  $\xi$* , the triples

$$(Z_n \circ b_2 \circ Z_n^{-1}, Z_n(B_2), Z_n(B'_2))$$

converge to some  $(c_2, C_2, C'_2)$ . Note that  $Yc_2 = Yc_1$  automatically. By the zooming principle again,  $Xc_2$  is linear and  $Xc_2 = Xb'_2(0)$ .

To summarize, we have elements  $(c_1, C_1, C'_1)$  and  $(c_2, C_2, C'_2)$  such that

1.  $Xc_j$  is linear,
2.  $Xc_j = Xa'(x_j)$ ,
3.  $Yc_1 = Yc_2$ .

### 7.5. Second limit

Let  $(c_j, C_j, C'_j) \in \mathcal{S}$  be as above. Let  $s$  be some point of differentiability for the (common) function  $Yc_j$ . (This function is bi-lipshitz.) Let  $I$  be the isometry of  $X$  which preserves  $\sigma$  and has

$$PI(x, y, t) = (x, y - s, t).$$

Let  $J$  be the isometry of  $X$  which preserves  $\sigma$  and has

$$PJ(x, y, t) = (x, y - Yc_j(s), t).$$

Define

1.  $D_j = I(C_j)$ ,
2.  $D'_j = J(C_j)$ ,
3.  $d_j = J \circ c_j \circ I^{-1}$ .

Note that

1.  $(d_j, D_j, D'_j) \in \mathcal{S}$ ,
2.  $Xd_j = Xa'(x_j)$ , and  $Xd_j$  is linear,

3.  $Yd_1 = Yd_2$ ,
4.  $Yd'_j(0)$  exists.

For each integer  $n > 0$ , consider the triple

$$(Z_n^{-1} \circ d_1 \circ Z_n, Z_n^{-1}(D_1), Z_n^{-1}(D'_1)).$$

(Note here that the inverse of  $Z_n$  appears on the other side of the composition. This is deliberate.) For some subsequence  $\xi$  of integers, this collection converges to a triple  $(e_1, E_1, E'_1)$ . Note that  $Xe_1 = Xd_1$ , since linear maps commute with the dilations  $m_n$ , and note also that  $Ye_1 = Yd'_1(0)$ , by the zooming principle.

On some subsequence of  $\xi$ , the pairs

$$(Z_n^{-1} \circ d_2 \circ Z_n, Z_n^{-1}(D_2), Z_n^{-1}(D'_2))$$

converge to some  $(e_2, E_2, E'_2)$ . By the zooming principle,  $Ye_2 = Yd'_2(0) = Yd'_1(0) = Ye_1$ , and also  $Xe_2 = Xd_2$ . (Once again, this is because linear maps commute with  $m_n$ .)

To summarize, we have elements  $(e_1, E_1, E'_1)$  and  $(e_2, E_2, E'_2)$  such that:

1.  $Xe_j$  is linear,
2.  $Xe_j = Xa'(x_j)$ ,
3.  $Ye_j$  is linear,
4.  $Ye_1 = Ye_2$ .

### 7.6. Modified commensurators

Recall that  $\Omega$  is the standard neutered space associated to  $SL_2(\mathcal{O}_d)$ . Let  $\tau$  be a peripheral horosphere of  $\Omega$ . Recall that the commensurator  $Comm(SL_2(\mathcal{O}_d))$  is isomorphic to the semi-direct product of  $PGL_2(\mathbf{Q}(\sqrt{d}))$  with a cyclic group of order two; the cyclic group switches the two factors of  $X$ . Let  $\phi$  be a commensurator of  $SL_2(\mathcal{O}_d)$ . Then  $\phi$  takes  $\tau$  to a horosphere of  $X$  which is parallel to some other horosphere  $\tau' \subset \Omega$ . The distance between  $\phi(\tau)$  and  $\tau'$  is uniformly bounded; the bound depends only on  $\phi$ . Let  $\pi_\tau$  be the horospherical projection onto  $\tau'$ . Then, composing  $\phi|_\tau$  with  $\pi_\tau$  gives a new quasi-isometry which takes  $\tau$  to  $\tau'$ . Doing this for every such  $\tau$ , we obtain a quasi-isometry  $\bar{\phi}$  of  $\Omega$  which is equivalent to  $\phi$ . In other words, every commensurator can be modified canonically so that it acts as a self map of  $\Omega$ .



**Lemma 7.3.** *Suppose that  $f$  is an affine map over  $\mathbf{Q}(\sqrt{d})$ . Then there is a modified commensurator of  $\phi : \Omega \rightarrow \Omega$  such that*

1.  $\phi(\sigma) = \sigma$ ,
2.  $X\phi = f$ ,
3.  $Y\phi = \bar{f}$ .

*Proof.* Suppose  $f(x) = \mu x + \beta$ . Consider the diagonal matrix

$$M = \begin{bmatrix} \mu & \beta \\ 0 & 1 \end{bmatrix}.$$

Then  $M \times \bar{M}$  is a commensurator, and has the desired effect on  $\sigma$ . q.e.d.

**7.7. Galois conjugation** In this section, we will set  $(e, E, E') = (e_1, E_1, E'_1)$ . The other case, that of  $j = 2$ , is treated similarly. Let  $I : \Omega \rightarrow E$  and  $I' : \Omega \rightarrow E'$  be isometries which preserve  $\sigma$ . Consider the quasi-isometry  $f : \Omega \rightarrow \Omega$  defined by the equation

$$f = (I')^{-1} \circ e \circ I.$$

Then  $f(\sigma) = \sigma$ , and both  $Xf$  and  $Yf$  are affine maps of  $\mathbf{R}$ . (It is not necessarily true that  $Pf(0, 0, 0) = (0, 0, 0)$ .)

**Lemma 7.4.**  *$Xf$  and  $Yf$  are affine maps defined over  $\mathbf{Q}(\sqrt{d})$ .*

*Proof.* Let  $L = L_d$  be the lattice  $\mathcal{O}_d \times \bar{\mathcal{O}}_d$ . The shadows on  $\sigma$ , of the horospheres of  $\Omega$  which are closest to  $\sigma$ , are vertical lines which intersect  $\mathbf{R}^2 \times \{0\}$  exactly in  $L$ .

Since  $f$  is a quasi-isometry which preserves peripheral horospheres of  $\Omega$ , we know that  $f$  takes these "closest horospheres" to other horospheres of  $\Omega$ , none of which are very far from  $\sigma$ . The shadows of such horospheres are vertical lines on  $\sigma$  which intersect  $\mathbf{R}^2 \times \{0\}$  in a subset contained in some lattice  $\alpha L$ , for some  $\alpha \in \mathbf{Q}(\sqrt{d})$ .

This states that  $Xe \times Ye$  takes  $L$  into (but not necessarily onto) the lattice  $\alpha L$ . This clearly implies the lemma. q.e.d.

From Lemma 7.3 and Lemma 7.4, we can find a modified commensurator  $\phi : \Omega \rightarrow \Omega$  such that

1.  $\phi(\sigma) = \sigma$ ,
2.  $Y\phi = Yf$ ,

3.  $X\phi = \overline{Y}f.$

Consider the quasi-isometry  $g = \phi \circ f^{-1}$ . Then  $g : \Omega \rightarrow \Omega$  is a quasi-isometry which preserves  $\sigma$ , and has the following properties:

1.  $Xg = \overline{Y}\phi \circ Xf^{-1} = \overline{Y}\phi \circ (Xf)^{-1}.$
2.  $Yg$  is the identity.

Suppose that  $l \subset \sigma$  is any shadow. Then  $Pg(l)$  is also a shadow. However,  $l$  and  $Pg(l)$  lie in the same leaf of one of the hyperbolic plane foliations of  $\sigma$ . This is only possible if  $l$  and  $Pg(l)$  coincide. This means that the  $x$  coordinate of  $l$  is a fixed point of  $Xg$ . However, choosing a different shadow produces a second fixed point of  $Xg$ . Since  $Xg$  is affine, the two distinct fixed points imply that it is the identity. This is to say that  $\overline{Y}\phi = Xf$ . Therefore  $Xf$  and  $Yf$  are conjugate maps.

Hence,  $Xe_1$  and  $Ye_1$  are conjugate (linear) maps. By switching indices,  $Xe_2$  and  $Ye_2$  are conjugate (linear) maps. Since  $Ye_1 = Ye_2$ , we have  $Xe_1 = Xe_2$ . Thus,  $Xa'(x_1) = Xa'(x_2)$ , as desired.

This result, together with the outline given above, completes the proof of the Rigidity Lemma.

### 8. Main Theorem and corollaries

In this chapter, we will prove the Main Theorem and Corollary 1. The other corollaries follow rather easily from these.

#### 8.1. Proof of the Main Theorem

Let  $\Omega$  be a canonical model for a Hilbert modular group. Let  $q : \Omega \rightarrow \Omega$  be a  $K$ -quasi-isometry. Let  $\sigma$  be a horosphere of  $\Omega$ . Since the modified commensurators of  $\Omega$  act transitively on cusps, we can find compose  $q$  by a modified commensurator so that  $q(\sigma) = \sigma$ .

The isometry  $\Delta : X \rightarrow X$  defined by

$$\Delta(x_1, y_1, x_2, y_2) = (x_2, y_2, x_1, y_1)$$

has the effect of switching the two factors of  $X$ .  $\Delta$  belongs to the commensurator of  $SL_2(\mathcal{O}_d)$ , and preserves  $\sigma$ . Composing with  $\Delta$  if necessary, we can assume that  $q$  is factor preserving on  $\sigma$ .

From the Rigidity Lemma, we see that  $Xq$  and  $Yq$  are conjugate affine maps defined over  $\mathbb{Q}(\sqrt{d})$ . From Lemma 7.3, we can compose by a modified commensurator so that  $Xq$  and  $Yq$  are the identity maps. In other words,  $Pq$  is the *identity*. This implies, in particular, that  $q$

preserves (one by one) all horospheres of  $\Omega$ . This easily implies that  $q$  is equivalent to the identity quasi-isometry of  $\Omega$  (after composing with a commensurator). Furthermore, the equivalence is uniform, and only depends on the quasi-isometry constant of  $q$ .

### 8.2. Proof of Corollary 1

Let  $f : \Gamma \rightarrow \Omega$  be a  $K$ -quasi-isometry of a finitely generated group  $\Gamma$  with a neutered space  $\Omega$  associated to an irreducible, nonuniform lattice  $\Lambda$  in  $G = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{R})$ . By the Main Theorem, in order to show that  $\Gamma$  is (a finite extension of a lattice) commensurable with  $\Lambda$ , it suffices to find a representation  $\rho : \Gamma \rightarrow G$  with finite kernel and  $\rho(\Gamma)$  a (nonuniform) lattice in  $G$ .

The group  $\Gamma$  acts on itself on the left by isometries (in the word metric). We may conjugate this action via  $f$  to obtain a uniform family of quasi-isometries of  $\Omega$  (N.B.: the action is NOT conjugated to an action, but to a “quasi-action”, i.e., a (uniformly) bounded distance from an action). By the Main Theorem, each of these quasi-isometries is a uniformly bounded distance from an isometry. Call the composition of the conjugation and the bounded alteration

$$\rho : \Gamma \rightarrow Isom(X).$$

Since any isometry of  $X$  which is a bounded distance from the identity isometry is itself the identity, it follows that  $\rho$  is in fact a homomorphism. We need only show that  $\rho(\Gamma)$  is a lattice and has finite kernel.

Firstly,  $\rho(\Gamma)$  acts cocompactly on  $\Omega$  since  $f(\Gamma)$  is a net in  $\Omega$ . Thus  $\rho(\Gamma)$  acts on  $X = \mathbf{H}^2 \times \mathbf{H}^2$  with cofinite volume.

Pick a basepoint  $x \in \Omega$ . Then, since  $\rho(g)$  was defined via the conjugated action of  $\Gamma$ , and  $f$  is a quasi-isometry, and by the uniformity of everything involved, there is a constant  $C' > 0$  so that  $d_X(x, \rho(g) \cdot x) \leq C'$  for only finitely many  $g \in \Gamma$ ; in particular this distance is 0 for only finitely many  $g \in \Gamma$ . But if  $d_X(x, \rho(g) \cdot x) > 0$ , then the isometry  $\rho(g)$  is not the identity isometry.

From this it follows that  $\rho$  has finite kernel and  $\rho(\Gamma)$  is discrete.

### References

- [1] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric Group Theory, Vol. 2, (G. Niblo and M. Roller, eds.), London Math. Soc. Ser. Lecture Notes No. 182, 1993.

- [2] G. Mostow, *Strong rigidity of locally symmetric spaces*, Ann. of Math. Stud. No. 78, Princeton Univ. Press, 1973.
- [3] G. Prasad, *Strong rigidity of  $\mathbb{Q}$ -rank one lattices*, Invent. Math. **21** (1973) 255-286.
- [4] R. Schwartz, *The quasi-isometry classification of rank-one lattices*, Publ. Math. Inst. Hautes Études Sci., to appear.
- [5] ———, in preparation.
- [6] A. Selberg, *Recent developments in the theory of discontinuous groups of motions of symmetric spaces*, Lecture Notes in Math. Vol. 118, Springer, Berlin, 1970.
- [7] G. van der Geer, *Hilbert modular surfaces*, Ergeb. Math. Grenzgeb. (3) Vol. 16, Springer, Berlin, 1987.

UNIVERSITY OF CHICAGO