AN APPLICATION OF METRIC DIOPHANTINE APPROXIMATION IN HYPERBOLIC SPACE TO QUADRATIC FORMS

S. L. VELANI

Abstract

For any real $\tau$, a lim sup set $W_G,y(\tau)$ of $\tau$-(well)-approximable points is defined for discrete groups $G$ acting on the Poincaré model of hyperbolic space. Here $y$ is a 'distinguished point' on the sphere at infinity whose orbit under $G$ corresponds to the rationals (which can be regarded as the orbit of the point at infinity under the modular group) in the classical theory of diophantine approximation.

In this paper the Hausdorff dimension of the set $W_G,y(\tau)$ is determined for geometrically finite groups of the first kind. Consequently, by considering the hyperboloid model of hyperbolic space, this result is shown to have a natural but non trivial interpretation in terms of quadratic forms.

1. Introduction

The Euclidean norm of a vector $x = (x_1, \ldots, x_k, x_{k+1})$ in $\mathbb{R}^{k+1}$ will be denoted by $||x||$, i.e. $||x|| = (x_1^2 + \cdots + x_k^2 + x_{k+1}^2)^{1/2}$. The $k+1$-dimensional unit ball

$$B^{k+1} = \{x \in \mathbb{R}^{k+1} : ||x|| < 1\}$$

is a model of $k+1$-dimensional hyperbolic space and supports a metric $\rho$ derived from the differential

$$d\rho = \frac{||dx||}{1 - ||x||^2}.$$
The unit ball model is usually referred to as the Poincaré model. Geodesics for the metric $\rho$ are arcs of circles orthogonal to the unit sphere $S^k$ and straight lines through the origin.

Denote by $M(B^{k+1})$ the group of orientation preserving Moebius transformations preserving the unit ball $B^{k+1}$. The group $M(B^{k+1})$ preserves the above metric $\rho$. Let $G$ be an arbitrary discrete subgroup of $M(B^{k+1})$ and denote by $L(G)$ the limit set of $G$, the set of cluster points in $S^k$ of any orbit of $G$ in $B^{k+1}$. The limit set is either empty or contains one, two or an infinite number of points. The group $G$ is said to be elementary if the number of points in $L(G)$ is at most two, and non-elementary otherwise. Also the group $G$ is said to be of the first kind if $L(G) = S^k$ and of the second kind otherwise. So all elementary groups are of the second kind. In this paper, $G$ will be restricted to non-elementary geometrically finite discrete subgroups of $M(B^{k+1})$. By definition, there exists some convex fundamental polyhedron with finitely many faces which implies that $G$ is finitely presented (see [7] and [13] for further details). The reader is referred to [1], [2] and [7] for a general introduction to the theory of discrete groups and hyperbolic geometry.

For each $z, w$ in $B^{k+1}$ let

$$L(z, w) = \frac{1}{2} + \frac{\|z - w\|^2}{(1 - \|z\|^2)(1 - \|w\|^2)}.$$ 

Clearly $L(z, w) \geq 1/2$, and $L(z, w)$ is related to the hyperbolic distance $\rho(z, w)$ between $z$ and $w$ by $\cosh 2\rho(z, w) = 2L(z, w)$. For $s \in \mathbb{R}$ define the exponent of convergence of $G$ as

$$\delta(G) = \inf \left\{ s > 0 : \sum_{g \in G} L(z, g(w))^{-s} < \infty \right\}.$$ 

The series in the above definition diverges if $s < \delta(G)$ and converges if $s > \delta(G)$. It is a classical fact that $\delta(G) \in [0, k]$ and that if $G$ is of the first kind $\delta(G) = k$, while if $G$ is of the second kind $\delta(G) \in [0, k)$. The reader is referred to [7] and [9] for further details.

The following definition of $\tau$-approximable is the natural extension of Definition 1 given in [14].

**Definition 1.** For each real number $\tau$, a point $x$ in the set $L(G)$ is said to be $\tau$-approximable with respect to a non-empty finite subset $A$ of $L(G)$ if for some $y$ in $A$ the inequality

$$\|x - g(y)\| \leq L(0, g(0))^{-\tau}$$ 

is satisfied for all $g \in G$. The set $L(G)$ is said to be $\tau$-approximable if it contains points that are $\tau$-approximable.

The series $\sum_{g \in G} L(z, g(w))^{-s}$ converges if $s < \delta(G)$, and diverges if $s > \delta(G)$.
is satisfied for infinitely many $g$ in $G$. Let,

$$W_{G;Y}(\tau) = \{x \in L(G) : \|x - g(y)\| \leq L(0, g(0))^{-\tau}$$

for infinitely many $g$ in $G$}

denote the set of $\tau$-approximable points with respect to $y$ in $L(G)$, and denote by $W_{G;A}(\tau)$ the set of points $x$ in $L(G)$ which are $\tau$-approximable with respect to $A$. Thus

$$W_{G;A}(\tau) = \bigcup_{y \in A} W_{G;Y}(\tau).$$

With reference to the above definition let $y$ be an arbitrary parabolic cusp of $G$ is there any, and a hyperbolic fixed point otherwise. Let $\dim F$ denote the Hausdorff dimension of a subset $F$ of $\mathbb{R}^{k+1}$. For the definition of Hausdorff dimension and for further details the reader is referred to [5]. The following results are a generalization of the metric results given in [14].

**Theorem 1.** For $\tau > 1$,

$$\dim W_{G;Y}(\tau) \leq \frac{\delta(G)}{\tau}.$$  

**Theorem 2.** If the group $G$ is of the first kind, then for $\tau > 1$

$$\dim W_{G;Y}(\tau) = \frac{k}{\tau}.$$  

Note that for groups of the first kind $\delta(G) = k$; so Theorem 2 just implies that for these groups equality holds in Theorem 1. Since $\dim(\bigcup_{y \in A} W_{G;Y}(\tau)) = \max_{y \in A}\{\dim W_{G;Y}(\tau)\}$ for any finite set $A$, the following result is a direct consequence of Theorem 2.

**Corollary 1.** Let $A$ be a non empty finite set of parabolic cusps if there are any, and a finite set of hyperbolic fixed points otherwise. For $\tau > 1$,

$$\dim W_{G;A}(\tau) = \frac{k}{\tau},$$

where $W_{G;A}(\tau)$ is the set of points $x$ in $S^k$ which are $\tau$-approximable with respect to $A$.

**Remark.** For groups with parabolic elements, Theorem 2 also appears in the work of M. V. Melián and D. Pestana [6]. They also give
the equivalent geometric result in terms of the 'rate' of excursions by geodesics into a cuspidal end of the associated hyperbolic manifold.

In Section 2, we provide the reader with the results required to prove the above theorems. In view of these results, the proofs of the theorems follow on using essentially the same arguments as in the Fuchsian case \((k = 1)\) \[14\]; and will therefore be merely sketched. On the whole, the results of Section 2 are generalizations of Patterson's results for Fuchsian groups stated in Section 2 of \[14\], to geometrically finite groups with the restriction that all parabolic cusps (if there any) are of maximal rank.

The main purpose of this article is to give an interpretation of Theorem 2 in terms of quadratic forms. For this, suppose that \(c \in \mathbb{R}\) and consider the \(k + 1\)-dimensional hyperboloid

\[
Q(c) = \{(x, y) \in \mathbb{R}^{k+1} \times \mathbb{R} : y^2 - \|x\|^2 = c\}.
\]

If \(c\) is positive \(Q(c)\) is two sheeted, if \(c\) is negative it is of connected and if \(c = 0\) it is a cone. Denote by \(Q^+(c)\) the upper half of \(Q(c)\), i.e.

\[
Q^+(c) = \{(x, y) \in Q(c) : y > 0\}.
\]

The map \(p : Q^+(c) \to B^{k+1}\) defined by

\[
p(x, y) = \frac{x}{y + \sqrt{c}}
\]

is a bijection for positive \(c\) and \(p : Q^+(0) \to S^k\) is a surjection. A ray on \(Q^+(0)\) passing through the origin is mapped by \(p\) onto a single point on \(S^k\), different rays are mapped onto different points. However the map \(p : \Pi \cap Q(0) \to S^k\), where \(\Pi\) is a \(k + 1\)-dimensional hyperplane such that \(\Pi \cap Q(0)\) is a subset of \(Q^+(0)\), is a bijection and moreover is easily seen to be bi-Lipschitz.

For \(z = (x, y)\), let

\[
q(z) = y^2 - \|x\|^2
\]

denote the quadratic form on \(\mathbb{R}^{k+1} \times \mathbb{R}\). The function \(q(z)\) of the single variable \(z\) determines the symmetric bilinear form \(q(z_1, z_2)\) by the formula

\[
q(z_1, z_2) = \frac{1}{4} \{q(z_1 + z_2) - q(z_1 - z_2)\}
\]

so that \(q(z, z) = q(z)\), see \[3\]. A lattice in \(\mathbb{R}^{k+2}\) is the set of points

\[
u_1 b_1 + v_2 b_2 + \cdots + v_{k+2} b_{k+2}
\]

where \(b_1, b_2, \ldots, b_{k+2}\) is a basis for the lattice (i.e. a set of \(k + 2\) linearly independent elements of \(\mathbb{R}^{k+2}\)), and \(u_1, u_2, \ldots, u_{k+2}\) run through \(\mathbb{Z}\). Let...
Let \( \Lambda \) be a lattice in \( \mathbb{R}^{k+1} \times \mathbb{R} \) on which the quadratic form \( q \) takes integral values, i.e. \( q : \Lambda \to \mathbb{Z} \).

The group \( O(k+1,1) \) acts on each \( Q(c) \) and leaves the quadratic form \( q(z) \) and the bilinear form \( q(z_1,z_2) \) invariant. The action is transitive if \( c \neq 0 \) and is also transitive on \( Q(0)\setminus \{0\} \). Let \( I_{k+1} \) denote the \( k+1 \)-dimensional identity matrix. Then in matrix notation \( A \in O(k+1,1) \) if \( A^T J A = J \) where

\[
J = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \ddots & 0 \\
0 & \cdots & & -I_{k+1}
\end{pmatrix},
\]

whence \((\det A)^2 = 1\). For further details see [1, page 37] and [2, Section 3.7]. For \( c \) positive, Patterson [10] has shown that \( M(B^{k+1}) \), the group of orientation preserving Moebius transformations preserving the unit ball \( B^{k+1} \), can be identified with the subgroup \( O^+(k+1,1) \) of \( O(k+1,1) \) of index 2 which preserves all the \( Q^+(c) \). More precisely, \( M(B^{k+1}) = p(O^+(k+1,1))p^{-1} \).

Let \( \Gamma \) be the maximal subgroup of \( O^+(k+1,1) \) which preserves the lattice \( \Lambda \) introduced above. By definition \( \Gamma \) preserves each \( Q(c) \). It is known that \( \Gamma \) acts discontinuously on \( Q(1) \) and that the quotient \( Q(1)/\Gamma \) has finite volume. Also, if \( q \) takes on the value 0 on \( \Lambda \setminus \{0\} \), so that \( (Q(0)\setminus\{0\})\cap\Lambda \neq \emptyset \), the quotient \( Q(1)/\Gamma \) is not compact. Let \( G = p\Gamma p^{-1} \), then \( G \) is a discrete subgroup of \( M(B^{k+1}) \) with finite covolume. Since \( G \) is of finite covolume it is certainly of the first kind, and by a theorem of W. P. Thurston ([11, Prop. 8.4.3]) is also geometrically finite. Hence \( G \) is a geometrically finite group of the first kind. Furthermore, if \( Q(1)/\Gamma \) is not compact then the quotient \( B^{k+1}/G \) is not compact and so \( G \) contains parabolic elements [13]. In this case, to each parabolic cusp of \( G \) there corresponds a ray on \( Q(0) \) such that its intersection with \( \Lambda \setminus \{0\} \) is non-empty; and any point of \( (Q(0)\setminus\{0\})\cap\Lambda(\neq \emptyset) \) corresponds to some parabolic cusp of \( G \).

In Section 3, we prove the following interpretation of Theorem 2 in terms of the quadratic form \( q \) and the lattice \( \Lambda \) introduced above.

**Theorem 3.** Let \( q, \Lambda \) be as above and let \( \alpha \) be a positive real number. Suppose that \( Q^+(0) \cap \Lambda \neq \emptyset \). For \( \alpha > 1 \), the set

\[
W(\alpha) = \{ \xi \in Q(0) : |q(\xi,z)| \leq ||\xi|| ||z||^{-\alpha} \text{ for infinitely many } z \text{ in } Q(0) \cap \Lambda \}
\]

has Hausdorff dimension \( \frac{2k}{\alpha+1} + 1 \).
Some results on discrete groups

In order to simplify notation, write $L_g$ for $L(0, g(0))$. It is a well known fact that for any positive real number $N$

$$
\sum_{\substack{g \in G \\
L_g < N}} 1 \ll N^{\delta(G)},
$$

where the implied positive multiplicative constant is dependent only on the group $G$ ([7], [9]). A direct consequence of inequality (1) is that for $t < \delta(G)$

$$
\sum_{\substack{g \in G \\
L_g < N}} L_g^{-t} \ll \frac{\delta(G)}{\delta(G) - t} \cdot N^{\delta(G) - t}.
$$

The reader is referred to [14] for the proof of (2). In fact inequalities (1) and (2) are valid for arbitrary discrete subgroups of $M(B^{k+1})$. Given these estimates, the proof of Theorem 1 now follows on using the same arguments as those in Section 4 of [14].

The proof of Theorem 2 relies on the concept of a 'ubiquitous system' ([4, Section 5] in [14]), and the following results are essential in setting up such a system.

**Theorem 4.** Suppose that $G$ has no parabolic elements. Let $\eta, \eta'$ be the set of fixed points of a hyperbolic subgroup of $G$. Then there is a positive constant $c$ with the following property: for each $x$ in $L(G)$, $N > 1$, there exist $y$ in $\{\eta, \eta'\}$, $g$ in $G$ so that

$$
L_g < N \quad \text{and} \quad \|x - g(y)\| < \frac{c}{N}.
$$

**Theorem 5.** Suppose $G$ has parabolic elements only of maximal rank $k$ and let $P$ be a complete set of representatives of the cusps of $G$. Then there is a positive constant $c$ with the following property: for each $x$ in $L(G)$, $N > 1$, there exist $y$ in $P$, $g$ in $G$ so that

$$
L_g < N \quad \text{and} \quad \|x - g(y)\| < \frac{c}{\sqrt{L_g N}}.
$$

These theorems are a generalization of Dirichlet's theorem in the theory of global diophantine approximation and were proved by Patterson in [8] for Fuchsian groups. It should be noted that for Fuchsian groups
the concepts of geometrically finite and finitely presented coincide and that all parabolic cusps are of maximal rank. The proofs of the above theorems make use of the fact that $G$ has a convex fundamental polyhedron with finitely many faces and that in the case of Theorem 5, there exists an orbit point of the origin within a bounded distance (dependent only on $G$ and $P$) of the summit of a horoball belonging to a complete set of horoballs [12] for $G$. This latter fact is guaranteed by considering only groups with cusps of maximal rank. However, it should be possible to remove this restriction. In any case, for the purpose of this paper, Theorem 5 is required in the proof of Theorem 2 in which $G$ is of the first kind, and so all cusps are automatically of maximal rank.

For a subgroup $H$ of $G$, let $G/H$ be a set of representatives of the cosets $\{gH : g \in G\}$ so chosen that if $g \in G/H$, $h \in H$ then $L_g \leq L_{gh}$. So in this notation the minimum of $L_{gh}$ ($h \in H$) occurs when $gh \in G/H$. The following geometrical results are extensions of Propositions 8.1 and 8.2 in [8], and can be proved using the same arguments as in the Fuchsian case [8].

**Proposition 1.** Suppose $G$ has parabolic elements only of maximal rank $k$ and let $P$ be a complete set of representatives of the parabolic cusps of $G$. Then there are constants $c_1$, $c_2$, $c_3$, depending only on $G$ and $P$ with the following property: if $p, q \in P$, $g \in G\{G_p$ there is an element $h$ in $G\{G_p$ so that

$$\|h(p) - g(q)\| < \frac{c_1}{L_g} \text{ and } c_2L_g \leq L_h \leq c_3L_g.$$  

**Proposition 2.** Suppose that has $G$ no parabolic elements. Let $\eta, \eta'$ be the set of fixed points of a hyperbolic subgroup of $G$. Then there are constants $c_4$, $c_5$, $c_6$, depending only on $G$ and $\{\eta, \eta'\}$ with the following property: if $g \in G\{G_{\eta\eta'}$ there is an element $h$ in $G\{G_{\eta\eta'}$ so that

$$\|g(\eta') - h(\eta)\| < \frac{c_4}{L_g} \text{ and } c_5L_g \leq L_h \leq c_6L_g.$$  

Providing the group $G$ has parabolic elements, it should be noted that the hypotheses of Theorem 5 and Proposition 1 are satisfied for all non-elementary geometrically finite groups of the first kind and for all non-elementary finitely generated Fuchsian groups. The two propositions play an important role in ‘decoupling’ the parabolic cusps or the hyperbolic fixed points in order to obtain results for a single parabolic cusp or a single hyperbolic fixed point.

Let $G$ be a group of the first kind and $y$ be an arbitrary parabolic cusp of $G$ if there are any, and a hyperbolic fixed point otherwise. By
Theorem 1, \( \dim W_{G,y}(\tau) \leq k/\tau \) for \( \tau > 1 \); so in order to prove Theorem 2 it is sufficient to show that \( \dim W_{G,y}(\tau) \geq k/\tau(\tau > 1) \). This lower bound inequality is obtained by using the concept of ubiquity and the reader is referred to [4] and/or [14] for the definition of ubiquity.

Let \( \Omega \) be an open subset of \( S^k \) and let \( R = \{g(y) \in \Omega : g \in G\} \) be a set of points in \( \Omega \). For a positive constant \( c_7 \), let

\[
\beta : G \to \mathbb{R}^+ : g \mapsto \frac{L_g}{c_7}
\]

be a positive function on \( G \) and let

\[
\psi : \mathbb{R}^+ \to \mathbb{R}^+ : x \mapsto (c_7 x)^{-\tau},
\]

be a positive decreasing function with \( \psi(x) \to 0 \) as \( x \to \infty \). On combining the results stated in this section with the arguments used to prove Theorem 2 in [14]; it can be verified that the system \( (R, \beta) \) is ubiquitous relative to the function

\[
\lambda(n) = 4c_8n^{-1} \log n,
\]

where \( c_8 \) is a positive constant. Basically, this implies that for each positive integer \( N \) there exists a Lebesgue measurable subset \( A(N) \) of \( \Omega \) such that (i) \( \lim_{n \to \infty} |\Omega \setminus A(n)| = 0 \), where \( |X| \) denotes the \( k \)-dimensional Lebesgue measure of a measurable set \( X \); and (ii) for any \( x \) in \( A(N) \), there exists a \( g \) in \( G \) with \( L_g \leq c_7 N \) such that the inequality \( \|x - g(y)\| < \lambda(N) \) is satisfied. The function \( \lambda \) associated with the system \( (R, \beta) \) above, arises naturally from Theorem 5 and Proposition 1 in the case when \( G \) has parabolic elements and from Theorem 4 and Proposition 2 otherwise. The set \( A(N) \) is an approximating set for \( \Omega \) in the measure theoretical sense and is obtained by ‘thickening’ each point \( g(y) \) in \( \Omega \) with \( L_g \leq c_7 N \) by a \( \lambda(N) \)-neighbourhood.

It follows from Theorem 6 in [14] that

\[
\dim W_{G,y}(\tau) \geq k\gamma,
\]

where

\[
\gamma = \min \left\{ 1, \limsup_{n \to \infty} \frac{\log \lambda(n)}{\log \psi(n)} \right\} = \min \left\{ 1, \frac{1}{\tau} \right\}.
\]

Hence for \( \tau > 1 \), the correct lower bound inequality follows.
3. Proof of Theorem 3

Let $\Gamma$ be the maximal subgroup of $O^+(k + 1, 1)$ preserving the lattice $\Lambda$ and consider the action of $\Gamma$ on $Q(1)$. Then $G = p\Gamma p^{-1}$ is a geometrically finite group of the first kind, and since $Q^+(0) \cap \Lambda \neq \emptyset$ the group $G$ contains parabolic elements. Let $A$ be a complete set of inequivalent parabolic cusps. Since $G$ is geometrically finite, the set $A$ is finite. Consider the set $W_{G;A}(\tau) = \bigcup_{y \in A} W_{G;Y}(\tau)$ where

$$W_{G;Y}(\tau) = \{ x \in S^k : \| x - g(y) \| \leq L_g^{-\tau} \text{ for infinitely many } g \in G \}.$$  

By Corollary 1, $\dim W_{G;A}(\tau) = k/\tau$ for $\tau > 1$.

For any $\gamma$ in $\Gamma$ let $g = p\gamma p^{-1}$ be the corresponding element in the group $G$. Patterson in [10] shows that there exist positive constants $c_9, c_{10}$ so that if $z \in Q^+(0)$, then for any $\gamma$ in $\Gamma$ there exists $\gamma^*$ in $\gamma \Gamma_z$ such that

$$\| \gamma(z) \| < c_9 L_g \| z \|$$

and

$$\| \gamma^*(z) \| > c_{10} L_g \| z \|.$$  

Here $\Gamma_z = \{ \gamma \in \Gamma : \gamma(z) = z \}$ is the stabilizer of $z$, and so $\gamma(z) = \gamma^*(z)$; and $g^* \in g G_{p(z)}$ where $G_{p(z)} = \{ g \in G : g(p(z)) = p(z) \}$ is the stabilizer of the point $p(z)$ in $S^k$. This shows that $L_g$ can be taken to be comparable with $\| \gamma(z) \|$ for the optimal choice of $g$, i.e. for $g$ in $G \| G_{p(z)}$ in terms of the notation introduced in Section 2.

In the set $W_{G;Y}(\tau)$ replace the $g(y)$ by $p(z)$ for some $z$ in $Q^+(0) \cap \Lambda$ and take $L_g$ to be comparable with $\| z \|$. On writing $g(y) = p(z)$, $x = p(\zeta)$ where $\zeta \in Q(0)$, it is easily verified that

$$(3) \quad \| p(z) - p(\zeta) \|^2 = \frac{4|q(\zeta, z)|}{\| z \| \| \zeta \|}.$$  

Consider the set

$$W_{II} = \{ \zeta \in \Pi \cap Q(0) : |q(\zeta, z)| \leq \| \zeta \| \| z \|^{-2\tau + 1} \text{ for i. m. } z \text{ in } Q^+(0) \cap \Lambda \},$$

where $\Pi$ is a $k + 1$-dimensional hyperplane of the form described earlier. In view of (3) and the discussion above, the map $p = W_{II} \to W_{G;A}(\tau)$ is bi-Lipschitz. Since Hausdorff dimension is invariant under bi-Lipschitz transformations, the set $W_{II}$ also has Hausdorff dimension $k/\tau$ for $\tau > 1$.

Consider a ray on $Q(0)$ passing through the origin and containing a point of the set $W_{II}$. Excluding the origin, each point on the ray belongs to the set $W(2\tau - 1)$, since $q(\tau \zeta, z) = r q(\zeta, z)$ for $\tau \in \mathbb{R}\setminus\{0\}$; and each
point in $W(2\tau - 1)$ lies on some ray passing through the origin and containing a point in $W_\Pi$. Hence

$$W(2\tau - 1) = \{ r\zeta : \zeta \in W_\Pi, r \in \mathbb{R}\setminus\{0\}\},$$

from which it follows that for $\tau > 1$

$$\dim W(2\tau - 1) = \dim W_\Pi + 1 = \frac{k}{\tau} + 1.$$


Department of Mathematics
University of York
Heslington, York YO1 5DD
ENGLAND

Rebut el 3 de Setembre de 1993