Durham E-Theses

Short Geodesics in Hyperbolic Manifolds

Thomson, Scott Andrew

How to cite:

Thomson, Scott Andrew (2012) *Short Geodesics in Hyperbolic Manifolds*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3604/

Short Geodesics in Hyperbolic Manifolds

Scott Andrew Thomson

(College of St Hild and St Bede, Durham)

February 2012

Durham University

Scott Andrew Thomson

Short Geodesics in Hyperbolic Manifolds

Abstract

Given a closed Riemannian n-manifold M, its shortest closed geodesic is called its *systole* and the length of this geodesic is denoted $\mathrm{syst}_1(M)$. For any $\varepsilon > 0$ and any $n \geqslant 2$ one may construct a closed hyperbolic n-manifold M with $\mathrm{syst}_1(M) \leqslant \varepsilon$. Constructions are detailed herein. The volume of M is bounded from below, by $A_n/\mathrm{syst}_1(M)^{n-2}$ where A_n is a positive constant depending only on n. There also exist sequences of n-manifolds M_i with $\mathrm{syst}_1(M_i) \to 0$ as $i \to \infty$, such that $\mathrm{vol}(M_i)$ may be bounded above by a polynomial in $1/\mathrm{syst}_1(M_i)$. When ε is sufficiently small, the manifold M is non-arithmetic, so that its fundamental group is an example of a non-arithmetic lattice in $\mathrm{PO}(n, 1)$. The lattices arising from this construction are also exhibited as examples of non-coherent groups in $\mathrm{PO}(n, 1)$.

Also presented herein is an overview of existing results in this vein, alongside the prerequisite theory for the constructions given.

TABLE OF CONTENTS

Copyright Notice

The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged.

Opening Remarks

Overview of results

In this thesis, I present some results concerning hyperbolic n-manifolds and their systoles; that is, their shortest closed geodesics (cf. §6.1). The main results are: that one may construct closed hyperbolic n-manifolds with systole lengths as short as desired, and that one may bound the volumes of these manifolds from below by a quantity that tends to infinity as the systole length tends to zero (and the behaviour of this is described quantitatively). An interesting by-product of the construction given here is that the manifolds thus obtained are non-arithmetic when the systole length is sufficiently short, and so one obtains new examples of non-arithmetic lattices in $\mathrm{PO}(n, 1)$. The non-arithmetic lattices so obtained may be used to exhibit new examples of non-coherent non-arithmetic lattices in $\mathrm{PO}(n, 1)$.

Existence of short systole n-manifolds

In 2006, I. Agol showed that for any $\varepsilon > 0$, there exists a closed hyperbolic 4-manifold whose systole $\mathrm{syst}_1(M)$ is at most ε. This had been known to hold in dimensions 2 and 3, as consequences of Teichmüller Theory and Thurston's Dehn Surgery Theorem respectively (cf. §6.2). Agol's proof is by construction, and uses a certain separability property (the 'GFERF' property) of a particular lattice in $\mathrm{Isom}(\mathbb{H}^4)$. (See §5.3 and §6.4 for details.) At that time it was not

clear that the GFERF property was held by higher dimensional lattices and so his proof was restricted to the case of dimension 4.

Theorem 6.1 is a generalisation of Agol's 4-dimensional result to every dimension $n \geqslant 2$, and its proof is a construction based on some of his ideas. However, the proof presented here does not use the GFERF property of the lattices in question; instead Theorem 5.1 is used, which states that one may always find a finite cover of a compact hyperbolic manifold, that admits immersions of given rational hypersurfaces in such a way that they do not intersect. A less general version of this theorem was given by Margulis and Vinberg and the proof is a generalisation of theirs.

Since 2006 it has emerged that the GFERF property of arithmetic lattices in $PO(n, 1)$ does indeed hold, so that Agol's proof may be directly generalised. Details of this are given in §6.4.

Volumes of short systole n-manifolds

It follows from a theorem of H.-C. Wang that the volumes of the manifolds in Agol's construction must tend to infinity as the systole length tends to zero (cf. §7.3), but no bounds have previously been given that explicitly describe the behaviour of the volume growth in terms of the systole when the systole is small. This is addressed in Theorem 7.4, where it is shown that the n-manifolds constructed in the proof of Theorem 6.1 have volumes bounded below by $A_n / \operatorname{syst}_1(M)^{n-2}$ (for $A_n > 0$ depending on n). This result is a consequence of recent work by M. Bridgeman and J. Kahn that gives volumes of hyperbolic manifolds with boundaries in terms of their orthospectrum (cf. §7.2).

One may exhibit examples of sequences of manifolds from the proof of Theorem 6.1 whose volumes grow no faster than a polynomial in $1/\operatorname{syst}_1(M)$, so as to establish the bound of Theorem 7.4 as an optimal one. (See the proof of Proposition 7.5 and the discussion following it for details.)

The inequalities of the type in Theorem 7.4 are similar in spirit to those by Gromov and Reznikov (cf. Theorem 7.1 and Theorem 7.6) and some discussion of their results is included (cf. §7.1 and §7.3).

NON-ARITHMETICITY OF SHORT SYSTOLE n-MANIFOLDS

It has been known since 1987 that there exist, for any $n \geqslant 2$, non-arithmetic lattices in $\mathrm{PO}(n,1)$. The proof was given by Gromov and Piatetski-Shapiro (cf. §8.1). The discussion in §8.2 concludes that if $\varepsilon > 0$ is small enough then any closed hyperbolic n-manifold from Theorem 6.1 with systole smaller than ε is non-arithmetic. For ε to be small enough one requires that it is less than some $\varepsilon_{n,d}$ which depends on both n and the $\mathbb{Q}$-degree d of the field K over which one works in the proof of Theorem 6.1. This proof is therefore a new construction of non-arithmetic lattices in $\mathrm{PO}(n,1)$. The conjecture that $\varepsilon_{n,d}$ is in fact independent of n and K is discussed in §8.2.

SUMMARY OF CHAPTERS

Before presenting the main results and their proofs, I give four chapters of background material.

Chapter 1 deals with the necessary algebraic number theory, some theory of quadratic forms, algebraic groups and orthogonal groups. Chapter 2 introduces hyperbolic space, its isometry groups, and some standard results concerning hyperplanes.

The theory of discrete groups of isometries of hyperbolic space is introduced in Chapter 3, with some general results concerning Lie groups, including the theorems of Každan and Margulis. In Chapter 4, the arithmeticity property for discrete groups is introduced, and the major theorems of Borel, Harish-Chandra and Margulis are presented. Examples of arithmetic lattices are given.

The main result of Chapter 5 (namely Theorem 5.1) is concerned with embedding quotients of hyperplanes in covers of arithmetic hyperbolic manifolds. The remainder of the chapter introduces separability properties of groups, including the GFERF property.

In Chapter 6 is detailed the construction of short systole manifolds, consti-

tuting the proof of Theorem 6.1. Some explanations of the previously known low-dimensional cases are given, including Agol's 4-dimensional construction. A more direct generalisation of his construction (concerning the GFERF property) is also outlined.

The volume growth results (Theorem 7.4 and Proposition 7.5) are given in Chapter 7. By way of context, Gromov's systolic inequality is introduced, along with the work of Bridgeman and Kahn. Reznikov's inequality is examined and compared with Theorem 7.4.

Chapter 8 deals with the non-arithmeticity of manifolds with short systole, as well as their non-coherence. The original construction of Gromov and Piatetski-Shapiro is presented, along with some auxilliary results on group presentations.

I give as many references to the literature as possible, especially in the introductory material, in the hope that this will increase the accessibility of the text to novices and non-experts.

Personal Remarks

This thesis concerns work undertaken under the supervision of Dr Mikhail (Misha) Belolpetsky, between October 2008 and August 2011, and I am especially grateful to Misha for his generosity of time, guidance and advice during that time. This thesis seems to me to have only minor consequence in comparison with some of his works and I have been privileged indeed to have had the benefit of a collaboration with such a leading mathematician. One outcome of the time working with Misha has been a joint article in 2011 [BT11], which is referred to in the text.

I am also grateful to Dr Norbert Peyerimhoff, Prof. John Parker and Dr John Bolton for advice of both a general and mathematical nature, and to Dr Jens Funke, for his guidance, both informal and in the formal capacity as supervisor following Misha's departure from Durham. My gratitude is also extended to

Dr James Anderson from the University of Southampton, who most kindly agreed to act as the external examiner for this thesis. In addition, I have had much to be thankful for in the advice and experience of Dr Vernon Armitage.

The encouragement one may take from one's postgraduate colleagues is indispensable: John Rhodes, Luke Stanbra, Ben Lambert, John Mcleod, Joey Oliver, Amani Saloom, Rachel Oxlade, Ramon Vera and Keith McCabe all have my thanks.

I am indebted to the Willmore Trust and to Dr Gillian Boughton of St Mary's College, Durham, for some financial support during my time here. I hope that the Trust will continue its work for some time to come.

In Durham I have been fortunate to have membership of the College of St Hild and St Bede: a place of residence (moreover, a home) for more than three years, and a community of a sort probably unknown elsewhere, even perhaps amongst the Durham colleges. To begin to list the names of those Hild-Bede members — student, staff or SCR — to whom I shall ever be grateful would be to unwillingly choose in a dichotomy of accidental omission and excessive length: needless to say I am indebted to very many indeed.

SCOTT THOMSON

Durham, February 2012

1

Quadratic Forms

§1.1 Algebraic number theory

Algebraic number fields

A summary of the number-theoretic facts given below can be found in Chapter 0 of the book of Maclachlan and Reid [MR03], and there they are stated in their most useful form for us. Most of what follows can also be found in (or at least easily deduced from) standard references on algebraic number theory (e.g., Lang [Lan70] or Neukirch [Neu99]). Some references to Lang's 'Algebra' [Lan02] are given, as this work more comprehensively covers the most foundational material.

Suppose α is a root of a polynomial equation with coefficients in $\mathbb{Z}$; i.e.,

$$a_m \alpha^m + a_{m-1} \alpha^{m-1} + \cdots + a_1 \alpha + a_0 = 0 \qquad \text{where } a_i \in \mathbb{Z} \text{ for each } i.$$

Such an α is said to be *algebraic*, and if the polynomial is monic (i.e., $a_m = 1$), then we call α an *algebraic integer*. We will have occasion to refer to the set of all algebraic integers, which will be denoted by $\mathbf{A}$. If α is algebraic then it satisfies a unique monic polynomial equation of minimal degree with coefficients in $\mathbb{Q}$, called the *minimal polynomial* of α [Lan02, p. 224].

Now supposing $\{\alpha_i\}_{i=1}^k$ to be any finite set of complex numbers, we denote by $\mathbb{Q}(\alpha_1, \ldots, \alpha_k)$ the smallest subfield of $\mathbb{C}$ containing both $\mathbb{Q}$ and $\{\alpha_i\}_i$. The field $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_k)$ can be viewed as a vector space over $\mathbb{Q}$, and if it has finite dimension d then K is said to be a *finite extension (of $\mathbb{Q}$)* and d is called the *degree* of the extension, denoted in general by $[K : \mathbb{Q}]$. If the α_i are all algebraic then the extension K will be finite, [Lan02, p. 227, Prop. 1.6] and in this case we call K an *algebraic number field*. In the case where $m = 1$ (i.e., extensions $\mathbb{Q}(\alpha)$ for algebraic α), then the degree $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is equal to the degree of the minimal polynomial of α [Lan02, p. 225, Prop. 1.4], and $\mathbb{Q}(\alpha)$ is said to be a *simple extension*.

Given an algebraic number field K, consider the set of all field automorphisms $\sigma \colon K \to K$ such that $\sigma|_{\mathbb{Q}} = \mathbf{id}$. This set forms a group, denoted $\mathrm{Gal}(K/\mathbb{Q})$, and called the *Galois group* of K over $\mathbb{Q}$ [Lan02, p. 262]. It can be shown that the order of $\mathrm{Gal}(K/\mathbb{Q})$ is equal to the degree $[K : \mathbb{Q}]$ [Lan02, p. 254, Theorem 1.8]. The non-trivial elements of $\mathrm{Gal}(K/\mathbb{Q})$ can be realised as field isomorphisms $\mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha_j)$ where the α_j are the $d-1$ other roots of the minimal polynomial, given by $\alpha \mapsto \alpha_j$ [MR03, p. 2]. We will have occasion to write x^σ in place of $\sigma(x)$ (and in particular in the proof of Theorem 5.1).

We may also regard the elements of the Galois group as embeddings of K in $\mathbb{C}$ (which is desirable from a formal point of view since K is *a priori* defined as an algebraic object); and, taking this stance, if for every embedding $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ we have $\sigma(K) \subseteq \mathbb{R}$ then we say that the extension K is *totally real*.

The subset $K \cap \mathbf{A} \subseteq K$ is called the *ring of integers* of K (and one can show that it is indeed a ring) [Lan70, p. 5] [Lan02, p. 336, Prop. 1.4]. It is denoted here by $\mathcal{O}_K$, and the set of non-zero elements of $\mathcal{O}_K$ will be denoted by $\mathcal{O}_K^*$. If $x \in \mathcal{O}_K$ then we may say that x is *K-integral* (and any $y \in K$ may be called *K-rational*).

EXAMPLE 1.1.

1. The field $\mathbb{Q}$ is trivially a degree-1 extension over itself. In this case, the ring of integers is $\mathbb{Z}$, and there is but one Galois conjugate, namely the

identity map.

2. The polynomial $x^2 - 2$ has two roots $+\sqrt{2}$ and $-\sqrt{2}$, so the field $\mathbb{Q}(\sqrt{2})$ is an algebraic number field of degree 2. The Galois conjugate different from the identity is the map $\sigma\colon a + b\sqrt{2} \mapsto a - b\sqrt{2}$. The ring of integers $\mathcal{O}_K$ is $\{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$. This is a totally real extension.

   In general we may consider $\mathbb{Q}(\sqrt{d})$ where d is a square-free positive integer (i.e., d has no squares as factors). This is a totally real degree-2 extension, but the ring of integers takes a form which depends on d:

   $$\mathcal{O}_{\mathbb{Q}(\sqrt{d})} = \begin{cases} \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\} & \text{if } d \not\equiv 1 \pmod 4 \\ \{\frac{1}{2}(2a + 1) + \frac{1}{2}b\sqrt{d} \mid a, b \in \mathbb{Z}\} & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

   Considering the ring of integers as a $\mathbb{Z}$-module, we see that a basis in the second case is $\{1, \frac{1}{2}(1 + \sqrt{d})\}$ [MR03, Examples 0.2.7], whilst the first case gives rise to the simpler $\mathbb{Z}$-basis $\{1, \sqrt{d}\}$.

3. Consider the polynomial $x^3 - 2$. This has the roots $\sqrt[3]{2}e^{2ik\pi/3}$ for $k \in \{0, 1, 2\}$ (where $\sqrt[3]{2}$ is understood to be real). The image in $\mathbb{C}$ of the field $\mathbb{Q}(\sqrt[3]{2})$ (under the obvious identity embedding) is real, but there is an isomorphism $\sigma\colon \mathbb{Q}(\sqrt[3]{2}) \to \mathbb{Q}(\sqrt[3]{2}e^{2i\pi/3})$ whose image is not contained in $\mathbb{R}$. Thus this extension is not totally real, even though the field $\mathbb{Q}(\sqrt[3]{2})$ may be regarded as contained in $\mathbb{R}$.

IDEALS IN NUMBER FIELDS

A subset $\mathfrak{a} \subset \mathcal{O}$ of a ring $\mathcal{O}$ is called an *ideal* if for every $x \in \mathcal{O}$, we have $x\mathfrak{a} \subseteq \mathfrak{a}$. A *prime ideal* is an ideal $\mathfrak{p}$ such that if $ab \in \mathfrak{p}$ then $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. If $\mathfrak{a}$ is a *maximal ideal* in $\mathcal{O}$, that is, no proper ideal of $\mathcal{O}$ contains $\mathfrak{a}$ (and, by convention, $\mathfrak{a} \neq \mathcal{O}$), then the quotient ring $\mathcal{O}/\mathfrak{a}$ is a field [Lan02, p. 93]. It turns out that the ring of integers $\mathcal{O}_K$ of an algebraic number field K has the property that every prime ideal is maximal [Neu99, p. 17, Theorem 3.1], which is a consequence of its being a so-called *Dedekind domain* or Dedekind ring [Lan70, p. 20] [Neu99, p.18] [Lan02, pp. 88, 116, 353]. We also find that for an algebraic number field K every non-zero ideal $\mathfrak{a} \subseteq \mathcal{O}_K$ gives a finite

quotient $\mathcal{O}_K/\mathfrak{a}$ [MR03, p. 12, Theorem 0.3.2] [Lan70, p. 17]. An ideal $\mathfrak{a} \subseteq \mathcal{O}$ is *principal* if it can be generated by a single element $a \in \mathcal{O}$; that is $\mathfrak{a} = (a)$ where the notation (a) denotes the set $\{ax \mid x \in \mathcal{O}\}$.

Norm and trace

If K is an algebraic number field, and $x \in K$ then we define the *norm* and *trace* of x to be, respectively, the numbers $N(x)$ and $\mathrm{tr}(x)$ given by

$$N(x) = \prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} \sigma(x) \qquad \text{and} \qquad \mathrm{tr}(x) = \sum_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} \sigma(x). \qquad (1.1)$$

The norm enjoys the following properties:

Lemma 1.2.

1. $N(x) \in \mathbb{Q}$ *for every* $x \in K$;

2. $N(xy) = N(x)N(y)$ *for every* $x, y \in K$;

3. $N(x) = x^{[K:\mathbb{Q}]}$ *if* $x \in \mathbb{Q}$; *and*

4. $N(x) \in \mathbb{Z}$ *if* $x \in \mathcal{O}_K$.

Note that the second and third statements above are immediate from the definitions of $\mathrm{Gal}(K/\mathbb{Q})$. The first follows from the fact that the norm must be invariant under $\mathrm{Gal}(K/\mathbb{Q})$ [MR03, p. 3]. The proof of the last assertion can be found in, e.g., the book of Lang [Lan02, p. 337, Cor. 1.6]

We will also need the following elementary result (cf. Belolipetsky and Thomson [BT11, p. 1462]), which will be used in the proof of Theorem 5.1:

Lemma 1.3. *Let K be a totally real algebraic number field.*

1. *Suppose that $x \in \mathcal{O}_K$ and that $\sqrt{x} \in K$. Then $\sqrt{x} \in \mathcal{O}_K$.*

2. *If $x, y \in K$ and $\sqrt{x/y} \in K$ then $\sqrt{xy} \in K$.*

Proof.   1. If x is an algebraic integer then it satisfies $x^m + a_{m-1}x^{m-1} + \cdots + a_0 = 0$ for some $a_i \in \mathbb{Z}$ and some $m \in \mathbb{N}$. It is then immediate that $\sqrt{x}$ satisfies $(\sqrt{x})^{2m} + a_{m-1}(\sqrt{x})^{2(m-1)} + \cdots + a_0 = 0$ and so is an algebraic integer. Since $\sqrt{x} \in K$ we have $\sqrt{x} \in K \cap \mathbf{A} = \mathcal{O}_K$.

2. We simply write $\sqrt{xy} = y\sqrt{x/y}$, which clearly lies in K.   □

The norm can also be defined for an ideal of a number field K. If $\mathfrak{a}$ is a non-zero ideal of $\mathcal{O}_K$ then the *norm* $N(\mathfrak{a})$ is defined by $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$, which is finite as noted above. This function is multiplicative and if $\mathfrak{a} = (a)$ is principal then $N(\mathfrak{a}) = N(a)$ [MR03, Theorem 0.3.15].

§1.2   Real Quadratic Forms

A fairly general treatment of quadratic forms is given by O'Meara [O'M71], but here we will give a slightly more specific presentation in order to make clear the results that will be used later. Cassels [Cas08] gives a general treatment for rational quadratic forms. For the background theory on bilinear forms the reader may wish to consult a standard text on linear algebra (e.g., Roman [Rom05]).

An n-ary *quadratic form* is a homogeneous polynomial over $\mathbb{R}$, of degree 2 in n variables X_i; i.e.,

$$f(X) = \sum_{i,j=1}^{n} a_{ij} X_i X_j, \qquad \text{where } a_{ij} \in \mathbb{R}. \qquad (1.2)$$

Given a real vector space V of dimension $m \geqslant n$, and a basis B for V, the form f defines a function $f_B \colon V \to \mathbb{R}$ by the rule $f_B(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ where the x_i are the coordinates of x with respect to B. Different quadratic forms may actually define the same function on a vector space. If f and g are two forms, then there may be bases B and C for V such that $f_B = g_C$, and in this case we say that the two forms are *equivalent*. We may also, given a form f and a basis B, find a form g and a basis C for V such that $g_C = f_B$ and g has the simple 'diagonal' form $g(X) = \sum_{i=1}^{n} a_i X_i^2$ [Rom05, p. 269].

Provided no confusion can arise, we will generally assume that a form f and the function f_B it defines on a vector space are one and the same, with the basis and vector space being implicit.

Associated with a quadratic form $f\colon V \to \mathbb{R}$ is a *bilinear form* $(\cdot,\cdot)_f\colon V \times V \to \mathbb{R}$ given by

$$(x,y)_f = \tfrac{1}{2}\big(f(x+y) - f(x) - f(y)\big). \tag{1.3}$$

We have $f(x) = (x,x)_f$, for

$$(x,x)_f = \tfrac{1}{2}\big(f(2x) - 2f(x)\big) = \tfrac{1}{2}\big(4f(x) - 2f(x)\big) = f(x).$$

Here we use the fact that f is defined by a degree-2 homogeneous polynomial; i.e., that $f(\lambda x) = \lambda^2 f(x)$ for any real λ. The form $(\cdot,\cdot)_f$ is *symmetric*; that is, $(x,y) = (y,x)$ for every $x, y \in V$. We say that the form $(\cdot,\cdot)_f$ is *non-degenerate* (on V) if, whenever $(x,y) = 0$ for all $x \in V$, we must have $y = 0 \in V$.

For a given basis $\{e_i\}_i$ one can compute the matrix F associated to $(\cdot,\cdot)_f$: by definition the entries of F are

$$F_{ij} = (e_i, e_j). \tag{1.4}$$

This matrix is symmetric, owing to the symmetry of the form $(\cdot,\cdot)$, and indeed any symmetric matrix can be used to define a bilinear form by assuming (1.4) to hold for a set of basis vectors for V, and extending linearly. The form $f(X) = \sum_{i,j=1}^n F_{ij} X_i X_j$ gives rise to the function f which is the same as the function defined by $f(x) = (x,x)$. Thus, the notions of quadratic form, bilinear form, and symmetric matrix are essentially equivalent. In considering the orthogonal groups of quadratic forms (cf. §1.5), it is often desirable to take the matrix point of view. Regarding $x, y \in V$ as column vectors, we have

$$(x,y) = x^{\mathsf{T}} F y, \tag{1.5}$$

where F is as in (1.4).

If V is an n-dimensional $\mathbb{R}$-vector space on which is defined a bilinear form $(\cdot,\cdot)$, then V is sometimes called a *quadratic space*. Any vector $x \in V$ such that $f(x) > 0$ is called *positive* (with respect to f); and similarly any $x \in V$ with $f(x) < 0$ is called *negative*. In determining co-compactness of arithmetic groups it is often necessary to consider whether or not the quadratic form $f(x) = (x,x)$ takes the value zero for any $x \neq 0$, and, if indeed there is

$x \in V \setminus \{0\}$ such that $f(x) = 0$ then we say that f (or V) is *isotropic*. (The vector x may be called an isotropic vector.) If $f(x) > 0$ for all $x \neq 0$ then f (or V) is called *positive definite*. This is equivalent to f having associated symmetric matrix whose eigenvalues are all positive. If f has some negative eigenvalues, say k of them, then we say that f has *signature* $(n - k, k)$. That this is independent of the choice of basis is justified by the following theorem:

THEOREM 1.4. [Rom05, Theorem 11.26] [Cas08, p. 27] *Let f be a quadratic form on a real vector space V, and suppose that the associated bilinear form is non-degenerate. Then f is equivalent to the form $\sum_{i=1}^{r} x_i^2 - \sum_{i=n-r}^{n} x_i^2$.*

We will be mainly concerned with quadratic forms that either have signature $(n, 1)$ or are positive definite.

CONVENTIONS

In most situations we will be considering quadratic forms from the point of view of functions on $\mathbb{R}^n$ or K^n where K is a number field. These will always be specified as homogeneous polynomials of degree 2 in the variables $x_1, \ldots, x_n$, where the variables are understood to correspond to the standard basis in the obvious way:

$$x_1 \leftrightarrow (1, 0, \ldots, 0)$$
$$x_2 \leftrightarrow (0, 1, \ldots, 0)$$
$$\vdots$$
$$x_n \leftrightarrow (0, 0, \ldots, 1).$$

Therefore, given a vector space, specifying a form f as a polynomial will be sufficient to determine the quadratic space in question.

§1.3 Quadratic Forms over Number Fields

Thus far, the quadratic forms under consideration have had coefficients in $\mathbb{R}$, but it will be important to utilise forms whose coefficients are restricted to lying in a given algebraic number field K. Supposing f to be such a form, and that K is minimal with respect to this property, we say that f is *defined over* K. The minimality requirement in this definition is necessary, as the following example shows: The form $x_1^2 + x_2^2$ is defined over $\mathbb{Q}$, but could also be viewed as defined over $\mathbb{Q}(\sqrt{2})$ (were the minimality requirement to be dropped), as indeed the coefficients of f do lie in $\mathbb{Q}(\sqrt{2})$. With this definition we find that we are in a position to differentiate between classes of quadratic forms based on their fields of definition. (cf. Example 1.5.)

As in §1.1, the extension $K/\mathbb{Q}$ has a Galois group $\mathrm{Gal}(K/\mathbb{Q})$. For every $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ we define the *conjugate form* f^σ by $f^\sigma = \sum_{i,j} a_{ij}^\sigma X_i X_j$ where f is defined using the notation of (1.2). It may be that the signature of f^σ is different from that of f when $\sigma \neq \mathbf{id}$. This is the case, for example, with the form $f = x_1^2 - \sqrt{2}x_2^2$. The field of definition for f is $\mathbb{Q}(\sqrt{2})$, and the (non-trivial) conjugate form of f is $f^\sigma = x_1^2 + \sqrt{2}x_2^2$.

Remark. It will later be convenient to compute with the Galois conjugates of field elements obtained as images under bilinear forms; that is, elements of the form $\sigma\big((x,y)_f\big)$. In keeping with the notation on p. 2 we will write $(x,y)_f^\sigma$ for such an element, and it should be noted that this is in general different from the element $(x,y)_{f^\sigma}$. Explicitly, we have $(x,y)_f^\sigma = (x^\sigma, y^\sigma)_{f^\sigma}$.

Equivalence of forms over number fields

Consider two quadratic forms $f = \sum_{i,j} a_{ij} X_i X_j$ and $g = \sum_{i,j} b_{ij} X_i X_j$ in n variables. Choosing a basis of some real vector space V allows one to apply Theorem 1.4 and obtain indices for these forms (provided that they are non-degenerate on V). If the indices are the same then the forms f and g are equivalent. This does not in general follow in the case of a k-vector space (for a number field k), since it may not be possible to find a k-basis of the space

such that a given form can be written as $\sum_{i=1}^{n} (-1)^{\delta_i} x_i^2$.

EXAMPLE 1.5. Let f and g be the two forms

$$f = 3X_1^2 + X_2^2 \qquad \text{and} \qquad g = X_1^2 + X_2^2.$$

Assigning X_1 and X_2 to the standard basis vectors for $\mathbb{R}^2$ we have the functions on $\mathbb{R}^2$ given by

$$f(a,b) = 3a^2 + b^2 \qquad \text{and} \qquad g(a,b) = a^2 + b^2. \tag{1.6}$$

But, in the basis $\{(\sqrt{3},0), (0,1)\}$ of $\mathbb{R}^2$, the form f represents the same function as g, for

$$f(a,b) = 3\left(\tfrac{1}{\sqrt{3}}a\right)^2 + b^2 = a^2 + b^2,$$

$$\text{whilst} \qquad g(a,b) = a^2 + b^2.$$

Thus over $\mathbb{R}$ the two forms can be seen explicitly to be equivalent.

Now consider the space $\mathbb{Q}^2$. In the standard basis the forms f and g define functions $\mathbb{Q}^2 \to \mathbb{Q}$ given by identical expressions to those in (1.6), but this time the set $\{(\sqrt{3},0), (0,1)\}$ is not a $\mathbb{Q}$-basis of $\mathbb{Q}^2$, and so over $\mathbb{Q}$ these forms are not equivalent as the $\mathbb{Q}$-basis required for f to define the same function as g does not exist.

In addition to dealing with equivalence of quadratic forms, we will also have occasion to consider whether or not two quadratic forms are similar. Two forms are similar, roughly speaking, if one is a multiple of the other. More precisely, suppose that f and g are forms defined over a number field K. We say that the forms f and g are *similar* over K if there exists $\lambda \in K$ for which f and λg are equivalent. This notion is used by, for example, Gromov and Piatetski-Shapiro [GPS87, §2.6].

ISOTROPY OF QUADRATIC FORMS

As will be seen later, quadratic forms play an important role in constructing arithmetic hyperbolic manifolds. The question of whether or not a particular manifold is compact translates into one of whether or not an associated quadratic form is isotropic or not.

THEOREM 1.6 (MEYER). [Cas08, p. 75] [Ser73, p. 43] *Suppose that f is an indefinite quadratic form in n variables, where $n \geqslant 5$, defined over $\mathbb{Q}$. Then f is isotropic over $\mathbb{Q}$.*

By an *indefinite form* is meant a form that is isotropic over $\mathbb{R}$; i.e., viewing f as attached to $\mathbb{R}^n$ there is a non-zero vector $v \in \mathbb{R}^n$ such that $f(v) = 0$. To then say that f is isotropic *over* $\mathbb{Q}$ means that there is a vector $v \in \mathbb{Q}^n$ such that $f(v) = 0$.

§1.4   ALGEBRAIC GROUPS

Algebraic groups are usually defined abstractly using the language of affine varieties; however for us it will be sufficient to use a more concrete definition. The book of Platonov and Rapinchuk serves as an excellent reference to the theory of algebraic groups [PR94, Ch. 2], which does use a relatively concrete definition of an algebraic group, whilst the book of Humphreys [Hum75] is a fairly accessible introduction to the more abstract theory.

Fix some $n \in \mathbb{N}$. If M is an $n \times n$ matrix then its n^2 entries are customarily denoted by $x_{11}, x_{12}, \ldots, x_{nn}$, and we may consider polynomials in these n^2 variables, with coefficients in a field $K \subseteq \mathbb{C}$. The collection of such polynomials is usually denoted $K[x_{11}, \ldots, x_{nn}]$ but here we will use $K[x_{ij}]_n$ for brevity. If $p \in K[x_{ij}]_n$, then denote by $\mathrm{Var}(p)$ the set $\{M \in \mathrm{GL}_n(\mathbb{R}) \mid p(M) = 0\}$ where $p(M)$ indicates substitution of the entries of M into the corresponding variables of p. For a finite collection $\mathcal{P} \subset K[x_{ij}]_n$ we write $\mathrm{Var}(\mathcal{P})$ for the set of common zeroes of all the $p \in \mathcal{P}$:

$$\mathrm{Var}(\mathcal{P}) = \bigcap_{p \in \mathcal{P}} \mathrm{Var}(p).$$

On $\mathrm{GL}_n(\mathbb{R})$ we can define a topology, called the *Zariski topology*, as follows: we call a set *elementary* if it is of the form $\mathrm{Var}(\mathcal{P})$ for some finite collection $\mathcal{P}$ of polynomials. The *Zariski-closed* sets are then (by definition) those sets that can be written as a combination of intersections and finite unions of elementary

sets. One can actually show that a Zariski-closed set is itself elementary so that any Zariski-closed set is defined by a finite collection $\mathcal{P}$ of polynomials [Hum75, 1.2].

If $G \subset \mathrm{GL}_n(\mathbb{R})$ is a Zariski-closed subset defined by polynomials p_m ($1 \leqslant m \leqslant \ell$), and K is a field, then we say that G is *defined over* K if K is the smallest field such that $p_m \in K[x_{ij}]_n$ for every m. If G is also a subgroup of $\mathrm{GL}_n(\mathbb{R})$ then we call G an *algebraic group*. Note that $\mathrm{GL}_n(\mathbb{R})$ is itself algebraic (and defined over $\mathbb{Q}$) if it is viewed as a subset of $\mathrm{GL}_{n+1}(\mathbb{R})$ via the embedding

$$\iota \colon M \mapsto \begin{pmatrix} M & 0 \\ 0 & \dfrac{1}{\det(M)} \end{pmatrix}.$$

Thus $\iota\big(\mathrm{GL}_n(\mathbb{R})\big)$ is defined by the polynomials $x_{n+1,i}$ ($i = 1, \ldots, n$), $x_{j,n+1}$ ($j = 1, \ldots, n$), and $\big(x_{n+1,n+1} \cdot \det(x_{kl})_{k,l=1}^n\big) - 1$, and is therefore Zariski-closed. (Note that this embedding in fact realises GL_n as a subset of SL_{n+1}.) A slightly simpler example of an algebraic subgroup of $\mathrm{GL}_n(\mathbb{R})$ is $\mathrm{SL}_n(\mathbb{R})$:

$$\mathrm{SL}_n(\mathbb{R}) = \big\{ M \in \mathrm{GL}_n(\mathbb{R}) \mid \det(M_{ij})_{i,j=1}^n - 1 = 0 \big\}.$$

This group is particularly useful since unit determinants allow integral matrices to have integral inverses.

It will later be useful to consider *Zariski dense* subsets of some algebraic group; that is, subsets (or subgroups) $\Gamma \subset G$ such that the closure of Γ with respect to the Zariski topology is the entirety of G.

RESTRICTION OF SCALARS

We will often be concerned with groups defined over $\mathbb{Q}$; but even when the group is defined over a (non-trivial) finite extension of $\mathbb{Q}$, it can still be shown to embed in a group of higher dimension that *is* defined over $\mathbb{Q}$, in such a way as to preserve certain 'rationality' properties [Zim84, pp. 115-116] [MR03, pp. 316-317]. In what follows this is described in more detail.

Suppose G is an algebraic subgroup of $\mathrm{GL}_n(\mathbb{R})$, defined over an algebraic number field K, having finite degree $d = [K : \mathbb{Q}]$. Then G is the set of common zeroes of some polynomials $p_1, \ldots, p_k \in K[x_{ij}]_n$.

Let $a \in K$. Regarding K as a $\mathbb{Q}$-vector space, we have associated with the number a a linear map $\phi_a \colon K \to K$ defined by $\phi_a(x) = ax$, and the map $\rho \colon a \mapsto \phi_a$ is known as the *left regular representation* of K over $\mathbb{Q}$. In fixing a $\mathbb{Q}$-basis $\{v_i\}_i$ for K, each of the linear maps ϕ_a can be written as a matrix with respect to $\{v_i\}$. The image $\rho(K) \subseteq M_d(\mathbb{Q})$ is defined by (say) ℓ equations $F_\ell(x^{\alpha\beta}) = 0$ in the d^2 variables $x^{\alpha\beta}$ ($\alpha, \beta = 1, \ldots, d$) (and these equations are linear since each entry of ϕ_a is a rational multiple of one of the coordinates of $a \in K$).

We now 'expand' each entry of a matrix $[x_{ij}]$ in $M_n(K)$, by replacing x_{ij} with a matrix $(x_{ij}^{\alpha\beta})_{\alpha\beta}$, and by using ρ we identify $M_{nd}(\mathbb{Q})$ with $M_n(K)$. In this way $M_{nd}(\mathbb{Q})$ can be viewed as partitioned into 'blocks', with the i, j specifying the block and the α, β the entry within a given i, j block.

Now the equations $p_i = 0$ ($i = 1, \ldots, k$) still need to be satisfied if their coefficients (which lie in K) are replaced by their images under ρ (which lie in $M_d(\mathbb{Q})$), and their variables x_{ij} by the blocks $x_{ij}^{\alpha\beta}$. Thus the group of rational points $H(\mathbb{Q})$ defined by these new polynomials $\tilde{p}_i$, subject to the $F_s(x_{ij}^{\alpha\beta})$ (for every block ij), is isomorphic to the group $G(K)$ of K-points of G.

The group H is often denoted by $\mathrm{Res}_{K/\mathbb{Q}}(G)$. The groups of real points of G and H are not in general isomorphic, of course.

There is an alternative view of restriction of scalars that is often more useful in practice. Suppose again that G is defined over an algebraic number field K of degree d, so that it is defined as a subgroup of $\mathrm{GL}_n(\mathbb{C})$ by a collection $\{p_i \mid i \in I\}$ of polynomials with coefficients in K. Now for each $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, denote by G^σ the algebraic group defined by the set of polynomials $\{\sigma(p_i) \mid i \in I\}$. Then we have an embedding

$$G \hookrightarrow \prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} G^\sigma \quad \text{given by} \quad g \mapsto (g, \sigma_2(g), \ldots, \sigma_d(g)). \tag{1.7}$$

Indeed, the product of the G^σ is isomorphic to $\mathrm{Res}_{K/\mathbb{Q}}(G)$ [PR94, (2.4), p. 50].

UNIPOTENCY

An element of an algebraic group G is *unipotent* if all its eigenvalues are equal to 1. In other words, if $g \in G$ is unipotent, then $(g - \mathbf{id})^\ell = 0$ for some $\ell \in \mathbb{N}$.

(As previously remarked, we are viewing the group G as a subgroup of some $\mathrm{GL}_n(\mathbb{C})$.) A subgroup of G is called *unipotent* if all its elements are unipotent. One thinks of unipotent elements (in particular in the case of GL_n) as being conjugate to upper-triangular matrices with all the diagonal entries equal to 1.

For an algebraic group G that is connected (in the Zariski topology), its *unipotent radical* $R_{\mathrm{u}}(G)$ is its maximal connected unipotent normal subgroup [PR94, p. 58] [Hum75, Sect. 19.5]. If $R_{\mathrm{u}}(G)$ is trivial then G is said to be *reductive*. As an algebraic group GL_n is connected and, in fact, reductive. Note also that semisimple groups are reductive (cf. p. 28).

§1.5   Orthogonal Groups

The notion of an orthogonal group associated to a quadratic form is examined by both O'Meara [O'M71] and Cassels [Cas08]. The draft book of David Witte Morris [WM08] also covers some properties of the orthogonal groups of quadratic forms.

Let $f\colon V \to \mathbb{R}$ be a quadratic form on a quadratic space as in §1.2. Suppose $T\colon V \to V$ is an invertible linear map such that $f\big(T(v)\big) = f(v)$ for all $v \in V$. Then T is said to be an *orthogonal transformation*, or an *isometry* of the quadratic space V. By abuse of notation let T also denote the matrix of the map T with respect to a basis $\mathcal{B}$, and let F be the matrix of the quadratic form f with respect to $\mathcal{B}$ as in (1.4). Then $T^{\mathsf{T}}FT = F$, for (viewing any v as a column vector)

$$f\big(T(v)\big) = f(v) \Leftrightarrow T(v)^{\mathsf{T}}FT(v) = v^{\mathsf{T}}Fv$$
$$\Leftrightarrow v^{\mathsf{T}}T^{\mathsf{T}}FTv = v^{\mathsf{T}}Fv;$$

and indeed this reasoning shows that $f\big(T(v)\big) = f(v)$ for every $v \in V$ if and only if $T^{\mathsf{T}}FT = F$. Thus we define the *orthogonal group* of f by

$$\mathrm{O}_f = \big\{T \in \mathrm{GL}(V) \mid T^{\mathsf{T}}FT = F\big\}. \tag{1.8}$$

Note that this definition shows O_f to be an algebraic group as in §1.4. Thus we can write $O_f(R)$ to refer to the subgroup of invertible elements of $O_f(\mathbb{R})$ whose entries lie in some ring R.

It is common to denote the orthogonal group of the form $x_1^2 + \cdots + x_n^2$ by $O(n)$, and that of the form $x_1^2 + \cdots + x_n^2 - x_{n+1}^2$ by $O(n, 1)$.

EXAMPLE 1.7. Let G be the orthogonal group of the quadratic form $f(X) = X_1^2 - \sqrt{2}X_2^2$, which has associated matrix

$$F = \begin{pmatrix} 1 & 0 \\ 0 & -\sqrt{2} \end{pmatrix}.$$

Thus the definition of O_f (as in (1.8)) becomes

$$\begin{pmatrix} x_{11}^2 - \sqrt{2}x_{21}^2 & x_{11}x_{12} - \sqrt{2}x_{21}x_{22} \\ x_{11}x_{12} - \sqrt{2}x_{21}x_{22} & x_{12}^2 - \sqrt{2}x_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\sqrt{2} \end{pmatrix}.$$

If G denotes the orthogonal group of f, then G is defined by the polynomials

$$p_1(x) = x_{11}^2 - \sqrt{2}x_{21}^2 - 1,$$
$$p_2(x) = x_{12}^2 - \sqrt{2}x_{22}^2 + \sqrt{2},$$
$$\text{and} \quad p_3(x) = x_{11}x_{12} - \sqrt{2}x_{21}x_{22} \qquad [\text{where } x = (x_{11}, x_{12}, x_{21}, x_{22})];$$

as above. The field K over which G is defined is $\mathbb{Q}(\sqrt{2})$, which has a $\mathbb{Q}$-basis consisting of vectors $v_1 = 1$ and $v_2 = \sqrt{2}$. With respect to this basis we have matrices for the left regular representation given by

$$\rho_{a+b\sqrt{2}} = \begin{pmatrix} a & 2b \\ b & a \end{pmatrix},$$

and the image ρ_K is defined by the equations

$$F_1(y) = y^{11} - y^{22}$$
$$\text{and} \quad F_2(y) = y^{21} - 2y^{12} \qquad [\text{where } y = (y^{11}, y^{12}, y^{21}, y^{22})].$$

We make the identification $M_4(\mathbb{Q}) \cong M_2\big(M_2(\mathbb{Q})\big)$. The final set of polynomials is then

$$\tilde{P}_1(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11}^{11} & x_{11}^{12} \\ x_{11}^{21} & x_{11}^{22} \end{pmatrix}^2 - \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{21}^{11} & x_{21}^{12} \\ x_{21}^{21} & x_{21}^{22} \end{pmatrix}^2 - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\tilde{P}_2(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{12}^{11} & x_{12}^{12} \\ x_{12}^{21} & x_{12}^{22} \end{pmatrix}^2 - \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{22}^{11} & x_{22}^{12} \\ x_{22}^{21} & x_{22}^{22} \end{pmatrix}^2 - \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix},$$

$$\text{and} \quad \tilde{P}_3(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{12}^{11} & x_{12}^{12} \\ x_{12}^{21} & x_{12}^{22} \end{pmatrix} \begin{pmatrix} x_{11}^{11} & x_{11}^{12} \\ x_{11}^{21} & x_{11}^{22} \end{pmatrix}$$
$$- \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{21}^{11} & x_{21}^{12} \\ x_{21}^{21} & x_{21}^{22} \end{pmatrix} \begin{pmatrix} x_{22}^{11} & x_{22}^{12} \\ x_{22}^{21} & x_{22}^{22} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Setting each of these equal to the zero matrix and multiplying out, we have a total of $3 \times 4 = 12$ equations. These define an alegbraic group over $\mathbb{Q}$, whose $\mathbb{Q}$-points forms a group isomorphic to the $\mathbb{Q}(\sqrt{2})$-points of G. The group G is a 1-dimensional Lie group; but note that as a real Lie group $\mathrm{Res}_{K/\mathbb{Q}}(G)(\mathbb{R})$ has dimension 4 (by considering the number of equations defining it as an algebraic variety in the 16-dimensional ambient space).

2

Hyperbolic Geometry

§2.1 The hyperboloid model of hyperbolic n-space

Let f_n denote the quadratic form on $\mathbb{R}^{n+1}$, which with respect to the standard basis has the diagonal form $f_n(x) = x_1^2 + x_2^2 + \cdots + x_n^2 - x_{n+1}^2$. The set $\mathbb{H}^n$ given by

$$\mathbb{H}^n = \{x \in \mathbb{R}^{n+1} \mid f_n(x) = -1, \ x_{n+1} > 0\} \tag{2.1}$$

is the upper sheet of a two-sheeted hyperboloid $\mathcal{H}_n$ in $\mathbb{R}^{n+1}$, and we will endow it with a metric, whereupon it is referred to as the *hyperboloid model* of *hyperbolic n-space*. Associated to f_n in the usual way (cf. §1.2) is the bilinear form $(\cdot, \cdot)_{f_n}$ on $\mathbb{R}^{n+1}$ given by

$$(x, y)_{f_n} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n - x_{n+1} y_{n+1},$$

and this in turn defines the metric $\mathrm{d}_{\mathbb{H}^n}$ on the set $\mathbb{H}^n$ when we declare

$$\cosh \mathrm{d}_{\mathbb{H}^n}(x, y) = -(x, y)_{f_n}. \tag{2.2}$$

The Special Theory of Relativity in physics is in some sense a study of the geometry of the quadratic form f_3 [Nab03]. From this theory arises the

17

following nomenclature: a vector $x \in \mathbb{R}^{n+1}$ is *space-like* if x is positive with respect to f_n, *time-like* if x is negative and *light-like* or *null* if x is isotropic (cf. §1.2). Figure 2.1 shows how the space $\mathbb{R}^{n+1}$ is partitioned into the three subsets of space-, time- and light-like vectors. The set of light-like vectors is often called the *light cone*.
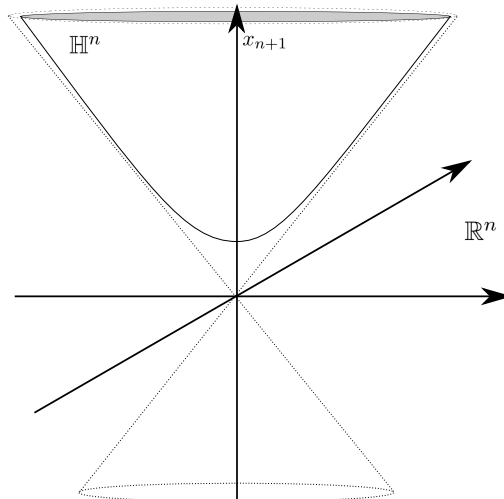
Figure 2.1: The hyperboloid in $\mathbb{R}^{n+1}$ with the light cone bounding it.

By (2.1), every point in $\mathbb{H}^n$ is a time-like vector in $\mathbb{R}^{n+1}$; but we can conversely associate to any time-like vector $x \in \mathbb{R}^{n+1}$ the (unique) point $(1/\sqrt{|f_n(x)|})x \in \mathbb{H}^n$. Thus $\mathbb{H}^n$ can also be viewed as the space of lines through the origin in $\mathbb{R}^{n+1}$, having timelike direction vectors.

The set $\mathbb{H}^n$ is a Riemannian manifold, and this can be seen by the following [BP92]: the tangent space to $\mathbb{H}^n$ at $x \in \mathbb{H}^n$ can be identified with the set $\{y \in \mathbb{R}^{n+1} \mid (x,y)_{f_n} = 0\}$, but this is simply the orthogonal complement $\langle x \rangle^{\perp_{f_n}}$. On this subspace f_n is positive definite and so defines a Riemannian metric on $\mathbb{H}^n$; although we shall have little occasion to consider the Riemannian metric on $\mathbb{H}^n$ as such. The interested reader will find further details on various forms that the Riemannian metric can take in the book of Bridson and Haefliger [BH99, pp. 92–96].

It has been asserted above that (2.2) implies that $(\mathbb{H}^n, \mathrm{d}_{\mathbb{H}^n})$ is a metric space, and we justify this claim now, following essentially the same argument of Bridson and Haefliger [BH99, pp. 20–21]. It is easy to see from (2.2) that the metric defined therein is positive definite, for all vectors in $\mathbb{H}^n$ are negative. To establish the triangle inequality we first introduce a convenient parameterisation of a geodesic segment $[x, y]$ between two distinct points $x, y \in \mathbb{H}^n$. Let u be the unit vector $(1/\sinh(\ell))\big(y + (x, y)_{f_n} x\big)$ where $\ell = \mathrm{d}_{\mathbb{H}^n}(x, y)$. Note that $u \in \langle x \rangle^{\perp}$. Furthermore, the curve

$$\beta \colon \mathbb{R} \to \mathbb{H}^n \colon t \mapsto (\cosh t)x + (\sinh t)u$$

satisfies $\beta(0) = x$ and $\beta(\ell) = y$.

Now if $[x, y]$ and $[x, z]$ are two hyperbolic line segments with unit vectors u and v as above, then we call the quantity $\alpha \in [0, \pi]$, such that $\cos \alpha = (u, v)_{f_n}$, the *hyperbolic angle* between $[x, y]$ and $[x, z]$. (This is called the *Lorentzian space-like angle* by Ratcliffe [Rat06, p. 68].)

LEMMA 2.1 (HYPERBOLIC COSINE RULE). *Let $\triangle$ be a triangle with vertices A, B and C, and side lengths a, b and c (with a opposite A as usual). Then*

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos C. \qquad (2.3)$$

PROOF. [BH99, p. 20] Let $[C, A]$ and $[C, B]$ have unit vectors u and v respectively, as above. Then

$$
\begin{aligned}
\cosh c &= -(A, B) \qquad \text{(by (2.2))} \\
&= -\big((\cosh b)C + (\sinh b)u \,,\, (\cosh a)C + (\sinh a)v\big) \\
&= -\cosh a \cosh b\,(C, C) - \sinh a \sinh b\,(u, v) \\
&= \cosh a \cosh b - \sinh a \sinh b \cos C. \qquad\qquad \square
\end{aligned}
$$

This leads to a proof of the triangle inequality for the hyperbolic metric:

PROPOSITION 2.2. *Let x, y and z be distinct points in $\mathbb{H}^n$. Then*

$$\mathrm{d}_{\mathbb{H}^n}(x, y) \leqslant \mathrm{d}_{\mathbb{H}^n}(x, z) + \mathrm{d}_{\mathbb{H}^n}(z, y). \qquad (2.4)$$

PROOF. [BH99, p. 21] Let $a = \mathrm{d}_{\mathbb{H}^n}(y, z)$, $b = \mathrm{d}_{\mathbb{H}^n}(z, x)$ and $c = \mathrm{d}_{\mathbb{H}^n}(x, y)$, and let α be the hyperbolic angle between $[x, z]$ and $[z, y]$. The function

$$\phi \colon t \mapsto \cosh a \cosh b - \sinh a \sinh b \cos t$$

is increasing on $[0, \pi]$ to its maximum at $t = \pi$, whence its value may be written as $\cosh(a + b)$ by standard formulae concerning hyperbolic functions. So, by (2.3),

$$\cosh c = \phi(\alpha) \leqslant \phi(\pi) = \cosh(a + b)$$

with equality if and only if $\alpha = \pi$. Since $\cosh$ is monotone increasing, $c \leqslant a + b$. $\qquad\square$

§2.2   ISOMETRIES OF $\mathbb{H}^n$

We will write $\mathrm{Isom}(\mathbb{H}^n)$ for the group of isometries of $\mathbb{H}^n$.

Clearly if $T \in \mathrm{O}(n, 1)$ (cf. p. 14), then T preserves distances according to (2.2). However not every element of $\mathrm{O}(n, 1)$ preserves $\mathbb{H}^n$, for such maps as $(x_1, \ldots, x_{n+1}) \mapsto (x_1, \ldots, x_n, -x_{n+1})$ will interchange the two sheets of the hyperboloid $\mathcal{H}_n$. In light of this, we consider the quotient group $\mathrm{PO}(n, 1) = \mathrm{O}(n, 1)/\{\pm\mathbf{id}\}$. This corresponds to identifying $x \in \mathbb{H}^n$ with $-x$ (which lies on the lower sheet of the hyperboloid $\mathcal{H}_n$), or equivalently thinking of $\mathbb{H}^n$ as a space of lines as described above (cf. p.18). One can show [Rat06, p. 63, Theorem 3.2.3] that any isometry of $\mathbb{H}^n$ extends to a linear map in $\mathrm{PO}(n, 1)$ so that $\mathrm{Isom}(\mathbb{H}^n) \cong \mathrm{PO}(n, 1)$.

It is frequently desirable to consider only those isometries that preserve the orientation of $\mathbb{H}^n$ (which form a subgroup denoted $\mathrm{Isom}^+(\mathbb{H}^n)$, and these are the maps in $\mathrm{PO}(n, 1)$ that are the image (under the projection $\mathrm{O}(n, 1) \to \mathrm{PO}(n, 1)$) of maps in $\mathrm{O}(n, 1)$ with determinant equal to 1. The subgroup of such maps is denoted $\mathrm{PSO}(n, 1)$ and has index 2 in $\mathrm{PO}(n, 1)$.

One also sees in the literature the notation $\mathrm{O}_0(n, 1)$, for the identity component of the group $\mathrm{O}(n, 1)$. This group may be identified with $\mathrm{Isom}(\mathbb{H}^n)$ (for it

is the group that preserves the upper sheet of $\mathcal{H}_n$. The group $\mathrm{SO}_0(n,1)$ denotes the intersection $\mathrm{O}_0(n,1) \cap \mathrm{SO}(n,1)$ and may be identified with $\mathrm{Isom}^+(\mathbb{H}^n)$.

It should be noted that in the low-dimensional cases $n = 2$ and $n = 3$, we have $\mathrm{Isom}(\mathbb{H}^n)$ isomorphic to $\mathrm{PSL}_2(\mathbb{R})$ and $\mathrm{PSL}_2(\mathbb{C})$ respectively [WM08, B.3].

CLASSIFICATION OF ISOMETRIES OF $\mathbb{H}^n$

It is well-known [Rat06, §4.7] that if $\gamma\colon \mathbb{H}^n \to \mathbb{H}^n$ is a non-trivial orientation-preserving isometry, then γ is one of the following types:

1. *elliptic*, if it fixes at least one point in $\mathbb{H}^n$;

2. *parabolic*, if it fixes exactly one point, and this lies on $\partial\mathbb{H}^n$;

3. *loxodromic*, if it fixes exactly two points, and these lie on $\partial\mathbb{H}^n$.

A loxodromic isometry preserves the geodesic whose endpoints are its two fixed points, and this is known as the *axis* of the transformation [Rat06, p. 140]. Additionally, a parabolic isometry preserves a horosphere tangent to $\partial\mathbb{H}^n$ at the fixed point of the transformation [Rat06, p. 139].

§2.3   CONSTRUCTION OF $\mathbb{H}^n$ FROM OTHER QUADRATIC FORMS

If $f = \sum_{i,j=1}^{n+1} a_{ij} x_i x_j$ is a quadratic form then it defines a function $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$. It is well-known [Rom05, p. 269] that there is a basis for $\mathbb{R}^{n+1}$ (depending on f) so that $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ can be written as $\sum_{i=1}^{n+1} b_i x_i^2$ with respect to this basis. Therefore without loss of generality we may consider only diagonal forms, and if we require the signature to be $(n,1)$ then we may assume all coefficients to be positive, except for the last which will be negative. Now suppose $g = \sum_{i=1}^{n} a_i x_i^2 - a_{n+1} x_{n+1}^2$ is such a form, with all $a_i > 0$ (and so

having signature $(n, 1)$). The map $\phi\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, given by

$$\phi\colon (x_1, \ldots, x_{n+1}) \mapsto \left(\sqrt{a_1}\, x_1, \ldots, \sqrt{a_{n+1}}\, x_{n+1}\right), \tag{2.5}$$

has the property that $g(x) = f_n\big(\phi(x)\big)$. So by a suitable isomorphism of $\mathbb{R}^{n+1}$ we find a basis in which g is represented by our standard form f_n.

The construction of $\mathbb{H}^n$ as above (cf. §2.1) can be carried out with a more general inner product $(\cdot, \cdot)_g$ in place of $(\cdot, \cdot)_{f_n}$, if g is a quadratic form of signature $(n, 1)$. Then the restriction $\tilde{\phi}\colon \mathbb{H}^n_g \to \mathbb{H}^n_{f_n}\colon x \mapsto \phi(x)$ is an isometry, for

$$\begin{aligned}
(x, y)_g &= \tfrac{1}{2}\big(g(x+y) - g(x) - g(y)\big) \\
&= \tfrac{1}{2}\big(f_n(\phi(x) + \phi(y)) - f_n(\phi(x)) - f_n(\phi(y))\big) = \big(\phi(x), \phi(y)\big)_{f_n}.
\end{aligned}$$

The isometry groups of $\mathbb{H}^n_g$ and $\mathbb{H}^n_{f_n}$ are related by conjugation. That is, there is a matrix $M \in \mathrm{GL}_{n+1}(\mathbb{R})$ such that $\mathrm{O}_g(\mathbb{R}) = M^{-1}\, \mathrm{O}_{f_n}(\mathbb{R}) M$.

NOTATION. Henceforth, we shall use the notation $(\cdot, \cdot)$ in place of $(\cdot, \cdot)_{f_n}$ whenever there can be no confusion. It will also be convenient to write $\|x\|$ instead of $\sqrt{f_n(x)}$ (allowing the possibility that $\|x\|$ be imaginary).

§2.4   VOLUME IN $\mathbb{H}^n$

The space $\mathbb{H}^n$ inherits a natural volume from its Riemannian metric [Rat06, §3.4]. Thus one may speak of volumes of measurable subsets of $\mathbb{H}^n$. If $A \subset \mathbb{H}^n$ is measurable, then we denote its volume by $\mathrm{vol}(A)$.

We will not be so concerned with calculating volumes of arbitrary subsets of $\mathbb{H}^n$, and so we simply remark here that the volume of a ball B_r of radius r in hyperbolic space is given by [Rat06, Ex. 3.4.1, p. 70]

$$\mathrm{vol}(B_r) = \mathrm{vol}(\mathbb{S}^{n-1}) \int_0^r \sinh^{n-1} t \, \mathrm{d}t \tag{2.6}$$

where $\mathbb{S}^{n-1}$ is the $(n-1)$-sphere. For convenience we note that the volumes of the n-spheres ($n = 2k - 1$ or $n = 2k$) are given by (abusing notation for

vol($\cdot$) slightly) [Rat06, Ex. 2.4.5, p. 46]

$$\text{vol}(\mathbb{S}^{2k-1}) = \frac{2\pi^k}{(k-1)!}$$

$$\text{and} \qquad \text{vol}(\mathbb{S}^{2k}) = \frac{2^{k+1}\pi^k}{(2k-1)(2k-3)\cdots 3 \cdot 1}. \tag{2.7}$$

§2.5   HYPERPLANES

If $e \in \mathbb{R}^{n+1}$ is a space-like vector (cf. p. 18), then its orthogonal complement $\langle e \rangle^\perp$ intersects $\mathbb{H}^n$. We call $\mathbb{H}^n \cap \langle e \rangle^\perp$ a *hyperplane* in $\mathbb{H}^n$. Clearly if $\lambda \in \mathbb{R} \smallsetminus \{0\}$ then λe defines the same hyperplane as e. In particular if e has K-rational entries then by multiplying by suitable $\lambda \in K$ we can use λe to define a hyperplane by a vector with K-*integral* entries instead. (We will use this fact in §5.1.) This definition of a hyperplane illustrates a particular advantage of the hyperboloid model, namely that hyperplanes are very simply described in terms of a single vector. The image of a hyperplane under an isometry of $\mathbb{H}^n$ can be easily computed by applying the isometry's corresponding matrix in $\mathrm{O}_0(n, 1)$ (or a representative in $\mathrm{PO}(n, 1)$) to the vector defining the hyperplane.

We will later need to deal with two related questions concerning hyperplanes in $\mathbb{H}^n$: whether or not two given hyperplanes intersect, and, if they do not, what the distance between them is. For intersection, the following will be of fundamental importance later [Rat06, pp. 68–70]:

THEOREM 2.3. *Suppose H_1 and H_2 are hyperplanes in $\mathbb{H}^n$, defined by vectors e_1 and e_2 respectively, where $e_1 \neq \lambda e_2$ for any $\lambda \in \mathbb{R}$.*

1. *If $\big|(e_1, e_2)\big| > \|e_1\|\|e_2\|$ then $H_1 \cap H_2 = \emptyset$ and the hyperplanes do not meet at infinity.*

2. *If $\big|(e_1, e_2)\big| = \|e_1\|\|e_2\|$ then $H_1 \cap H_2 = \emptyset$ but the hyperplanes meet at infinity.*

3. *If $\big|(e_1, e_2)\big| < \|e_1\|\|e_2\|$ then H_1 and H_2 intersect in $\mathbb{H}^n$.*

Supposing two hyperplanes do not intersect, we have the following means of establishing the distance between them [Rat06, p. 69]:

THEOREM 2.4. *Suppose $H_1 = \mathbb{H}^n \cap \langle e_1 \rangle^\perp$ and $H_2 = \mathbb{H}^n \cap \langle e_2 \rangle^\perp$ are two disjoint hyperplanes, and denote by ℓ the hyperbolic line segment $[a, b]$ orthogonal to both of them. Then the hyperbolic distance $\mathrm{d}_{\mathbb{H}^n}(H_1, H_2)$ between the endpoints a and b is equal to the length of ℓ and is given by*

$$\cosh \mathrm{d}_{\mathbb{H}^n}(H_1, H_2) = \frac{\left|(e_1, e_2)\right|}{\|e_1\| \|e_2\|}. \tag{2.8}$$

A hyperplane H in hyperbolic space is an example of a *totally geodesic* subspace; that is, one in which for every pair of distinct points $x, y \in H$ there is a geodesic in H containing both x and y [Rat06, p. 26].

3

DISCRETE SUBROUPS OF LIE GROUPS

Most of what is described below may be found in references such as the works of Vinberg and Shvartsman [VS93], and a second work by those authors and Gorbatsevich [VGS00]. The book of Raghunathan [Rag72] contains the proofs of many results listed below. The unpublished book of Morris [WM08] remains a good reference in addition.

§3.1   LIE GROUPS

For the more specifically 'Lie-theoretic' aspects of Lie group theory, the reader may find that such books as that of Onishchik and Vinberg [OV93] serve as a good introduction, along with such works as that of Fulton and Harris [FH91], where representations of Lie groups form the main theme of the exposition.

   A (real) *Lie group G* is a group that is also a (real) differentiable manifold, with the property that the multiplication and inversion maps

$$(\, \cdot \, , \, \cdot \,) \colon G \times G \to G \colon (g, h) \mapsto gh \tag{3.1}$$

$$\cdot^{-1} \colon G \to G \colon g \mapsto g^{-1} \tag{3.2}$$

are both differentiable. A Lie group is an example of a topological group, and in light of this one may regard subgroups of a Lie group as being open (and/or closed) if they are open (and/or closed) subsets of the topological space G. Note that if $H \leqslant G$ is an open subgroup of a Lie group G then H is also closed, since its (set-theoretic) complement in G is open, being a union of its cosets gH, which are open owing to the continuity of the map in (3.1). If $H \leqslant G$ is closed and of finite index in G, then it is also open, for its complement is a finite union of closed (co)sets. One may also make reference to other topological properties of subgroups of a Lie group, such as compactness, discreteness, density etc.

Haar measure and symmetric spaces

Lie groups have the property that they are Hausdorff, locally compact (every element has a compact neighbourhood) and second-countable (the topology on the group has a countable base). A significant consequence of having this collection of properties is the following theorem [WM08, Prop. A.19]:

THEOREM 3.1. *On any locally compact topological group G there is a measure μ (unique up to a multiplicative constant) such that*

1. *$\mu(C)$ is finite if $C \subseteq G$ is compact;*

2. *$\mu(gS) = \mu(S)$ for every $g \in G$ and all Borel measurable sets $S \subseteq G$; and*

3. *μ is a σ-finite Borel measure.*

(By a σ-finite measure is meant that G is a countable union of sets of finite μ-measure.) Such a measure μ is known as a *Haar measure* on G. Since a Haar measure is unique only up to multiplication by a constant, one must fix a normalisation for the measure.

Before considering the normalisation of Haar measure for $\mathrm{Isom}(\mathbb{H}^n)$, we introduce the concept of a symmetric space. (The reader may find further details in the book of D. Witte Morris [WM08, §1A & §1B].) Let X be a connected Riemannian manifold. The space X is called a *symmetric space* if for every

$x \in X$, there is a symmetry τ of X such that $\tau(x) = x$ and $\mathrm{d}_x\tau = -\mathbf{id}$. We call a symmetric space *irreducible* if its universal cover is not isometric to a non-trivial product $X_1 \times X_2$. It is well-known that if G is a connected Lie group and $K < G$ is a maximal compact subgroup of G, and σ is an automorphism of G with $\sigma^2 = \mathbf{id}$ such that $\mathrm{Fix}(\sigma)$ contains K as an open subgroup, then G/K can be given the structure of a symmetric space; and, moreover, every symmetric space arises in this way.

By way of example, consider $\mathbb{H}^n$, and fix a basepoint $x_0 \in \mathbb{H}^n$. Since $\mathrm{Isom}(\mathbb{H}^n)$ acts transitively on $\mathbb{H}^n$, for any $x \in \mathbb{H}^n$ there is $g \in \mathrm{Isom}(\mathbb{H}^n)$ such that $g(x_0) = x$. The coset space $\mathrm{Isom}(\mathbb{H}^n)/\mathrm{Stab}(x_0)$ may be identified homeomorphically with $\mathbb{H}^n$ and this identification may be acheived using any continuous map $\eta\colon \mathrm{Isom}(\mathbb{H}^n) \to \mathbb{H}^n$ such that $\eta^{-1}(\eta(g)) = g\,\mathrm{Stab}(x_0)$ [Rat06, Theorem 5.1.5]. One may prescribe an explicit homeomorphism $\Psi\colon \mathbb{B}^n \times \mathrm{O}(n) \to \mathrm{Isom}(\mathbb{H}^n)$, where $\mathbb{B}^n$ is the Poincaré Ball model of hyperbolic space (cf. p. 33). The map Ψ is given by $\Psi(b, A) = \tau_b A$ where

$$\tau_b(x) = \frac{(1 - |b|^2)x + 2(1 + x \cdot b)b}{|x + b|^2}.$$

(See Ratcliffe [Rat06, p. 155].)

Now, we consider the normalisation of the Haar measure on $\mathrm{Isom}(\mathbb{H}^n)$ such that volumes in $\mathrm{Isom}(\mathbb{H}^n)$ in some sense correspond to volumes in $\mathbb{H}^n$. We follow the exposition of Ratcliffe [Rat06, p. 560]. Denote $G = \mathrm{Isom}(\mathbb{H}^n)$ and let H be the compact subgroup $H = \mathrm{Stab}_G(x_0)$ for some $x_0 \in \mathbb{H}^n$. (Note that H is isomorphic to $\mathrm{SO}(n)$.) Note that by identifying $\mathbb{B}^n$ with $\mathbb{H}^n$ via an isometry $\zeta\colon \mathbb{B}^n \to \mathbb{H}^n$, and composing with the map $\Psi(\cdot, \mathbf{id})$, we obtain a homeomorphism $\eta_0\colon \mathrm{Isom}(\mathbb{H}^n) \to \mathbb{H}^n$ satisfying the properties of η above. Let ν be a left-invariant Haar measure on H, and let ω be the left-invariant Haar measure on G/H, normalised via the homeomorphism η_0, so that $\omega(C/H) = \mathrm{vol}_{\mathbb{H}^n}(\eta_0(C))$ whenever $\eta_0(C) \subseteq \mathbb{H}^n$ is a measurable set. Then, we normalise the Haar measure on H so that $\int_H \mathrm{d}\nu(h) = 1$. The Haar measure of a measurable subset $B \subseteq G$ (with indicator function χ_B) is then given by the integral

$$\int_G \chi_B(g)\,\mathrm{d}\mu(g) = \int_{G/H} \left(\int_H \chi_B(gh)\,\mathrm{d}\nu(h) \right) \mathrm{d}\omega(gH).$$

SEMISIMPLICITY

There are several notions of semisimplicity appearing in the literature. Broadly speaking, these mean that an object is 'nearly simple', in that it is a direct product (or sum) of 'simple' objects. For us, semisimplicity appears in the theory of Lie groups, as well as algebraic groups.

Suppose that G_1 and G_2 are Lie groups. We say that G_1 and G_2 are *isogenous* if they both finitely cover some common connected Lie group.

A Lie group is *semisimple* if it is isogenous to a (finite) direct product of simple Lie groups, where by a *simple Lie group* is meant a non-Abelian Lie group with no non-trivial, proper, connected, closed normal subgroups.

By way of example, note that $\mathrm{SL}_n(\mathbb{R})$ is a semisimple Lie group, whereas $\mathrm{GL}_n(\mathbb{R})$ is not. Of particular interest is the fact that $\mathrm{SO}(n,m)$ is semisimple for $n + m \geqslant 3$ [WM08, 3.22]. So, $\mathrm{SO}(n,1)$ is semisimple for $n \geqslant 2$.

Now let G be an algebraic group. One may associate with G an algebra (called the Lie algebra of G) in a manner analogous to that for Lie groups [Hum75, Sect. 9.1]. If G is connected (in the Zariski topology) and has a Lie algebra containing no non-zero commutative ideals, then we call G *semisimple* [Hum75, Sect. 13.5]. (This definition implies that G has no non-trivial, connected, Abelian normal subgroups.)

§3.2   DISCRETE SUBGROUPS OF LIE GROUPS

Let $\Gamma \leqslant G$ be a subgroup of a Lie group. If Γ is a discrete subset of the topological space G then it is called a *discrete subgroup* (of G).

A *fundamental domain* for Γ in G is a set $F \subseteq G$ such that there is a bijection $F \to \Gamma \backslash G$, where $\Gamma \backslash G$ is the space of right cosets of Γ in G. One can show that if $U \subseteq G$ is an open set such that $UU^{-1} \cap \Gamma = \{1\}$ then for some sequence $\{g_n\}_n \subset G$ we have

$$\bigcup_n U g_n = G$$

and the set B given by

$$B = \bigcup_{n=1}^{\infty} \left(U g_n \smallsetminus \bigcup_{i<n} \Gamma U g_i \right)$$

is a Borel fundamental domain for Γ in G [WM08, Lemma 4.1]. If F is now any Borel fundamental domain then the measure ν on $\Gamma\backslash G$ given by

$$\nu(A) = \mu\big(F \cap \pi_\Gamma^{-1}(A)\big)$$

is G-invariant and σ-finite, where $\pi_\Gamma\colon G \to \Gamma\backslash G$ is the natural projection map. If $\nu(\Gamma\backslash G)$ is finite then we call Γ a *lattice* in G. If $\Gamma\backslash G$ is a compact set (equivalently if there exists a compact fundamental domain for Γ in G) then we say that Γ is *co-compact* or *uniform*.

§3.3 Discrete subgroups of $\mathrm{Isom}(\mathbb{H}^n)$

The group $\mathrm{Isom}(\mathbb{H}^n)$ can be given the structure of a Lie group, and in its realisation as $\mathrm{PO}(n,1)$ its topology is naturally inherited from the Euclidean norm topology on $\mathrm{O}(n,1)$, in turn coming from $M_{n+1}(\mathbb{R})$ regarded as an $(n+1)^2$-dimensional vector space.

We are concerned with discrete subgroups of $\mathrm{Isom}(\mathbb{H}^n)$. The notions described in §3.2 (fundamental domains, co-volume, etc.) of course apply to such discrete groups. However they can be translated into analogous concepts in terms of the action of $\mathrm{Isom}(\mathbb{H}^n)$ and its discrete subgroups on $\mathbb{H}^n$, and these are described presently.

Let $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ be a discrete group. Then Γ acts on $\mathbb{H}^n$ by isometries; i.e., we have for each $\gamma \in \Gamma$ a map

$$\gamma\colon \mathbb{H}^n \to \mathbb{H}^n\colon x \mapsto \gamma(x),$$

and the multiplication in the group is compatible with the action in the sense that

1. $1(x) = x$ for every $x \in \mathbb{H}^n$; i.e., the identity element in G acts as the identity map; and

2. $(\gamma_1 \gamma_2)(x) = \gamma_1 \circ \gamma_2(x)$ for every $x \in \mathbb{H}^n$.

Suppose $F \subseteq \mathbb{H}^n$ is an open subset of hyperbolic space such that

1. $\displaystyle\bigcup_{\gamma \in \Gamma} \gamma(\overline{F}) = \mathbb{H}^n$ and

2. $\gamma_1(F) \cap \gamma_2(F) \neq \emptyset \;\Leftrightarrow\; \gamma_1 = \gamma_2$ for every γ_1 and γ_2 in Γ.

(Here, $\overline{F}$ denotes the topological closure of F in $\mathbb{H}^n$.) The set F is then called a *fundamental domain* for Γ in $\mathbb{H}^n$. The space $\Gamma \backslash \mathbb{H}^n$ of orbits for the action, or *quotient space*, may be regarded as the closure of a fundamental domain, with points on the boundary identified if they are in the same Γ-orbit. Equivalent to the previous definition of Γ being a lattice in $\mathrm{Isom}(\mathbb{H}^n)$ (cf. §3.2) is the condition that Γ must have a finite-volume fundamental domain in $\mathbb{H}^n$. If Γ has a compact fundamental domain in $\mathbb{H}^n$ then Γ is called *co-compact*, and the volume of the fundamental domain is called the *co-volume* of Γ.

Note that we have two similar notions of 'fundamental domain'. This pertains to the discussion on p. 27 concerning symmetric spaces and the possibility of choosing to work either in the group $\mathrm{Isom}(\mathbb{H}^n)$ or in the space $\mathbb{H}^n$ itself.

Actions of discrete groups on $\mathbb{H}^n$

Suppose that $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ is any subgroup (and not necessarily discrete). We say that the action of Γ on $\mathbb{H}^n$ is *discontinuous* (or in some nomenclature *properly discontinuous*) if for every compact subset $C \subset \mathbb{H}^n$

$$\text{the set} \quad \{\gamma \in \Gamma \mid \gamma(C) \cap C \neq \emptyset\} \quad \text{is finite;} \tag{3.3}$$

equivalently, if the collection $\{\gamma(C) \mid \gamma \in \Gamma\}$ is *locally finite*. We have the following characterisations [Rat06, Theorem 5.3.4, Theorem 5.3.5]:

THEOREM 3.2. *Let $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$. Then the following are equivalent:*

1. *Γ is discrete;*

2. *Γ acts discontinuously on $\mathbb{H}^n$;*

  3. for any $x \in \mathbb{H}^n$ the stabiliser $\mathrm{Stab}_\Gamma(x)$ *is finite and the orbit* $\Gamma(x)$ *is closed and discrete.*

If Γ has a discontinuous action, then the quotient space $\Gamma \backslash \mathbb{H}^n$ is known as a hyperbolic *orbifold*. It is often desirable that the quotient might be, in fact, a manifold, and for this, we require the action of Γ on $\mathbb{H}^n$ to be *free*; that is, we require that for every $x \in \mathbb{H}^n$ the stabiliser $\mathrm{Stab}_\Gamma(x)$ is trivial (and not merely finite). This assertion is a consequence of the following theorem:

THEOREM 3.3. [Rat06, Theorem 8.1.3, Theorem 8.2.1] *Let* $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$.

  1. The group Γ *acts freely if and only if* Γ *has no elements of finite order.*

  2. Supposing the action of Γ *is free and discontinuous, the quotient map* $\pi \colon \mathbb{H}^n \to \Gamma \backslash \mathbb{H}^n$ *is a local isometry and covering projection, and* Γ *is the group of covering transormations of* π.

The elements of finite order in a group are known as the *torsion elements*, and a group without non-trivial torsion elements is said to be *torsion-free*. If $\Gamma < \mathrm{Isom}^+(\mathbb{H}^n)$ is discrete then the non-trivial torsion elements of Γ are precisely the elliptic elements of Γ [Rat06, p. 177]).

In light of Theorem 3.3 it is desirable to be able to obtain torsion-free discrete groups. The following well-known result shows that this is easier than might be expected, at least in principle:

THEOREM 3.4 (SELBERG'S LEMMA). [Rat06, Cor. 4, p. 331] *Suppose that* Γ *is a finitely-generated subgroup of* $\mathrm{GL}_n(\mathbb{C})$ *for some* $n \in \mathbb{N}$. *Then* Γ *has a torsion-free subgroup of finite index.*

It should be remarked that the usual method of proving Selberg's Lemma is by exhibiting a congruence subgroup $\Gamma(\mathfrak{p})$ of Γ (cf. §4.2) and showing that it is torsion-free for appropriate choice of ideal $\mathfrak{p}$.

Since $\mathrm{O}(n, 1)$ may be regarded as a subgroup of $\mathrm{GL}_{n+1}(\mathbb{C})$, Theorem 3.4 also applies to finitely generated subgroups of $\mathrm{PO}(n, 1)$. ($\mathrm{PO}(n, 1)$ can even be viewed as a subgroup of $\mathrm{GL}_{n+1}(\mathbb{C})$ by viewing it as the group of so-called 'positive' Lorentzian matrices [Rat06, p. 58].) Thus for any finitely-generated

group $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ we can always find a subgroup of finite index, $\Gamma_1 < \Gamma$ say, such that the space $\Gamma_1\backslash\mathbb{H}^n$ is a manifold. The subgroup Γ_1 corresponds to a (finite) cover of the orbifold $\Gamma\backslash\mathbb{H}^n$. In some sense, one does not lose too much from a geometric point of view in passing from the orbifold to the manifold. For instance, we have the simple formula

$$\mathrm{vol}(\Gamma_1\backslash\mathbb{H}^n) = |\Gamma : \Gamma_1|\,\mathrm{vol}(\Gamma\backslash\mathbb{H}^n), \qquad (3.4)$$

meaning that the volume of the cover is an integer multiple of the 'base' orbifold.

For later reference we note that in a hyperbolic n-manifold M, a closed geodesic corresponds to a loxodromic element γ (cf. §2.2) of the discrete group giving rise to the manifold M: the closed geodesic is the quotient of the axis of the loxodromic transformation γ by the group $\langle\gamma\rangle$ generated by that transformation.

§3.4   First examples of discrete subgroups of $\mathrm{Isom}(\mathbb{H}^n)$

The hyperbolic space $\mathbb{H}^n$ has isometry group $\mathrm{PO}(n,1)$. One can look at the integral matrices in that group, namely the subgroup $\mathrm{PO}(n,1)(\mathbb{Z})$ obtained by considering the projection of $\mathrm{O}(n,1) \cap \mathrm{GL}_{n+1}(\mathbb{Z})$ to $\mathrm{PO}(n,1)$. Clearly this group is discrete, since $\mathbb{Z}^{(n+1)^2}$ is a discrete set in $\mathbb{R}^{(n+1)^2}$ in the usual topology, inside which lies $\mathrm{O}(n,1)$; and so the group of invertible integral elements in $\mathrm{PO}(n,1)$ forms a discrete subgroup. It is not immediately obvious whether or not this group is a lattice, but the fact that it is follows from the discussion in §§4.3–4.4.

A similar and well-known example of a discrete group is the group $\mathrm{SL}_2(\mathbb{Z})$, which is the group of invertible integral 2×2 matrices of determinant 1. This is a subgroup of $\mathrm{SL}_2(\mathbb{R})$, which acts on $\mathbb{H}^2$, using the alternative realisation of $\mathbb{H}^2$ as the upper half-plane $\mathbb{H}^2_{\mathrm{U}} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\} \subset \mathbb{C}$ and using the

metric [Kat92, Theorem 1.2.6]

$$\exp\big(\mathrm{d}_{\mathbb{H}^2_{\mathrm{U}}}(z,w)\big) = \frac{|z - \overline{w}| + |z - w|}{|z - \overline{w}| - |z - w|}$$

and the $\mathrm{SL}_2(\mathbb{R})$-action [Kat92, Theorem 1.1.2]

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad g(z) = \frac{az + b}{cz + d}.$$

The group $\mathrm{SL}_2(\mathbb{Z})$ also turns out to be a lattice in $\mathrm{SL}_2(\mathbb{R})$. One may show that the set $\big\{z \in \mathbb{H}^2_{\mathrm{U}} \,\big|\, |z| > 1, \, -\frac{1}{2} < \mathrm{Re}(z) < \frac{1}{2}\big\}$ is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ in $\mathbb{H}^2_{\mathrm{U}}$ [Kat92, p. 55], and that it has hyperbolic area $\pi/3$.

Clearly there are some fairly trivial examples of discrete groups. For example, one might take a hyperbolic transformation γ (cf. §2.2) and consider the group $\Gamma = \langle \gamma \rangle$. This will be discrete, but not a lattice. Such groups are known as *elementary* groups [Rat06, §5.5]. A group is said to be elementary if it has a finite orbit in the closure of hyperbolic space $\overline{\mathbb{H}^n}$. (In considering the closure of hyperbolic space one tends to make use of models such as the Poincaré ball model

$$\mathbb{B}^n = \{x \in \mathbb{R}^n \mid |x| < 1\} \quad \text{with} \quad \cosh \mathrm{d}_{\mathbb{B}^n}(x,y) = 1 + \frac{2|x - y|^2}{\big(1 - |x|^2\big)\big(1 - |y|^2\big)},$$

in which the boundary has the convenient realisation as $\mathbb{S}^{n-1}$ [Rat06, §4.5].)

§3.5   The Každan-Margulis Theorem and Margulis' Lemma

A significant early work by D. Každan and G. Margulis [KM68a] contains as one of its corollaries the following result:

THEOREM 3.5. [KM68a, Corollary to Theorem 1] *Let G be the connected component of an algebraic semisimple group with no compact factors. There exists a contsant $\mu_0 > 0$ such that for any discrete subgroup $\Gamma < G$, we have $\mu(\Gamma\backslash G) > \mu_0$ (where μ is the Haar measure on $\Gamma\backslash G$).*

For us, the following particular case is most relevant:

COROLLARY 3.6. *Let $n \geqslant 2$, with $n \in \mathbb{N}$. Then there is $\mu_n > 0$ such that any orientable hyperbolic n-orbifold $\Gamma \backslash \mathbb{H}^n$ has volume $\mathrm{vol}(\Gamma \backslash \mathbb{H}^n) > \mu_n$.*

This result does not say whether or not a minimum volume exists (i.e., that there is an orbifold attaining it), but results of M. Belolipetsky and V. Emery (and others such as Siegel before) address this question, in the case of arithmetic orbifolds (which will be defined in §4.2):

THEOREM 3.7. [Bel04, Theorem 4.1] [BE11, Theorems 1 & 2] *Let $n \in \mathbb{N}$ with $n \geqslant 4$.*

1. *There exists a unique orientable arithmetic compact hyperbolic n-orbifold of minimal volume.*

2. *There exists a unique orientable arithmetic non-compact hyperbolic n-orbifold of minimal volume.*

Thus we see that the minimum is indeed attained in a particular class. The compact orbifold in Theorem 3.7 is defined over $\mathbb{Q}(\sqrt{5})$ whilst the non-compact orbifold is defined over $\mathbb{Q}$. Note that this theorem does not say anything about the situation for non-arithmetic hyperbolic manifolds, and in this case the question is still open.

Fundamental to the study of the structure of hyperbolic manifolds is the lemma of Margulis:

THEOREM 3.8 (MARGULIS' LEMMA). *Let $n \geqslant 2$ (with $n \in \mathbb{N}$). Then there is $\varepsilon_n > 0$ such that for any positive $\varepsilon \leqslant \varepsilon_n$, for any discrete $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$, and for any $x \in \mathbb{H}^n$, the subgroup $\Gamma_\varepsilon(x) < \Gamma$, generated by the set*

$$\left\{ \gamma \in \Gamma \mid \mathrm{d}_{\mathbb{H}^n}\big(x, \gamma(x)\big) \leqslant \varepsilon \right\}$$

is elementary.

The proof of this lemma can be found in the book of Ratcliffe [Rat06, Theorem 12.6.1], and also in that of Benedetti and Petronio [BP92, Theorem D.1.1]

where it is stated in its slightly different form; namely that there is $\delta_n > 0$ such that $\Gamma_\delta(x)$ is almost nilpotent for any positive $\delta \leqslant \delta_n$, any discrete $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ and any $x \in \mathbb{H}^n$. (A group G is *nilpotent* if the sequence of subgroups generated by the commutators

$$\big[G,G\big], \quad \big[G,[G,G]\big], \quad \ldots, \quad \big[G,[G,\ldots,[G,G]]\big], \quad \ldots$$

terminates after a finite number of steps in the trivial subgroup $\{1\} \leqslant G$. A group G is *almost nilpotent* if it has a nilpotent subgroup $H \leqslant G$ of finite index.)

Recall that if M is a Riemannian manifold, and $x \in M$, then the *injectivity radius* of M at x, denoted $\mathrm{InjRad}_x(M)$, is the biggest distance in M for which the map $\exp \colon T_x M \to M$ is a diffeomorphism. The injectivity radius at x may be regarded as the largest real number r for which a ball of radius r, centred at x, may be embedded in M. The geometric significance (especially to us) of Theorem 3.8 is the following result:

THEOREM 3.9. [BP92, Theorem D.3.3] *Let $M = \Gamma \backslash \mathbb{H}^n$ be a closed hyperbolic n-manifold, and let $\varepsilon < \varepsilon_n$. Denote by $M_{(0,\varepsilon]}$ the* thin part *of M, namely the set*

$$\{x \in M \mid \mathrm{InjRad}_x(M) \leqslant \varepsilon\}.$$

Then $M_{(0,\varepsilon]}$ consists of either closed geodesics of length ε, or regions of the form $\overline{D^{n-1}} \times S^1$.

The latter regions described in this theorem are known as *Margulis tubes*. One also notes that the complement $M \smallsetminus M_{(0,\varepsilon)}$ is compact if M is of finite volume. In the case where M is not compact (but still of finite volume), the thin part will also contain cusp parts homeomorphic to $V \times [0, \infty)$ for an oriented Euclidean manifold V without boundary [BP92, Theorem D.3.3].

4

ARITHMETIC LATTICES

In this chapter we introduce some general theory concerning arithmetic lattices in $\mathrm{Isom}(\mathbb{H}^n)$, and look at some specific examples. One reference for this is the book of Margulis [Mar91], but for an introduction the reader may find useful the books of Maclachlan and Reid [MR03], Zimmer [Zim84], and Witte Morris [WM08]. The book of Klopsch, Nikolov and Voll is also worthwhile [KNV11] for its emphasis on the algebraic group approach.

§4.1 COMMENSURABILITY OF SUBGROUPS

We begin with two elementary facts, which will be used later. One can find similar results in W. R. Scott's book on group theory [Sco87, Sect. 1.7].

LEMMA 4.1. *Let G be a group.*

  1. *Let H and K be subgroups of G, both having finite index in G. Then*

  $$|G : H \cap K| \leqslant |G : H||G : K|. \qquad (4.1)$$

  2. *Suppose H_1, H_2 and H_3 are subgroups of G. Then we have*

  $$|H_1 \cap H_2 : H_1 \cap H_2 \cap H_3| \leqslant |H_2 : H_2 \cap H_3|. \qquad (4.2)$$

Proof.

1. Denote, for any subgroup $S \leqslant G$, the collection of (left) cosets of S in G by G/S: this set need not form a group. Define a map

$$\phi \colon G/(H \cap K) \to G/H \times G/K \quad \text{by} \quad \phi \colon a(H \cap K) \mapsto (aH, aK).$$

Now, to show that this is well-defined and injective, suppose that $a(H \cap K) = b(H \cap K)$. We have

$$a(H \cap K) = b(H \cap K) \Leftrightarrow ab^{-1} \in H \cap K$$
$$\Leftrightarrow aH = bH \text{ and } aK = bK$$
$$\Leftrightarrow (aH, aK) = (bH, bK)$$
$$\Leftrightarrow \phi\big(a(H \cap K)\big) = \phi\big(b(H \cap K)\big)$$

and so ϕ is well-defined. However this also shows that ϕ is injective, and so there must be at most $|G/H \times G/K|$ elements in the domain of ϕ.

2. Define a map

$$\phi \colon (H_1 \cap H_2)/(H1 \cap H_2 \cap H_3) \to H_2/(H_2 \cap H_3)$$
$$\text{by} \quad \phi \colon x(H_1 \cap H_2 \cap H_3) \mapsto x(H_2 \cap H_3). \quad (4.3)$$

Then, as above

$$a(H_1 \cap H_2 \cap H_3) = b(H_1 \cap H_2 \cap H_3) \Leftrightarrow ab^{-1} \in H_1 \cap H_2 \cap H_3$$
$$\Leftrightarrow ab^{-1} \in H_2 \cap H_3$$
$$\Leftrightarrow a(H_2 \cap H_3) = b(H_2 \cap H_3).$$

Thus ϕ is both well-defined and injective, and the same argument as above applies. $\qquad\square$

Now, suppose that we have two subgroups Γ_1 and Γ_2 of a group G. We say that Γ_1 and Γ_2 are *commensurable* if

$$|\Gamma_i : \Gamma_1 \cap \Gamma_2| < \infty \qquad \text{for both } i = 1 \text{ and } i = 2. \qquad (4.4)$$

It should be noted that commensurability is an equivalence relation on the class of all subgroups of G. That the relation of commensurability is symmetric and reflexive is trivial: any group has finite index in itself, and by definition the relation is symmetric as it deals with a common intersection of two groups. To see that it is transitive, let Γ_1 and Γ_2 be commensurable subgroups of G, and suppose that Γ_3 is a further subgroup (of G), commensurable with Γ_2. Then on the one hand

$$|\Gamma_1 \cap \Gamma_3 : \Gamma_1 \cap \Gamma_2 \cap \Gamma_3| \cdot |\Gamma_3 : \Gamma_1 \cap \Gamma_3| = |\Gamma_3 : \Gamma_1 \cap \Gamma_2 \cap \Gamma_3|, \qquad (4.5)$$

whereas on the other

$$|\Gamma_2 \cap \Gamma_3 : \Gamma_1 \cap \Gamma_2 \cap \Gamma_3| \cdot |\Gamma_3 : \Gamma_2 \cap \Gamma_3| = |\Gamma_3 : \Gamma_1 \cap \Gamma_2 \cap \Gamma_3|. \qquad (4.6)$$

(Note that both of the right-hand side expressions are the same.) Now by Lemma 4.1, the leftmost factors are finite since Γ_3 and Γ_2 are commensurable, and the middle factor in (4.6) is again finite by the same commensurability relation. Thus $|\Gamma_3 : \Gamma_1 \cap \Gamma_3| < \infty$. Interchanging Γ_1 and Γ_3, we have $|\Gamma_1 : \Gamma_1 \cap \Gamma_3| < \infty$, completing the proof of transitivity.

Supposing Γ to be a discrete subgroup of some Lie group G, it will often be the case that for some $g \in G$ we find that Γ and $g^{-1}\Gamma g$ are commensurable. With this in mind, we define the *commensurator* of Γ in G to be the set

$$\mathrm{Comm}_G(\Gamma) = \{g \in G \mid \Gamma \text{ is commensurable with } g^{-1}\Gamma g\}.$$

Note that this set actually forms a subgroup of G, since commensurability is an equivalence relation. Indeed, if $g \in \mathrm{Comm}_G(\Gamma)$ and $h \in \mathrm{Comm}_G(\Gamma)$, then — denoting commensurability by $\sim$ for brevity — we have $\Gamma \sim g^{-1}\Gamma g$ and $\Gamma \sim h^{-1}\Gamma h$, so that

$$h^{-1}g^{-1}\Gamma gh \sim h^{-2}\Gamma h^2 \sim h^{-1}\Gamma h \sim \Gamma.$$

Hence the product gh lies in $\mathrm{Comm}_G(\Gamma)$.

§4.2 Arithmetic Groups
and Congruence Subgroups

Let G be an algebraic group defined over an alegbraic extension $K/\mathbb{Q}$. (So we are considering G to be a Zariski-closed subgroup of some $\mathrm{GL}_n(\mathbb{C})$.) A subgroup $\Gamma \subseteq G$ is said to be *arithmetic* if it is commensurable with the group of K-integral points $G(\mathcal{O}_K)$ (cf. §4.1). The group of K-integral points $G(\mathcal{O}_K)$ will depend on the embedding of G in $\mathrm{GL}_n(\mathbb{C})$, but we do have the following [PR94, Prop. 4.1]:

PROPOSITION 4.2. *Let* $\phi\colon G \to G'$ *be an isomorphism of algebraic groups, with the groups and the morphism* ϕ *defined over* $\mathbb{Q}$. *Suppose* $\Gamma \leqslant G$ *is an arithmetic subgroup. Then* $\phi(\Gamma)$ *is an arithmetic subgroup of* G'.

This proposition only applies directly to groups defined over $\mathbb{Q}$ (in which case $\mathcal{O}_K = \mathbb{Z}$), but as Platonov and Rapinchuk point out [PR94, p. 175], one can simply use restriction of scalars (cf. p. 11) to obtain a group defined over $\mathbb{Q}$.

EXAMPLE 4.3. A standard example of an arithmetic group is the group $\mathrm{SL}_2(\mathbb{Z})$ of integral matrices that lie in the group $\mathrm{SL}_2(\mathbb{R})$ of 2×2 matrices of determinant 1. We note that $\mathrm{SL}_2(\mathbb{Z})$ is a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$, and we saw on p. 33 that it is an example of a lattice in $\mathrm{SL}_2(\mathbb{R})$.

Congruence subgroups

The example of an arithmetic group given above is a fairly trivial one, for it is the group $G(\mathbb{Z})$ and is trivially commensurable with itself. It is often desirable to obtain groups that are commensurable with, but not equal to, $G(\mathbb{Z})$, and one extremely useful means of obtaining such groups is by reduction modulo some ideal in a number field. To be more precise, suppose that we have an algebraic group G defined over an algebraic extension $K/\mathbb{Q}$. The group $G(\mathcal{O}_K)$ is an arithmetic group. As mentioned above, we regard G as a Zariski-closed subgroup of some $\mathrm{GL}_n(\mathbb{C})$, and so have a concrete realisation of $G(\mathcal{O}_K)$ as a matrix group. Now let $\mathfrak{a}$ be an ideal in $\mathcal{O}_K$. If $m, n \in \mathcal{O}_K$ then we say that m

is *congruent* to n modulo $\mathfrak{a}$, and write $m \equiv n \pmod{\mathfrak{a}}$, if $m - n \in \mathfrak{a}$. We can define the *principal congruence subgroup* $G(\mathcal{O}_K)(\mathfrak{a})$ by

$$G(\mathcal{O}_K)(\mathfrak{a}) = \{M \in G(\mathcal{O}_K) \mid M_{ij} \equiv \delta_{ij} \pmod{\mathfrak{a}}\} \tag{4.7}$$

where δ_{ij} is the Kronecker delta. Thus $G(\mathcal{O}_K)(\mathfrak{a})$ is the kernel of the homomorphism $G(\mathcal{O}_K) \to G(\mathcal{O}_K/\mathfrak{a})$. By the remarks on p. 4 we find that this kernel must have finite index in $G(\mathcal{O}_K)$ if $\mathfrak{a}$ is non-zero, since $\mathcal{O}_K/\mathfrak{a}$ must be finite, which implies that $G(\mathcal{O}_K/\mathfrak{a})$ must be finite. It is well-known that the kernel of a group homomorphism is a normal subgroup in the domain of the map, and so the principal congruence subgroups defined here are normal in $G(\mathcal{O}_K)$. In general a *congruence subgroup* of an arithmetic group is one that contains a principal congruence subgroup.

It has already been remarked (cf. p. 31) that principal congruence subgroups are torsion-free for the appropriate choice of $\mathfrak{a}$. In fact, for all but finitely many prime ideals $\mathfrak{p}$, one finds that the groups $G(\mathcal{O}_K)(\mathfrak{p})$ are torsion-free [Mil76, p. 239]. Therefore, when constructing hyperbolic manifolds, one often chooses a congruence subgroup $\Gamma(\mathfrak{a})$ of some arithmetic lattice Γ such that $\Gamma(\mathfrak{a})$ is torsion-free, and then invokes Theorem 3.3.

§4.3   ARITHMETIC SUBGROUPS OF Isom($\mathbb{H}^n$)

The isometry groups of hyperbolic spaces are closely related to linear algebraic groups and so it is natural to carry over the notion of arithmeticity to these groups. There are several major results concerning arithmetic subgroups of semisimple Lie groups, which will be outlined below.

Switching from the algebraic group point of view to the Lie group one, we may regard arithmetic subgroups as discrete subgroups of Lie groups. In this case, one asks in the first instance whether or not these groups are in fact lattices, and in the second whether or not they are co-compact. The following result of Borel and Harish-Chandra, also given by Mostow and Tamagawa [MT62], is fundamental:

THEOREM 4.4. [BHC62, Theorem 7.8] *Let G be a semisimple algebraic Lie group, defined over $\mathbb{Q}$. Then $G(\mathbb{Z})\backslash G(\mathbb{R})$ has finite volume; i.e., $G(\mathbb{Z})$ is a lattice in $G(\mathbb{R})$.*

Thus we see that we have a source of examples of lattices in semisimple Lie groups, and in particular in $\mathrm{Isom}(\mathbb{H}^n)$. Co-compact lattices are characterised in the following way (cf. Witte Morris [WM08, Prop. 5.30]):

THEOREM 4.5 (GODEMENT COMPACTNESS CRITERION). *Suppose that G is a semisimple Lie group with finite centre, and defined over $\mathbb{Q}$. Then $G(\mathbb{Z})\backslash G$ is compact if and only if $G(\mathbb{Z})$ contains no non-trivial unipotent elements.*

(Unipotency is explained on p. 12.) Again, this deals with groups defined over $\mathbb{Q}$, but by restriction of scalars one may consider groups defined over algebraic number fields, although the assumption that G has no compact factors is required [WM08, Theorem 5.31].

Note that whilst $\mathbb{Z}$ is discrete in $\mathbb{R}$, the ring of integers $\mathcal{O}_K$ (of some number field K of degree $d \geqslant 2$) need not be. Thus, Theorem 4.4 (along with restriction of scalars) is important in that it tells us we do in fact have a discrete group in the K-integral points (cf. §4.4). On the other hand, it might not be obvious that the set of (K-)integral points of an algebraic group is infinite, which it would need to be in a non-compact setting if it were to constitute a lattice.

We will be concerned mainly with orthogonal groups of quadratic forms with coefficients in number fields. In this setting, one has the following characterisation:

PROPOSITION 4.6. *Suppose that f is a quadratic form in more than two variables with coefficients in an alegbraic number field K. The lattice $PO_f(\mathcal{O}_K)$ is co-compact in $\mathrm{PO}_f(\mathbb{R})$ if and only if the form f is not isotropic.*

(This is stated in Platonov and Rapinchuk [PR94, p. 212] and Witte Morris [WM08, Prop. 5.32] for quadratic forms over $\mathbb{Q}$.) This proposition is related to Theorem 4.5, and follows from it when one shows that unipotent elements only arise from isotropic quadratic forms [Cas08, p. 300]. The result also follows

from the so-called Mahler Compactness Criterion [WM08, Prop. 4.34; cf. also pp. 79–80] [PR94, Prop. 4.8].

It is not *a priori* clear whether or not every lattice in a semisimple Lie group might be arithmetic, so we make the following definition, making use of a homomorphism from an algebraic group onto a Lie group [Zim84, p. 114] [MR03, p. 316]: suppose that G is a connected semisimple Lie group with no compact factors. A lattice Γ in G is called an *arithmetic lattice* if there is a surjective homomorphism $f \colon H(\mathbb{R})^0 \to G$ such that

1. H is an algebraic group defined over $\mathbb{Q}$;

2. $\ker f$ is compact; and

3. $f\big(H(\mathbb{Z}) \cap H(\mathbb{R})^0\big)$ is commensurable with Γ.

Note that the notation $H(\mathbb{R})$ refers to the (Lie) group of real points of the (algebraic) group H. The superscript $\cdot^0$ denotes the identity component.

It turns out, after all, that lattices in Isom($\mathbb{H}^n$) are not always arithmetic:

THEOREM 4.7 (GROMOV-PIATETSKI-SHAPIRO [GPS87]). *For any* $n \geqslant 2$, *there exists a non-arithmetic lattice in* PO($n, 1$).

(For a description of the proof of this theorem, see §8.1.)

Before turning to the theorem of Margulis, we briefly recall some definitions concerning lattices in Lie groups. Supposing G to be an algebraic group, a *torus* is a connected diagonalisable subgroup T of G, where by 'diagonalisable' is meant that there is an injective homomorphism $\phi \colon G \to \mathrm{GL}_m(\mathbb{C})$ such that $\phi(T)$ consists of diagonal matrices. A torus is $\mathbb{R}$-*split* if g actually has real entries (i.e., $g \in \mathrm{GL}_m(\mathbb{R})$). Now, the R-*rank* of G is the dimension of a maximal R-split torus in G.

The situation described in Theorem 4.7 does not occur for groups of higher rank, it turns out:

THEOREM 4.8 (MARGULIS). [Mar91, Theorem A][Zim84, Theorem 6.1.2] *Let* G *be a connected semisimple real Lie group of rank at least* 2, *with trivial centre and no compact factors. Then any irreducible lattice in* G *is arithmetic.*

By an *irreducible lattice* Γ in G is meant one such that ΓN is dense in G for every non-compact closed normal subgroup N of G [WM08, 4.23]. Raghunathan shows that if G is connected, semisimple and without compact factors, then this is equivalent to saying that for any proper connected normal subgroup $H \lhd G$, the intersection $H \cap \Gamma$ is not a lattice in Γ [Rag72, 5.21].

Theorem 4.8 (in light of Theorem 4.7) marks out the case of rank 1 as one for investigation, and it is here that much attention is focussed in the literature. As well as the orthogonal groups, one also has the unitary groups $\mathrm{SU}(n,1)$ (for $n \geqslant 2$), the symplectic groups $\mathrm{Sp}(n,1)$ (for $n \geqslant 2$), and an exceptional group F_4^{-20} [WM08, §7F]. K. Corlette proved the 'superrigidity' of lattices in the groups $\mathrm{Sp}(n,1)$ and F_4^{-20} [Cor92], which implies arithmeticity of these lattices (cf. §12C of Witte Morris [WM08]). This leaves the question of whether or not there exist any non-arithmetic lattices in $\mathrm{SU}(n,1)$ for any $n \geqslant 2$. This is not known for $n \geqslant 4$, but examples do exist in dimensions 2 and 3, notably those given by Deligne and Mostow [Mos78] [DM93]. A survey on lattices in complex hyperbolic space has been given by J. Parker [Par09].

§4.4   EXAMPLES OF ARITHMETIC LATTICES IN $\mathrm{Isom}(\mathbb{H}^n)$

The simplest example of an arithmetic lattice in $\mathrm{Isom}(\mathbb{H}^n)$ is one obtained by taking the integral points of an appropriate algebraically defined subgroup.

Let $f \colon \mathbb{R}^{n+1} \to \mathbb{R}$ be a quadratic form of signature $(n,1)$, with coefficients in an algebraic number field K, such that every conjugate form f^σ (for $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$) is positive definite. (For example, one might take a form such as $f = x_1^2 + \cdots + x_n^2 - \sqrt{3}x_{n+1}^2$, defined over $\mathbb{Q}(\sqrt{3})$.) Then the group O_f is an algebraic group defined over K, and it is isomorphic to the Lie group $\mathrm{O}(n,1)$. The group PO_f is then isomorphic to $\mathrm{PO}(n,1)$ and so may be identified with $\mathrm{Isom}(\mathbb{H}^n)$. Now, by definition, $\mathrm{O}_f(\mathcal{O}_K)$ is an algebraic subgroup of $\mathrm{O}_f(\mathbb{R})$, and by restriction of scalars we could view this as defined over $\mathbb{Q}$. Indeed, using (1.7) we find that $\mathrm{O}_f(\mathcal{O}_K)$ is an arithmetic lattice in

$\prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} \mathrm{O}_{f^\sigma}$ (cf. Morris [WM08, 5.45, p. 86]). However since each f^σ is positive definite for $\sigma \neq \mathbf{id}$, we find that all the factors corresponding to $\sigma \neq \mathbf{id}$ are compact groups. Therefore, projecting onto the first factor leaves us with an arithmetic lattice in O_f. There is an isomorphism $\phi \colon \mathrm{O}_f(\mathbb{R}) \to \mathrm{O}(n,1)$, and $\phi\big(\mathrm{O}_f(\mathcal{O}_K)\big)$ is a discrete group in $\mathrm{O}(n,1)$. Composing ϕ with the natural projection $\pi \colon \mathrm{O}(n,1) \to \mathrm{PO}(n,1)$, we have an arithmetic lattice in $\mathrm{PO}(n,1)$.

The question of whether or not $\mathrm{O}_f(K)$ is co-compact has been examined above (cf. Proposition 4.6) and, combining that discussion with Theorem 1.6, one finds that for $n \geqslant 4$, the group $\mathrm{O}_f(\mathcal{O}_K)$ is co-compact if and only if $[K : \mathbb{Q}] \geqslant 2$. In lower dimensions, it is still true that for $[K : \mathbb{Q}] \geqslant 2$ the form f is non-isotropic, but for $K = \mathbb{Q}$ one must examine each form individually. For example, the form $x_1^2 + x_2^2 - x_3^2$ is defined over $\mathbb{Q}$ and has the vector $x = (3, 4, 5)$ as a non-trivial solution to $f(x) = 0$. Thus the lattice associated to this quadratic form is non-co-compact. On the other hand, the form $f = x_1^2 + x_2^2 - 3x_3^2$ has no non-trivial solutions (in $\mathbb{Q}^3$) to the equation $f(x) = 0$, and so it will produce a co-compact lattice.

A GEOMETRIC EXAMPLE

A specific example of an arithmetic lattice in $\mathrm{Isom}(\mathbb{H}^n)$ will be of interest to us later, and so it is presented here.

Let ϕ denote the algebraic integer $\frac{1}{2}(1 + \sqrt{5})$ (arising from the polynomial $x^2 - x - 1$), and let q_4 be the quadratic form $-\phi x_0^2 + x_1^2 + \cdots + x_4^2$, which is defined over $\mathbb{Q}(\sqrt{5})$. Now $\mathrm{PO}_{q_4}(\mathbb{R}) \cong \mathrm{Isom}(\mathbb{H}^4)$, and in the same manner as above, $\mathrm{PO}_{q_4}(\mathcal{O}_{\mathbb{Q}(\sqrt{5})})$ is a co-compact arithmetic lattice in $\mathrm{PO}_{q_4}(\mathbb{R})$.

From a geometric point of view, this lattice is interesting since it is commensurable with a group generated by reflections in the faces of the 120-cell in $\mathbb{H}^4$ [ALR01, Lem. 3.3]. The 120-cell is an object of much interest from various geometrical viewpoints, and the article of J. Stillwell [Sti01] makes for an excellent overview of the object's properties from these different perspectives.

§4.5   Some properties of arithmetic lattices

Arithmetic lattices provide a convenient way of obtaining finite-volume orbifolds and manifolds, but are interesting objects in themselves, and in particular from a group-theoretic point of view. It is natural to ask whether or not a group is finitely presented, and in turns out that the answer to this in the case of arithmetic groups is affirmative [PR94, Theorem 4.2, p. 195]:

Theorem 4.9. *Let Γ be an arithmetic subgroup of an algebraic group G defined over $\mathbb{Q}$. Then Γ is finitely presented.*

In proving this theorem one often considers fundamental domains for Γ, and in particular one may show that Γ has an open, connected fundamental domain [WM08, 4.57]. It can then be shown that this implies that Γ is finitely generated, and moreover finitely presented [WM08, 4.54]. The proof given by Platonov and Rapinchuk considers an open connected set $\Omega \subset G/K$ (where K is a maximal compact subgroup of G and G/K is the symmetric space associated to G), such that $\Gamma\Omega = G/K$, and such that $\Delta = \{\gamma \in \Gamma \mid \gamma(\Omega) \cap \Omega = \emptyset\}$. The existence of such a set is given by the results of Borel and Harish-Chandra. The set Δ is a set of generators for the group Γ [PR94, Lemma 4.9]. For the generators, one considers the free group F_Δ on the elements of Δ, and considers the natural homomorphism $f \colon F \to \Gamma$. Relators for Γ are then given by $F(\delta_i)F(\delta_j)F(\delta_i\delta_j)^{-1}$ for every i, j for which $\delta_i\delta_j \in \Delta$.

The notion of the commensurator of a subgroup of some ambient group G has already been introduced (cf. p. 39), and this turns out to be of great interest in the theory of arithmetic groups, because of the following dichotomy:

Theorem 4.10 (Margulis). [MR03, Theorem 10.3.5, p. 318][Mar91, Theorem B, p. 298] *Suppose that G is a connected semisimple Lie group with trivial centre and no compact factors, and that Γ is an irreducible lattice in G. Exactly one of the following two cases occurs:*

1. *Γ is a finite-index subgroup of $\mathrm{Comm}_G(\Gamma)$, and Γ is non-arithmetic; or*

2. *$\mathrm{Comm}_G(\Gamma)$ is dense in G, and Γ is arithmetic.*

Thus we arrive at a characterisation of arithmeticity of lattices in semisimple Lie groups based on group-theoretic properties of the lattice.

5

HYPERPLANE AND SUBGROUP SEPARABILITY

Suppose that H_0 and H_1 are hyperplanes in $\mathbb{H}^n$. If M is a hyperbolic manifold given as a quotient $\Gamma \backslash \mathbb{H}^n$ by a discrete group Γ, then each of H_0 and H_1 project under the quotient map to immersed totally geodesic submanifolds of M (cf. p. 24). In general the images of H_0 and H_1 may intersect even when H_0 and H_1 do not, and moreover each of the submanifolds may not be embedded (i.e., they may have self-intersections). Nevertheless it is sometimes possible to find a finite cover of M such that the images of H_0 and H_1 are both embedded, and do not intersect each other if H_0 and H_1 are disjoint in $\mathbb{H}^n$.

We will examine two approaches to finding such covers, one being based on arguments concerning congruence covers and some hyperbolic geometry, the other relating to so-called separability properties of the fundamental groups of the manifolds in question.

§5.1 CONGRUENCE COVERS AND HYPERPLANES

Suppose that f is a quadratic form over a totally real number field K, having signature $(n, 1)$, with positive definite conjugate forms f^σ for each $\sigma \in$

$\mathrm{Gal}(K/\mathbb{Q}) \smallsetminus \{\mathbf{id}\}$ (cf. p. 8) . (We allow the possibility of $K = \mathbb{Q}$, whereupon $\mathrm{Gal}(K/\mathbb{Q})$ is trivial and there are no conjugate forms.) As in §4.4, the group $\Gamma = \mathrm{PO}_f(\mathcal{O}_K)$ is an arithmetic lattice in $\mathrm{PO}_f(\mathbb{R})$, and it is co-compact if and only if the form f is not isotropic over K. By Meyer's Theorem (Theorem 1.6), if $n + 1 \geqslant 5$ then the form will be isotropic if it is defined over $\mathbb{Q}$; but in lower dimensions one finds that groups defined by rational forms may be either co-compact or non-co-compact (cf. §4.4). If the form f is defined over a number field of degree at least 2 then the lattice so obtained will be co-compact.

THEOREM 5.1 (BELOLIPETSKY-THOMSON, 2011). *Suppose that f and K are as above, and let $H_0, \dots, H_k$ be pairwise disjoint hyperplanes in $\mathbb{H}^n$ given by $H_i = \langle e_i \rangle^{\perp_f} \cap \mathbb{H}^n$, for some $e_i \in K^{n+1}$ $(i = 0, \dots, k)$. Then there exists a finite-index subgroup $\Gamma' < \mathrm{PO}_f(\mathcal{O}_K)$ such that for every $h \in \Gamma'$*

$$either \quad h(H_0) = H_0 \quad or \quad h(H_0) \cap (H_0 \cup H_1) = \emptyset. \qquad (5.1)$$

This theorem was given for $K = \mathbb{Q}$ by Margulis and Vinberg [MV00] and again by Kapovich, Potyagailo and Vinberg [KPV08]. The extension to the case $K \neq \mathbb{Q}$ (i.e., $d \geqslant 2$) is necessary to produce co-compact groups Γ' satisfying (5.1). The proof in the case $K = \mathbb{Q}$ is relatively straightforward, owing to the fact that $\mathcal{O}_K = \mathbb{Z}$ is discrete in $\mathbb{R}$. This discreteness is not manifest if $K \neq \mathbb{Q}$, and so the norm on the field extension is used in order to work with integers. (Examination of the following proof indicates that the case $|\alpha| < 1$ (p. 52) does not occur when $K = \mathbb{Q}$ and so the result still follows from this proof.)

PROOF OF THEOREM 5.1. First of all, let us assume without loss of generality that the form f has coefficients in $\mathcal{O}_K$ and that $e_i \in \mathcal{O}_K^{n+1}$ for each i (cf. §2.5). (If f does not have coefficients in $\mathcal{O}_K$ then we can multiply f by some suitable constants until it does, and this will give an equivalent form as in §1.3, and the arguments in §2.3 apply.) For brevity we will write $(\cdot, \cdot)$ for $(\cdot, \cdot)_f$. Thus, if $h \in \Gamma$, then $\big(h(e_0), e_i\big) \in \mathcal{O}_K$.

Throughout this proof, inner products and orthogonal complements are understood to be with respect to f as in (1.3).

Assume first that $k = 1$.

Let $\mathfrak{p}$ be the principal ideal generated by the K-integer β:

$$\mathfrak{p} = (\beta) \subseteq \mathcal{O}_K, \qquad \text{where } \beta = 2C(e_0, e_1), \tag{5.2}$$

and where $C > 1$ is an integer determined by (5.10) (cf. p. 53); and let Γ_1 denote the principal congruence subgroup $\Gamma(\mathfrak{p}) < \Gamma$ as defined in (4.7). (Here we use the fact that $\mathrm{PO}_f(\mathbb{R})$ can be identified with the matrix group $\mathrm{O}_f(\mathbb{R})_0$, the subgroup of the orthogonal group $\mathrm{O}_f(\mathbb{R})$ which preserves the upper half-space (cf. §2.2).) Hence, for $h \in \Gamma_1 = \Gamma(\mathfrak{p})$, we have $h \equiv \mathbf{id} \pmod{\mathfrak{p}}$, so that

$$\big(h(e_0), e_1\big) = \big(e_0, e_1\big) + \alpha\beta \tag{5.3}$$

for some $\alpha \in \mathcal{O}_K$ (where α depends on h). We wish to show that for every $h \in \Gamma_1$ we have $h(H_0) \cap H_1 = \emptyset$.

To be able to examine the intersections of the hyperplanes we use Theorem 2.3, by which we find that hyperplanes defined by vectors v_0 and v_1 are disjoint in $\mathbb{H}^n$ if

$$\big|(v_0, v_1)\big| \geqslant \sqrt{(v_0, v_0)(v_1, v_1)}, \tag{5.4}$$

where $|\cdot|$ denotes the standard absolute value on $\mathbb{R}$. Note that (5.4) is an equality if the two hyperplanes coincide.

If $\alpha = 0$ in (5.3) then

$$\big|(h(e_0), e_1)\big| = \big|(e_0, e_1)\big| \geqslant \sqrt{(e_0, e_0)(e_1, e_1)} = \sqrt{\big(h(e_0), h(e_0)\big)(e_1, e_1)} \tag{5.5}$$

(where the inequality follows from (5.4) and the initial condition $H_0 \cap H_1 = \emptyset$), and hence the hyperplanes $h(H_0)$ and H_1 are either disjoint or equal. We will eliminate the possibility of equality later in the proof.

If $|\alpha| \geqslant 1$, then

$$\begin{aligned}
\big|(h(e_0), e_1)\big| &= \big|(e_0, e_1) + \alpha\beta\big| \\
&\geqslant \big||\alpha||\beta| - |(e_0, e_1)|\big| && \text{(by the triangle inequality for } |\cdot|) \\
&= \big|(e_0, e_1)\big| \cdot \big|2C|\alpha| - 1\big| \\
&> \big|(e_0, e_1)\big| && \text{(since } C > 1).
\end{aligned}$$

The case $|\alpha| < 1$ requires consideration of the norm $N(x)$ of x as defined in §1.1. The motivation for this comes from Lemma 1.2: if $x \in \mathcal{O}_K$ then $N(x) \in \mathbb{Z}$ so that $|N(x)| \geqslant 1$ for $x \in \mathcal{O}_K^*$ (cf. p. 2). Since $\alpha \in \mathcal{O}_K^*$ and $|\alpha| < 1$, the definition of the norm (as the product of conjugates) implies that $|\alpha^{\sigma_j}| > 1$ for some $j \in \{2, \ldots, d\}$. For this j, we get

$$
\begin{aligned}
\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right| &= \left| (e_0, e_1)^{\sigma_j} + \alpha^{\sigma_j} \beta^{\sigma_j} \right| \\
&\geqslant \left| |\alpha^{\sigma_j}||\beta^{\sigma_j}| - |(e_0, e_1)^{\sigma_j}| \right| = |\beta^{\sigma_j}| \left| |\alpha^{\sigma_j}| - \frac{1}{2C} \right| \\
&> \tfrac{1}{2} |\beta^{\sigma_j}| = C \left| (e_0, e_1)^{\sigma_j} \right|.
\end{aligned}
$$

We have

$$
\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right| > C \left| (e_0, e_1)^{\sigma_j} \right| = C \left| (e_0, e_1)^{\sigma_j} \right| \frac{\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right|}{\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right|}. \tag{5.6}
$$

Now since $(\cdot, \cdot)_{f^{\sigma_j}}$ is positive definite (for $j \geqslant 2$), the Cauchy-Schwarz inequality

$$
\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right| \leqslant \sqrt{\left(h(e_0), h(e_0) \right)^{\sigma_j} \left(e_1, e_1 \right)^{\sigma_j}}
$$

applies and we can use it to bound the denominator from above. Thus (5.6) becomes

$$
\left| \left(h(e_0), e_1 \right)^{\sigma_j} \right| > C \frac{\left| (e_0, e_1)^{\sigma_j} \right| \left| \left(h(e_0), e_1 \right)^{\sigma_j} \right|}{\sqrt{\left(h(e_0), h(e_0) \right)^{\sigma_j} \left(e_1, e_1 \right)^{\sigma_j}}} = C \frac{\left| (e_0, e_1)^{\sigma_j} \right| \left| \left(h(e_0), e_1 \right)^{\sigma_j} \right|}{\sqrt{(e_0, e_0)^{\sigma_j} (e_1, e_1)^{\sigma_j}}},
$$

where for the equality we note that h is an isometry of the quadratic space on which $(\cdot, \cdot)$ is defined (cf. §1.2). Multiplying each side of this inequality by all the $\left| \left(h(e_0), e_1 \right)^{\sigma_k} \right|$ for which $k \neq j$ gives

$$
\left| N\left(\left(h(e_0), e_1 \right) \right) \right| > C \underbrace{\left| N\left(\left(h(e_0), e_1 \right) \right) \right|}_{(*)} \frac{\left| (e_0, e_1)^{\sigma_j} \right|}{\sqrt{(e_0, e_0)^{\sigma_j} (e_1, e_1)^{\sigma_j}}}. \tag{5.7}
$$

We can replace $(*)$ by $\left| N\left((e_0, e_1) \right) \right|$, for

$$
\begin{aligned}
\left| N\left(\left(h(e_0), e_1 \right) \right) \right| &= \left| N\left((e_0, e_1) + 2C\alpha (e_0, e_1) \right) \right| \\
&= \left| N\left((e_0, e_1) \right) \right| \cdot \left| N(1 + 2C\alpha) \right| \quad \text{(by Lemma 1.2)} \\
&\geqslant \left| N\left((e_0, e_1) \right) \right| \quad\quad\quad\quad\quad \text{(since } 1 + 2C\alpha \in \mathcal{O}_K^* \text{).}
\end{aligned}
$$

$$\tag{5.8}$$

Writing both norms in (5.7) as products of Galois conjugates, (5.7) and (5.8) give

$$\prod_{i=1}^{d}\left|(h(e_0),e_1)^{\sigma_i}\right| > C\left(\prod_{i=1}^{d}\left|(e_0,e_1)^{\sigma_i}\right|\right)\frac{\left|(e_0,e_1)^{\sigma_j}\right|}{\sqrt{(e_0,e_0)^{\sigma_j}(e_1,e_1)^{\sigma_j}}},$$

so that by rearranging,

$$\left|(h(e_0),e_1)\right| > C\frac{\left|(e_0,e_1)^{\sigma_j}\right|}{\sqrt{(e_0,e_0)^{\sigma_j}(e_1,e_1)^{\sigma_j}}}\left(\prod_{i=2}^{d}\frac{\left|(e_0,e_1)^{\sigma_i}\right|}{\left|(h(e_0),e_1)^{\sigma_i}\right|}\right)\left|(e_0,e_1)\right|.$$

Applying the Cauchy-Schwarz inequality to the denominator of the product gives the final estimate

$$\left|(h(e_0),e_1)\right| \geqslant C\frac{\left|(e_0,e_1)^{\sigma_j}\right|}{\sqrt{(e_0,e_0)^{\sigma_j}(e_1,e_1)^{\sigma_j}}}\left(\prod_{i=2}^{d}\frac{\left|(e_0,e_1)^{\sigma_i}\right|}{\sqrt{(e_0,e_0)^{\sigma_i}(e_1,e_1)^{\sigma_i}}}\right)\left|(e_0,e_1)\right|. \tag{5.9}$$

At this point we see that having chosen C to be sufficiently large we have ensured that $\left|(h(e_0),e_1)\right| > \left|(e_0,e_1)\right|$. Notice that since α depends on h, so too does j; however, C is independent of h having assumed

$$C \geqslant \prod_{\sigma_j\in S}\frac{(e_0,e_0)^{\sigma_j}(e_1,e_1)^{\sigma_j}}{[(e_0,e_1)^{\sigma_j}]^2}, \tag{5.10}$$

where $S \subseteq \{\sigma_2,\ldots,\sigma_d\}$ is the set of all σ_j for which the corresponding factors in (5.10) are greater than 1.

Thus we get $\left|(h(e_0),e_1)\right| \geqslant \left|(e_0,e_1)\right| \geqslant \sqrt{(h(e_0),h(e_0))(e_1,e_1)}$ as in the other two cases for α. This means that $h(H_0)$ either coincides with, or does not intersect H_1.

To avoid the possibility of $h(H_0)$ coinciding with H_1, we have to ensure that $h(e_0) \neq \pm\omega e_1$ for some $\omega \in \mathbb{R}_{>0}$. If it exists, then this ω would be given by $\omega = \sqrt{(e_0,e_0)/(e_1,e_1)}$ and there are two possible cases:

(i) $\omega \notin K$, whence $h(e_0) = \pm\omega e_1$ is impossible.

(ii) $\omega \in K$.

For the second case, let e_1' be the vector obtained by scaling ωe_1 by $\sqrt{(e_0,e_0)}\cdot\sqrt{(e_1,e_1)}$, so that $e_1' = (e_0,e_0)e_1$. Similarly, define $e_0' = \sqrt{(e_0,e_0)(e_1,e_1)}\,e_0$.

Thus we have $h(e_0) = \pm\omega e_1$ if and only if $h(e_0') = \pm e_1'$. Lemma 1.3 shows that e_0' and e_1' are in fact in $\mathcal{O}_K^{n+1}$.

Now for either of the equalities $h(e_0') = \pm e_1'$ to hold we must have $e_0' + v = \pm e_1'$ for some v, where $v \equiv 0$ modulo $\mathfrak{p}$. If $e_0' + e_1'$ and $e_0' - e_1'$ are not congruent to 0 modulo $\mathfrak{p}$, then this coincidence will not occur, and we can ensure this by choosing C sufficiently large.

It remains to check that $h(H_0)$ and H_0 either coincide or are disjoint. One can repeat all of the above argument as far as (5.9), with e_0 in place of e_1, and we find that the ideal $\mathfrak{p}' = 2(e_0, e_0)$ actually suffices in place of $\mathfrak{p}$ to ensure that we have $|(h(e_0), e_0)| \geqslant \sqrt{(h(e_0), h(e_0))(e_0, e_0)}$ for every $h \in \Gamma(\mathfrak{p}')$. Denote $\Gamma(\mathfrak{p}')$ by Γ_0.

If $k \geqslant 2$, then to separate all hyperplanes we apply the above argument to all other e_i, $(i = 2, \ldots, k)$ so that we get $\Gamma_2, \ldots, \Gamma_k$ which are also finite-index subgroups of Γ. The group $\Gamma' = \Gamma_0 \cap \Gamma_1 \cap \cdots \cap \Gamma_k$ will then satisfy the conclusion of the theorem, and is still of finite index in Γ (cf. Lemma 4.1).

Remark. If we assume that the hyperplanes $H_0, \ldots, H_k$ are not only disjoint but also do not meet at infinity then the inequality in (5.5) becomes strict and the coincidence of $h(H_0)$ and H_i (for $i = 1, \ldots, k$) is automatically avoided. $\quad\square$

Remark. One may prove Theorem 5.1 for the case of complex hyperbolic space; however it could not assist in proving a complex version of Theorem 6.1, due to the absence of real co-dimension-1 hypersurfaces in complex hyperbolic space. (See also the remarks on p. 68 and p. 85.)

§5.2   SUBGROUP SEPARABILITY AND
THE LERF PROPERTY

Let G be a group. We say that G is *residually finite* if for every non-trivial $g \in G$, there exists $G_1 \leqslant G$ with $|G : G_1| < \infty$ and $g \notin G_1$. Now let $H \leqslant G$ be a subgroup. By definition the statements 'G is *H-separable*', 'G is *H-residually*

finite', or 'H is *separable* in G' are all taken to be synonymous and mean that for every $g \in G \smallsetminus H$, there is $K \leqslant G$ such that

$$|G : K| < \infty, \qquad H \leqslant K, \quad \text{and } g \notin K.$$

Additionally, we say that G is LERF (for 'locally extended residually finite') if G is H-separable for every finitely generated $H \leqslant G$, and that G is *subgroup separable* if it is H-separable for *all* $H \leqslant G$.

The term 'H-residually finite' appears in Scott's 1978 article [Sco78] but the more recent literature [Ago06, ALR01] uses 'H-separable'. An alternative view which we can see to be equivalent is used by authors such as Agol [Ago06]:

LEMMA 5.2. *G is H-separable if and only if*

$$H = \bigcap_{\substack{H \leqslant B \leqslant G \\ |G:B| < \infty}} B. \tag{5.11}$$

PROOF. (Only if): Assume G is H-separable, and let $g \in H^{\complement}$ (where $\cdot^{\complement}$ denotes the set-theoretic complement in G). Denote by $\mathcal{B}$ the set of finite index subgroups of G that contain H. Now there exists $B \in \mathcal{B}$ with $g \notin B$ (by the separability assumption), so $g \in \left(\cap_{B \in \mathcal{B}} B \right)^{\complement}$. Conversely let $g \in \left(\cap_{B \in \mathcal{B}} B \right)^{\complement}$. Then for some $B \in \mathcal{B}$, we have $g \notin B$, and so $g \in H^{\complement}$. This establishes the equality in (5.11).

(If): Suppose that (5.11) holds and let $g \in H^{\complement}$. Then we must find that there exists $B \in \mathcal{B}$ with $g \notin B$; otherwise $g \in H$. So, H is separable in G.   □

Residual finiteness has this consequence:

LEMMA 5.3 (SCOTT [SCO78]). *Let X be a Hausdorff topological space with a regular covering $\tilde{X}$, and covering group G. Then the following are equivalent:*

1. *G is residually finite.*

2. *If $C \subset \tilde{X}$ is compact, then there exists $G_1 \leqslant G$ of finite index, such that $gC \cap C = \emptyset$ for every $g \in G_1 \smallsetminus \{1\}$.*

3. *If $C \subset \tilde{X}$ is compact then there is a finite covering X_1 of X such that C projects homeomorphically into X_1.*

Proof. (Scott)

(2)⇔(3): If G_1 is as in (2) then $X/G_1 = X_1$ works in (3). Conversely let G_1 be the covering group in (3), so that (2) is satisfied.

(1)⇒(2): The set

$$G_C = \{g \in G \mid g \cap gC \neq \emptyset\}$$

is finite (cf. (3.3) on p. 30). For each of the $g \in G_C$, there is $G_g \leqslant G$ of finite index such that $g \notin G_g$. Then if $G_1 = \cap_{g \in G_C} G_g$, we find, using Lemma 4.1, that G_1 satisfies (2).

(2)⇒(1): Let $g \in G \smallsetminus \{1\}$. Let $x \in \tilde{X}$ and $C = gx \cup x$. Then (2) gives $G_1 \leqslant G$ of finite index with $g_1(C) \cap C = \emptyset$ for all $g_1 \in G_1 \smallsetminus \{1\}$. Then $g \notin G_1$, for if it were in G_1,

$$(g^{-1}x \cup x) \cap (x \cup gx) = \emptyset. \qquad \lightning$$

Thus G is residually finite. $\qquad\square$

We will also use a corollary of the following lemma, also due to Scott. The 'only if' part of the proof is given, as it will provide the conclusion of Corollary 5.5.

Lemma 5.4 (Scott [Sco78]). *Let X be a Hausdorff topological space with regular covering $\tilde{X}$ and covering group G. Then G is LERF if and only if for every finitely generated $S \leqslant G$ and any compact subset $C \subseteq \tilde{X}/S$, there exists a finite covering X_1 of X such that the projection $p \colon \tilde{X}/S \to X$ factors through X_1 and C projects homeomorphically into X_1.*

Proof of 'only if' [Sco78]. Assume G is LERF and that $S \leqslant G$ is finitely generated; and also let $C \subseteq \tilde{X}/S$ be compact. The inverse image $p^{-1}(C)$ has a compact subset $D \subseteq p^{-1}(C)$ such that $p(D) = C$, and the set $\{g \in G \mid g(D) \cap D \neq \emptyset\}$ is finite (cf. (3.3)). Since S is separable in G there exists $G_1 \leqslant G$, of finite index, and containing S, such that whenever $g(D) \cap D \neq \emptyset$ (for $g \in G_1$) we must have $g \in S$. For, suppose that $g \in G \smallsetminus S$ is such that $g(D) \cap D \neq \emptyset$: there are only finitely many such g. By the separability of S in G one finds a finite-index subgroup $G_g \subseteq G$ such that $g \notin G_g$. The finite intersection $\cap_g G_g$ has finite index in G (by Lemma 4.1) and provides the required G_1. This gives the required finite cover $\tilde{X}/G_1$. $\qquad\square$

The following will be used in §6.4, and we use the notation of that section:

COROLLARY 5.5. *Suppose that Γ is a discrete subgroup of* $\mathrm{Isom}(\mathbb{H}^n)$, *and that* $G < \Gamma$ *is separable in* Γ. *Let* $C \subset G\backslash\mathbb{H}^n$ *be compact. Then there exists a finite-index subgroup* $\Gamma_1 \leqslant \Gamma$ *such that* C *embeds homeomorphically in the quotient* $\Gamma_1\backslash\mathbb{H}^n$.

PROOF. One follows the proof of Lemma 5.4, with Γ in place of G, and G in place of S, with the *a priori* assumption that G is separable in Γ.   □

§5.3   THE GFERF PROPERTY

We turn now to groups that are geometrically finite. If $n = 2$ or $n = 3$ then a discrete group $\Gamma < \mathrm{Isom}(\mathbb{H}^n)$ is called *geometrically finite* if it admits a convex finite-sided polyhedron as a fundamental domain. In higher dimensions this definition can be used, but it is customary to use a slight variation on the definition. A group Γ is said to be geometrically finite if it has a fundamental domain F that satisfies the following [Rat06, p. 251, p. 627, p. 637]:

1. F is a convex polyhedron;

2. for every side S of F there is $\gamma \in \Gamma$ with $S = F \cap \gamma(F)$ (i.e., F is an *exact polyhedron* for Γ);

3. for every $x \in \overline{F} \cap \partial\mathbb{H}^n$ there is a neighbourhood $U \ni x$ in $\overline{\mathbb{H}^n}$ such that U only meets the sides of F incident to x (F is a *geometrically finite polyhedron*).

In particular, finite-sided convex polyhedra are geometrically finite. Geometrical finiteness has important consequences for a group, such as its being finitely generated [Rat06, Theorem 12.4.9]. A full discussion of this is given by B. Bowditch, and he gives five equivalent definitions [Bow93, Sect. 4]. Bowditch also shows that the lattices Γ (i.e., where $\Gamma\backslash\mathbb{H}^n$ is a finite-volume orbifold) are geometrically finite [Bow93, Prop. 4.7].

We now introduce another separability notion: a group Γ is GFERF (for 'geometrically finite extended residually finite') if Γ is H-subgroup separable for every geometrically finite subgroup $H < \Gamma$. Since geometrically finite groups are finitely generated, any LERF group is also GFERF. The GFERF property has significance in the construction of short-systole manifolds, as will be seen later (cf. §6.4). Certain specific examples of discrete subgroups of $\mathrm{Isom}(\mathbb{H}^n)$ were known to be GFERF, and notably I. Agol, D. Long and A. Reid showed that the Bianchi groups and groups generated by the reflections in the faces of so-called all right polyhedra are GFERF [ALR01, Theorem 1.1, Theorem 3.1]. The following represents a well-known example of a lattice that is GFERF:

EXAMPLE 5.6. [ALR01, Theorem 3.1, Lemma 3.2] Let P be the 120-cell in $\mathbb{H}^4$ (cf. p. 45). Then the group generated by the reflections in the faces of P is GFERF. Moreover, this group is commensurable with $\mathrm{PO}_f(\mathcal{O}_{\mathbb{Q}(\sqrt{5})})$ where f is the quadratic form $x_1^2 + \cdots x_4^2 - \frac{1}{2}(1 + \sqrt{5})x_5^2$.

It is also noted in a 2008 article of M. Kapovich, L. Potyagailo and E. Vinberg that every non-co-compact lattice in $\mathrm{Isom}(\mathbb{H}^n)$ is GFERF if $n \leqslant 5$ [KPV08, Theorem C]. We record here the following recent result of N. Bergeron, F. Haglund and D. Wise, which generalises the results of Agol, Long and Reid, and Kapovich, Potyagailo and Vinberg:

THEOREM 5.7. [BHW11, Cor. 1.12] *Let Γ be an arithmetic (congruence) lattice in $\mathrm{SO}(n, 1)$. Then Γ is GFERF.*

The article of Bergeron, Haglund and Wise is concerned with generalising J. Millson's famous results on separating hypersurfaces [Mil76], but Theorem 5.7 is a corollary to some of the work there, concerning embeddings of arithmetic lattices into so-called right-angled Coxeter groups. For later reference, we give Millson's result in the form stated by Bergeron, Haglund and Wise, along with a useful lemma from Millson's article:

THEOREM 5.8. *Suppose that Γ is an arithmetic torsion-free discrete subgroup of $\mathrm{Isom}^+(\mathbb{H}^n)$, and let $\pi_\Gamma \colon \mathbb{H}^n \to \Gamma \backslash \mathbb{H}^n$ be the natural projection to the mani-*

fold $\Gamma\backslash\mathbb{H}^n$. *Suppose* $H \subset \mathbb{H}^n$ *is a hyperplane such that* $\pi_\Gamma(H)$ *is an immersed totally geodesic compact submanifold of* $\Gamma\backslash\mathbb{H}^n$. *Then there is a subgroup* $\Gamma_1 \leqslant \Gamma$ *of finite index, such that*

1. *the image* $\pi_{\Gamma_1}(H)$ *is an embedded submanifold of* $\Gamma_1\backslash\mathbb{H}^n$ *and*
2. $[\pi_{\Gamma_1}(H)] \neq 0$ *in* $H_{n-1}(\Gamma_1\backslash\mathbb{H}^n)$.

Here, $H_{n-1}(M)$ denotes the $(n-1)$-th homology group of M. A lemma, to be found in Millson's article, illustrates the consequences of having trivial (or non-trivial) homology class:

LEMMA 5.9. *Let S be an oriented totally geodesic co-dimension* 1 *submanifold of a connected orientable n-dimensional Riemannian manifold. Then* $[S] = 0$ *in* $H_{n-1}(M)$ *if and only if S separates M into two parts.*

This lemma therefore characterises in an intuitive manner the trivial $(n-1)$-homology classes.

6

Existence of hyperbolic manifolds with short closed geodesics

The notions of 'systole' and 'systole length' for a compact Riemannian manifold are defined below (cf. §6.1). In this chapter is outlined a construction by which one can produce a closed hyperbolic n-manifold with a systole as short as desired. These are interesting examples since they represent extremal cases of spaces studied in systolic geometry.

A good survey of what is known in systolic geometry is provided by the book of M. Katz [Kat07], and another is given by M. Gromov [Gro96].

§6.1   Systoles of Riemannian Manifolds

Let M be a non-contractible compact Riemannian manifold, and define the *systole* of M, denoted $\mathrm{syst}_1(M)$, to be the length of the shortest non-contractible curve in M. Since M is compact, this value is positive [Kat07].

When M is a surface (i.e., when it is 2-dimensional), one sometimes uses the notation $\mathrm{syst}\,\pi_1(M)$ for the systole to indicate that this is the shortest loop in the fundamental group of M, and usually this is done when one wishes

to distinguish from the *homology systole* of M, (denoted syst $H_1(M)$) which is the length of the shortest curve in M not homologous to 0 (in $H_1(M, \mathbb{Z})$). (By Lemma 5.9, 'non-homologous to zero' is equivalent to stating that the curve does not separate M into two parts.) We of course have syst $\pi_1(M) \leqslant$ syst $H_1(M)$ when M is a surface [Gro96, §2.A].

§6.2   Euclidean and low-dimensional hyperbolic manifolds with short systoles

'Thin' flat tori

Let $\varepsilon > 0$. The additive Abelian group $\mathbb{Z} \oplus \overset{n-1}{\cdots} \oplus \mathbb{Z} \oplus \varepsilon\mathbb{Z}$ acts on $\mathbb{R}^n$ by

$$(m_1, \ldots, m_{n-1}, \varepsilon m_n) \cdot (x_1, \ldots, x_n) = (x_1 + m_1, \ldots, x_n + \varepsilon m_n),$$

and the space of orbits is the n-torus $T_\varepsilon^n = S^1 \times \overset{n-1}{\cdots} \times S^1 \times \varepsilon S^1$. Here, the curve $\alpha \colon [0,1] \to T_\varepsilon^n$ given by $\alpha(t) = (0, \ldots, 0, \varepsilon t)$ is a simple closed geodesic of length ε. Thus it is very easy to construct a closed Euclidean manifold with a systole as short as one would like.

It is somewhat less straightforward to provide such constructions in hyperbolic space.

Hyperbolic surfaces with 'thin' parts

It is shown, for example, in the notes of W. Thurston [Thu80], that a closed hyperbolic surface S of genus g may be decomposed into $2g - 2$ pieces called *pairs of pants*, each of which is homeomorphic to a sphere with three open discs removed. These pieces are obtained as pairs of isometric hyperbolic hexagons with a common edge e identified, and two other non-adjacent edges, not meeting e, also identified. The pairs of pants are separated by $3g - 3$ closed curves L_i (for $i = 1, \ldots, 3g - 3$) of lengths $l_1, \ldots, l_{3g-3}$. Each curve L_i is a boundary component of a pair of pants, and two pairs of pants are glued

together by an isometry α_i between the two copies of L_i. The isometry may be changed by 'twisting' one of the L_i by an angle $\tau_i \in \mathbb{R}$.

The *Teichmüller space* $\mathcal{T}(S)$ of S is the space of all marked hyperbolic structures on S; that is the space of all hyperbolic structures on S up to isotopy equivalence. For the surface S of genus g, the space $\mathcal{T}(S)$ is homeomorphic to $\mathbb{R}^{6g-6}$, and explicit co-ordinates are given by

$$(\log l_1, \tau_1, \ldots, \log l_{3g-3}, \tau_{3g-3}),$$

with the l_i and τ_i as above [Thu80, Theorem 5.3.5] (and these are known as Fenchel-Nielsen co-ordinates). Thus we see that it is possible to have simple closed curves of length as small (or large) as we would like by choosing one of the l_i to be sufficiently small.

A more detailed exposition of the theory of hyperbolic surfaces, including the main aspects of Teichmüller theory, may be found in Chapter 9 of the book of Ratliffe [Rat06].

HYPERBOLIC 3-MANIFOLDS

Producing manifolds with short systole in three dimensions can be acheived using Dehn filling.

It is well-known that an orientable non-compact hyperbolic 3-manifold has finitely many ends, and these are isometric to $T^2 \times [0, \infty)$ [MR03, Theorem 1.3.2]. For a given end, the process of *Dehn filling* is one whereby a solid torus is glued into the end, and the non-compact part discarded. That is, we remove the end $T^2 \times (t, \infty)$ and glue a solid torus V into the resulting boundary torus W. The gluing is described by an identification of a meridian curve on the solid torus V with a simple closed curve on W corresponding to an element $m^p \ell^q \in \pi_1(W)$ for some coprime integers p, q, and generators m and ℓ of $\pi_1(W)$ that intersect in only one point. By taking large p and q one may obtain small systole length for the resulting manifold. A discussion of this is given by W. Neumann and D. Zagier where the lengths of closed geodesics arising from the filling process are indicated [NZ85, Prop. 4.3].

That the process of Dehn filling indeed produces a manifold admitting a hyperbolic structure follows from Thurston's Hyperbolic Dehn Surgery Theorem [Thu80, Theorem 5.8.2].

Hyperbolic 4-manifolds

In 2006, Ian Agol showed [Ago06] that it is possible to construct closed hyperbolic 4-manifolds with short systoles. His construction involves immersing two hyperplanes as totally geodesic submanifolds of an arithmetic compact manifold, cutting along these hyperplanes, and taking the double of the connected component containing the geodesic segment between them. This geodesic segment becomes a closed loop in the double, and by taking the two hyperplanes close together we make this loop as short as desired.

In order to achieve embedding of the two submanifolds a certain property of the ambient lattice is required; namely the GFERF property (as defined in §5.3).

Instead of giving all the details of Agol's construction at this point, we postpone their examination until after the general one given below; and we will also see how his construction more directly generalises using the results of Bergeron, Haglund and Wise (Theorem 5.7 of this thesis) [BHW11].

§6.3   Hyperbolic n-manifolds with short systoles

In this section we prove the following theorem [BT11, Lemma 3.1]:

Theorem 6.1 (Belolipetsky-Thomson, 2011). *Let $\varepsilon > 0$ and let n be an integer at least 2. Then there exist closed hyperbolic n-manifolds with systole length at most ε.*

The proof of this theorem relies on Theorem 5.1 and thus avoids the need to generalise the results on subgroup separability used by Agol. Nevertheless the proof does use some of Agol's main ideas: see §6.4 for further discussion of his

proof versus the one given presently. The proof presented here is that given by Belolipetsky and Thomson [BT11].

Let us for the remainder of this chapter fix $\varepsilon > 0$ and $n \geqslant 2$, where $\varepsilon \in \mathbb{R}$ and $n \in \mathbb{N}$.

OVERALL STRATEGY OF THE PROOF OF THEOREM 6.1

We seek two hyperplanes in hyperbolic space that are at most a distance $\varepsilon/2$ apart, and that admit compact quotients embedding with empty intersection as totally geodesic submanifolds in a compact hyperbolic manifold. By cutting along these embedded submanifolds, and doubling the connected component containing the geodesic segment orthogonal to them, we obtain a compact manifold in which this geodesic doubles to a closed loop of length at most ε.

INITIAL CONFIGURATION

We fix the following:

1. a totally real algebraic number field K, with degree $[K : \mathbb{Q}]$ at least 2 denoted by d, and ring of integers $\mathcal{O}_K$, as defined in §1.1;

2. a quadratic form $f \colon \mathbb{R}^{n+1} \to \mathbb{R}$ with coefficients in K and signature $(n, 1)$, such that f^σ has signature $(n + 1, 0)$ for each $\sigma \in \mathrm{Gal}(K/\mathbb{Q}) \smallsetminus \{\mathbf{id}\}$ (i.e., every non-trivial conjugate form is positive definite), and with associated bilinear form $(\cdot, \cdot)_f$ which for brevity we will denote $(\cdot, \cdot)$ (cf. §1.2);

3. a vector $e_0 \in K^{n+1}$, which we regard as lying in $\mathbb{R}^{n+1}$ via the identity embedding $\sigma_{\mathbf{id}} \colon K \hookrightarrow \mathbb{R}$, such that $f(e_0) > 0$.

In adopting this scenario we arrive at a model of hyperbolic space as in the construction in §2.1 and §2.3, and the isometry group of this resulting hyperbolic space is $\mathrm{PO}_f(\mathbb{R})$. We may assume without loss of generality that the form f has coefficients in $\mathcal{O}_K$ (as with the start of the proof of Theorem 5.1).

Since the degree of the field is at least 2, the discrete group $\mathrm{PO}_f(\mathcal{O}_k)$ is a co-compact arithmetic lattice in $\mathrm{PO}_f(\mathbb{R})$. By Selberg's lemma (Theorem 3.4)

there is a finite-index torsion-free subgroup $\Gamma \leqslant \mathrm{PO}_f(\mathcal{O}_K)$, so that by Theorems 3.2 and 3.3 the quotient space $\Gamma \backslash \mathbb{H}^n$ is a compact arithmetic hyperbolic manifold.

The vector e_0 defines a hyperplane $H_0 = \langle e_0 \rangle^\perp \cap \mathbb{H}^n$. Note that scaling the vector e_0 produces the same hyperplane and so we may assume without loss of generality that e_0 is in fact K-integral; i.e., $e_0 \in \mathcal{O}_K^{n+1}$. Since this vector has entries in K, the stabiliser $\mathrm{Stab}_\Gamma(H_0)$ (denoted Γ_0) is actually a co-compact subgroup of $\mathrm{Isom}(H_0)$; that is, the quotient $\Gamma_0 \backslash H_0$ is compact.

We will also suppose that we have (see Figure 6.1)

4. a vector $e_1 \in \mathcal{O}_K^{n+1}$ with $f(e_1) > 0$, that defines a hyperplane $H_1 = \langle e_1 \rangle^\perp \cap \mathbb{H}^n$ such that $H_0 \cap H_1 = \emptyset$ but $\mathrm{d}_{\mathbb{H}^n}(H_0, H_1) < \varepsilon/2$ (where $\mathrm{d}_{\mathbb{H}^n}$ denotes the distance between the two hyperplanes as in Theorem 2.4).
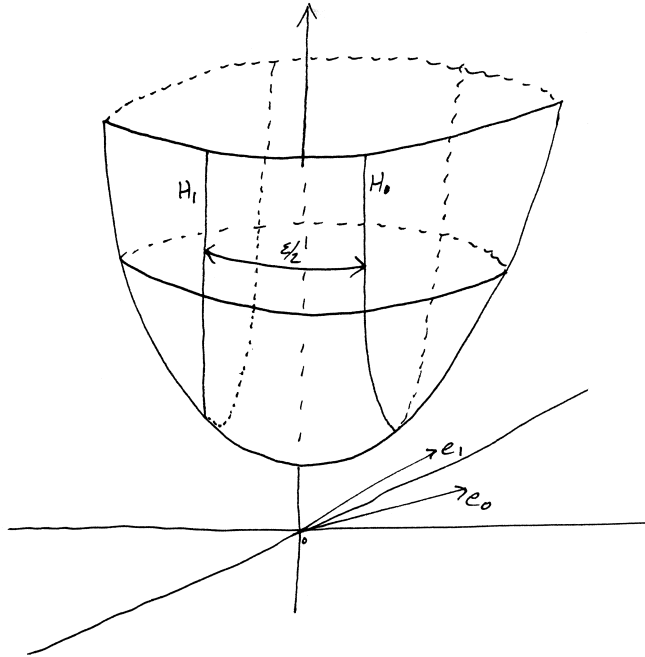

Figure 6.1: Initial configuration of hyperplanes in $\mathbb{H}^n$.

Since H_1 is also K-rational, the stabiliser $\Gamma_1 = \mathrm{Stab}_\Gamma(H_1)$ is co-compact in $\mathrm{Isom}(H_1)$. Note that it is occasionally convenient to regard H_1 as the image of H_0 under some map $\gamma \in \mathrm{PO}_f(K)$. This group is dense in $\mathrm{PO}_f(\mathbb{R})$ (cf. The-
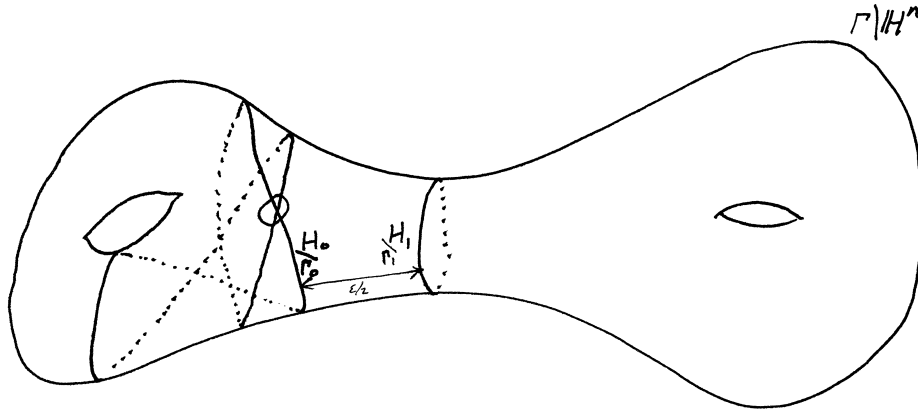
Figure 6.2: The manifold $\Gamma\backslash\mathbb{H}^n$. In this instance the hyperplanes do not intersect each other but the projection of H_0 has a self-intersection at the point circled.

orem 4.10), and so the shortest distance between H_0 and H_1 can be realised: for, if ϕ is any isometry moving H_0 to a desired H_1 (that is $\varepsilon/2$-close to H_0), then there is $\gamma \in \mathrm{PO}_f(K)$ that is as close as we would like to ϕ. (See the examples beginning on p. 80 for an illustration of finding suitable transformations γ.)

Embedding Submanifolds

The quotients $\Gamma_0\backslash H_0$ and $\Gamma_1\backslash H_1$ each immerse into $\Gamma\backslash\mathbb{H}^n$ as totally geodesic submanifolds: see Figure 6.2. The two immersions may have nonempty intersection, however, and they may not be embeddings. Nevertheless Theorem 5.1 implies that there exists a finite-index subgroup $\Gamma' \leqslant \mathrm{PO}_f(\mathcal{O}_K)$ such that each $h \in \Gamma'$ satisfies (5.1); i.e.,

$$\text{either} \quad h(H_0) = H_0 \quad \text{or} \quad h(H_0) \cap (H_0 \cup H_1) = \emptyset. \tag{6.1}$$

Similarly for H_1 in place of H_0 (and vice-versa) we obtain $\Gamma'' \leqslant \mathrm{PO}_f(\mathcal{O}_K)$ of finite index. Now we set $\Lambda = \Gamma \cap \Gamma' \cap \Gamma''$, so that Λ satisfies (6.1) for each of the H_i and is of finite index in Γ (cf. Lemma 4.1). Thus we have empty intersection $\pi_\Lambda(H_0) \cap \pi_\Lambda(H_1)$ of the images of the H_i under the natural projection $\pi_\Lambda \colon \mathbb{H}^n \to \Lambda\backslash\mathbb{H}^n$. We also have an embedding $\Lambda_i\backslash H_i \to \Lambda\backslash\mathbb{H}^n$ for
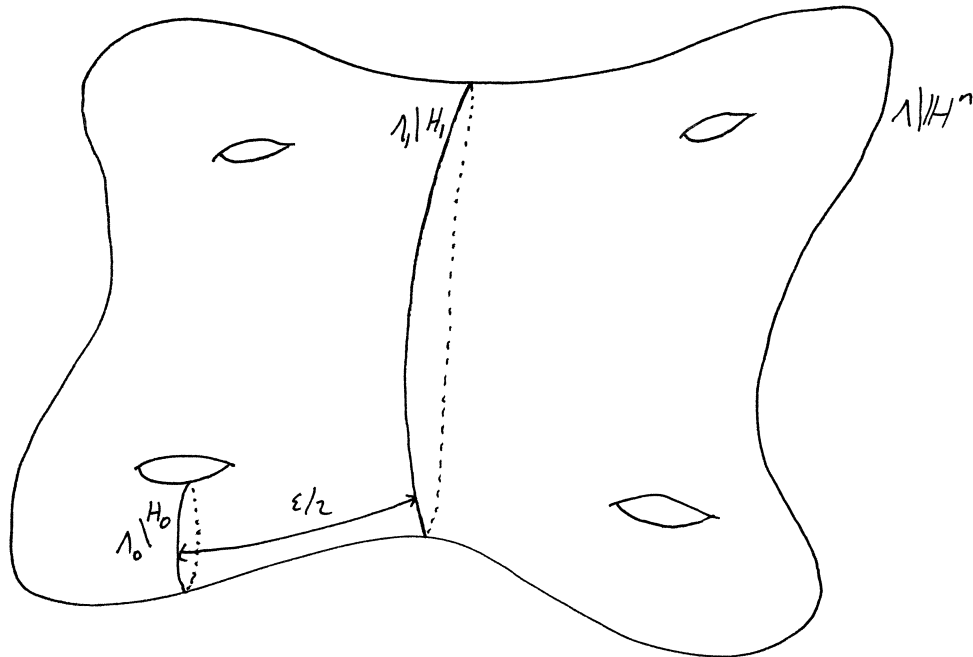
Figure 6.3: The cover $\Lambda\backslash\mathbb{H}^n$.

each H_i, where $\Lambda_i = \mathrm{Stab}_\Lambda(H_i)$, again by (5.1): see Figure 6.3. Since the H_i were $\varepsilon/2$-close, so too are the $\Lambda\backslash H_i$.

Let g be a geodesic segment orthogonal to both submanifolds, so that g has length at most $\varepsilon/2$.

Cutting and doubling

We cut $\mathbb{H}^n$ along the two embedded submanifolds $\Lambda_1\backslash H_1$ and $\Lambda_2\backslash H_2$. Then, the double of the resulting manifold, along the boundary arising from these cuts, will contain the double of g, which is a closed geodesic of length at most ε: see Figure 6.4. If the cutting procedure separates $\Lambda\backslash\mathbb{H}^n$ into multiple parts then we may consider only the connected component containing g.

This concludes the proof of Theorem 6.1. $\square$

Remark. As already remarked (cf. p. 54), Theorem 5.1 applies in the complex hyperbolic case too, but the hypersurfaces are *complex* co-dimension-1, which are not suitable for use in a cut-and-paste construction such as the one above.
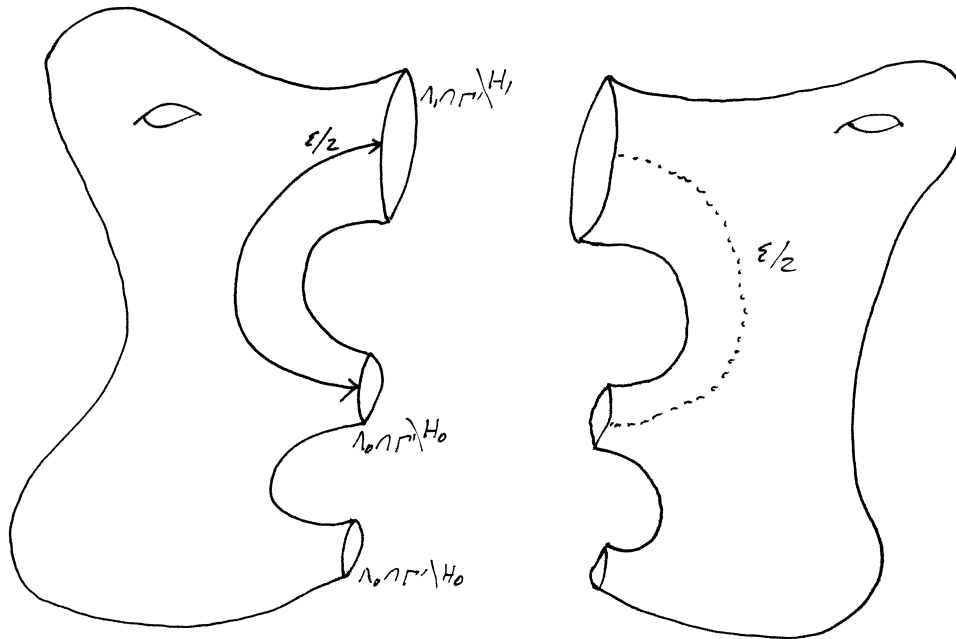
Figure 6.4: Cutting and doubling the manifold. In this case, the projection of H_1 separates $\Lambda\backslash\mathbb{H}^n$ into two, and the piece not containing g is simply disregarded.

Therefore, the proof of this theorem does not directly generalise to the case of complex hyperbolic space.

§6.4   An alternative approach to constructing short systole manifolds

In his 4-dimensional construction (mentioned in §6.2), Agol produces a geometrically finite (in fact, finite-sided) fundamental domain for a group whose quotient has a short systole but is of infinite volume [Ago06, p. 3]. The systole, being compact, is shown to embed in some finite cover of an initial compact manifold using Corollary 5.5. We give some details:

Let e_0 and e_1 be rational vectors in $\mathbb{Q}^5 \hookrightarrow \mathbb{R}^5$, such that their respective orthogonal hyperplanes H_0 and H_1 in $\mathbb{H}^4$ are at most $\varepsilon/2$ apart but not inter-

secting. Let q denote the geodesic segment between the H_i $(i = 0, 1)$, and for both H_i let p_i denote the endpoint of q on H_i. Let Γ denote a torsion-free subgroup of $\mathrm{PO}_f(\mathcal{O}_K)$, where f is the quadratic form $x_1^2 + \cdots + x_4^2 - \frac{1}{2}(1 + \sqrt{5})x_5^2$ and $K = \mathbb{Q}(\sqrt{5})$. It is known that Γ is GFERF, independently of Theorem 5.7 (cf. Example 5.6). The stabilisers in Γ of the hyperplanes H_0 and H_1 respectively have subgroups, G_0 and G_1, say, such that $\mathrm{d}_{\mathbb{H}^n}\big(p_i, g(p_i)\big) > \delta(\varepsilon)$ for every $g \in G_i \smallsetminus \{1\}$, where $\delta(q)$ is a constant depending on the length of the geodesic segment q. The group $G = \langle G_0 \cup G_1 \rangle$ can be shown to be geometrically finite, and since Γ is GFERF, the subgroup $G < \Gamma$ must be separable in Γ. Then, by Corollary 5.5, one may find $\Gamma_1 < \Gamma$, of finite index, such that the compact set $q \cup (G_0 \backslash H_0) \cup (G_1 \backslash H_1)$ embeds in $\Gamma_1 \backslash \mathbb{H}^4$. Then, cutting along $G_0 \backslash H_0$ and $G_1 \backslash H_1$ and taking the double of the remaining part containing q, one completes the construction.

The major obstacle to directly generalising this procedure, prior to the knowledge of Theorem 5.7, was that it was not proven that lattices in any given dimension were GFERF, and so Scott's lemma (Lemma 5.4) could not be used. Theorem 5.1 removes this obstacle and allows a slightly different path to be taken, that, incidentally does not rely on the lemma of Scott.

It is clear that in light of Theorem 5.7, the lattice Γ could be taken to be any torsion-free subgroup of an arithmetic lattice in higher dimensions, and Agol's argument may be directly generalised to these dimensions.

7

Systolic inequalities and volumes of short systole manifolds

Systolic geometry studies inequalities of the type in (7.1) below, which relate the volume of a Riemannian manifold to its systole length. (As mentioned in Chapter 6, the book of Katz [Kat07] and the notes of Gromov [Gro96] make for good introductions to the area.) In this chapter we examine inequalities of this type for the manifolds produced in Theorem 6.1, and we also examine some explicit examples towards the end of the chapter.

§7.1   Gromov's systolic inequality

The 'systolic inequality' of M. Gromov gives a lower bound for the volume of a Riemannian manifold (or at least a certain type of Riemannian manifold) in terms of the systole.

A closed Riemannian n-manifold is called *essential* if, in the homology of its fundamental group, its fundamental class (i.e., generator of $H_n(M)$) is non-zero [Kat07, p. 95]. We have the following [Kat07, Theorem 12.2.2]:

Theorem 7.1 (Gromov's systolic inequality). *Any essential, compact Riemannian n-manifold M satisfies*

$$\operatorname{syst}_1(M)^n \leqslant C_n \operatorname{vol}(M), \tag{7.1}$$

where $C_n > 0$ and depends only on the dimension n.

Gromov's inequality is perhaps surprising in that it applies to a very general class of manifolds (which includes, for example, those manifolds M with $\pi_j(M)$ trivial for $j \geqslant 2$). We will see later (cf. Proposition 7.5) how a similar inequality can be obtained for manifolds with short systoles constructed as in §6.3. If it were the case that there was a lower bound on the systole of a hyperbolic n-manifold, then (7.1) would imply that the volume of any n-manifold were bounded below; but, Theorem 6.1 of course tells us that there is no lower bound for the systole length (cf. Theorem 3.7).

Similar results to those of Gromov existed previously, but only for certain special cases. For example, C. Loewner showed that if T is a 2-torus with any Riemannian metric, then $\operatorname{syst}_1(T) \leqslant (2/\sqrt{3}) \operatorname{vol}(T)^{1/2}$ [Gro96, 1.B], and a similar inequality exists for the real projective plane, due to P. Pu [Kat07] [Gro96].

Non-essential manifolds

The inequality (7.1) fails for some manifolds, and by way of example, consider an essential compact Riemannian n-manifold M with positive systole $\operatorname{syst}_1(M)$, and form the closed manifold $M \times \varepsilon \mathbb{S}^2$, where $\varepsilon > 0$ and where by $\varepsilon \mathbb{S}^2$ is meant $\mathbb{S}^2$ with a metric such that $\operatorname{vol}(\varepsilon \mathbb{S}^2) = \varepsilon$. The systole of $M \times \varepsilon \mathbb{S}^2$ is equal to that of M but we have $\operatorname{vol}(M \times \varepsilon \mathbb{S}^2)$ as small as desired so as to violate (7.1). We thus find an example of a non-essential manifold (i.e., $M \times \mathbb{S}^2$). (This argument is similar to one in Gromov's original article on systolic inequalities [Gro83].)

§7.2   Orthospectra and orthogeodesics of hyperbolic manifolds

If M is a compact hyperbolic n-manifold with totally geodesic boundary, then an *orthogeodesic* for M is a geodesic segment with endpoints lying on the boundary of M and orthogonal to that boundary at both endpoints. The *orthospectrum* for M is the set

$$\Lambda_M = \{\ell \mid \ell \text{ is the length of an orthogeodesic}\} \quad \text{with multiplicities.}$$

M. Bridgeman and J. Kahn use this terminology in a recent article [BK10], where they attribute the first appearance of 'orthospectrum' to A. Basmajian in 1993. For us a relevant result from their article is the following:

THEOREM 7.2 (BRIDGEMAN-KAHN). *Let $n \geqslant 2$. Then there exists a function $F_n \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ that is continuous and monotonically decreasing, such that for any compact hyperbolic n-manifold with totally geodesic boundary,*

$$\mathrm{vol}(M) = \sum_{l \in \Lambda_M} F_n(l). \tag{7.2}$$

Furthermore, there exists $K_n > 0$ such that

$$\lim_{l \to 0} l^{n-2} F_n(l) = K_n. \tag{7.3}$$

This theorem is proven by using the observation that for a hyperbolic n-manifold (or indeed any Riemannian n-manifold), we have

$$\mathrm{vol}(M) = \frac{1}{\mathrm{vol}(\mathbb{S}^{n-1})} \mathrm{vol}(T_1 M)$$

where $T_1 M$ is the unit tangent bundle of M and $\mathbb{S}^{n-1}$ is the $(n-1)$-sphere. To each tangent vector $v \in T_1 M$ is assigned a geodesic arc α_v of maximal length with v as tangent. The set $T_1^f M$, of all $v \in T_1 M$ such that α_v has endpoints in ∂M, has full measure in $T_1 M$ by ergodicity of the geodesic flow on the double DM. The set $T_1^f M$ is partitioned into equivalence classes by considering homotopy of orthogeodesics, and this partitioning allows the decomposition into a sum. The function F_n is given explicitly in the article and for some small n is presented in convenient forms [BK10, p. 1221].

§7.3 Volumes of short systole hyperbolic manifolds

Whilst Theorem 6.1 asserts that it is possible to have hyperbolic manifolds with a systole as short as we would like, it seems reasonable that there should be some restriction on the geometry of the resulting manifolds. We turn to the following result of H.-C. Wang [Wan72, 8.1]:

THEOREM 7.3 (WANG). *Let G be a connected semisimple Lie group with neither compact nor three-dimensional factor, and let $c > 0$. Then there is a finite collection $\Gamma_1, \ldots, \Gamma_{m(c)}$ of lattices in G such that any other lattice Γ with co-volume at most c is conjugate (in G) to one of the Γ_i $(i = 1\ldots, m)$.*

(Recall that the dimension of $\mathrm{SL}_2(\mathbb{R})$ is equal to 3, so it is excluded from the above theorem, and hence the result does not apply to hyperbolic surfaces; i.e., to hyperbolic 2-manifolds.) It turns out that this result is not quite correct, and contradicts the discussion in §6.6 of Thurston's notes [Thu80], which asserts (among other things) that for hyperbolic 3-manifolds, the set of manifolds with volume bounded by a given $c > 0$ is infinite. The Lie group associated with the symmetric space $\mathbb{H}^3$ is $\mathrm{SL}_2(\mathbb{C})$ (cf. p. 21 and p. 27), which also needs to be excluded from Theorem 7.3. A. Borel gives a discussion of this [Bor81, 8.3], pointing out that Wang's mistake is to mis-quote a theorem of H. Garland and M. S. Raghunathan concerning rigidity. The upshot is that for $n \geqslant 4$ and for any $c > 0$, there are only finitely many isometry classes of hyperbolic n-manifold with volume at most c.

If we construct a sequence of hyperbolic n-manifolds M_m (for $n \geqslant 4$) with systole $\mathrm{syst}_1(M_m) \to 0$ as $m \to \infty$, then the above discussion implies that we must have $\mathrm{vol}(M_m) \to \infty$ as $m \to \infty$. In the particular case of manifolds constructed as in the proof of Theorem 6.1, we have the following [BT11]:

THEOREM 7.4 (BELOLIPETSKY-THOMSON). *Let $n \geqslant 3$. Then there exists $C_n > 0$ (depending only on n), such that for any n-manifold M constructed in*

§*6.3 we have*

$$\mathrm{vol}(M) \geqslant \frac{C_n}{\mathrm{syst}_1(M)^{n-2}}. \tag{7.4}$$

Despite the above discussion concerning the case $n = 3$, shortening the systole does indeed imply volume growth for 3-manifolds constructed in §6.3, as well as for n-manifolds with $n \geqslant 4$. The bound given in (7.4) is optimal in the sense that one may find a sequence $\{M_i\}$ of manifolds with at most polynomial volume growth in $1/\mathrm{syst}_1(M_i)$ [BT11]:

PROPOSITION 7.5 (BELOLIPETSKY-THOMSON). *Let $n \geqslant 2$. There exists a sequence $(M_i)_i$ of n-manifolds as in §6.3, with $\mathrm{syst}_1(M_i) \to 0$ as $i \to \infty$, and positive constants B_n and γ_n such that*

$$\mathrm{vol}(M_i) \leqslant \frac{B_n}{\mathrm{syst}_1(M_i)^{\gamma_i}}. \tag{7.5}$$

Let us return presently to the Theorem:

PROOF OF THEOREM 7.4. Let M be a manifold as constructed in §6.3 with systole of length at most ε. Then M is a double of a manifold N with boundary, which has as one of its orthogeodesics a curve of length at most $\varepsilon/2$. Thus $\ell = \frac{1}{2} \mathrm{syst}_1(M)$ is in the orthospectrum of N, and by Theorem 7.2 we have

$$\mathrm{vol}(N) = \sum_{l \in \Lambda_N} F_n(l) \geqslant F_n(\ell);$$

that is $\mathrm{vol}(M) \geqslant 2F_n(\ell)$. (Here, F_n is of course the function given in Theorem 7.2.) By the same theorem we also have $\lim_{l \to 0} l^{n-2} F_n(l) = K_n$ for some constant $K_n > 0$. Noting that F_n is positive we find another constant $K_n' > 0$ such that for $\ell < 1$

$$F_n(\ell) \geqslant K_n'/l^{n-2}.$$

Thus when $\mathrm{syst}_1(M) < 2$, we have

$$\mathrm{vol}(M) = 2\,\mathrm{vol}(N) \geqslant \frac{2^{n-1} K_n'}{\mathrm{syst}_1(M)^{n-2}}. \tag{7.6}$$

Supposing $\mathrm{syst}_1(M) \geqslant 2$, we find that the expression in (7.6) could be very small, but by Theorem 3.5 (Každan-Margulis) the volume of M must be

bounded below by some constant α_n. If we set $C_n = \min\{2^{n-1}K'_n, 2^{n-2}\alpha_n\}$, then for $\mathrm{syst}_1(M) \geqslant 2$ we have

$$\mathrm{vol}(M) \geqslant \alpha_n \geqslant \frac{2^{n-2}\alpha_n}{\mathrm{syst}_1(M)^{n-2}} \geqslant \frac{C_n}{\mathrm{syst}_1(M)^{n-2}}. \tag{7.7}$$

Together, (7.6) and (7.7) imply the conclusion of Theorem 7.4. $\qquad\square$

Remark. One might hope to apply the results of Bridgeman and Kahn to a closed hyperbolic manifold, by using Theorem 5.1 to obtain an embedded totally geodesic submanifold of co-dimension 1, along which a cut can be made so as to obtain an orthogeodesic. Doing this would require taking the finite covers as described in Theorem 5.1, and one would need to be able to estimate the degree of such a cover and hence bound the volume of the initial manifold M.

A similar lower bound to that in Theorem 7.4, for the volume of a hyperbolic manifold in terms of its systole, has already been given by A. Reznikov [Rez95]. In full generality it is as follows:

THEOREM 7.6 (REZNIKOV). *Suppose M is a compact hyperbolic manifold of dimension $n \geqslant 4$. Then we have a positive constant C_n (depending only on n) such that*

$$\mathrm{InjRad}(M) \geqslant \frac{C_n}{\mathrm{vol}(M)^{1+4/(n-3)}}. \tag{7.8}$$

On noting that $\mathrm{InjRad}(M) = \frac{1}{2}\mathrm{syst}_1(M)$, one immediately obtains $C'_n > 0$ such that

$$\mathrm{vol}(M) \geqslant \frac{C'_n}{\mathrm{syst}_1(M)^{\frac{n-3}{n+1}}}. \tag{7.9}$$

However, the proof given by Reznikov can be slightly simplified in the case where the geodesic realising the systole corresponds to an isometry of $\mathbb{H}^n$ with no rotational part. What follows is a description of this, and we keep the notation similar to that of Reznikov's.

Let M be a compact hyperbolic n-manifold $\Gamma \backslash \mathbb{H}^n$. Let γ denote the closed geodesic in M that realises the systole of M, and let $\tilde{\gamma}$ denote the geodesic line in $\mathbb{H}^n$ that projects to M under the natural map $\mathbb{H}^n \to \Gamma \backslash \mathbb{H}^n$. Assume that the systole of M is smaller than $2\varepsilon_n$ where ε_n is the nth Margulis constant (cf.

Theorem 3.8 on p. 34). The closed geodesic γ corresponds to an isometry ϕ of $\mathbb{H}^n$ that leaves γ invariant: we consider the special case where this has no rotational part.

Given any positive $\varepsilon < \varepsilon_n$, the '$\varepsilon$-thin' part of the manifold M, i.e., $\{x \in M \mid \mathrm{InjRad}_x(M) < \varepsilon\}$ (cf. §3.5), contains the systole of M: let Q denote the connected component of the thin part in which γ lies. The estimate of the volume of M is in fact an estimate for the volume of the Margulis tube Q. The length of γ will be denoted by ℓ, and the width $w(z, Z)$ of Q will be the distance between $z \in \gamma$ and the first intersection of a geodesic in the Z direction with ∂Q ($Z \in T_z(M)$), where Z is assumed to be orthogonal to $\dot{\gamma}(z)$.

Now suppose that we have chosen $z \in \gamma$ and that ζ is the geodesic in some Z-direction. Denote its first intersection with the boundary ∂Q by q. The points z and q lift to points $\tilde{z}$ and $\tilde{q}$ in $\mathbb{H}^n$. By the definition of Q, we have $\mathrm{d}_{\mathbb{H}^n}\big(\tilde{q}, \phi(\tilde{q})\big) \geqslant \varepsilon$. Reznikov also uses the estimate $\mathrm{d}_{\mathbb{H}^n}\big(\tilde{q}, \phi(\tilde{q})\big) \leqslant C_n \exp\big(w(z, Z)\big)\ell$, so that we have

$$\exp\big(w(z, Z)\big) \geqslant \frac{C_{n,\varepsilon}}{\ell}. \tag{7.10}$$

Now, we have, considering the volume of an $(n-1)$-ball,

$$\begin{aligned}
\mathrm{vol}(Q) &\geqslant C_n' \ell \exp\big((n-1)w(z, Z)\big) \\
&\geqslant \frac{C_{n,\varepsilon}'}{\ell^{n-2}}.
\end{aligned}$$

The volume $\mathrm{vol}(Q)$ is a lower bound for $\mathrm{vol}(M)$, so we establish the same volume growth as in (7.4), noting that ℓ is the systole length of M.

The upper bound in (7.5) is proven by taking a concrete example that may be constructed in any dimension:

PROOF OF PROPOSITION 7.5. [BT11] We provide the sequence $\{M_i\}_{i \in \mathbb{Z}}$ for the conclusion of the proposition by the construction that follows. Let $K = \mathbb{Q}(\sqrt{5})$ and $f = -\sqrt{5}x_0^2 + x_1^2 + \cdots + x_n^2$. We first claim that the sequence

$\{A_i\}_{i \in \mathbb{Z}}$ of matrices

$$
A_i = \begin{pmatrix}
\frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}} & 0 & \cdots & 0 & \frac{-2i}{i^2 - \sqrt{5}} \\
0 & 1 & & & 0 \\
\vdots & & \ddots & & \vdots \\
0 & & & 1 & 0 \\
\frac{-2i\sqrt{5}}{i^2 - \sqrt{5}} & 0 & \cdots & 0 & \frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}}
\end{pmatrix} \qquad (i \in \mathbb{Z})
$$

lies in $\mathrm{O}_f(K)$. Indeed, we have

$$
\begin{aligned}
f(A_i x) &= -\sqrt{5} \left(\frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}} x_0 + \frac{-2i}{i^2 - \sqrt{5}} x_n \right)^2 + x_1^2 + \cdots + x_{n-1}^2 \\
&\qquad + \left(\frac{-2i\sqrt{5}}{i^2 - \sqrt{5}} x_0 + \frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}} x_n \right)^2 \\
&= -\sqrt{5} \frac{(i^2 + \sqrt{5})^2 - 4i^2\sqrt{5}}{(i^2 - \sqrt{5})^2} x_0^2 + x_1^2 + \cdots + x_{n-1}^2 \\
&\qquad + \frac{(i^2 + \sqrt{5})^2 - 4i^2\sqrt{5}}{(i^2 - \sqrt{5})^2} x_n^2 \\
&= f(x),
\end{aligned} \tag{7.11}
$$

so each A_i is an isometry of the quadratic space of f. Clearly $A_i \to \mathbf{id}$ as $i \to \infty$. Let $e_0 = (0, 0, \ldots, 0, 1)$, so that

$$
A_i(e_0) = \left(\frac{-2i}{i^2 - \sqrt{5}}, 0, \ldots, 0, \frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}} \right) \in K^{n+1}.
$$

Rescaling e_0 and $A_i(e_0)$, we define

$$
e_0^{(i)} = (0, 0, \ldots, i^2 - \sqrt{5}) \quad \text{and} \quad e_1^{(i)} = (-2i, 0, \ldots, 0, i^2 + \sqrt{5}),
$$

which give

$$
\left(e_0^{(i)}, e_1^{(i)} \right) = i^4 - 5 \quad \text{and} \quad \left(e_0^{(i)}, e_0^{(i)} \right) = \left(e_1^{(i)}, e_1^{(i)} \right) = (i^2 - \sqrt{5})^2.
$$

Then $e_0^{(i)}$ and $e_1^{(i)}$ can be seen to define disjoint hyperplanes in $\mathbb{H}^n$ by Theorem 2.3 (as in (5.4)): note that the inequality is strict.

For our choice of K there is only one non-trivial Galois automorphism $\sigma \colon a + b\sqrt{5} \mapsto a - b\sqrt{5}$, so we also compute

$$
\left(e_0^{(i)}, e_1^{(i)} \right)^\sigma = i^4 - 5 \quad \text{and} \quad \left(e_0^{(i)}, e_0^{(i)} \right)^\sigma = \left(e_1^{(i)}, e_1^{(i)} \right)^\sigma = (i^2 + \sqrt{5})^2.
$$

The proof of Theorem 5.1 gives two ideals $\mathfrak{p}_0^{(i)} = \left(2\big(e_0^{(i)}, e_0^{(i)}\big)\right)$ and $\mathfrak{p}_1^{(i)} = \left(2C\big(e_0^{(i)}, e_1^{(i)}\big)\right)$, and in order to satisfy (5.10) we require that

$$C \geqslant \frac{(i^2 + \sqrt{5})^2}{(i^2 - \sqrt{5})^2}.$$

We also need $e_0^{(i)} \pm e_1^{(i)}$ to be nonzero modulo $\mathfrak{p}_1^{(i)}$. That is,

$$(-2i, 0, \ldots, 0, 2i^2) \quad \text{and} \quad (2i, 0, \ldots, 0, -2\sqrt{5})$$

must not be zero modulo $\mathfrak{p}_1^{(i)}$. Since $\mathfrak{p}_1^{(i)} = \left(2C(i^4 - 5)^2\right)$, this holds automatically. Observe that if i is large, then $C = 2$ is sufficient. Note also that since $\mathfrak{p}_0^{(i)}$ divides $\mathfrak{p}_1^{(i)}$, we need only consider $\mathfrak{p}_1^{(i)}$ and can take $\Gamma_i' = \Gamma(\mathfrak{p}_1^{(i)})$. (Actually by the remark on p. 54 we needn't make this justification but it is included here for completeness of exposition.)

Note that the proof of Theorem 6.1 requires Theorem 5.1 to be applied a second time, with $e_0^{(i)}$ and $e_1^{(i)}$ interchanged. However, since both vectors are of the same length, the ideal $\mathcal{Q}_1^{(i)} = \left(4\big(e_0^{(i)}, e_1^{(i)}\big)\right)$ is equal to $\mathfrak{p}_1^{(i)}$ anyway, and so we can effectively ignore this step.

Now by (2.8) we have

$$\cosh \mathrm{d}_{\mathbb{H}^n}\big(H_0^{(i)}, H_1^{(i)}\big) = \frac{\left|\big(e_0^{(i)}, e_1^{(i)}\big)\right|}{\|e_0^{(i)}\|\|e_1^{(i)}\|} = \frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}} \tag{7.12}$$

where $\mathrm{d}_{\mathbb{H}^n}\big(H_0^{(i)}, H_1^{(i)}\big)$ is the distance between the hyperplanes $H_0^{(i)}$ and $H_1^{(i)}$ defined by $e_0^{(i)}$ and $e_1^{(i)}$ respectively. We see that $\mathrm{d}_{\mathbb{H}^n}\big(H_0^{(i)}, H_1^{(i)}\big) \to 0$ as $i \to \infty$.

In the manifold M_i obtained by the inbreeding construction, we have $\varepsilon_i = \mathrm{syst}_1(M_i) = 2\rho_i$ where $\rho_i = \mathrm{d}_{\mathbb{H}^n}\big(H_0^{(i)}, H_1^{(i)}\big)$. Now, by (7.12),

$$\cosh(\varepsilon_i/2) = \left(\frac{i^2 + \sqrt{5}}{i^2 - \sqrt{5}}\right),$$

and by using a Taylor expansion for $\cosh(\varepsilon_i/2)$ we obtain (for large i)

$$\varepsilon_i \sim 2\sqrt{\frac{\sqrt{5}}{i^2 + \sqrt{5}}}, \tag{7.13}$$

so that for some constant $\delta > 0$

$$\varepsilon_i \sim \frac{\delta}{i} \quad \text{for large } i.$$

Writing $\mathfrak{p}_1^{(i)} = (\beta)$ with $\beta = 4(i^4 - 5)^2$, we have $|N(\mathfrak{p}_1^{(i)})| = 16(i^4 - 5)^2 \sim B(\delta/\varepsilon_i)^8$ for some constant $B > 0$.

Now for a given $\mathfrak{p}^{(i)}$, note that $|N(\mathfrak{p}^{(i)})|$ is the number of elements in the residue class ring $\mathcal{O}_K/\mathfrak{p}^{(i)}$ [Lan70, Ch. I, Sect. 7], so $|\Gamma : \Gamma_i'| \leqslant |N(\mathfrak{p}^{(i)})|^{(n+1)^2}$ since $|\Gamma : \Gamma_i'|$ is the order of a matrix group over $\mathcal{O}_K/\mathfrak{p}^{(i)}$. Thus for some positive constant D,

$$\mathrm{vol}(\Gamma_i' \backslash \mathbb{H}^n) = \mathrm{vol}(\Gamma \backslash \mathbb{H}^n) \cdot |\Gamma : \Gamma_i'| \leqslant D\big(B(\delta/\varepsilon_i)^8\big)^{(n+1)^2}, \tag{7.14}$$

which is a polynomial in $1/\mathrm{syst}_1(M_i)$ of degree $8(n+1)^2$. $\qquad \square$

FINDING RATIONAL HYPERPLANES WITH SMALL DISTANCE APART

Producing examples such as those in the proof of Proposition 7.5 essentially amounts to choosing matrices that lie in $\mathrm{O}_f(K)$ for the field K over which the quadratic form f is defined. *A priori* this might seem like a highly non-trivial task (especially in higher dimensions), for it amounts to finding K-rational solutions to systems of quadratic equations in several variables over K, namely the equations that define the orthogonal group of f. However, by considering products of reflections in hyperplanes defined by integral vectors, one easily arrives at explicit examples. (By way of motivation, one considers the result that every element of $\mathrm{O}_f(K)$ must be a product of reflections [O'M71, 43:3; also §42E]. This sort of approach has been used by authors seeking solutions to this problem for classical orthogonal groups [Sch08].)

Let us consider a slightly more general example than that in the proof of Proposition 7.5. Let $n \in \mathbb{N}$ with $n \geqslant 2$.

Suppose that for every $i \in \mathbb{N}$, we denote by u_i the vector $(F_i, 0, \ldots, 0, 1) \in \mathbb{R}^{n+1}$, where $F_i \in \mathbb{Q}_{>0}$ and $F_i \to 0$ as $i \to \infty$. Thus $u_i \to (0, 0, \ldots, 0, 1)$ as $i \to \infty$. Now let f be the quadratic form $-\varphi x_0^2 + x_1^2 + \cdots + x_n^2$, where φ is a positive algebraic integer, thus giving f signature $(n, 1)$.

With respect to f one has the reflection R_{u_i} in the hyperplane defined by u_i:

$$R_{u_i} \colon x \mapsto x - 2 \frac{(x, u_i)_f}{(u_i, u_i)_f} u_i$$

$$\text{i.e.,} \qquad R_{u_i} \colon \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} - 2 \frac{-\varphi x_0 F_i + x_n}{-\varphi F_i^2 + 1} \begin{pmatrix} F_i \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \qquad (7.15)$$

Note that R_{u_i}, being a reflection, is not orientation-preserving, but by composing with the map

$$R \colon \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_0 \\ \vdots \\ -x_n \end{pmatrix} \qquad (7.16)$$

one obtains the map $R \circ R_{u_i}$ with matrix

$$A_i = \begin{pmatrix} \dfrac{1 + \varphi F_i^2}{1 - \varphi F_i^2} & 0 & \cdots & 0 & \dfrac{-2F_i}{1 - \varphi F_i^2} \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 1 & 0 \\ \dfrac{-2\varphi F_i}{1 - \varphi F_i^2} & 0 & \cdots & 0 & \dfrac{1 + \varphi F_i^2}{1 - \varphi F_i^2} \end{pmatrix}.$$

One may verify directly that A_i lies in $\mathrm{SO}_f(K)$ (as in (7.11)), but this is clear anyway since it is the product of two reflections.

In the short systole manifold construction (of §6.3) one only requires to be able to find two positive K-integral vectors e_0 and e_1. The matrix A_i provides a convenient way of doing this as $A_i(e_0)$ is clearly positive whenever e_0 is. (One may need to scale $A_i(e_0)$ so as to ensure that it is integral, of course.)

A FURTHER VOLUME GROWTH EXAMPLE

Using the notation above, let $\varphi = \sqrt{d}$ where d is a square-free positive integer. Let $K = \mathbb{Q}(\sqrt{d})$ (cf. p. 3), and let f be the quadratic form defined in the

previous subsection.

Now denote by u_i (for $i \in \mathbb{Z}$) the vector $\big(1/i, 0, \ldots, 0, (i+1)/i\big) \in K^{n+1}$. Again, let $e_0 = (0, \ldots, 0, 1) \in K^{n+1}$, so that $u_i \to e_0$ as $i \to \infty$. For every i, the map $R \circ R_{u_i}$ (cf. (7.15) and (7.16)), applied to e_0, gives a vector e_i in K^{n+1}, such that $e_i \to e_0$ as $i \to \infty$. By scaling the vectors e_0 and e_i we have, for every i, the following pair of f-positive K-integral vectors in $\mathbb{R}^{n+1}$:

$$e_0^{(i)} = \big(0, \ldots, 0, (i+1)^2 - \varphi\big) \quad \text{and} \quad e_1^{(i)} = \big(-2(i+1), 0, \ldots, 0, \varphi + (i+1)^2\big).$$
(7.17)

The inner products computed in the proof of Proposition 7.5 are

$$(e_0^{(i)}, e_0^{(i)}) = (e_1^{(i)}, e_1^{(i)}) = \big((i+1)^2 - \varphi\big)^2 \quad \text{and} \quad (e_{0,i}, e_{1,i}) = (i+4)^4 - \varphi^2,$$
(7.18)

and so their conjugates are

$$\big(e_0^{(i)}, e_1^{(i)}\big)^\sigma = \big(e_1^{(i)}, e_1^{(i)}\big)^\sigma = \big((i+1)^2 + \varphi\big)^2 \quad \text{and} \quad \big(e_0^{(i)}, e_1^{(i)}\big)^\sigma = (i+4)^4 + \varphi^2.$$
(7.19)

The vectors in (7.17) give rise to hyperplanes $H_0^{(i)}$ and $H_1^{(i)}$ respectively. The distance between these hyperplanes is given by

$$\cosh \mathrm{d}_{\mathbb{H}^n}(H_0^{(i)}, H_1^{(i)}) = \frac{(i+1)^4 - \varphi^2}{\big((i+1)^2 - \varphi\big)^2},$$
(7.20)

so that we indeed have the distance between $H_0^{(i)}$ and $H_1^{(i)}$ tending to zero as i tends to infinity. Analogously to (7.13), we have for large i

$$\varepsilon_i \approx \frac{4\varphi(i+1)^2}{\big((i+1)^2 - \varphi\big)^2} \approx \frac{A}{i^2} \qquad \text{for some } A > 0.$$
(7.21)

We also need to consider the norm of the ideal $\mathfrak{p}_0^{(i)}$ generated by $2C(e_0^{(i)}, e_1^{(i)})$, where C is an integer satisfying (5.10). As with the proof of Proposition 7.5, we can assume that $C \geqslant 2$; in fact, $C = 2$ will suffice for all but finitely many i. So, we compute $N(\mathfrak{p}_0^{(i)}) = 16\big((i+1)^{16} - \varphi^4\big)$. By using this, along with (7.21) and the estimates for (7.14), we find that the volume again grows like a polynomial in $1/\varepsilon_i$ (i.e., $1/\operatorname{syst}_1(M_i)$), of at most degree $8(n+1)^2$.

8

(Non-)arithmeticity and (non-)coherence of lattices in $\mathrm{PO}(n,1)$

§8.1 The existence of non-arithmetic lattices in $\mathrm{PO}(n,1)$

It has been known for some time that not all lattices in $\mathrm{PO}(n,1)$ are arithmetic, in contrast to the case of a Lie group of real rank at least 2 (cf. Theorem 4.8). More precisely, a construction of non-arithmetic groups in every dimension $n \geqslant 2$ was given by Gromov and Piatetski-Shapiro in 1987 [GPS87]. Their basic method is to obtain two arithmetic manifolds with a common totally geodesic submanifold as boundary, and by an isometry between the two copies of the submanifold glue the manifolds together. If the manifolds are not commensurable, then the resulting manifold is non-arithmetic [GPS87, 0.2]. It is shown to be easy to arrange non-commensurability of the two manifolds [GPS87, 2.6, 2.7].

The construction of Gromov and Piatetski-Shapiro is also described by Margulis [Mar91], and by Vinberg and Shvartsman [AVS93, p. 228]. Both Margulis, and also Vinberg and Shvartsman, describe in addition the existence

of non-arithmetic reflection groups in lower dimensions using a criterion of Vinberg [AVS93, Theorem 3.1, p. 226].

We illustrate the construction of Gromov and Piatetski-Shapiro by giving an example. Fix $n \in \mathbb{N}_{\geqslant 2}$, let $\varphi = \frac{1}{2}(1 + \sqrt{5})$ and let f_0 be the quadratic form $x_1^2 + \cdots + x_{n-1}^2 - \varphi x_n^2$. This has signature $(n-1, 1)$ and is defined over $K = \mathbb{Q}(\sqrt{5})$, whence the only non-trivial conjugate form of f is $f^\sigma = x_1^2 + \cdots + \frac{1}{2}(\sqrt{5} - 1)x_n^2$, which is positive definite. Thus $\mathrm{PO}_{f_0}(\mathcal{O}_K)$ is an arithmetic lattice in $\mathrm{PO}_{f_0}(\mathbb{R})$. We also define forms $f_1 = x_0^2 + f_0$, and $f_2 = 7x_0^2 + f_0$. Let $\Gamma_i(\mathfrak{p})$ denote the principal congruence subgroup (cf. §4.2) $\mathrm{PO}_{f_i}(\mathcal{O}_K)(\mathfrak{p})$ (for $i = 0, 1, 2$), where $\mathfrak{p}$ is such that all $\Gamma_i(\mathfrak{p})$ are torsion-free.

Hyperbolic $(n-1)$-space $\mathbb{H}^{n-1}$ can be identified with the hyperplane

$$\{(x_0, \ldots, x_{n-1}, 0)\} \subseteq \mathbb{H}^n.$$

In this way, we have an embedding $\iota \colon \mathbb{H}^{n-1} \to \mathbb{H}^n$. Then the map $\pi_i \circ \iota$, where $\pi_i \colon \mathbb{H}^n \to \Gamma_i(\mathfrak{p})\backslash\mathbb{H}^n$ $(i = 1, 2)$ is the natural projection, gives an immersion of $\mathbb{H}^{n-1}$ into the two quotients. It turns out that we can in fact achieve an embedding $\iota_i \colon \Gamma_0(\mathfrak{p})\backslash\mathbb{H}^{n-1} \hookrightarrow \Gamma_i(\mathfrak{p})\backslash\mathbb{H}^n$ $(i = 1, 2)$ so that the diagram

$$
\begin{array}{ccc}
\mathbb{H}^{n-1} & \overset{\iota}{\lhook\joinrel\longrightarrow} & \mathbb{H}^n \\
\downarrow{\scriptstyle \pi_0} & & \downarrow{\scriptstyle \pi_i} \\
\Gamma_0(\mathfrak{p})\backslash\mathbb{H}^{n-1} & \overset{\iota_i}{\lhook\joinrel\longrightarrow} & \Gamma_i(\mathfrak{p})\backslash\mathbb{H}^n
\end{array}
$$

commutes for each i [GPS87, 2.8.A]. The quotients $\Gamma_i(\mathfrak{p})\backslash\mathbb{H}^n$ will be denoted by $V_i(\mathfrak{p})$, and $V_0(\mathfrak{p})$ will denote the image $\iota_i\big(\Gamma_0(\mathfrak{p})\backslash\mathbb{H}^{n-1}\big)$ in $V_i(\mathfrak{p})$ for $i = 1, 2$. There is a double cover of each $V_i(\mathfrak{p})$, denoted $\tilde{V}_i(\mathfrak{p})$, such that $V_0(\mathfrak{p})$ lifts to a separating hypersurface $\tilde{V}_0(\mathfrak{p})$ which is the union of two copies of $V_0(\mathfrak{p})$. Then there is a connected submanifold $V_i^+ \subseteq \tilde{V}_i(\mathfrak{p})$ with $\partial V_i^+ = \tilde{V}_0(\mathfrak{p})$. We can identify V_1^+ and V_2^+ along $\partial V_1^+ = \partial V_2^+$ so as to obtain a new manifold.

The new manifold V obtained in this way is non-arithmetic, for in order for it to have arithmetic fundamental group we would need Γ_1 and Γ_2 to be commensurable [GPS87, 0.2]. Commensurability is only achieved if the forms f_0 and f_1 are similar over $\mathbb{Q}(\sqrt{5})$ (cf. §1.3) [GPS87, 2.6], and for this to happen we would need 7 to be a square in $\mathbb{Q}(\sqrt{5})$ [GPS87, 2.7], which is not the case.

The technique used here — that of gluing together non-commensurable manifolds — is known as *interbreeding*. In §6.3 manifolds are constructed by taking a double of a manifold with boundary, and so by contrast this is known as *inbreeding* [BT11, p. 1467] [Ago06].

Remark. This geometric approach relies on the existence of real co-dimension 1 totally geodesic submanifolds in a given closed hyperbolic manifold. One finds in complex hyperbolic geometry that these submanifolds do not exist and so this approach is not immediately applicable to $\mathrm{SU}(n,1)$ and hence cannot be directly adapted to produce examples of non-arithmetic lattices in those groups.

It has been shown by Bergeron, Haglund and Wise that any non-arithmetic lattice $\Gamma < \mathrm{SO}(n,1)$, constructed in the above way, may be virtually embedded as a 'quasi-convex' subgroup of an arithmetic lattice in $\mathrm{SO}(n+1,1)$ [BHW11, Prop. 9.1]. It is shown in §8.2 that for small enough ε, the manifolds constructed in Theorem 6.1 are also non-arithmetic, and so it is natural to ask whether or not this result also holds for these lattices. (This is something for future investigation, but it is expected that the result of Bergeron-Haglund-Wise would not hold for short systole lattices.)

We remark that the 'technology' used by Gromov and Piatetski-Shapiro — cutting and pasting covers — has recently been put to use in a preprint of J. Raimbault on maximal lattice growth in $\mathrm{SO}(n,1)$ [Rai11].

§8.2   NON-ARITHMETICITY OF

SHORT SYSTOLE MANIFOLDS

T. Gelander has shown that non-compact arithmetic hyperbolic manifolds cannot have systoles of length shorter than some ε_n where this ε_n depends on the dimension n [Gel04, Rem. 5.7]. He also shows that compact arithmetic manifolds have a shortest possible systole, but the ε_n in this case depends not only on the dimension, but also on the field of definition of the manifold. The

arguments used are outlined below, and we also allude to the implications for the manifolds obtained in §6.3.

Non-compact arithmetic manifolds

For non-compact lattices, we have the following lemma [Gel04, Lem. 5.1]:

Lemma 8.1 (Gelander). *Fix $n \geqslant 2$. There are numbers $\varepsilon_n > 0$ and $m \in \mathbb{N}$ such that if $\Gamma \leqslant \mathrm{Isom}(\mathbb{H}^n)$ is a non-co-compact torsion-free lattice then for every $x \in \mathbb{H}^n$ the following is satisfied:*

- *Write $\Gamma_{\varepsilon_n}(x) = \big\langle \gamma \in \Gamma \mid \mathrm{d}_{\mathbb{H}^n}\big(x, \gamma(x)\big) \leqslant \varepsilon_n \big\rangle$. The group of real points of the Zariski closure $\overline{\Gamma_{\varepsilon_n}(x)}$ has at most m connected components and its identity component is unipotent (cf. p. 12).*

Thus, we cannot have, in a non-compact arithmetic hyperbolic n-manifold, a closed geodesic of length less than ε_n, for this would correspond to a hyperbolic element γ of the fundamental group of m, with $\mathrm{d}_{\mathbb{H}^n}\big(x, \gamma(x)\big) \leqslant \varepsilon_n$ for every $x \in \mathbb{H}^n$. Along with Theorem 4.5 (the Godement criterion), Lemma 8.1 tells us that this cannot happen, as such an element would need to have a unipotent power [Gel04, Rem. 5.7]. (To see this, suppose γ is not unipotent. There are only finitely many cosets of the unipotent subgroup so some powers γ^i and γ^j lie in the same coset; that is the product $\gamma^i \gamma^{-j}$ lies in the subgroup, and this is of course a power of γ.)

Compact arithmetic manifolds

Consideration of compact manifolds involves the notion of *Mahler measure*. If $p(X)$ is a monic polynomial with coefficients in $\mathbb{Z}$ and n roots $\alpha_1, \ldots, \alpha_n$ then we define its Mahler measure $m(p)$ by

$$m(p) = \prod_{i=1}^{n} \max\big\{1, |\alpha_i|\big\}.$$

(What is called Mahler measure here might more correctly be called exponential Mahler measure as it is the exponential of another quantity which is sometimes used to define Mahler measure. On the other hand, some authors

simply use exponentiation in their definitions.) As noted by Smyth [Smy08], it follows from a result of Kronecker that the Mahler measure of p equals 1 if and only if either p or $-p$ is of the form $X^k \Phi_{n-k}(X)$ (for some $0 \leqslant k \leqslant n$), where Φ_ℓ denotes the ℓth cyclotomic polynomial

$$\Phi_\ell(X) = \prod_{\substack{\omega^\ell = 1 \\ \omega \text{ primitive}}} (X - \omega).$$

(Recall that the root ω is *primitive* if for every $k = 1, \ldots, \ell - 1$ we have $\omega^k \neq 1$.) If α is an algebraic integer, then it is a root of a unique monic integral polynomial p of smallest degree (cf. §1.1), and so for any such α we define $m(\alpha)$ to equal $m(p)$. We will need the following [SZ65]:

THEOREM 8.2 (SCHINZEL-ZASSENHAUS). *Let α be an algebraic integer, with $\alpha \neq 0$, and α not a root of unity. Let its conjugates (including α itself) be denoted by $\alpha_1, \ldots, \alpha_n$. Suppose that $2s$ of its conjugates are complex (i.e., $0 \leqslant s \leqslant n/2$). Then*

$$\max_{1 \leqslant i \leqslant n} |\alpha_i| > 1 + \frac{1}{4^{s+2}}.$$

This has the immediate consequence that $m(\alpha) > 1 + 1/4^{s+2}$, and hence the Mahler measure of any irreducible monic polynomial with integer coefficients and bounded degree is bounded away from 1. In Smyth's survey he explains that the smallest known Mahler measure is $m(p) = 1.176280818$, where p is the polynomial

$$p(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$$

[Smy08, Sect. 2]. *Lehmer's problem* is to find polynomials of smaller Mahler measure, or even to find polynomials with Mahler measure arbitrarily close to 1. It is conjectured that one cannot find such polynomials (and this is sometimes known as *Lehmer's Conjecture*). That is, it is conjectured that there exists $\beta > 1$ such that $m(p) \geqslant \beta$ for any monic integral non-cyclotomic polynomial p.

If M is a compact arithmetic hyperbolic n-manifold, then it is a quotient of the space $\mathbb{H}^n$ by a torsion-free group of isometries Γ, whose elements may

be regarded as K-integral matrices for some degree d number field $K/\mathbb{Q}$; and using restriction of scalars (cf. p. 11) we can view elements of Γ as $\mathbb{Q}$-integral matrices. A matrix $\gamma \in \Gamma$ has a monic characteristic polynomial p_γ with integral coefficients and degree at most $d(n+1)$. Hence by Theorem 8.2 there exists a constant $\delta_{n,d} > 0$ such that $m(p_\gamma) \geqslant 1 + \delta_{n,d}$ for every $\gamma \in \Gamma$, and this $\delta_{n,d}$ depends on n and d. If the roots of p_γ are bounded away from 1, then so too must be the translation length of γ; thus we see that there is some $\varepsilon_{n,d}$ such that for every $x \in \mathbb{H}^n$ and every $\gamma \in \Gamma$, $\mathrm{d}(x, \gamma x) \geqslant \varepsilon_{n,d}$.

To put this more precisely into the context of the construction of §6.3, suppose that M is a manifold produced by the proof of Theorem 6.1, arising as the double of some manifold M'. It follows from the argument used by Gromov and Piatetski-Shapiro [GPS87, §1.6, §1.7] that if M is arithmetic then its fundamental group is commensurable with $\mathrm{PO}_f(\mathcal{O}_K)$ (with f and K as in §6.3). Thus we can apply the above argument to establish non-arithmeticity in the case of K being fixed (cf. Gelander [Gel04, Prop. 10.5]). Summarising this discussion, we have the following proposition:

PROPOSITION 8.3. *Let $n \geqslant 2$ and let K be an algebraic number field of degree $d = [K : \mathbb{Q}]$. Then there is a constant $\varepsilon_{n,d} > 0$ such that the following holds:*

- *If M is a closed hyperbolic manifold from the inbreeding construction in Theorem 6.1, obtained as the double of some manifold M' with boundary, where in turn M' is obtained by cutting an arithmetic manifold commensurable with some $\mathrm{PO}_f(\mathcal{O}_K)$; and if $\mathrm{syst}_1(M) < \varepsilon_{n,d}$, then M is non-arithmetic.*

It is conjectured that there exists $L > 0$ such that for any arithmetic hyperbolic 2-manifold (or 3-manifold) M, we have $\mathrm{syst}_1(M) \geqslant L$. This is known as the *Short Geodesic Conjecture*, and if Lehmer's conjecture is true then so is the Short Geodesic Conjecture [Bel10]. Note in particular that if Lehmer's Conjecture were true then the $\varepsilon_{n,d}$ in the argument above would be independent of the field K or the dimension n.

Some of the above arguments also appear in the 2011 article of Belolipetsky and Thomson [BT11].

§8.3   Non-coherence of some lattices in $\mathrm{PO}(n,1)$

Coherence is concerned with whether of not groups are finitely presented: we say that a group Γ is *coherent* if every finitely generated subgroup of Γ is also finitely presented. In other words, Γ is *non-coherent* if it contains a subgroup $A \leqslant \Gamma$ such that A is finitely generated but not finitely presented.

We first examine some elementary results that will be of use later. A good survey of what is known of (non)-coherence of lattices in $\mathrm{O}(n,1)$ and $\mathrm{SU}(n,1)$ is given by M. Kapovich in a preprint of 2010 [Kap10].

Finite-index subgroups of finitely presented groups

Some of the following results can also follow from Reidemeister-Schreier rewriting [Joh80, p. 106], but we give elementary proofs below.

Theorem 8.4 (Schreier's Theorem and Index Formula). [Sco87, Theorem 8.4.13] *Let F be a free group and $G \leqslant F$ a subgroup. Then:*

1. *G is free; and*

2. *assuming H has finite index in F and F has finite rank, the formula*

$$\mathrm{rank}(G) = |F : G|\big(\mathrm{rank}(F) - 1\big) + 1$$

   holds.

Since every finitely generated group is a quotient of a free group of finite rank [Sco87, Theorem 8.22], this theorem implies that a subgroup of finite index in a finitely generated group is also finitely generated.

Lemma 8.5. *Let G be a group and $H \trianglelefteq G$ be a normal subgroup. Assume that $|G : H|$ is finite and that G is finitely generated. Then H is finitely presentable only if G is.*

Proof. Note that H is finitely generated by Theorem 8.4. Write $H = \langle h_1, \ldots, h_m \mid r_1, \ldots, r_n \rangle$, and let $a_1, \ldots, a_N$ be a right transversal for H in

G; that is $N = |G : H|$ and

$$G = \bigsqcup_{i=1}^{N} Ha_i.$$

Clearly if $g \in G$ then $g = ha_i$ for some $h \in H$ and some $i \in \{1, \dots, N\}$. Thus we find that H is generated by $\{h_1, \dots, h_m, a_1, \dots, a_n\}$. Now note that for any h_i and any a_j as above, we have $a_j h_i a_j^{-1} = w_{ij}(h_1, \dots, h_m)$ for some word w_{ij} in the generators of H: this follows since H is normal in G. For each $i = 1, \dots, m$ and each $j = 1, \dots, N$, denote the word $a_j h_i a_j^{-1}(w_{ij})^{-1}$ by s_{ij} and note that this is equal to the identity in g. Note also that for any $i, j = 1, \dots, N$ we have $a_i a_j = v_{ij}(h_1, \dots, h_m) a_{k(i,j)}$ for some word $v_{ij} \in H$ and some $a_{k(i,j)}$ as above, since $a_i a_j \in G$. Denote each $a_i a_j a_{k(i,j)}^{-1}(v_{ij})^{-1}$ by t_{ij} and again note that this equals the identity in G.

It is claimed that the group

$$G' = \left\langle h_1, \dots, h_m; a_1, \dots, a_N \;\middle|\; \right.$$
$$\left. r_i\ (i = 1, \dots, n),\ s_{ij}\ (i = 1, \dots, m,\ j = 1, \dots, N),\ t_{ij}\ (i, j = 1, \dots, N) \right\rangle$$

is isomorphic to G and so the expression in $\langle \cdots \rangle$ is a (finite) presentation for G. To see this, first note that if $g_1, g_2 \in G$ then $g_1 = h_{i_1} \cdots h_{i_\alpha} a_i$ and $g_2 = h_{j_1} \cdots h_{j_\beta} a_j$. Then the multiplication in G' is the same as that in G if $g_1 g_2$ can be brought into the form ha_k for some $h \in H$ and some $k \in \{1, \dots, N\}$. Indeed,

$$
\begin{aligned}
g_1 g_2 &= h_{i_1} \cdots h_{i_\alpha} a_i \cdot h_{j_1} \cdots h_{j_\beta} a_j \\
&= h_{i_1} \cdots h_{i_\alpha} a_i \cdot h_{j_1} a_i^{-1} a_i h_{j_2} a_i^{-1} a_i h_{j_3} \cdots a_i^{-1} a_i h_{j_\beta} a_i^{-1} a_i \, a_j \\
&= h_{i_1} \cdots h_{i_\alpha} w_{j_1 i} w_{j_2 i} \cdots w_{j_\beta i} \cdot a_i a_j \\
&= h' \cdot v_{ij} \, a_{k(i,j)} \qquad \text{since the } w_{ij} \text{ are in } H;\ v_{ij} \text{ as above} \\
&= h'' a_k \qquad \text{since the } v_{ij} \text{ are in } H.
\end{aligned}
$$

Thus the multiplication in G' is the same as that in G (since the relations for G' came from G in the first place). $\qquad\square$

Remark. It appears that this is the same argument as one given by P. Hall [Hal54].

One also has the following converse:

LEMMA 8.6. *Let $K \trianglelefteq H$, and assume $|H : K|$ is finite. Furthermore assume H is finitely presented. Then K is finitely presented.*

PROOF. Again note that Theorem 8.4 implies that K is finitely generated. First write

$$K = \langle k_1, \ldots, k_m \mid u_1, \ldots, u_j, \ldots \rangle$$
$$\text{and} \quad H = \langle k_1, \ldots, k_m; b_1, \ldots, b_M \mid r_1, \ldots, r_n \rangle$$

where $\{b_i\}_i$ is a (finite) transversal for K in H. Now clearly a system of relators of the form in the proof of Lemma 8.5, along with relators for K, is a system of relators for H. There are only finitely many such relations, and since n is finite there can only be finitely many relations among the k_i. (Here we use the fact that if a group is finitely presented in one set of generators then it can be finitely presented in any finite set of generators.) Hence K is finitely presented. $\qquad\square$

This lemma has the following consequence:

COROLLARY 8.7. *Let $H \leqslant G$ with finite index and assume G is finitely generated. Then H is finitely presented if and only if G is.*

PROOF. (Only if): Let $a_1, \ldots, a_N$ be a transversal for H in G. Now the group

$$K = \bigcap_{i=1}^{N} a_i^{-1} H a_i$$

has finite index in H (as it is a finite intersection of finite index subgroups: cf. Lemma 4.1), and is normal in G. To see this, choose $g \in G$ and $k \in K$: we wish to show that $gkg^{-1} \in K$. For every $i = 1, \ldots, N$ we can write $g = a_i^{-1} h' a_j$ for some j and some $h' \in H$. We can then write k as $a_j^{-1} h a_j$ for some $h \in H$, by definition of K. Then

$$gkg^{-1} = a_i^{-1} h' a_j \cdot a_j^{-1} h a_j \cdot a_j^{-1} h'^{-1} a_i = a_i^{-1} h'' a_i \in a_i^{-1} H a_i.$$

Since this holds for every i, we have $gkg^{-1} \in K$, proving normality.

By Lemma 8.6 we find that K is finitely presented, and then by Lemma 8.5 we find that G is finitely presented.

(If): Supposing G to have a finite presentation, we see by Lemma 8.6 that the group K defined above also has a finite presentation. Then Lemma 8.5 implies that H has a finite presentation. $\square$

We will need the following result in the final subsection:

PROPOSITION 8.8. *Let $\Gamma' \leqslant \Gamma$ and suppose that the index $|\Gamma : \Gamma'|$ is finite. Then Γ is coherent if and only if Γ' is.*

PROOF. We will actually prove an equivalent statement, namely that Γ is non-coherent if and only if Γ' is.

(If): This is fairly trivial. If $A \leqslant \Gamma'$ is finitely generated but not finitely presented, then A is also a subgroup of Γ with the same property.

(Only if): Let $A \leqslant \Gamma$ be finitely generated but not finitely presented. We claim that $A \cap \Gamma'$ also has this property. Indeed, since $|A : A \cap \Gamma'|$ is finite, Corollary 8.7 asserts that $A \cap \Gamma'$ is not finitely presented if it is finitely generated, and it is finitely generated by Theorem 8.4. Thus Γ' has a subgroup (i.e., $A \cap \Gamma'$) which is finitely generated but not finitely presented. $\square$

NON-COHERENCE OF LATTICES IN $\mathrm{Isom}(\mathbb{H}^n)$

Suppose Γ is a lattice in $\mathrm{Isom}(\mathbb{H}^n)$ for some $n \geqslant 2$. One may ask whether or not Γ is coherent, and whether or not its properties of (non-)arithmeticity and (non-)co-compactness might imply its (non-)coherence; indeed, one may ask whether or not there is any possibility of Γ being coherent. It is known that if $n = 2$ or $n = 3$ then every lattice in $\mathrm{Isom}(\mathbb{H}^n)$ is coherent. D. Wise has asked if there are any coherent lattices in $\mathrm{Isom}(\mathbb{H}^n)$ for $n \geqslant 4$.

In 2008, M. Kapovich, L. Potyagailo and E. Vinberg (hereinafter abbreviated as 'K.P.V.') showed the following, which we summarise as one theorem:

THEOREM 8.9 (KAPOVICH-POTYAGAILO-VINBERG). [KPV08, Theorems B & D, Corr. 1.1]

1. *If $n \geqslant 4$, then there are infinitely many commensurability classes of non-co-compact non-coherent lattices in* Isom($\mathbb{H}^n$).

2. *If $n \geqslant 6$, then every non-co-compact arithmetic lattice in* Isom($\mathbb{H}^n$) *is non-coherent.*

3. *If $n \geqslant 4$, then there exist both co-compact and non-co-compact non-arithmetic non-coherent lattices in* Isom($\mathbb{H}^n$).

This theorem suggests that the question of Wise is likely not to have an answer in the affirmative.

In proving the third part of Theorem 8.9, K.P.V. use the construction of Gromov and Piatetski-Shapiro (cf. §8.1) to produce a non-arithmetic lattice in Isom($\mathbb{H}^n$), and they show that a specific example of a non-coherent subgroup may be embedded in the resulting lattice. The fundamental example used in their proof is that of the lattice Γ in Isom($\mathbb{H}^4$) generated by reflections in the faces of the 120-cell (cf. Example 5.6 and p. 45), whose non-coherence is originally due to B. Bowditch and G. Mess. The proof of this involves embedding an example of a non-coherent group (due to B. Neumann) into Γ [KPV08, Proof of Theorem 3.1]. For ease of reference we record this example as a lemma:

LEMMA 8.10. *Let Γ be the group in* Isom($\mathbb{H}^4$) *generated by reflections in the faces of the* 120-*cell. Then Γ is non-coherent.*

In order to produce non-arithmetic examples of lattices in Isom($\mathbb{H}^n$), K.P.V. begin with a quadratic form f of signature $(n-1, 1)$, over an algebraic number field K, such that $\mathrm{PO}_f(\mathcal{O}_K)$ is a lattice in $\mathrm{PO}_f(\mathbb{R})$ (cf. §4.4). They then consider the quadratic form $h_a = f + ax_n^2$, where $a \in K$, and a is positive and such that $\mathrm{PO}_{h_a}(\mathcal{O}_K)$ is a lattice in $\mathrm{PO}_{h_a}(\mathbb{R})$. Denoting $\Gamma_a = \mathrm{PO}_{h_a}(\mathcal{O}_K)$ and $\Gamma_0 = \mathrm{PO}_f(\mathcal{O}_K)$; they choose some torsion-free finite-index subgroups $\Gamma_i' < \Gamma_i$ (for $i = 1$ and $i = a$), such that $\Gamma_1' \cap \Gamma_0 = \Gamma_a' \cap \Gamma_0$. By following the Gromov-Piatetski-Shapiro construction (cf. §8.1), and assuming a is not a square in K, they find a non-arithmetic lattice Γ with embedded submanifold $\Gamma_0' \backslash \mathbb{H}^{n-1}$ (where Γ_0' is torsion-free and of finite index in Γ_0).

The non-coherent examples of non-arithmetic lattices are acheived by considering $K = \mathbb{Q}(\sqrt{5})$ and choosing the form f above as $f = -\varphi x_0^2 + x_1^2 + \cdots + x_{n-1}^2$, where $\varphi = \frac{1}{2}(1 + \sqrt{5})$ (cf. Example 5.6). By Lemma 8.10, Γ_0 (in the notation above) is non-coherent, and so by applying the construction above to f, K.P.V. obtain examples of non-arithmetic non-coherent lattices in $\mathrm{Isom}(\mathbb{H}^n)$ [KPV08, Sect. 4].

Note that in moving from one group G to a finite-index subgroup $H < G$, one does not 'lose non-coherence', since if G is non-coherent then so too is H by Proposition 8.8.

NON-COHERENT SHORT-SYSTOLE MANIFOLDS

We construct a non-arithmetic non-coherent lattice in $\mathrm{PO}(n, 1)$ in a manner similar to K.P.V. (see above), but using the construction of non-arithmetic lattices as described in §6.3 (recalling the discussion in §8.2 concerning the (non-)arithmeticity of these groups).

Let q_n denote the quadratic form $-\varphi x_0^2 + x_1^2 + \cdots + x_n^2$ where $\varphi = \frac{1}{2}(1 + \sqrt{5})$, and let K denote the field $\mathbb{Q}(\sqrt{5})$. The group $\mathrm{PO}_{q_4}(\mathcal{O}_K)$, where $\mathcal{O}_K$ is the ring of integers of K, is commensurable to the group generated by reflections in the faces of the 120-cell in hyperbolic 4-space, and this in turn is known to be non-coherent (see above). Fixing n for the time being, let Γ be a torsion-free finite-index subgroup of $\mathrm{PO}_{q_n}(\mathcal{O}_K)$.

Now consider the hyperplane H_1 in $\mathbb{H}^n$ given by $x_n = 0$. This is stabilised by $\mathrm{PO}_{q_n'}(\mathbb{R})$ where $q_n' = -\varphi x_0^2 + \cdots + x_{n-1}^2$. In particular we have $\mathrm{PO}_{q_n'}(\mathcal{O}_K) \subseteq \mathrm{PO}_{q_n}(\mathcal{O}_K)$ in a natural way (i.e., if $\gamma \in \mathrm{PO}_{q_n'}(\mathcal{O}_K)$ then γ extends to act on $\mathbb{H}^n$ by fixing the nth coordinate). Thus the group Γ' defined by $\Gamma' = \Gamma \cap \mathrm{PO}_{q_n}(\mathcal{O}_K)$ is a torsion-free finite-index subgroup of $\mathrm{PO}_{q_n'}(\mathcal{O}_K)$. So, $\Gamma' \backslash \mathbb{H}^{n-1}$ is an $(n-1)$-manifold. By taking a suitable congruence cover $\Gamma(\mathfrak{p})$, one can embed $\Gamma' \backslash \mathbb{H}^{n-1}$ in $\Gamma(\mathfrak{p}) \backslash \mathbb{H}^n$ [Mil76].

Now as in §6.3 we can arrange to have a hyperplane H_2 which is $\varepsilon/2$-close to H_1, but with $H_1 \cap H_2 = \emptyset$. There is then a congruence cover $\Gamma(\mathfrak{p}') \backslash \mathbb{H}^n$ which allows these hyperplanes to embed without overlap (cf. Theorem 5.1).

Now $\Gamma(\mathfrak{p}) \cap \Gamma(\mathfrak{p}')$ is again a finite cover so that the fundamental group Λ of the embedded hyperplane $(\Gamma(\mathfrak{p}) \cap \Gamma(\mathfrak{p}')) \backslash H_1$ is commensurable with Γ'.

To see that Γ' is non-coherent, one uses Lemma 8.10; and since this group injects into Γ', the non-coherence is also manifest in Γ'.

In the construction of Theorem 6.1 we cut the manifold $M = (\Gamma(\mathfrak{p}) \cap \Gamma(\mathfrak{p}')) \backslash \mathbb{H}^n$ along the two embedded hyperplanes, and then take the double of the connected component containing both of them. By the Seifert-van Kampen Theorem, the fundamental group of this double will contain Λ as an amalgamated subgroup, and so the lattice associated with this manifold is non-coherent.

This construction does not add to the result of K.P.V. inasmuch as non-arithmetic non-coherent groups are already known from Theorem 8.9, but the lattices obtained from the short systole construction do constitute new examples from those previously obtained, and thus further suggest that the question of D. Wise has an answer in the negative.

Bibliography

[Ago06]   Ian   Agol,   *Systoles   of   hyperbolic   4-manifolds*,   preprint,
           arXiv:math/0612290v1, 2006.

[ALR01]   I. Agol, D. D. Long, and A. W. Reid, *The Bianchi groups are separa-
           ble on geometrically finite subgroups*, Ann. of Math. (2) **153** (2001),
           no. 3, 599–621. MR 1836283 (2002e:20099)

[AVS93]   D. V. Alekseevskij, È. B. Vinberg, and A. S. Solodovnikov, *Geom-
           etry of spaces of constant curvature*, Geometry, II, Encyclopaedia
           Math. Sci., vol. 29, Springer, Berlin, 1993, pp. 1–138. MR 1254932
           (95b:53042)

[BE11]   Mikhail Belolipetsky and Vincent Emery, *On volumes of arithmetic
           quotients of* $\mathrm{PO}(n,1)$*, n odd*, preprint, arXiv:1001.4670v2, (to ap-
           pear in Proc. L.M.S.), 2011.

[Bel04]   Mikhail   Belolipetsky,   *On   volumes   of   arithmetic   quotients   of*
           $\mathrm{S}O(1,n)$, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **3** (2004), no. 4,
           749–770, See also [Bel07]. MR 2124587 (2005k:11080)

[Bel07]   _____, *Addendum to: "On volumes of arithmetic quotients of*
           $\mathrm{S}O(1,n)$*" [Bel04]*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **6** (2007),
           no. 2, 263–268. MR 2352518 (2008f:11048)

97

[Bel10]   _____, *Geodesics, volumes and Lehmer's conjecture*, Oberwolfach
            Reports (2010), 2136–2139.

[BH99]   Martin R. Bridson and André Haefliger, *Metric Spaces of
            Non-Positive Curvature*, Grundlehren der Mathematischen Wis-
            senschaften [Fundamental Principles of Mathematical Sciences], vol.
            319, Springer-Verlag, Berlin, 1999. MR 1744486 (2000k:53038)

[BHC62]   Armand Borel and Harish-Chandra, *Arithmetic subgroups of alge-
            braic groups*, Ann. of Math. (2) **75** (1962), 485–535. MR 0147566
            (26 #5081)

[BHW11]   Nicolas Bergeron, Frédéric Haglund, and Daniel T. Wise,
            Hyperplane sections in arithmetic hyperbolic manifolds,
            J. Lond. Math. Soc. (2) **83** (2011), no. 2, 431–448. MR 2776645

[BK10]   Martin Bridgeman and Jeremy Kahn, *Hyperbolic volume of mani-
            folds with geodesic boundary and orthospectra*, Geom. Funct. Anal.
            20 (2010), no. 5, 1210–1230. MR 2746952

[Bor81]   A. Borel, *Commensurability classes and volumes of hyperbolic 3-
            manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **8** (1981), no. 1,
            1–33. MR 616899 (82j:22008)

[Bow93]   B. H. Bowditch, *Geometrical finiteness for hyperbolic groups*,
            J. Funct. Anal. **113** (1993), no. 2, 245–317. MR 1218098 (94e:57016)

[BP92]   R. Benedetti and C. Petronio, *Lectures on Hyperbolic Geometry*,
            Universitext, Springer-Verlag, 1992.

[BT11]   Mikhail V. Belolipetsky and Scott A. Thomson, *Systoles of hyper-
            bolic manifolds*, Algebraic & Geometric Topology **11** (2011), no. 3,
            1455–1469. MR 2821431

[Cas08]   J. W. S. Cassels, *Rational quadratic forms*, Dover Publications, Mi-
            neola, NY, 2008, Unabridged republication of 1968 edition, Aca-
            demic Press, London and New York.

[Cor92]   Kevin Corlette, *Archimedean superrigidity and hyperbolic geometry*, Ann. of Math. (2) **135** (1992), no. 1, 165–182. MR 1147961 (92m:57048)

[DM93]   Pierre Deligne and G. Daniel Mostow, *Commensurabilities among lattices in* $\mathrm{PU}(1, n)$, Annals of Mathematics Studies, vol. 132, Princeton University Press, Princeton, NJ, 1993. MR 1241644 (95a:22013)

[FH91]   William Fulton and Joe Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics. MR 1153249 (93a:20069)

[Gel04]   Tsachik Gelander, *Homotopy type and volume of locally symmetric manifolds*, Duke Math. J. **124** (2004), no. 3, 459–515. MR 2084613 (2005i:53050)

[GPS87]   M. Gromov and I. Piatetski-Shapiro, *Non-arithmetic groups in Lobachevsky spaces*, Publications mathématiques de l'I.H.É.S. **66** (1987), no. 3, 93–103. MR 0932135 (89j:22019)

[Gro83]   Mikhael Gromov, *Filling Riemannian manifolds*, J. Differential Geom. **18** (1983), no. 1, 1–147. MR 697984 (85h:53029)

[Gro96]   ———, *Systoles and intersystolic inequalities*, Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992), Sémin. Congr., vol. 1, Soc. Math. France, Paris, 1996, pp. 291–362. MR 1427763 (99a:53051)

[Hal54]   P. Hall, *Finiteness conditions for soluble groups*, Proc. London Math. Soc. (3) **4** (1954), 419–436. MR 0072873 (17,344c)

[Hum75]   James E. Humphreys, *Linear algebraic groups*, Springer-Verlag, New York, 1975, Graduate Texts in Mathematics, No. 21. MR 0396773 (53 #633)

[Joh80]   D. L. Johnson, *Topics in the theory of group presentations*, London Mathematical Society Lecture Note Series, vol. 42, Cambridge University Press, Cambridge, 1980. MR 695161 (84f:20003)

[Kap10]   Michael Kapovich, *Non-coherence of arithmetic hyperbolic lattices*, preprint, arXiv:1005.4135v1, 2010.

[Kat92]   Svetlana Katok, *Fuchsian Groups*, University of Chicago Press, 1992.

[Kat07]   Mikhail G. Katz, *Systolic geometry and topology*, Mathematical Surveys and Monographs, vol. 137, American Mathematical Society, Providence, RI, 2007, With an appendix by Jake P. Solomon. MR 2292367 (2008h:53063)

[KM68a]   D. A. Každan and G. A. Margulis, *A proof of Selberg's hypothesis*, Math. USSR Sb. **4** (1968), no. 1, 147–152, English translation of [KM68b].

[KM68b]   ———, *A proof of Selberg's hypothesis*, Mat. Sb. (N.S.) **75 (117)** (1968), 163–168. MR 0223487 (36 #6535)

[KNV11]   Benjamin Klopsch, Nikolay Nikolov, and Christopher Voll, *Lectures on profinite topics in group theory*, London Mathematical Society Student Texts, vol. 77, Cambridge University Press, Cambridge, 2011, Lectures from the course on Asymptotic Methods in Infinite Group Theory held in Oxford, September 2007, Edited by Dan Segal. MR 2797110 (2011m:20002)

[KPV08]   Michael Kapovich, Leonid Potyagailo, and Ernest Vinberg, *Noncoherence of some lattices in* $\mathrm{I}som(\mathbb{H}^n)$, The Zieschang Gedenkschrift, Geom. Topol. Monogr., vol. 14, Geom. Topol. Publ., Coventry, 2008, pp. 335–351. MR 2484708 (2010g:22018)

[Lan70]   Serge Lang, *Algebraic number theory*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London-Don Mills, Ont., 1970. MR 0282947 (44 #181)

[Lan02] _____, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR 1878556 (2003e:00003)

[Mar91] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, Berlin, 1991. MR 1090825 (92h:22021)

[Mil76] John J. Millson, *On the first Betti number of a constant negatively curved manifold*, Ann. of Math. (2) **104** (1976), no. 2, 235–247. MR 0422501 (54 #10488)

[Mos78] G. D. Mostow, *Existence of a nonarithmetic lattice in* SU(2, 1), Proc. Nat. Acad. Sci. U.S.A. **75** (1978), no. 7, 3029–3033. MR 499658 (80f:22007)

[MR03] Colin Maclachlan and Alan W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics, vol. 219, Springer-Verlag, New York, 2003. MR 1937957 (2004i:57021)

[MT62] G. D. Mostow and T. Tamagawa, *On the compactness of arithmetically defined homogeneous spaces*, Ann. of Math. (2) **76** (1962), 446–463. MR 0141672 (25 #5069)

[MV00] G. A. Margulis and È. B. Vinberg, *Some linear groups virtually having a free quotient*, J. Lie Theory **10** (2000), no. 1, 171–180. MR 1748082 (2001h:22016)

[Nab03] Gregory L. Naber, *The Geometry of Minkowski Spacetime*, Dover Publications, 2003, originally published by Springer-Verlag (1992).

[Neu99] Jürgen Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999, Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. MR 1697859 (2000m:11104)

[NZ85]   Walter D. Neumann and Don Zagier, *Volumes of hyperbolic three-manifolds*, Topology **24** (1985), no. 3, 307–332. MR 815482 (87j:57008)

[O'M71]   O. T. O'Meara, *Introduction to quadratic forms*, second ed., Springer-Verlag, New York, 1971, Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 117. MR 0347768 (50 #269)

[OV93]   A. L. Onishchik and È. B. Vinberg, *Foundations of Lie theory [see MR0950861 (89m:22010)]*, Lie groups and Lie algebras, I, Encyclopaedia Math. Sci., vol. 20, Springer, Berlin, 1993, pp. 1–94, 231–235. MR 1306738

[Par09]   John R. Parker, *Complex hyperbolic lattices*, Discrete groups and geometric structures, Contemp. Math., vol. 501, Amer. Math. Soc., Providence, RI, 2009, pp. 1–42. MR 2581913 (2011g:22018)

[PR94]   Vladimir Platonov and Andrei Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics, vol. 139, Academic Press Inc., Boston, MA, 1994, Translated from the 1991 Russian original by Rachel Rowen. MR 1278263 (95b:11039)

[Rag72]   M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, New York, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68. MR 0507234 (58 #22394a)

[Rai11]   Jean Raimbault, *A note on maximal lattice growth in* $\mathrm{SO}(1, n)$, preprint, arXiv:1112.2484v1, 2011.

[Rat06]   John G. Ratcliffe, *Foundations of Hyperbolic Manifolds*, Second ed., Graduate Texts in Mathematics, no. 149, Springer, 2006.

[Rez95]   Alexander Reznikov, *The volume and the injectivity radius of a hyperbolic manifold*, Topology **34** (1995), no. 2, 477–479. MR 1318888 (96a:53049)

[Rom05]   Steven Roman, *Advanced Linear Algebra*, Second ed., Graduate Texts in Mathematics, no. 135, Springer, 2005.

[Sch08]   Eric Schmutz, *Rational points on the unit sphere*, Cent. Eur. J. Math. **6** (2008), no. 3, 482–487. MR 2425007 (2009c:11112)

[Sco78]   Peter Scott, *Subgroups of surface groups are almost geometric*, J. London Math. Soc. (2) **17** (1978), no. 3, 555–565. MR 0494062 (58 #12996)

[Sco87]   W. R. Scott, *Group theory*, second ed., Dover Publications Inc., New York, 1987. MR 896269 (88d:20001)

[Ser73]   J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7. MR 0344216 (49 #8956)

[Smy08]   Chris Smyth, *The Mahler measure of algebraic numbers: a survey*, Number theory and polynomials, London Math. Soc. Lecture Note Ser., vol. 352, Cambridge Univ. Press, Cambridge, 2008, (Preprint version at http://arxiv.org/abs/math/0701397v2), pp. 322–349. MR 2428530 (2009j:11172)

[Sti01]   John Stillwell, *The story of the 120-cell*, Notices Amer. Math. Soc. **48** (2001), no. 1, 17–24. MR 1798928 (2001k:52019)

[SZ65]   A. Schinzel and H. Zassenhaus, *A refinement of two theorems of Kronecker*, Michigan Math. J. **12** (1965), 81–85. MR 0175882 (31 #158)

[Thu80]   W. P. Thurston, *The geometry and topology of 3-manifolds*, 1980, available at: `http://library.msri.org/books/gt3m/`.

[VGS00]   È. B. Vinberg, V. V. Gorbatsevich, and O. V. Shvartsman, *Discrete subgroups of Lie groups [MR0968445 (90c:22036)]*, Lie groups and Lie algebras, II, Encyclopaedia Math. Sci., vol. 21, Springer, Berlin, 2000, pp. 1–123, 217–223. MR 1756407

[VS93]   È. B. Vinberg and O. V. Shvartsman, *Discrete groups of motions of spaces of constant curvature*, Geometry, II, Encyclopaedia Math. Sci., vol. 29, Springer, Berlin, 1993, pp. 139–248. MR 1254933 (95b:53043)

[Wan72]   Hsien Chung Wang, *Topics on totally discontinuous groups*, Symmetric spaces (Short Courses, Washington Univ., St. Louis, Mo., 1969–1970), Dekker, New York, 1972, pp. 459–487. Pure and Appl. Math., Vol. 8. MR 0414787 (54 #2879)

[WM08]   David Witte Morris, *Introduction to arithmetic groups*, Available at: `http://people.uleth.ca/~dave.morris/books/IntroArithGroups.html`, 2008.

[Zim84]   Robert J. Zimmer, *Ergodic theory and semisimple groups*, Monographs in Mathematics, vol. 81, Birkhäuser Verlag, Basel, 1984. MR 776417 (86j:22014)

Index