Uniformly Diophantine Numbers in a Fixed Real Quadratic Field

Citation
McMullen, Curtis T. 2009. Uniformly Diophantine numbers in a fixed real quadratic field. Compositio Mathematica 145(4): 827-844.

Published Version
doi:10.1112/S0010437X09004102

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:9925390

Uniformly Diophantine numbers in a fixed real quadratic field

Curtis T. McMullen*

17 June, 2008

Abstract

The field $\mathbb{Q}(\sqrt{5})$ contains the infinite sequence of uniformly bounded continued fractions $[\overline{1,4,2,3}], [\overline{1,1,4,2,1,3}], [\overline{1,1,1,4,2,1,1,3}]\ldots$, and similar patterns can be found in $\mathbb{Q}(\sqrt{d})$ for any $d > 0$. This paper studies the broader structure underlying these patterns, and develops related results and conjectures for closed geodesics on arithmetic manifolds, packing constants of ideals, class numbers and heights.

Contents

1 Introduction

It is well-known that any periodic continued fraction defines a real number which is quadratic over $\mathbb{Q}$. Remarkably, it is also true that any fixed real quadratic field $\mathbb{Q}(\sqrt{d})$ contains infinitely many *uniformly bounded* periodic continued fractions. For example, $\mathbb{Q}(\sqrt{5})$ contains the infinite sequence of continued fractions

$$[\overline{1,4,2,3}], [\overline{1,1,4,2,1,3}], [\overline{1,1,1,4,2,1,1,3}]\ldots, \tag{1.1}$$

and similar patterns can be found for any $d > 0$ [Wil] (see also [Wd] and §4 below).

*Research supported in part by the NSF. 2000 Mathematics Subject Classification: Primary 11, Secondary 37.

In this paper we study the broader structure underlying these patterns, give a conceptual construction of them, and develop related results and conjectures for closed geodesics on arithmetic manifolds, packing constants of ideals, class numbers and heights on finite projective spaces.

Continued fractions. Every real number x can be expressed uniquely as a continued fraction

$$x = [a_0, a_1, \cdots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

with $a_i \in \mathbb{Z}$ and $a_i \geq 1$ for $i \geq 1$. If the continued fraction is periodic ($a_{i+p} = a_i$), we write $x = [\overline{a_0, \ldots, a_{p-1}}]$. In §2 we give a new proof of the following result of Wilson:

Theorem 1.1 *Any real quadratic field $\mathbb{Q}(\sqrt{d})$ contains infinitely many periodic continued fractions $x = \overline{[a_0, \ldots, a_{p-1}]}$ with $1 \leq a_i \leq M_d$.*

Here M_d denotes a constant that depends only on d; for example, by (1.1) we can take $M_5 = 4$.

Closed geodesics. Theorem 1.1 can be formulated geometrically as follows. Let $L(\gamma)$ denote the length of a closed geodesic γ on a Riemannian manifold (or orbifold) M. We say γ is *fundamental* if there is no shorter geodesic whose length divides $L(\gamma)$.

Theorem 1.2 *For any fundamental geodesic $\gamma \subset M = \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$, there is a compact subset of M that contains infinitely many primitive, closed geodesics whose lengths are integral multiples of $L(\gamma)$.*

(A geodesic is *primitive* if it is indivisible in $\pi_1(M)$.)

Measure-zero phenomena. To give some perspective on this result, fix a compact set $Z \subset \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$. Then the complete geodesics that lie entirely in Z form a closed set $G(Z) \subset Z$ of measure zero. On the other hand, the geodesics of length $mL(\gamma)$ become uniformly distributed on $\mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ as $m \to \infty$ [Du] (see also [Lin, Ch. 7]).

Thus most geodesics whose lengths are multiples of $L(\gamma)$ are not contained in Z. Theorem 1.2 shows that, nevertheless, there are infinitely many such geodesics once Z is sufficiently large.

It is also known that the Hausdorff dimension of $G(Z)$ can be made arbitrarily close to 2 by taking Z large enough [Ja] (see also [Sch] and [Hen]).

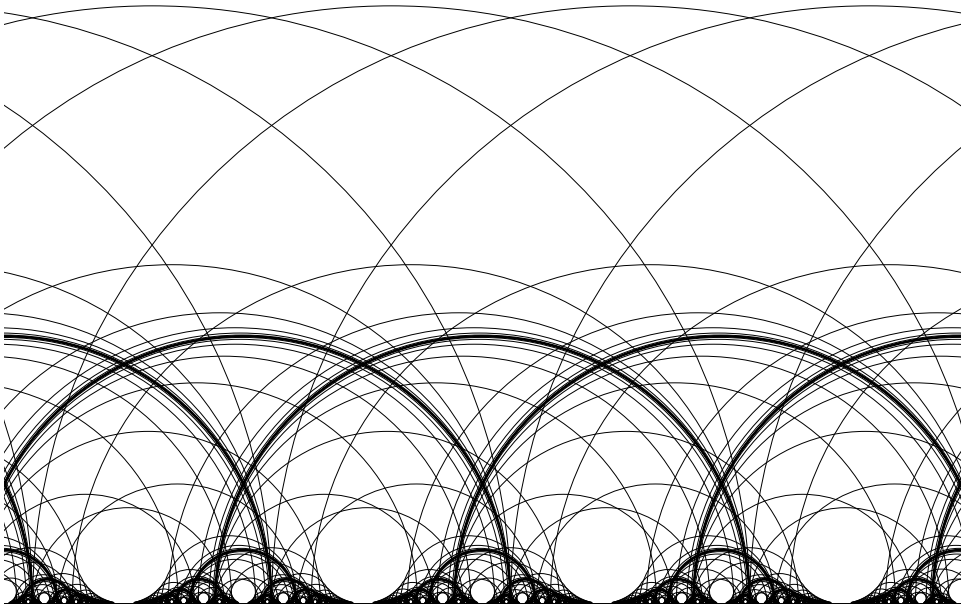

Figure 1. A long, bounded geodesic on $\mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ defined over $\mathbb{Q}(\sqrt{5})$.

A corresponding conjecture on the *number* of geodesics in $G(Z)$ of length $mL(\gamma)$ will be formulated (in terms of ideals) in §6.

Dynamics and laminations. An example of Theorem 1.2 is provided by the closed geodesics $\gamma_m \subset M = \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ associated to the periodic continued fractions given by equation (1.1). The preimage of one such geodesic on $\mathbb{H}$, for $m \gg 0$, is shown in Figure 1. As can be seen in the Figure, γ_m spends most of its time spiraling close to the golden mean geodesic ξ, defined by the continued fraction $[1, 1, 1, \ldots]$. This behavior is also apparent from the long strings of 1's that dominate the continued fraction expansion of x_m. At the same time γ_m stays well-away from the cusp of M; note the horoballs along the real axis that its lift avoids.

As $m \to \infty$, γ_m converges to a compact, immersed lamination γ_∞ consisting of the closed geodesic ξ and two infinite geodesics spiraling towards it. Conversely, it follows from general principles in dynamics that γ_∞ can be approximated by a sequence of closed geodesics γ_m (see e.g. [Sm]). What is unusual is that, in the case at hand, the geodesics γ_m can be chosen so their lengths are all multiples of a single number.

Hyperbolic 3-manifolds. Theorem 1.2 also holds for the Bianchi groups $\operatorname{SL}_2(\mathcal{O}_d)$, where $\mathcal{O}_d \subset \mathbb{Q}(\sqrt{-d})$ is the ring of integers in a quadratic imagi-

nary field; in §5 we show:

Theorem 1.3 *For any fundamental geodesic γ on the hyperbolic orbifold $\mathbb{H}^3/\operatorname{SL}_2(\mathcal{O}_d)$, there is a compact set that contains infinitely many primitive closed geodesic whose lengths are integral multiples of $L(\gamma)$.*

Ideals. To formulate a third variant of Theorem 1.1, let $K/\mathbb{Q}$ be a number field of degree d, and let $\mathrm{N}_{\mathbb{Q}}^K$ and $\mathrm{tr}_{\mathbb{Q}}^K$ denote the norm and the trace to $\mathbb{Q}$. Let $I(K)$ denote the set of lattices $I \subset K$ (meaning additive subgroups isomorphic to $\mathbb{Z}^d$), modulo rescaling by elements of K^*. Every $[I] \in I(K)$ represents an ideal class for some order in K [BoS, Ch 2.2].

Recall that the discriminant of $I = \oplus \mathbb{Z} x_i$ is given with respect to an integral basis by $\operatorname{disc}(I) = \det(\mathrm{tr}_{\mathbb{Q}}^K x_i x_j)$. We define the *packing density* of I by

$$\delta(I) = \frac{N^*(I)}{\det(I)},$$

where $\det(I) = \sqrt{|\operatorname{disc}(I)|}$ and

$$N^*(I) = \min\{|\mathrm{N}_{\mathbb{Q}}^K(x)| \,:\, x \in I, \mathrm{N}_{\mathbb{Q}}^K(x) \neq 0\}.$$

The packing density depends only on the class of I; in the case of a quadratic imaginary field, it measures the quality of the sphere packing defined by the lattice $I \subset K \subset \mathbb{C}$.

In these terms, Theorem 1.1 is equivalent to:

Theorem 1.4 *In any real quadratic field K, there are infinitely many ideal classes with $\delta(I) > \delta_K > 0$.*

It is easy to verify that the same result holds for quadratic imaginary fields. More generally, we propose:

Conjecture 1.5 *If K is a number field whose unit group $\mathcal{O}_K^*$ has rank one, then there are infinitely many ideal classes I whose packing density satisfies $\delta(I) > \delta_K > 0$.*

The remaining cases are cubic fields with one complex place and quartic fields with two complex places.[1] Conjecture 1.5 is meant to complement:

[1] The special case of quartic fields with quadratic subfields follows from Theorems 1.2 and 1.3.

Conjecture 1.6 *Up to isomorphism, there are only finitely many totally real cubic fields K and ideal classes $[I] \in I(K)$ with $\delta(I) \geq \delta > 0$.*

This conjecture was formulated in 1955 (in terms of products of linear forms) by Cassels and Swinnerton-Dyer [CaS, Thm. 5]; it is open even when K is fixed. A general rigidity conjecture of Margulis [Mg, Conj. 9] implies Conjecture 1.6 (cf. [ELMV, Conj. 1.3]).

Heights and densities. In §6 we show packing densities of ideals are related to heights on finite projective spaces. This perspective suggests a quantitative lower bound on the number of ideals with $\delta(I) > \delta$. It also connects the discussion to Zaremba's conjecture on *rationals* that are far from other rationals, and leads to a strategy for the cubic and quartic cases of Conjecture 1.5.

Arithmetic groups. As one final generalization Theorem 1.1, we propose:

Conjecture 1.7 *Given $U \in \mathrm{GL}_N(\mathbb{Z})$, either:*

1. *U has finite order;*
2. *The characteristic polynomial of U is reducible in $\mathbb{Z}[x]$; or*
3. *There exists a compact, U-invariant subset of $\mathrm{PGL}_N(\mathbb{R})/\mathrm{GL}_N(\mathbb{Z})$ containing U-periodic points of arbitrarily large period.*

(These alternatives are not mutually exclusive.) Theorem 1.2 establishes this conjecture for $N = 2$. More generally, in §5 we will show:

Theorem 1.8 *Conjecture 1.7 holds if U is conjugate to U^{-1} in $\mathrm{GL}_N(\mathbb{Q})$.*

Notes and references. The classical theory of continued fractions is presented in [HW]; for the geometric approach see e.g. [Po], [Ser] and [KU]. More on packing densities and the geometry of numbers can be found in [GL]. For a survey on bounded continued fractions, see [Sha].

I would like to thank N. Elkies, B. Gross and B. Kra for useful conversations, and A. Venkatesh for bringing the earlier work [Wil] to my attention.

Notation. The notations $A = O(B)$ and $A \asymp B$ mean $A < CB$ and $B/C < A < CB$, for an implicit constant $C > 0$.

2 Lattices and quadratic fields

In this section we prove Theorem 1.1 and its variants for real quadratic fields.

Matrices. Let $\mathrm{M}_2(\mathbb{R})$ denote the ring of 2×2 real matrices with identity I. Let $\|x\|$ denote the Euclidean norm on $\mathbb{R}^2$, and let $\|A\| = \sup \|Ax\|/\|x\|$ denote the operator norm on $\mathrm{M}_2(\mathbb{R})$. There is a unique involution $A \mapsto A^\dagger$ such that $A + A^\dagger = \mathrm{tr}(A)I$, given explicitly by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^\dagger = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. We have $(AB)^\dagger = B^\dagger A^\dagger$ and $AA^\dagger = (\det A)I$, which implies the useful identity:

$$\det(A + B) = \det(A) + \det(B) + \mathrm{tr}(AB^\dagger). \tag{2.1}$$

Lattices. Every lattice in $\mathbb{R}^2$ can be presented in the form $\Lambda = L(\mathbb{Z}^2)$ with $L \in \mathrm{GL}_2(\mathbb{R})$. The choice of L gives a basis for Λ, and multiplying L by a scalar changes Λ by a similarity. Since any two bases for $\mathbb{Z}^2$ are related by $\mathrm{GL}_2(\mathbb{Z})$, the moduli space of lattices up to similarity is given by

$$\mathrm{PGL}_2(\mathbb{R})/\,\mathrm{GL}_2(\mathbb{Z}).$$

We let $[L]$ denote the point in moduli space represented by L. There is a natural left action of $\mathrm{GL}_2(\mathbb{R})$ on $\mathrm{PGL}_2(\mathbb{R})/\,\mathrm{GL}_2(\mathbb{Z})$, sending $[L]$ to $[AL]$.

Real quadratic fields. Let $\epsilon \in \mathbb{R}$ be an algebraic unit of degree two over $\mathbb{Q}$, with $\epsilon > 1$. Then $\epsilon^2 = t\epsilon - n$, where $t = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon) > 0$ and $n = \mathrm{N}_{\mathbb{Q}}^K(\epsilon) = \pm 1$. The discriminant of the order $\mathbb{Z}[\epsilon]$ in the field $K = \mathbb{Q}(\epsilon)$ is given by

$$D = t^2 - 4n > 0.$$

We will use $(1, \epsilon)$ as a basis for $\mathbb{Z}[\epsilon]$. The action of multiplication by ϵ with respect to this basis is given by

$$U = \begin{pmatrix} 0 & -n \\ 1 & t \end{pmatrix}. \tag{2.2}$$

Similarly, the action of $\sqrt{D}$ is given by $S = 2U - tI = \left(\begin{smallmatrix} -t & -2n \\ 2 & t \end{smallmatrix}\right)$.

Traces. Galois conjugation in K stabilizes $\mathbb{Z}[\epsilon]$ and will be denoted by $x \mapsto x'$. We use the same notation for Galois conjugation on the entries of vectors in K^2 and matrices in $\mathrm{M}_2(K)$. In particular we have an entrywise trace map

$$\mathrm{tr}_{\mathbb{Q}}^K : \mathrm{M}_2(K) \to \mathrm{M}_2(\mathbb{Q})$$

sending A to $A + A'$.

Eigenprojections. Note that $v = (\epsilon', -1)$ and $v' = (\epsilon, -1)$ are eigenvectors for $U|K^2$ with eigenvalues ϵ and ϵ'. The projections $\widetilde{U}$ and $\widetilde{U}'$ onto these eigenspaces are given by

$$\widetilde{U} = \frac{1}{2}\left(I + \frac{S}{\sqrt{D}}\right) \quad \text{and} \quad \widetilde{U}' = \frac{1}{2}\left(I - \frac{S}{\sqrt{D}}\right) \tag{2.3}$$

respectively; they satisfy $\widetilde{U}\widetilde{U}' = \widetilde{U}'\widetilde{U} = 0$, $\widetilde{U} + \widetilde{U}' = I$, and $\widetilde{U}^\dagger = \widetilde{U}'$. For any $x \in K$, the matrix $\mathrm{tr}_{\mathbb{Q}}^K(x\widetilde{U})$ gives the action of multiplication by x on $K \cong \mathbb{Q}^2$ with respect to the basis $(1, \epsilon)$; in particular, $U^m = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^m \widetilde{U})$.

Fibonacci numbers. The unit ϵ determines a generalized Fibonacci sequence by $f_0 = 0$, $f_1 = 1$ and

$$f_{m+1} = tf_m - nf_{m-1}$$

for $m > 1$. (For $\epsilon = (1 + \sqrt{5})/2$ we obtain the usual Fibonacci sequence.) One can check that

$$f_m = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^m/\sqrt{D}); \tag{2.4}$$

in particular, $f_m \asymp \epsilon^m$ for large m.

By induction we find $\epsilon^m = f_m\epsilon - nf_{m-1}$, and hence the ring

$$\mathbb{Z}[\epsilon^m] = \mathbb{Z} + f_m\mathbb{Z}[\epsilon]$$

has discriminant $f_m^2 D$. Similarly we have

$$U^m = f_m U - nf_{m-1}I, \tag{2.5}$$

and hence

$$U^m = \begin{pmatrix} -n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{pmatrix} \equiv f_{m+1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod f_m. \tag{2.6}$$

These relations also hold for $m < 0$, and lead to the following useful fact.

Proposition 2.1 *If $L \in \mathrm{M}_2(\mathbb{Z})$ satisfies $\det(L) = \pm f_m$, then the lattice $[L] \in \mathrm{PGL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z})$ is fixed by U^m.*

Proof. Using the identity $L^{-1} = \pm f_m^{-1} L^\dagger$ and (2.5), we find $U^m L = LV_m$ where

$$V_m = L^{-1}U^m L = \pm L^\dagger U L - nf_{m-1}I$$

visibly lies in $\mathrm{GL}_2(\mathbb{Z})$. ∎

Main construction. We can now explicitly construct lattices with uniformly bounded orbits under the action of $\langle U \rangle$.

Theorem 2.2 *Given $A \in \mathrm{GL}_2(\mathbb{Z})$ such that $A^2 = I$, $\mathrm{tr}(A) = 0$ and $\mathrm{tr}(A^\dagger U) = \pm 1$, let*

$$L_m = U^m + U^{-m} A.$$

Then for all $m \geq 0$:

1. *$|\det(L_m)| = f_{2m}$ is a generalized Fibonacci number;*
2. *The lattice $[L_m]$ is fixed by U^{2m};*
3. *We have $L_{-m} = L_m A$;*
4. *For $0 \leq i \leq m$ we have:*

$$\|U^i L_m U^{-i}\|, \|U^{-i} L_{-m} U^i\| \ \leq C\sqrt{|\det L_m|}, \tag{2.7}$$

where C depends only on A and U.

Proof. Our assumptions imply $\det(A) = -1$. Since $UU^\dagger = \pm I$ and $U^{2m} = f_{2m} U - n f_{2m-1} I$, (2.5) gives

$$\begin{aligned} \det(L_m) &= \det(U^m) + \det(U^{-m} A) + \mathrm{tr}(U^m A^\dagger (U^{-m})^\dagger) \\ &= \pm\,\mathrm{tr}(A^\dagger U^{2m}) = \pm f_{2m} \end{aligned}$$

establishing (1). By construction L_m is integral, so Proposition 2.1 implies (2). Since $A^2 = I$ we have (3). For (4) first recall that $f_i \asymp \epsilon^i$ for $i > 0$; in particular, $\|U^{\pm i}\| \leq \epsilon^i$ by (2.6). Thus for $0 \leq i \leq m$ we have

$$\|U^i L_m U^{-i}\| = \|U^m + U^{i-m} A U^{-i}\| = O(\epsilon^m) = O(\sqrt{f_{2m}}) = O(\sqrt{|\det L_m|}).$$

A similar bound holds for $U^i L_{-m} U^{-i}$, which gives (4). ∎

Corollary 2.3 *There is a compact subset of $\mathrm{PGL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z})$ which contains the lattices $[U^i L_m]$ for all $i, m \in \mathbb{Z}$.*

Proof. Since $A, U \in \mathrm{GL}_2(\mathbb{Z})$ and $[U^{2m} L_m] = [L_m]$, the lattices $[U^i L_m]$ are represented in $\mathrm{GL}_2(\mathbb{R})$ by the matrices

$$\frac{U^i L_m U^{-i}}{\sqrt{|\det L_m|}} \quad \text{and} \quad \frac{U^{-i} L_{-m} U^i}{\sqrt{|\det L_m|}}$$

with $0 \leq i \leq m$. These matrices in turn lie in a compact subset of $\mathrm{GL}_2(\mathbb{R})$, since they have determinant ± 1 and their norms are uniformly bounded by (2.7). Projecting, we obtain a compact set in $\mathrm{PGL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z})$ containing the lattices $[U^i L_m]$. ∎

Theorem 2.4 *The size of the orbit of* $[L_m]$ *under* $\langle U\rangle$ *tends to infinity as* $m \to \infty$.

Proof. Let $V_m = L_m^{-1} U L_m$. Then the size $k(m)$ of the orbit of $[L_m]$ under $\langle U\rangle$ is the same as the least positive integer such that $V_m^{k(m)} \in \mathrm{GL}_2(\mathbb{Z})$.

Replacing U by U^2 if necessary, we can assume $\det(U) = 1$. Let $\widetilde{U}$ and $\widetilde{U}'$ (given by (2.3)) denote projection onto the ϵ and ϵ' eigenspaces of U, spanned by $v = (\epsilon', -1)$ and $v' = (\epsilon, -1)$ respectively. It then easy to see that

$$L = \lim_{m\to\infty} \epsilon^{-m} L_m = \widetilde{U} + \widetilde{U}' A, \tag{2.8}$$

and $\det(L) = \pm \lim \epsilon^{-2m} f_{2m} \neq 0$. Consequently

$$V_m \to V = L^{-1} U L$$

in $\mathrm{GL}_2(\mathbb{R})$. Since L^{-1} is a scalar multiple of $L^\dagger = \widetilde{U}' - A\widetilde{U}$, an eigenbasis for V is given by

$$(w, w') = (L^\dagger v, L^\dagger v') = (-Av, v').$$

Now suppose $V^k \in \mathrm{GL}_2(\mathbb{Z})$ for some $k > 0$. Then v' and $-A(v)$ are eigenvectors for V^k as well. Since V^k is integral, v is also an eigenvector for V^k, and hence $-A(v)$ is scalar multiple of v. But the eigenvalues of A are -1 and $+1$, so its eigenspaces are rational, contradicting the fact v and v' are linearly independent.

It follows that $V^k \notin \mathrm{GL}_2(\mathbb{Z})$ for all $k > 0$, and hence $k(m) \to \infty$. ∎

Existence. The matrix

$$A = \begin{pmatrix} 1 & t-1 \\ 0 & -1 \end{pmatrix} \tag{2.9}$$

satisfies the conditions of Theorem 2.2 with $\mathrm{tr}(A^\dagger U) = 1$. Thus lattices L_m of the type just described exist for any unit $\epsilon > 1$. For example, when $N(\epsilon) = 1$ this value of A gives

$$L_m = \begin{pmatrix} f_{m+1} - f_{m-1} & f_{m+2} - f_{m+1} - f_m \\ 0 & f_m \end{pmatrix}.$$

It is now straightforward to establish Theorem 1.1 and its variants, Theorems 1.2 and 1.4.

Geodesics: Proof of Theorem 1.2. Let $\gamma \subset M = \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ be a fundamental geodesic, corresponding to an element $U \in \operatorname{SL}_2(\mathbb{Z})$. Since U and $-U$ represent the same geodesic, we may assume the largest eigenvalue of U is a quadratic unit $\epsilon > 1$ with norm one. Changing γ to another geodesic of equal length, we can also assume U is given by equation (2.2).

Since U is semisimple, its centralizer H in $\operatorname{PSL}_2(\mathbb{R})$ is conjugate to the subgroup of diagonal matrices. Thus we can identify the unit tangent bundle $\mathrm{T}_1(M)$ with $\operatorname{PSL}_2(\mathbb{R})/\operatorname{SL}_2(\mathbb{Z})$ in such a way that H represents the geodesic flow, and the compact orbit $H \cdot [I] \cong H/\langle U\rangle$ projects to γ.

Now let $[L_m]$ be the sequence of lattices furnished by Theorem 2.2, e.g. with A given by (2.9). Normalize so that $\det(L_m) = 1$. Let $v_m \in \mathrm{T}_1(M)$ be the corresponding unit vectors, which lie in a compact, U-invariant set $Z \subset T_1(M)$. Since $H/\langle U\rangle$ is compact, we can also assume Z is H-invariant.

By Theorem 2.4, the orbit of v_m under U has length $k(m) \to \infty$. Since U is fundamental, the stabilizer of v_m in H is generated by $U^{k(m)}$ (else $\epsilon > 1$ would be a power of a smaller, norm one unit $\eta > 1$ in K). Thus $Hv_m \subset \mathrm{T}_1(M)$ projects to a closed geodesic $\gamma_m \subset M$ with $L(\gamma_m) = k(m)L(\gamma)$, and all these geodesics lie in the compact set obtained by projecting $Z \subset \mathrm{T}_1(M)$ to M. ∎

Continued fractions: Proof of Theorem 1.1. Let $K \subset \mathbb{R}$ be a real quadratic field. By Dirichlet's theorem, $K = \mathbb{Q}(\epsilon)$ for some unit $\epsilon > 1$ which arises as an eigenvalue of a matrix $U \in \operatorname{SL}_2(\mathbb{Z})$. The previous argument then gives an infinite sequence of bounded geodesics $\gamma_m \subset \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ with lifts $\widetilde{\gamma}_m \subset \mathbb{H}$ stabilized by conjugates of powers of U in $\operatorname{SL}_2(\mathbb{Q})$. It follows that the endpoints ξ, ξ' of $\widetilde{\gamma}$ in $\mathbb{R}$ are in fact a pair of Galois conjugate points in K.

Since the geodesic defined by $|z| = 1$ cuts $\mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$ into simply-connected pieces, the lifts $\widetilde{\gamma}_m$ can be chosen so they cross it; that is, we can assume $|\xi_m| > 1$ and $|\xi'_m| < 1$. The group $\operatorname{SL}_2(\mathbb{Z})$ is normalized by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, so we can also assume $\xi_m > 1$. With this normalization, ξ_m is a 'reduced' quadratic number, and hence its continued fraction expansion $[a_0, a_1, a_2, \ldots]$ is periodic (e.g. by [Ser, Thm. 5.23]); and the partial quotients a_i are uniformly bounded since $\overline{\bigcup \gamma_m}$ is compact. ∎

Ideals: Proof of Theorem 1.4. Let $\|a+b\epsilon\|^2 = (a^2+b^2)$ be the Euclidean norm on $K \cong \mathbb{Q}^2$ with respect to the basis $\langle 1, \epsilon\rangle$. Then it is easy to check that for all $x \in K$ we have

$$|\operatorname{N}_{\mathbb{Q}}^{K}(x)| \asymp \inf\{\|\epsilon^i x\|^2 \ : \ i \in \mathbb{Z}\}. \tag{2.10}$$

Let U be given by (2.2) and let $L_m \in \mathrm{M}_2(\mathbb{Z})$ be the matrices furnished by Theorem 2.2. Then we can regard

$$I_m = L_m(\mathbb{Z}) \subset \mathbb{Z}^2 \cong \mathbb{Z} \oplus \mathbb{Z}\epsilon$$

as fractional ideals in K. The smallest power $k(m)$ of ϵ stabilizing I_m tends to infinity with m, and hence the sequence $[I_m] \in I(K)$ ranges through infinitely many different ideal classes.

By (2.7), the norm squared $\|v\|^2$ of the shortest nonzero vector $v \in U^i L_m(\mathbb{Z}^2)$ is comparable to $|\det(L_m)|$. By (2.10) this implies $N^*(I_m) \asymp |\det(L_m)|$. But it is easy to see that $\det(I_m) \asymp |\det(L_m)|$, and hence

$$\delta(I_m) = \frac{N^*(I_m)}{\det(I_m)} \asymp 1$$

for all $m > 0$. In particular, the packing constants of the ideal classes I_m are uniformly bounded away from zero. ∎

Remark: Poincaré's periodic portrait. The iterates of a picture of Poincaré under the ergodic toral automorphism $U = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ appear in the popular article [CFPS]; the portrait becomes highly distorted, but then returns nearly to its original form after 240 iterates. This near–return illustrates, not Poincaré recurrence, but rather the case $m = 120$ of the identity $U^{2m} = \pm I \bmod f_m$ (which follows from (2.6), using the fact that $f_{m+1}^2 = \pm 1 \bmod f_m$). See [DF] and [Ghys] for more details.

3 Loop generators

Next we develop a more flexible mechanism for producing lattices with bounded orbits.

Definition. A matrix $\widetilde{L} \in \mathrm{M}_2(K)$ is a *loop generator* for ϵ if

$$L_m = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^m \widetilde{L}) \in \mathrm{M}_2(\mathbb{Q})$$

is invertible for all $m > 0$, and the collection of all lattices of the form

$$[U^i L_m] \in \mathrm{PGL}_2(\mathbb{R})/\mathrm{PGL}_2(\mathbb{Z}),$$

$i \in \mathbb{Z}$, $m > 0$ has compact closure. In this section we show:

Theorem 3.1 *Let $\widetilde{L} = X + \sqrt{D}Y$ where $X, Y \in \mathrm{M}_2(\mathbb{Q})$ have determinant zero. Suppose $\det(\widetilde{L}) \neq 0$ and $\det(X + SY) \neq 0$. Then $\widetilde{L}$ is a loop generator.*

(Recall from §2 that the matrix $S = 2U - tI$ represents multiplication by $\sqrt{D}$ on $\mathbb{Z}[\epsilon]$.)

Example. The matrix $\widetilde{L} = \left(\begin{smallmatrix} 1/\sqrt{D} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is a loop generator; the corresponding sequence of lattices is defined for $m > 0$ by

$$L_m = \begin{pmatrix} f_m & 0 \\ 0 & f_{m+1} - nf_{m-1} \end{pmatrix}. \tag{3.1}$$

Hecke correspondences. Given an integer $\ell > 0$, the multivalued *Hecke correspondence*

$$T_\ell : \mathrm{PGL}_2(\mathbb{R})/\,\mathrm{PGL}_2(\mathbb{Z}) \to \mathrm{PGL}_2(\mathbb{R})/\,\mathrm{PGL}_2(\mathbb{Z})$$

sends a lattice to its sublattices of index ℓ. In terms of matrices, we have

$$T_\ell([L]) = \{[LA] \ : \ A \in \mathrm{M}_2(\mathbb{Z}), \det(A) = \ell\}.$$

Since $\mathbb{Z}^2$ has only finitely many subgroups of index ℓ, T_ℓ sends compact sets to compact sets.

A key property of the Hecke correspondence is that it commutes with the left action of $\mathrm{GL}_2(\mathbb{R})$; in particular, we have

$$T_\ell([UL]) = U(T_\ell([L]))$$

for all $L \in \mathrm{GL}_2(\mathbb{R})$. It is also easy to see that $[L] \in T_\ell^2([L])$.

Proposition 3.2 *If $\widetilde{L} \in \mathrm{M}_2(K)$ is a loop generator, then so is $\widetilde{L}A$ for any $A \in \mathrm{GL}_2(\mathbb{Q})$.*

Proof. Since $[L] = [\lambda L]$ for any $\lambda \in \mathbb{R}^*$, we can assume A has integer entries. Let $\ell = \det(A)$. By assumption, the lattices $[U^i L_m]$ range in a compact subset $Z \subset \mathrm{PGL}_2(\mathbb{R})/\,\mathrm{GL}_2(\mathbb{Z})$. Thus the lattices $[U^i L_m A] \in T_\ell([U^i L_m])$ lie in the compact set $T_\ell(Z)$. Since $L_m A = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^m \widetilde{L}A)$, this shows $\widetilde{L}A$ is a loop generator. ∎

Proof of Theorem 3.1. Since the set of loop generators is invariant under the right action of $\mathrm{GL}_2(\mathbb{Q})$, we are free to replace (X, Y) with (Xg, Yg) where $g = (X + SY)^{-1}$; thus we can assume $X + SY = I$. A calculation (using 2.3) then shows

$$\widetilde{L} = \widetilde{U} + \widetilde{U}'A, \tag{3.2}$$

where $A = X - SY$. This implies, by the determinant identity (2.1), that

$$\det(A) = -\operatorname{tr}(XY^\dagger S^\dagger) = -\det(X + SY) = -1,$$

and hence $A \in \mathrm{GL}_2(\mathbb{Q})$. Letting

$$L_m = \operatorname{tr}_{\mathbb{Q}}^K(\epsilon^m \widetilde{L}) = \operatorname{tr}_{\mathbb{Q}}^K(\epsilon^m)X + \operatorname{tr}_{\mathbb{Q}}^K(\epsilon^m\sqrt{D})Y,$$

we find

$$\det(L_m) = \operatorname{tr}_{\mathbb{Q}}^K(\epsilon^m)\operatorname{tr}_{\mathbb{Q}}^K(\epsilon^m\sqrt{D})\operatorname{tr}(X^\dagger Y) = Df_{2m}\operatorname{tr}(X^\dagger Y), \tag{3.3}$$

using (2.4) and the fact that $\operatorname{tr}_{\mathbb{Q}}^K(x)\operatorname{tr}_{\mathbb{Q}}^K(x\sqrt{D}) = \operatorname{tr}_{\mathbb{Q}}^K(x^2\sqrt{D})$. By assumption, $\det(\widetilde{L}) = \sqrt{D}\operatorname{tr}(X^\dagger Y) \neq 0$, so L_m is invertible for all $m > 0$.

By (3.2) for $m > 0$ we can also write

$$L_m = U^m + n^m U^{-m} A$$

where $n = N(\epsilon)$, and hence obtain the bound

$$\|U^i L_m U^{-i}\| = O(\epsilon^m)$$

for $0 \le i \le m$, just as in the proof of Theorem 2.2. Similarly, if we *define*

$$L_{-m} = L_m A^{-1} = U^m A^{-1} + n^m U^{-m},$$

then we have

$$\|U^{-i} L_{-m} U^i\| = O(\epsilon^m)$$

as well. Since $|\det(L_m)| \asymp \epsilon^{2m}$ by (3.3), we find there is a compact set $Z \subset \mathrm{PGL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z})$ containing

$$[U^i L_m] \quad \text{and} \quad [U^{-i} L_{-m}]$$

for all $m > 0$ and $0 \le i \le m$.

Unfortunately, the period of $[L_m]$ under U might be greater than $2m$; and we need not have $[L_{-m}] = [L_m]$. However, since $L_{-m} = L_m A^{-1}$ and A is a fixed rational matrix, there is an $\ell > 0$ such that $[L_{-m}] \in T_\ell([L_m])$ for all m. Similarly, increasing ℓ if necessary, the fact that $\det(L_m)$ is a fixed rational multiple of f_{2m} implies there are integral matrices $[M_m] \in T_\ell(L_m)$ with $\det(M_m) = f_{2m}$ on the nose.

We claim the orbit of $[M_m]$ under $\langle U \rangle$ is contained in $T_\ell(Z) \cup T_\ell^2(Z)$. Indeed, for $0 \le i \le m$ we have

$$[U^i M_m] \in T_\ell([U^i L_m]) \subset T_\ell(Z),$$

and

$$[U^{-i}M_m] \in T_\ell([U^{-i}L_m]) \subset T_\ell(U^{-i}T_\ell([L_{-m}])) = T_\ell^2(U^{-i}L_{-m}) \subset T_\ell^2(Z),$$

and these lattices comprise the full orbit of $[M_m]$ since $[U^{2m}M_m] = [M_m]$ (Proposition 2.1). It follows that the orbit of $[L_m] \in T_\ell([M_m])$ under $\langle U \rangle$ is contained in the compact set $T_\ell^2(Z) \cup T_\ell^3(Z)$, which is independent of m. ■

Special case. We remark that if $A \in \mathrm{GL}_2(\mathbb{Z})$ and its eigenvalues are -1 and $+1$, then

$$\widetilde{L} = \widetilde{U} + \widetilde{U}'A = \frac{1}{2}(A+I) + \frac{\sqrt{D}}{2D}S(I-A),$$

clearly has the form $X + \sqrt{D}Y$ with $\det(X) = \det(Y) = 0$ and $X + SY = I$. If $\mathrm{tr}(A^\dagger U) \neq 0$ then $\det(\widetilde{L}) \neq 0$, and thus $\widetilde{L}$ is a loop generator by Theorem 3.1. The corresponding sequence of lattices are given by

$$L_m = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^m \widetilde{L}) = U^m + n^m U^{-m} A$$

where $n = N(\epsilon)$. Thus the construction of lattices with bounded orbits given in Theorem 2.2 is a special case of the loop-generator construction. In this case $V_m = L_m^{-1} U^{2m} L_m$ can also be given by the trace expression

$$V_m = \mathrm{tr}_{\mathbb{Q}}^K(\epsilon^{2m}\widetilde{L}^{-1}\widetilde{U}\widetilde{L}) + n^m(A+S).$$

4 Patterns of continued fractions

In this section we give a second, short proof of Theorem 1.1. It is based on the following Proposition, which is readily verified by induction.

Proposition 4.1 *For any $s > 0$, the periodic continued fractions*

$$x_m = [\overline{(1,s)^m, 1, s+1, s-1, (1,s)^m, 1, s+1, s+3}] \tag{4.1}$$

lie in $\mathbb{Q}(\sqrt{s^2+4s})$ for all $m \geq 0$.

(Here $(1,s)^m$ indicates that the pattern $1, s$ is repeated m times.) Similar patterns appear in [Wil] and [Wd].

Direct proof of Theorem 1.1. Let K be a real quadratic field. By Dirichlet's theorem, there exists a unit $\epsilon \in K$ with norm 1 and trace $t > 3$ (namely a suitable power of a fundamental unit). Then $K = \mathbb{Q}(\sqrt{t^2-4}) = \mathbb{Q}(\sqrt{s^2+4s})$ where $s = t-2 > 1$, and the sequence x_m above provides infinitely many periodic continued fractions in K with $1 \leq a_i \leq s+3$. ■

This pattern of continued fractions can be connected to the loop generator $\widetilde{L} = \begin{pmatrix} 1/\sqrt{D} & 0 \\ 0 & 1 \end{pmatrix}$, as follows.

Proposition 4.2 *For any quadratic unit $\epsilon > 1$, the numbers defined by*

$$y_m = \left(\frac{f_{m+1} - n f_{m-1}}{f_m} \right) \epsilon$$

for $m > 0$ have uniformly bounded continued fraction expansions.

(Here f_m is defined by (2.4) and $n = \mathrm{N}_{\mathbb{Q}}^{K}(\epsilon)$.)

Proof. Let L_m, given by (3.1), be the sequence of diagonal matrices determined by the loop generator $\widetilde{L}$. Then in terms of the usual action of $\mathrm{PGL}_2(\mathbb{R})$ on $\mathbb{P}^1(\mathbb{R})$ by $A(z) = (az+b)/(cz+d)$, we have $y_m = L_m^{-1}(\epsilon)$. Since $-(\epsilon, \epsilon')$ are the fixed points of $U(z) = -n/(z+t)$, the geodesics $\widetilde{\gamma}_m$ joining y_m to y'_m lie over a compact subset of $\gamma_m \subset \mathbb{H}/\mathrm{SL}_2(\mathbb{Z})$. Since $\lim y_m \neq \lim y'_m$, this compactness implies a uniform bound on the continued fraction expansion of y_m. ■

Cf. [Wd], which treats the case $\mathbb{Q}(\sqrt{5})$. Evaluating the continued fraction expansion of y_m quickly suggests (4.1); for example, when $\epsilon = (3+\sqrt{5})/2$ and $m = 10$ we have

$$y_m = \frac{15127(3+\sqrt{5})}{13530} = [5, \overline{1, 5, 1, 5, 1, 5, 1, 6, 8, 1, 5, 1, 5, 1, 5, 1, 6, 4}].$$

Many other patterns can be produced by varying the choice of the loop generator $\widetilde{L}$.

5 More general quadratic extensions

In this section we show the construction of §2 can be applied to $U \in \mathrm{SL}_2(\mathcal{O}_d)$ and, more generally, to $U \in \mathrm{GL}_N(\mathbb{Z})$ whenever U is conjugate to U^{-1} in $\mathrm{GL}_N(\mathbb{Q})$.

$\mathrm{SL}_2(\mathcal{O}_d)$: Proof of Theorem 1.3. Choosing a particular complex embedding of $k = \mathbb{Q}(\sqrt{-d}) \subset \mathbb{C}$, we can regard $\mathrm{SL}_2(\mathcal{O}_d)$ as a discrete subgroup of $\mathrm{SL}_2(\mathbb{C})$. Let $U \in \mathrm{SL}_2(\mathcal{O}_d)$ be a hyperbolic element corresponding to a fundamental geodesic γ, with eigenvalues $\epsilon^{\pm 1}$. We may assume $|\epsilon| > 1$. Then $K = k(\epsilon)$ is a quadratic extension of k, and up to conjugation in $\mathrm{GL}_2(k)$ we can assume U is given by (2.2), where $t = \mathrm{tr}_k^K(\epsilon)$ and

$n = \mathrm{N}_k^K(\epsilon) = \det(U) = 1$. (By a Hecke correspondence argument similar to the proof of Proposition 3.2, conjugating U by an element $\mathrm{GL}_2(k)$ does not affect the conclusions of of the theorem.)

Given $m > 0$, let $L_m = U^m + U^{-m}A$ with $A \in \mathrm{GL}_2(\mathcal{O}_d)$ given by (2.9), and let $f_m = \mathrm{tr}_k^K(\epsilon^m\sqrt{D})$. Then we have $|f_m| \asymp |\epsilon|^m$, $|\det(L_m)| \asymp |\epsilon|^{2m}$ and $\|U^{-m}\|, \|U^m\| = O(|\epsilon|^m)$ so the bounds (2.7) still hold; and $[U^{2m}L_m] = [L_m]$ by the same proof as before. Thus $[U^iL_m]$, $i \in \mathbb{Z}$ ranges in a compact subset of $\mathrm{PGL}_2(\mathbb{C})/\mathrm{SL}_2(\mathcal{O}_d)$. The periods of these orbits go to infinity by an immediate generalization of Theorem 2.4, and hence elements $L_m^{-1}U^{2m}L_m \in \mathrm{SL}_2(\mathcal{O}_d)$ correspond to an bounded, infinite sequence of geodesics $\gamma_m \subset \mathbb{H}^3/\mathrm{SL}_2(\mathcal{O}_d)$ whose lengths are multiples of $L(\gamma)$. ■

$\mathrm{GL}_N(\mathbb{Z})$: Proof of Theorem 1.8. This case has an additional twist, since for $N > 2$ the eigenvalues of U outside the unit circle may have different absolute values.

Let $U \in \mathrm{GL}_N(\mathbb{Z})$ be an element of infinite order with irreducible characteristic polynomial, such that U is conjugate to U^{-1} in $\mathrm{GL}_N(\mathbb{Q})$. Then the algebra $K \cong \mathbb{Q}(U) \subset \mathrm{M}_N(\mathbb{Q})$ is a field. Let $k = \mathbb{Q}(U + U^{-1}) \subset K$ and let $d = \deg(k/\mathbb{Q})$. Since $U \neq U^{-1}$, K/k is a quadratic field extension and hence $N = 2d$.

The ring of integers $\mathcal{O}_k \subset k$ embeds as a lattice in $\mathbb{R}^r \times \mathbb{C}^s$, where $r + 2s = d$ and r and s denote the number of real and complex places of k. Similarly we obtain a discrete subgroup

$$\Gamma = \mathrm{GL}_2(\mathcal{O}_k) \subset G = \mathrm{GL}_2(\mathbb{R})^r \times \mathrm{GL}_2(\mathbb{C})^s.$$

The projection of Γ to $PG = G/\mathbb{R}^*$ is a lattice.

Choosing an integral basis for $\mathcal{O}_k$, we obtain an embedding $\mathrm{GL}_2(\mathcal{O}_k) \to \mathrm{GL}_{2d}(\mathbb{Z})$ whose image contains U. Thus we can regard U as an element of $\mathrm{GL}_2(\mathcal{O}_k)$, with eigenvalues $\epsilon^{\pm 1} \in K$. Let $t = \mathrm{tr}_k^K(\epsilon)$ and note that $n = \mathrm{N}_k^K(\epsilon) = 1$. After conjugation by an element of $\mathrm{GL}_2(k)$ (which does not affect the conclusions of the theorem), we can assume that $U = \left(\begin{smallmatrix} 0 & -n \\ 1 & t \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathcal{O}_k)$.

We will show that $L_m = U^m + U^{-m}A$, $m > 0$, defines a sequence $[L_m] \in PG/\Gamma$ providing infinitely many $\langle U \rangle$-orbits ranging in a fixed compact set Z.

Let $|x|_v$ denote the absolute value on k associated to the place v (using $|z|^2$ at the complex places), and let $d_v = 1$ or 2 according to whether v is real or complex. Then $\sum d_v = d$, and

$$\|x\| = \max |x|_v^{1/d_v}$$

defines a norm on k whose completion is $\mathbb{R}^r \times \mathbb{C}^s$. Similarly we obtain a norm on k^2 and an operator norm on $\mathrm{M}_2(k)$. Given $L \in \mathrm{M}_2(k)$, we let $\mathrm{Det}(L) = \mathrm{N}_{\mathbb{Q}}^k(\det L)$. Clearly for any $C > 0$, the set of lattices

$$Z(C) = \{[L] \,:\, \|L\|^{2d} \leq C|\operatorname{Det} L|\} \subset PG/\Gamma$$

is compact.

Extend each valuation v to K in such a way that $|\epsilon|_v \geq 1$; then the definition of $\|x\|$ also extends to K.

Let $M(\epsilon) = \prod_{|\epsilon_i|\geq 1} |\epsilon_i|$ denote the *Mahler measure* of ϵ — the product of its conjugates outside the unit circle. Let $f_m = \mathrm{tr}_k^K(\epsilon^m/\sqrt{D})$ as before. We then have

$$|\,\mathrm{N}_{\mathbb{Q}}^k(f_m)| \asymp \prod |\epsilon^m|_v = M(\epsilon)^m.$$

As before, we have $\det(L_m) = f_m^2$, and thus $|\operatorname{Det} L_m| \asymp M(\epsilon)^{2m}$. We also have $\|L_m\| = O(\|U^m\|)$. Since $\|\epsilon\|$ gives spectral radius of U (the size of the largest eigenvalue of U acting on $\mathbb{R}^r \times \mathbb{C}^s$), we have

$$\|U^m\|^d \asymp \|\epsilon\|^{md}.$$

But in general we only have the inequality

$$\|\epsilon\|^d = (\max |\epsilon|_v^{1/d_v})^d \geq \prod |\epsilon|_v = M(\epsilon).$$

In other words, $\|L_m\|^{2d}$ may be much larger than $|\operatorname{Det}(L_m)|$ because some eigenvalues of U are much larger than others.

To remedy this, we correct $[L_m]$ by units in $\mathcal{O}_k$. By Dirichlet's theorem [BoS, §2.4.3], the quotient

$$\mathbb{R}_0^{r+s}/\,\mathcal{O}_k^* = \Big\{(x_v) \,:\, \sum x_v = 0\Big\}\Big/\Big\{\log|\eta|_v \,:\, \eta \in \mathcal{O}_k^*\Big\}$$

is compact. Thus we can find a unit $\eta \in \mathcal{O}_k^*$ such that

$$|\eta\epsilon^m|_v^{1/d_v} \asymp M(\epsilon)^{m/d}$$

for all v. Then

$$\|\eta\epsilon^m\|^d = O(M(\epsilon)^m).$$

By examining the eigenspaces of U, we find the same bound holds for $\|\eta U^{\pm m}\|$. Since η is a unit, ηI belongs to $\Gamma = \mathrm{GL}_2(\mathcal{O}_k)$, and thus we have

$$[L_m] = [\eta U^m + \eta U^{-m} A]$$

in PG/Γ; and since

$$\|\eta U^m + \eta U^{-m} A\|^{2d} = O(M(\epsilon)^{2m}) = O(|\operatorname{Det} L_m|),$$

$[L_m]$ now ranges in a compact subset of the form $Z(C) \subset PG/\Gamma$. A similar argument shows $[U^i L_m]$ and $[U^{-i} L_m]$ range in a compact set for all $m > 0$ and $0 \le i \le m$.

Noting that Proposition 2.1 and Theorem 2.4 generalize immediately to this setting, we conclude that the full $\langle U \rangle$-orbit of $[L_m]$ is contained in Z and that the length $k(m)$ of this orbit tends to infinity. Finally reduction of scalars provides a finite-to-one projection

$$\pi : PG/\Gamma \to \mathrm{PGL}_N(\mathbb{R})/\,\mathrm{GL}_N(\mathbb{Z}),$$

and the proof is completed by taking the images of $[L_m]$ under this projection. ■

6 Class numbers and heights on $\mathbb{P}^1$

Let $\operatorname{Pic} \mathcal{O}_D$ denote the group of invertible ideal classes for the quadratic order of discriminant D, and let $h(D) = |\operatorname{Pic} \mathcal{O}_D|$ denote the corresponding class number.

In this section we relate the packing densities of ideals to heights on $\mathbb{P}^1(\mathbb{Z}/f)$ and the computation of $h(f^2 D)$. This perspective suggests the following strengthening of Theorem 1.4. As usual, suppose $\epsilon > 1$ is a quadratic unit and $f_m^2 D$ is the discriminant of $\mathbb{Z}[\epsilon^m]$.

Conjecture 6.1 *Given $\alpha > 0$, there is a $\delta > 0$ such that*

$$|\{I \in \operatorname{Pic} \mathcal{O}_{f_m^2 D} \,:\, \delta(I) > \delta\}| \ge f_m^{1-\alpha} \tag{6.1}$$

for all m sufficiently large.

It also connects our results to Zaremba's conjecture, and provides an approach to Conjecture 1.5 for cubic and quartic fields.

The projective line. Given $f > 0$, we define the projective line over $\mathbb{Z}/f$ in terms of lattices in $\mathbb{Z}^2$ by

$$\mathbb{P}^1(\mathbb{Z}/f) = \{L \subset \mathbb{Z}^2 \,:\, \mathbb{Z}^2/L \cong \mathbb{Z}/f\}.$$

Given $a, b \in \mathbb{Z}$ with $\gcd(a, b, f) = 1$, we use $[a : b]$ as shorthand for the lattice

$$L_{[a:b]} = \mathbb{Z}(a, b) + f\mathbb{Z}^2 \subset \mathbb{Z}^2.$$

The number of points on $\mathbb{P}^1(\mathbb{Z}/f)$ is given by $f \prod_{p|f}(1 + 1/p)$.

Heights. We define the *height* of a point on $\mathbb{P}^1(\mathbb{Z}/f)$ by

$$H(L) = \inf\{\|x\|^2 \ : \ x \in L, x \neq 0\}. \tag{6.2}$$

Since $\mathrm{vol}(\mathbb{R}^2/L) = f$ we have $H(L)/f \leq 2/\sqrt{3}$ (the maximum comes from an hexagonal lattice), and $H(L)/f$ is small $\iff$ $[L]$ is near infinity in $\mathrm{PGL}_2(\mathbb{R})/\mathrm{PGL}_2(\mathbb{Z})$. It easy to see that the proportion of $L \in \mathbb{P}^1(\mathbb{Z}/f)$ with $H(L)/f > \delta > 0$ tends to 1 (uniformly in f) as $\delta \to 0$.

In the case where f is prime, the height also satisfies

$$H(L) = \inf\{|a|^2 + |b|^2 \ : \ L = L_{[a:b]}\};$$

thus it measures the minimal complexity of an arithmetic description of L. (A somewhat different height is considered in [NS].)

Ideals. Now let $\epsilon > 1$ be a quadratic unit, and identify $\mathbb{Z}[\epsilon]$ with $\mathbb{Z}^2$ using the basis $(1, \epsilon)$ as before. We will denote the order $\mathbb{Z}[f\epsilon] \subset \mathbb{Z}[\epsilon] \subset K = \mathbb{Q}(\epsilon)$ by $\mathcal{O}_{f^2D}$, since its discriminant is f^2D.

Given $f > 0$, every $x \in \mathcal{O}_D$ determines an ideal

$$I(x, f) = \mathbb{Z}x + f\,\mathcal{O}_D$$

for the order $\mathcal{O}_{f^2D}$. Clearly $I(x, f)$ only depends on the class $[x]$ of x in $(\mathcal{O}_D/f\,\mathcal{O}_D)$. Let

$$I(f) = \{I(x, f) \ : \ \mathcal{O}_D/I(x, f) \cong \mathbb{Z}/f\},$$

and let

$$I^*(f) = \{I(x, f) \ : \ [x] \in (\mathcal{O}_D/f\,\mathcal{O}_D)^*\}.$$

It can be shown that $I^*(f)$ consists of the ideals $I \in I(f)$ which are invertible as $\mathcal{O}_{f^2D}$-modules.

The basis $(1, \epsilon)$ for $\mathcal{O}_D$ determines a bijection

$$\pi : I(f) \to \mathbb{P}^1(\mathbb{Z}/f)$$

sending $I(a + b\epsilon, f)$ to $[a : b]$. The matrix U given by (2.2) acts naturally on $\mathbb{P}^1(\mathbb{Z}/f)$, and we have

$$\pi(\epsilon \cdot I(x, f)) = U \cdot \pi(I(x, f)).$$

Density and height. For $I \in I(f)$ with $L = \pi(I)$, we have $\det(I) = f\sqrt{D}$ and

$$N^*(I) = \inf\{|\,\mathrm{N}_{\mathbb{Q}}^K(x)| \,:\, x \in I, \mathrm{N}_{\mathbb{Q}}^K(x) \neq 0\} \asymp \inf\{H(U^i L) \,:\, i \in \mathbb{Z}\},$$

by the same reasoning as in the proof of Theorem 1.4. Thus the packing density of I satisfies

$$\delta(I) = N^*(I)/\det(I) \asymp \inf_{i\in\mathbb{Z}} H(U^i L)/f, \tag{6.3}$$

where the implicit constants depend only on U.

Class numbers. To put this discussion in context, we recall the calculation of $h(f^2 D)$ (cf. [Lang], [Sa]).

It is known that the natural map $\operatorname{Pic}\mathcal{O}_{f^2 D} \to \operatorname{Pic}\mathcal{O}_D$ is surjective, and that every ideal class in the kernel has a representative in $I^*(f)$. Moreover, $I, J \in I^*(f)$ represent the same ideal class iff $I = \eta J$ for some unit $\eta \in \mathcal{O}_D$. In other words, we have an exact sequence

$$0 \to (\mathcal{O}_D\,/f\,\mathcal{O}_D)^*/((\mathbb{Z}/f)^*\,\mathcal{O}_D^*) \to \operatorname{Pic}\mathcal{O}_{f^2 D} \to \operatorname{Pic}\mathcal{O}_D \to 0$$

whose second term is in bijection with the orbits of

$$\pi(I^*(f)) \subset \mathbb{P}^1(\mathbb{Z}/f)$$

under the action of $\langle U\rangle$. It follows that the class number of $\mathcal{O}_{f^2 D}$ is given by

$$h(f^2 D) = \frac{h(D)}{[\mathcal{O}_D^* : \mathcal{O}_{f^2 D}^*]}|I^*(f)| = \frac{h(D)R(D)}{R(f^2 D)}|I^*(f)|,$$

where $R(D)$ denotes the regulator of $\mathcal{O}_D$.

When D is a fundamental discriminant, one can compute $|I^*(f)|$ in terms of primes dividing f to obtain the formula:

$$h(f^2 D) = \frac{h(D)R(D)f}{R(f^2 D)} \prod_{p|f}\left(1 - \left(\frac{K}{p}\right)\frac{1}{p}\right);$$

see [Lang, Ch. 8.1, Thm 7.]. (Here $(K/p) = 1$ if p splits in K, 0 if it ramifies and -1 if it remains prime.)

For $f > 1$ the product on the right, and its reciprocal, are both $O(\log f)$. Thus the class number is controlled primarily by the regulator of $\mathcal{O}_{f^2 D}$: it satisfies

$$\frac{C_1 f}{R(f^2 D)\log f} \le h(f^2 D) \le \frac{C_2 f \log f}{R(f^2 D)},$$

where $C_1, C_2 > 0$ depend only on D. (A bound of this type holds whether D is fundamental or not.)

Fibonacci orders. As an example, note that the orders $\mathbb{Z}[\epsilon^m] = \mathcal{O}_{f_m^2 D}$ satisfy $R(f_m^2 D) = mR(D)$ and $f_m \asymp \epsilon^m$, and hence

$$h(f_m^2 D) \geq C_3 f_m/(\log f_m)^2. \tag{6.4}$$

In other words, the orders generated by powers of ϵ have large class numbers.[2]

Arithmetic independence. It is now straightforward to give a rationale for Conjecture 6.1.

Consider the uniform probability measure on $\mathbb{P}^1(\mathbb{Z}/f_m)$, assigning equal mass to each point. Fix a small $\delta > 0$; then the probability p that the height of a random $L \in \mathbb{P}^1(\mathbb{Z}/f_m)$ satisfies $H(L) > \delta f_m$ is close to one. Suppose that the events $H(L) > \delta f_m$, $H(UL) > \delta f_m$, $H(U^2 L) > \delta f_m$, etc. are essentially independent. Since $U|\mathbb{P}^1(\mathbb{Z}/f_m)$ has period m, the probability that all these events occur is roughly p^m. But m is comparable to $\log f_m$, so p^m is comparable to $f_m^{-\alpha}$ for some small $\alpha > 0$. Since $|\mathbb{P}^1(\mathbb{Z}/f_m)| \geq f_m$, the total number of $L \in \mathbb{P}^1(\mathbb{Z}/f_m)$ with $\inf H(U^i L)/f_m > \delta$ is at least $f_m^{1-\alpha}$, where $\alpha \to 0$ as $\delta \to 0$.

By (6.3), the same type of estimate holds for the number of ideals $I \in I(f_m)$ with $\delta(I) > \delta$. The probability that a random ideal lies in $I^*(f_m)$ is roughly $1/\log f_m$; assuming independence again, this introduces a negligible correction, and we now obtain ideal classes in $\operatorname{Pic} \mathcal{O}_{f_m^2 D}$. At most $m \asymp \log f_m$ ideals in $I^*(f_m)$ map to the same class, so we again obtain on the order of $f_m^{1-\alpha}$ distinct ideal classes with $\delta(I) > \delta$.

Counting geodesics. Let $L = \log \epsilon^2$ denote the length of the closed geodesic represented by $U \in \mathrm{SL}_2(\mathbb{Z})$. Then Conjecture 6.1 implies that for any $\alpha > 0$, there is a compact set $Z \subset \mathbb{H}/\mathrm{SL}_2(\mathbb{Z})$ that contains at least $\exp((1/2 - \alpha)mL)$ primitive geodesics of length mL for all $m \gg 0$. (For comparison, the total number of geodesics of length ℓ is $O_\eta(\exp((1/2+\eta)\ell))$ for all $\eta > 0$, and the number of length $\leq \ell$ is $\sim \exp(\ell)/\ell$; cf. [Sar, §2].)

Orders in $\mathbb{Q} \times \mathbb{Q}$. Similar phenomena can be studied for the algebra $K = \mathbb{Q} \times \mathbb{Q}$, whose orders are

$$\mathcal{O}_{f^2} = \{(a, b) \in \mathbb{Z}^2 \ : \ a \equiv b \bmod f\}.$$

[2] Orders with small class numbers can also be exhibited, e.g. $h(5^{2m+1}) = 1$ for all m; cf. [Lag, Lemma A-1]. This fact is compatible with (6.4) because for $m > 1$, 5^m is not a Fibonacci number.

With the trace and norm given by $a+b$ and ab, the packing density can be defined just as for a quadratic field, and one can also formulate:

Conjecture 6.2 *Given any $\alpha > 0$, there is a $\delta > 0$ such that*

$$|\{I \in \operatorname{Pic} \mathcal{O}_{f^2} \,:\, \delta(I) > \delta\}| \geq f^{1-\alpha} \tag{6.5}$$

for all f sufficiently large.

(Since $\mathcal{O}_1^*$ is finite, all orders should behave equally well.)

This conjecture implies:

Conjecture 6.3 (Zaremba) *There exists an $N > 0$ such that every $f > 0$ arises as the denominator of a rational number $a/f = [a_0, a_1, \ldots, a_n]$ with $1 \leq a_i \leq N$.*

Zaremba's conjecture is stated in [Zar]; it is plausible that it holds for $N = 5$, and even for $N = 2$ if finitely many f are excluded (see [Hen, §3, Conj. 3]). Explicit constructions show one can take $N = 3$ when f is a power of 2 or 3 [Nie].

To see Conjecture 6.2 implies Zaremba's conjecture, observe that $\operatorname{Pic}(\mathcal{O}_{f^2})$ is in bijection with $(\mathbb{Z}/f)^*$ via the map

$$a \mapsto I_a = \{(q, r) \in \mathbb{Z}^2 \,:\, r = aq \bmod f\} \subset \mathbb{Z} \times \mathbb{Z}.$$

Since $\det(I_a) = f$, the condition $\delta(I_a) > \delta$ is equivalent to

$$N^*(I_a) = \inf\{|q| \cdot |aq - pf| \,:\, q \neq 0, aq - pf \neq 0\} > \delta f,$$

which means exactly that

$$\left| \frac{a}{f} - \frac{p}{q} \right| > \frac{\delta}{q^2}$$

whenever $p/q \neq a/f$. This Diophantine condition implies that the continued fraction of a/f satisfies $a_i = O(1/\delta)$, and hence the ideals furnished by Conjecture 6.2 (say with $\alpha = 1/2$) determine the numerators required for Zaremba's conjecture.

Question. In Theorem 1.1, can one take $M_d = 2$ for all d? That is, does every real quadratic field contain infinitely many periodic continued fractions with $1 \leq a_i \leq 2$?

Cubic fields. The same approach can be applied to fields of higher degree. For concreteness, suppose K is a cubic field generated by a unit $\epsilon > 1$ whose

conjugates are complex. The discriminant of the ring $\mathbb{Z}[\epsilon^m]$ can be expressed in the form

$$Df_m^2 = \det \operatorname{tr}_{\mathbb{Q}}^K \begin{pmatrix} 1 & \epsilon^m & \epsilon^{2m} \\ \epsilon^m & \epsilon^{2m} & \epsilon^{3m} \\ \epsilon^{2m} & \epsilon^{3m} & \epsilon^{4m} \end{pmatrix},$$

with $f_1 = 1$.

As before, the matrix $U \in \mathrm{GL}_3(\mathbb{Z})$ for multiplication by ϵ acts on the projective space $\mathbb{P}^2(\mathbb{Z}/f_m)$. In the cubic case, however, $U^m|\mathbb{P}^2(\mathbb{Z}/f_m)$ need not be the identity. As a substitute, we know that the resultant of the minimal polynomial $p_m(x)$ for ϵ^m is divisible by f_m. For simplicity, suppose f_m is prime; then we have a factorization $p_m(x) = (x-a)^2(x-b) \bmod f_m$, and $\mathrm{Ker}(U^m - aI)$ determines a U-invariant *line* $P_m \subset \mathbb{P}^2(\mathbb{Z}/f_m)$ such that $U^m|P_m$ is the identity. Since the orbits of $U|P_m$ are small, there is a reasonable chance that many of them have large height; if so, they furnish ideals whose densities are bounded away from zero.

Example. Let $\epsilon > 1$ be the Pisot number satisfying $\epsilon^3 = \epsilon + 1$. Then $D = -23$. For $m = 10$ we have $p_m(x) = (4+x)^2(13+x) \bmod f_m = 19$; for $m = 41$ we have $p_m(x) = (4679681 + x)^2(5436593 + x) \bmod f_m = 7448797$. The vectors v_m given by

$$v_{10} = [5 : 9 : 1] \quad \text{and} \quad v_{41} = [5514143 : 5170633 : 7378397]$$

have period m and satisfy $\min H(U^i v_m)/f_m^2 \approx 0.267$ and 0.249 respectively, versus a maximum possible value of $\sqrt{2} \approx 1.4142$. (Here the associated lattices $L_m = \mathbb{Z}v_m + f_m\mathbb{Z}^3$ have determinant f_m^2, and we take $\|x\|^3$ in the definition (6.2) of the height.) Experimentally, it appears that such U-orbits of large height can be found for arbitrarily large m.

References

[BoS] Z. I. Borevich and I. R. Shafarevich. *Number Theory.* Academic Press, 1966.

[CaS] J. W. S. Cassels and H. P. F. Swinnerton-Dyer. On the product of three homogeneous linear forms and the indefinite ternary quadratic forms. *Philos. Trans. Roy. Soc. London. Ser. A.* **248**(1955), 73–96.

[CFPS] J. P. Crutchfield, J. D. Farmer, N. H. Packard, and R. S. Shaw. Chaos. *Scientific American* **255**(1986), 46–57.

[Du] W. Duke. Hyperbolic distribution problems and half-integral weight Maass forms. *Invent. Math.* **92**(1988), 73–90.

[DF] F. J. Dyson and H. Falk. Period of a discrete cat mapping. *Amer. Math. Monthly* **99**(1992), 603–614.

[ELMV] M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh. The distribution of periodic torus orbits on homogeneous spaces. *Preprint, 2006.*

[Ghys] E. Ghys. Variations autour du théorème de récurrence de Poincaré. *J. de maths des élèves (l'ENS de Lyon)* **1**(1994), 3–12.

[GL] P. M. Gruber and C. G. Lekkerkerker. *Geometry of Numbers.* Elsevier, 1987.

[HW] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, 1979.

[Hen] D. Hensley. A polynomial time algorithm for the Hausdorff dimension of continued fraction Cantor sets. *J. Number Theory* **58**(1996), 9–45.

[Ja] V. Jarnik. Zur metrischen Theorie der diophantischen Approximationen. *Prace Mat.-Fiz.* **36**(1928), 91–106.

[KU] S. Katok and I. Ugarcovici. Symbolic dynamics for the modular surface and beyond. *Bull. Amer. Math. Soc.* **44**(2007), 87–132.

[Lag] J. C. Lagarias. On the computational complexity of determining the solvability or unsolvability of the equation $X^2 - DY^2 = -1$. *Trans. Amer. Math. Soc.* **260**(Aug., 1980), 485–508.

[Lang] S. Lang. *Elliptic Functions.* Springer-Verlag, 1987.

[Lin] Yu. V. Linnik. *Ergodic Properties of Algebraic Fields.* Springer-Verlag, 1968.

[Mg] G. A. Margulis. Problems and conjectures in rigidity theory. In *Mathematics: Frontiers and Perspectives*, pages 161–174. Amer. Math. Soc., 2000.

[NS] M. B. Nathanson and B. D. Sullivan. Heights in finite projective space, and a problem on directed graphs. *Integers* **8**(2008), A13, 9 pp.

[Nie] H. Niederreiter. Dyadic fractions with small partial quotients. *Monatsh. Math.* **101**(1986), 309–315.

[Po] M. Pollicott. Distribution of closed geodesics on the modular surface and quadratic irrationals. *Bull. Soc. Math. de France* **114**(1986), 431–446.

[Sa] J. W. Sands. Generalization of a theorem of Siegel. *Acta Arith.* **58**(1991), 47–57.

[Sar] P. Sarnak. Class numbers of indefinite binary quadratic forms. *J. Number Theory* **15**(1982), 229–247.

[Sch] W. M. Schmidt. Badly approximable systems of linear forms. *J. Number Theory* **1**(1969), 139–154.

[Ser] C. Series. Geometric methods of symbolic coding. In *Ergodic Theory, Symbolic Dynamics, and Hyperbolic Spaces*, pages 125–152. Oxford University Press, 1991.

[Sha] J. Shallit. Real numbers with bounded partial quotients: a survey. *Enseign. Math.* **38**(1992), 151–187.

[Sm] S. Smale. Diffeomorphisms with many periodic points. In *Differential and Combinatorial Topology*, pages 63–80. Princeton University Press, 1965.

[Wil] S. M. J. Wilson. Limit points in the Lagrange spectrum of a quadratic field. *Bull. Soc. Math. France* **108**(1980), 137–141.

[Wd] A. C. Woods. The Markoff spectrum of an algebraic number field. *J. Austral. Math. Soc. Ser. A* **25**(1978), 486–488.

[Zar] S. K. Zaremba. La méthode des "bons treillis" pour le calcul des intégrales multiples. In *Applications of Number Theory to Numerical Analysis (Proc. Sympos., Univ. Montreal, Montreal, Que., 1971)*, pages pp. 39–119. Academic Press, 1972.

Mathematics Department
Harvard University
1 Oxford St
Cambridge, MA 02138-2901