OK

ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Scuola Internazionale Superiore di Studi Avanzati
International School for Advanced Studies

Equidistribution of parabolic fixed points in the limit set of Kleinian groups.

Thesis submitted for the degree of
"Doctor Philosophiæ"

CANDIDATE
Salvatore Cosentino

SUPERVISORS
Prof. Livio Flaminio
Prof. Alberto Verjovsky

October 1996

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

a Lalù

DÉFINITION. – *Dieu est le plus court chemin de zéro à l'infini.*

Dans quel sens ? dira-t-on.

– Nous répondrons que Son prénom n'est pas Jules, mais *Plus-et-Moins.* Et l'on doit dire :

$\pm$ *Dieu est le plus court chemin de* 0 *à* $\sim$*, dans un sens ou dans l'autre.*

Ce qui est conforme à la croyance aux deux principes ; mais il est plus exact d'attribuer le signe $+$ à celui de la croyance du sujet.

Mais Dieu, étant inétendu, n'est pas une ligne.

– Remarquons en effet que, d'après l'identité

$$\sim - 0 - a + a + 0 = \sim$$

la longueur a est nulle, a n'est pas une ligne, mais un point.

Donc, *définitivement* :

DIEU EST LE POINT TANGENT DE ZÉRO ET DE L'INFINI.

La Pataphysique est la science...

. . . .

(Alfred Jarry, *Gestes et Opinions du Docteur Faustroll*)

EQUIDISTRIBUTION OF PARABOLIC FIXED POINTS IN THE LIMIT SET OF KLEINIAN GROUPS

SALVATORE COSENTINO

Contents

§0. Introduction

0.1. *History.* Hyperbolic surfaces are quotients of the Poincaré half-plane $\mathbb{H} = \{x+iy \in \mathbb{C} \mid y > 0\}$ by discrete torsion-free subgroups Γ of $PSL(2,\mathbb{R})$ acting as fractional linear transformations and as isometries with respect to the metric $ds^2 = (dx^2 + dy^2)/y^2$. They provide some of the richest examples of dynamical systems.

The unit tangent bundle $S\Sigma$ of an hyperbolic surface Σ is canonically identified with the homogeneous space $PSL(2,\mathbb{R})/\Gamma$ where $\Gamma \simeq \pi_1(\Sigma)$. Geodesics and horospheres are the orbits of the one parameter subgroups

$$g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \qquad h_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} .$$

For compact Σ, the geodesic flow g_t on $S\Sigma$ is the prototype of an Anosov flow, and the orbits of h_x are the expanding (or unstable) leaves for g_t. If the surface is compact the geodesic flow is known to be ergodic, mixing and even Bernoulli with respect to the probability measure on $S\Sigma$ coming from the Haar measure on $PSL(2,\mathbb{R})$ (Hopf 1936, Ornstein and Weiss 1973). By contrast, the horocyclic flow h_t on $S\Sigma$ is known to be minimal and uniquely ergodic (Hedlund 1936, Furstenberg 1972).

Both minimality and unique ergodicity of the horocyclic flow fail when the surface Σ is not compact. By the Margulis lemma we know that non-compact finite area hyperbolic surfaces have a thick and thin-part decomposition, i.e. can be decomposed into a compact part and a finite number of ends diffeomorphic to cylinders $S^1 \times \mathbb{R}_+$. The fundamental group of an end of Σ is a cyclic subgroup $\mathbb{Z} \subset \pi_1(\Sigma)$, and corresponds to a conjugacy class of parabolic elements. It preserves each horosphere centered at its fixed point on the boundary of $\mathbb{H}$ producing closed orbits of the horocyclic flow.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

For example, we could think of a closed horocycle H as a wave front coming from an end of the surface. The geodesic flow acting on H produces a one parameter family of closed horocycles g_tH of increasing length $\exp(t)$ embedded in the unit tangent bundle. They project to immersed circles in the surface.

Sarnak was the first to show the following *asymptotic equidistribution* result. Let μ_t be the probability measure on the closed horocycle g_tH of uniform density with respect to the arc-length, then μ_t converges vaguely to the "Haar" probability measure on $S\Sigma$ as t goes to infinity [Sa1]. His proof uses the analytic properties of Eisenstein series and their Fourier expansions, and also gives error estimates in terms of certain distributions in the dual of compactly supported smooth functions.

Projecting the measures μ_t to the surface Σ, asymptotic equidistribution is a consequence of the Rankin-Selberg unfolding trick, and error terms are related to small eigenvalues of the Laplace-Beltrami operator in the Hilbert subspace orthogonal to cusp forms (Zagier [Za1,2]).

Equidistribution of horospheres is a particular case of a much more general result about the structure and ergodic properties of orbits of unipotent subgroups in homogeneous spaces. There is now a large amount of work originating from Ratner's proof of Raghunathan conjecture in 1990. The Séminaire Bourbaki by Ghys [Gh1] gives an account of work by Dani, Margulis, Mozes, Raghunathan, Ratner, Shah, Starkov and others.

In particular, for finite area hyperbolic surfaces Dani obtained a complete description of the space of σ-finite measures invariant for the horospheric flow. It is a closed set (in the vague topology) composed of: the "Haar" probability measure, the probability measures supported on the families of closed horospheres from the cusps, the zero measure (Dirac masses on the cuspidal fixed points in a natural compactification of $S\Sigma$). When the surface has infinite volume, it follows from the results of Ratner that the only invariant ergodic probability measures for the horocyclic flow are the ones concentrated on the closed horospheres (if there are any).

There is a well known relationship between mixing of the geodesic flow on a finite volume hyperbolic manifold (possibly non compact), equidistribution of wave fronts and asymptotic counting. Both spheres of increasing radius around a fixed center and closed horospheres (closed expanding leaves) from a cuspidal end have the interpretation of closed fronts acted upon by the geodesic flow. In particular, the mixing property of the geodesic flow on the unit tangent bundle of a finite volume hyperbolic surface can be used to obtain equidistribution, of spheres emanating from one point, as well as of closed horospheres from a cuspidal end. This in turn gives certain lattice point counting estimates for Γ-orbits inside $\mathbb{H}$. See the last pages in the book of Nicholls [Ni], or the paper by Eskin and McMullen for a more recent account [EM].

In the case of the (unit tangent bundle of the) modular orbifold, i.e. the quotient of $PSL(2,\mathbb{R})$ by the lattice $PSL(2,\mathbb{Z})$, Verjovsky gave a very simple geometrical proof of a part of Sarnak's result, also estimating the optimal error for mean values of characteristic functions. He reinterpreted the asymptotics of the measures supported on the closed horospheres from the unique cusp as a lattice point counting problem, and indeed a very well known one: the sum of the values of Euler totient function, known as Mertens formula. It is worth noting that fine results on the errors are a very delicate matter, as they are connected with the Riemann hypothesis (Verjovsky [Ve1,2], Zagier [Za1,2]).

0.2. *Equidistribution of parabolic fixed points.* Our purpose is to generalize asymptotic equidistribution results for closed horospheres to surfaces and higher dimensional hyperbolic manifolds of infinite volume.

We cannot look for ergodic invariant finite measures because of Ratner's above mentioned result. The counting problem version given in [Ve1,2] shows us the dynamical meaning of equidistribution. We illustrate our results for Kleinian groups, keeping in mind that they hold more generally for geometrically finite discrete groups of hyperbolic motions in any dimension.

A Kleinian group is a discrete subgroup Γ of

$$PSL(2,\mathbb{C}) = \{z \mapsto (az+b)/(cz+d) \mid a,b,c,d \in \mathbb{C}\ ,\ ad-bc=1\}\ .$$

The latter is the connected component of the isometry group of the three-dimensional hyperbolic space $\mathbb{H}^3 \simeq \mathbb{C} \times \mathbb{R}_+$, as well as the group of conformal automorphisms of the ideal boundary $\mathbb{C} \cup \infty \simeq S^2$. The group Γ is said to be geometrically finite if it admits a finite sided fundamental polyedron in $\mathbb{H}^3$. Such groups arise naturally as fundamental groups of convex hyperbolic manifolds with boundary.

A cusp of Γ is a conjugacy class of a parabolic subgroup P of Γ. It is always possible, after conjugating by an element of $PSL(2,\mathbb{C})$, to obtain a representative P of a cusp which fixes the point at infinity on the boundary of $\mathbb{H}^3$. A cusp P may be a finite extension of the cyclic group generated by $z \mapsto z+1$, in such a case we say it has rank one, or it may contain a finite index subgroup generated by $z \mapsto z+1$ and $z \mapsto z+\tau$ (for some τ with positive imaginary part), here we say it has rank two.

The limit set of Γ is the closed subset Λ of the Riemann sphere where Γ does not act discontinuously. The group Γ acts minimally on its limit set, and indeed Λ is the set of accumulation points of any Γ-orbit. If the quotient $\mathbb{H}^3/\Gamma$ has infinite volume then the limit set has zero Lebesgue measure inside S^2 and it is nowhere dense. The Hausdorff dimension δ of the limit set is called the critical exponent of Γ, since it is also the critical exponent of the classical Poincaré series of Γ.

The limit set of a Kleinian group has a remarkable dynamical significance, as it supports a non-trivial measure which restores all the nice statistical properties of the geodesic flow on the unit tangent bundle of the infinite volume manifold $\mathbb{H}^3/\Gamma$. This measure is the *Patterson-Sullivan measure* μ [Pa1] [Su1]. Let $|\gamma'|$ be the dilatation coefficient of an element γ of Γ computed with respect to the standard metric on the Riemann sphere. The measure μ is the unique probability measure supported on the limit set $\Lambda \subset S^2$ such that, for any $\gamma \in \Gamma$, we have $\gamma^*\mu = |\gamma'|^\delta \mu$.

The Patterson-Sullivan measure is equivalent to both the Hausdorff (covering) measure and the packing measure restricted to the limit set if the group Γ has no cusps (convex-cocompact). The presence of cusps forces the ratio between the packing and the Hausdorff measures of Λ to be infinite, unless Γ has finite volume quotient.

The action of Γ on Λ is ergodic and even weakly mixing (in the sense that $\Gamma \times \Gamma$ acts ergodically on $\Lambda \times \Lambda$) with respect to the measure class of μ. In analogy with the fact that the invariant measure of a uniquely ergodic dynamical system can be recovered as a weak limit of Dirac masses placed along any orbit of increasing length, we show that for Kleinian groups with cusps there is a natural way to select an "orbit" of the group Γ which simulates the Patterson-Sullivan measure, and indeed one orbit for each cusp. This happens because the Γ-orbit of any of the parabolic fixed points is dense in the limit set, and appropriately chosen balls around them are well behaved with respect to the Patterson-Sullivan measure.

Our first task is to recast the asymptotic equidistribution result for expanding horospheres centered at parabolic fixed points in the limit set of Γ.

Let μ be the Patterson-Sullivan measure supported on the limit set $\Lambda \subset S^2$ of Γ. Sullivan [Su1] constructs a measure m on the set $(\Lambda \times \Lambda - \mathrm{diag}) \times \mathbb{R}$ defined locally as

$$dm = \frac{d\mu(\xi) \cdot d\mu(\eta)}{|\xi - \eta|^{2\delta}} \cdot dr$$

where $|\xi - \eta|$ is the chordal distance in the euclidean metric of the unit closed ball. The measure m descends to a finite measure on the Γ-quotient, the *Bowen-Margulis measure*. The Bowen-Margulis measure is supported on the non-wandering set for the geodesic flow g_t on the unit tangent bundle of $\mathbb{H}^3/\Gamma$, it is g_t-invariant, and by the work of Rudolph the geodesic flow is Bernoulli with respect to m [Ru].

Assume that Γ contains a cusp P fixing ∞, and consider a horizontal horosphere H as an expanding horosphere (unstable leaf for the geodesic flow) in the unit tangent bundle. The conditional measure $d\mu^u = dm|_H$ is supported on a compact region of the quotient $H/P \simeq \mathbb{C}/P$ (this is trivial if the cusp has rank two, otherwise it follows from the fact that the group is geometrically finite). The geodesic flow produces the family of leaves g_tH, and the conditionals of dm are expanded as $d(g_t^*\mu^u) = e^{\delta t} d\mu^u$.

We generalize a result in [KM] obtaining asymptotic equidistribution for compact Borel subsets of unstable leaves under the geodesic flow. A particular case gives the following

Theorem 1. *The conditionals of the Bowen-Margulis measure on a family $g_t H$ of unstable fronts from a cusp, normalized by the factor $e^{-\delta t}$ to have constant mass, converge weakly to the Bowen-Margulis measure as time goes to infinity.*

We can state our main result in the following form. Let ∞ be the fixed point of a cusp P of Γ and take a sufficiently high horizontal hyperplane $H = \partial B$ bounding the horoball B centered at ∞. The images of B under the group Γ are disjoint horoballs (euclidean balls) resting at the points $b_\gamma = \gamma(\infty) \in \mathbb{C}$ with euclidean diameters s_γ. The *orbit* of the parabolic fixed point ∞ at time t is the collection of those resting points b_γ of the balls which have diameters s_γ bigger than e^{-t}. We construct the measures ν_t by placing Dirac masses at the t-orbit and normalizing with the factor $e^{-\delta t}$

$$\nu_t \;=\; e^{-\delta t} \cdot \sum_{s_\gamma > e^{-t}} \delta_{b_\gamma} \;.$$

Identify the fixed point ∞ of the cuspidal subgroup with the north-pole of the Riemann sphere S^2. The stereographic projection $\pi : (S^2 - \{\text{north-pole}\}) \to \mathbb{C}$ is a conformal map $\xi \mapsto \pi(\xi)$ with linear distorsion $|\pi'|$ proportional to $|\infty - \xi|^{-2}$ (norm in euclidean metric of the unit ball). The natural conformal image of μ on $\mathbb{C}$ is the push forward of μ by π times the conformal factor $|\infty - \xi|^{-2\delta}$

$$d\nu \;=\; |\pi'|^{\delta}\, d(\pi_* \mu) \;.$$

The measure $d\nu$ is proportional to $d\mu^u$ for the natural identification $H/P \simeq \mathbb{C}/P$ by vertical projection. Note that this construction would give the standard Lebesgue measure on $\mathbb{C}$ if the starting measure μ were the spherical one and $\delta = 2$. From Theorem 1 we get

Theorem 2. *For a geometrically finite Kleinian group with cusps and critical exponent δ the measures ν_t converge vaguely to a constant times ν.*

We refer to this as the *generalized Mertens formula*, as the statement is connected with the asymptotic behavior of a certain counting problem which reduces to Mertens formula for the Fuchsian group $PSL(2,\mathbb{Z})$.

Indeed, the collection of balls $\Gamma \cdot B$ is left invariant by the parabolic subgroup P and B itself is left invariant by P. Let $\mathcal{B} = (P\backslash\Gamma/P) \cdot B$ be the quotient collection. It is made of the images of B by a choice of representatives of the double sided coset $P\backslash\Gamma/P$. With no loss of generality we fix the starting horosphere $H = \partial B$ as the horizontal horosphere at unit height. A corollary of Theorem 2 is (the symbol $\sim$ stands for "asymptotic equivalence", i.e. the limit of the ratio is one)

Theorem 3. *Given any cusp P of a geometrically finite Kleinian group Γ with critical exponent δ there exists a positive constant $C(\Gamma, P)$ such that for t going to infinity*

$$\mathit{card}\left\{\gamma \in P\backslash\Gamma/P \;\mid\; s_\gamma > e^{-t}\right\} \;\sim\; C(\Gamma,P) \cdot e^{\delta t} \;.$$

The left-hand term above is the natural counting function that we attach to the pair (Γ, P). It is strictly related to the Eisenstein series of Γ with respect to the parabolic subgroup P.

Originally this kind of estimates dates back to Margulis' thesis for fundamental groups of compact negatively curved manifolds. Similar sharp counting estimates are due to Lalley and rely on a renewal theorem in the framework of symbolic dynamics [La]. He proved that for a convex-cocompact Kleinian group Γ with critical exponent δ

$$\mathrm{card}\{\gamma \in \Gamma \mid \mathrm{dist}(x, \gamma y) < R\} \;\sim\; C(\Gamma, x, y) \cdot e^{\delta R}$$

for $x, y \in \mathbb{H}^3$ and some constant $C(\Gamma, x, y)$. Convex-cocompact groups Γ are examples of Kleinian groups without cusps, and constitute the prototype of finitely generated groups that are hyperbolic in the sense of Gromov. In presence of cusps hyperbolicity fails, and the natural substitute is *relative hyperbolicity* with respect to the parabolic subgroups [Gr]. Theorem 3 explains the relative hyperbolicity since the parameter $-\log s_\gamma$ is the natural hyperbolic distance between the horospheres B and γB.

0.3. *Arithmetical examples.* There is an interesting relation between asymptotic equidistribution of close horocycles in the modular orbifold and the Riemann hypothesis [Ve1,2] [Za1,2]. We find an analogous relation for some arithmetic hyperbolic three-manifolds.

Let k be an imaginary quadratic number field. Its ring of integers ϑ is a lattice $\mathbb{Z}+\mathbb{Z}\tau$ in $\mathbb{C}$. For example, we could take $k=\mathbb{Q}(i)$ whose integers are the Gaussian integers $\mathbb{Z}+\mathbb{Z}i$.

Form the group of matrices $\Gamma_k = PSL(2,\vartheta)$, whose rows and columns are made of couples of relatively prime integers. The quotient $\mathbb{H}^3/\Gamma_k$ is a hyperbolic three-orbifold of finite volume with a number of cusps equal to the class number of the field k.

The parabolic subgroup P fixing the point at infinity is naturally identified with the lattice of the integers of k. A collection of $(P\backslash\Gamma_k/P)$-images of the horizontal horosphere at unit height is made of balls with resting points $a/b\in\mathbb{C}$ and diameters $1/|b|^2$ with $a,b\in\vartheta$ relatively prime and a mod b. We call the resting points of such balls with diameter bigger than x^{-1} the *Farey sequence* $\mathcal{F}_x$ of order x. They are torsion points of the elliptic curve $\mathbb{C}/\vartheta$, though not all of them.

The analogue of Mertens formula gives the asymptotics for $x\to\infty$

$$\mathrm{card}\{\mathcal{F}_x\} \sim C\cdot x^2$$

for some constant C related to the volume of the orbifold. This is Theorem 3 above, and Theorem 2 can be sharpened to include an estimate of the error as

Theorem 4. *The Farey sequence $\mathcal{F}_x$ is asymptotically equidistributed on the torus $\mathbb{C}/\vartheta$ with respect to the Lebesgue measure. The error for test characteristic functions of domains bounded by smooth curves can be estimated to be of order $x^{-1/2}$.*

The above estimates holds as well for the equidistribution of the closed horospheres in the oriented frame bundle $\mathcal{F}(\mathbb{H}^3/\Gamma_k)\simeq PSL(2,\mathbb{C})/PSL(2,\vartheta)$.

As explained by Zagier [Za2] the Rankin-Selberg unfolding trick can be used to investigate the way closed horospheres from the cusps fill the manifold when enlarged by the geodesic flow. Consider the closed horizontal horosphere H at unit height as an expanding front in the unit tangent bundle. The geodesic flow produces the closed horosphere g_tH. We define m_t to be the probability measure uniformly supported on the front of g_tH inside $M_k=\mathbb{H}^3/\Gamma_k$, and denote by m the probability Liouville measure on the hyperbolic manifold.

For mean values of smooth functions f with compact support on the manifold there holds the estimate $|m_t(f)-m(f)| = \mathcal{O}(\mathrm{e}^{-(1-\varepsilon)t})$. We generalize Zagier's theorem as

Theorem 5. *The Riemann hypothesis for the Dedekind zeta function ζ_k of an imaginary quadratic number field k is true if and only if for any smooth function f with compact support on M_k and any strictly positive ε*

$$|m_t(f)-m(f)| = \mathcal{O}(\mathrm{e}^{-(3/2-\varepsilon)t}).$$

0.4. *Organization of the material.* Sections 1 and 2 contain a review on the geometry of Kleinian groups, the Patterson-Sullivan measure on their limit sets, and the construction of the Bowen-Margulis measure on the nonwandering set for the geodesic flow. In section 3 we discuss the metric structure of the Patterson-Sullivan measure, we prove a result about the fluctuation of its density in the convex-cocompact case, and we discuss certain packing and covering results due to the presence of cusps. Section 4 contain a discussion of a rough version of Theorem 3 in the spirit of Sullivan's bounded ratio estimates, and its consequences. In section 5 we use Rudolph's result about the mixing property of the geodesic flow and derive Theorems 1-3 together with slightly finer versions of them. Section 6 contains a discussion of some arithmetic hyperbolic 3-manifolds and the proofs of Theorem 4-5.

0.5. *Acknowledgements.* I whish to thank my advisors Alberto Verjovsky and Livio Flaminio, who introduced me to the beautiful interplay between geometry and dynamics and generously shared with me their time and ideas. I benefited of a six months visit to the Centre international de recherche Emile Borel in the Institute Henri Poincaré of Paris on the occasion of a "Semestre de Geometrie Symplectique" in 1994, I thank the organizers for their hospitality. During the last two years I've been frequently a guest at the U.F.R. de Mathématiques of the Université des Sciences et Technologie de Lille, I thank all of them for this opportunity, and I am particularly indebted to Abdelghani Mouhallil for his hospitality. It is a pleasure to thank Alexander Starkov and Nimish Shah, who explained to me their recent results connected with the Raghunathan conjecture, Alexei Venkov, who kindly explained to me the theory of automorphic functions and Eisenstein series, Sidney Frankel, Youssef Hantout, and Anton Zorich for lots of interesting discussions.

§1. Geometric preliminaries

Here we introduce our notation. recall some basic hyperbolic geometry, and describe the class of discrete groups of hyperbolic motions we are going to be interested in. General references are the notes by Thurston [Th1], the historical monograph by Milnor [Mi], and the recent books [BP], [Ma] and [Ra].

1.1. *Hyperbolic geometry.* We have in mind two models of the $n+1$ dimensional hyperbolic space, which from now on will be denoted $\mathbb{H}$. If $\mathbb{R}^{n+1}$ denotes the euclidean space with metric ds^2, the *unit ball model* $\mathbb{B}^{n+1}$ is $\{x \in \mathbb{R}^{n+1} \mid |x| < 1\}$ equipped with the Riemannian metric $4 \cdot ds^2/(1-|x|^2)^2$, and the *upper half-space model* (abbreviated as u.h.s.) $\mathbb{H}^{n+1}$ is $\{x \in \mathbb{R}^{n+1} \mid x_{n+1} > 0\}$ equipped with the metric ds^2/x_{n+1}^2.

The hyperbolic metric on the open unit ball in $\mathbb{R}^{n+1}$ is the unique Riemannian metric such that the group of its isometries is the group of its conformal automorphisms (w.r.t. the euclidean metric), normalized as to have constant sectional curvature -1. As follows from a theorem by Liouville, this group is generated by orthogonal transformations and inversions along spheres perpendicular to the unit sphere. It is isomorphic to the Möbius group $\mathbb{M}^n = \mathrm{Con}(S^n)$, the group of conformal automorphisms of the standard unit sphere, i.e. the action of a conformal map γ from the unit ball to itself extends to the boundary sphere, and its trace on the sphere uniquely determines the action on the ball. The Möbius group acts transitively on $\mathbb{B}^{n+1}$ and even on its tangent bundle via the differential of the action.

The u.h.s. model is obtained from the unit ball model via the Cayley transform. Namely, composing the inversion along the sphere centered at $e_{n+1} = (0, ..., 0, 1)$ with radius $\sqrt{2}$ with the reflection w.r.t. the hyperplane $x_{n+1} = 0$ (to preserve orientation), we get a conformal map between the unit ball and the upper half-space. It extends to the boundary sending e_{n+1} to the point at infinity and indeed, when restricted to the sphere, it is a variant of the stereographic projection. It is clear that, composing the above map with a rotation, for any point ξ in the unit sphere we can construct an isometry between $\mathbb{B}^{n+1}$ and $\mathbb{H}^{n+1}$ that sends ξ to the point at infinity and the origin to e_{n+1}.

In both models, geodesics are arcs of circles perpendicular to the boundary, including also diameters in $\mathbb{B}^{n+1}$ and vertical lines in $\mathbb{H}^{n+1}$.

The *ideal boundary* (or sphere at infinity) S of $\mathbb{H}$ can be intrinsically defined as the space of equivalence classes of oriented geodesics that are at bounded hyperbolic distance from a certain time on. It is naturally identified with the unit sphere S^n in the unit ball model and with $\mathbb{R}^n \cup \infty$ in the u.h.s. model (e.g., to the oriented geodesic $x(t)$, $t \in \mathbb{R}$, there correspond the point at infinity $x_\infty = \lim_{t \to \infty} x(t)$, limit in the metric topology of euclidean unit ball). By what was said above, the group of orientation preserving isometries of $\mathbb{H}$ is as well the group of conformal automorphisms of the ideal boundary S.

In the following we will often switch from one model to another according to convenience of the arguments, keeping the same notation for abstract objects in $\mathbb{H}$ or S and representatives in the models. For this reason we introduce some invariant notation.

Hyperbolic distance between points x and y in the hyperbolic space will be indicated by (x, y).

Horospheres are limits of hyperbolic spheres as the centers approach ideal points, and in both models they are euclidean spheres tangent to the boundary, including horizontal hyperplanes in the u.h.s. model. It is customary to call *center* of the horosphere its point of tangency with the ideal boundary. Horospheres are hyperplanes inside $\mathbb{H}$, and inherit from the hyperbolic metric an euclidean structure. *Horoballs* are the interior of the horospheres.

Horospheres are level sets of *Busemann functions*, defined as follows: fixing a reference point $x \in \mathbb{H}$, for any ideal point $\xi \in S$ we construct a function $\rho_x(\cdot, \xi)$ on $\mathbb{H}$ normalized to be zero at the point x, and such that if $z(t)$ is any geodesic $z(t) \to \xi \in S$, then

$$\rho_x(y,\xi) \;=\; \lim_{t\to\infty} (y, z(t)) - (x, z(t)) \;\; .$$

That is, $\rho_x(y, \xi)$ is the signed hyperbolic distance between the two horospheres centered at ξ and passing through x and y, also referred as the "horospheric" distance. For example, if ξ is the point at infinity in the u.h.s. model (hence horospheres centered at it are horizontal hyperplanes), and x a point at unit height, $\exp(-\rho_x(y, \xi))$ is the height of the point y.

The *Busemann cocycle* based at x is

$$\beta_x(\xi, \eta) \;=\; \rho_x(y, \xi) + \rho_x(y, \eta)$$

where y is any point lying on the geodesic between ξ and η (it does not depend on the chosen y, and it is the hyperbolic length of the portion of the geodesic lying between the two horospheres based at ξ and η and passing from x).

We recall the classification of the elements of the Möbius group. An isometry γ is called *elliptic* if it fixes a point in $\mathbb{H}$, *parabolic* if it fixes exactly one point in S, *loxodromic* otherwise. A parabolic isometry fixes the collection of horospheres centered at its fixed point, acting on each of them as an euclidean motion.

If γ is an hyperbolic isometry, hence a conformal automorphism of the sphere at infinity, we denote by $|\gamma'|$ the *linear distorsion* from the standard spherical metric of the boundary of the unit ball model. This latter is the visual metric on the boundary as seen from the origin $o \in \mathbb{B}^{n+1}$, and we can map the unit ball to the u.h.s. model sending o to the point e_{n+1} and any ξ to ∞. In the u.h.s. model one has the infinitesimal formula (hyperbolic metric) $\cdot$ (height) $=$ (euclidean metric), hence, if x is the image of o by γ, the linear distorsion at the point ξ is the inverse of the height of x

$$|\gamma'(\xi)| \;=\; e^{\rho_o(x,\xi)} \;\; .$$

We end this section recalling the lowest dimensional examples of Möbius groups. The group of orientation preserving isometries of the hyperbolic plane $\mathbb{H}^2 \simeq \{z \in \mathbb{C} \;\text{ s.t. }\; \mathrm{Im}(z) > 0\}$ is canonically isomorphic to the group $PSL(2, \mathbb{R})$, acting as fractional linear transformations. Identifying the three-dimensional hyperbolic space with a subspace of the skew-field of quaternions we see its group of orientation preserving isometries is isomorphic to $PSL(2, \mathbb{C})$. This is also the group of conformal automorphisms of the boundary sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \infty$ (fractional linear transformations).

1.2. *Geometrically finite groups.* We use the term *Kleinian group* to denote any discrete subgroup Γ of the Möbius group $\mathbb{M}^n$ of some dimension n. This name is traditionally reserved for discrete subgroups of $PSL(2, \mathbb{C})$ with non-empty ordinary set on the Riemann sphere, see below, but for our purpose to investigate certain ergodic properties of their action on the limit set the dimension plays no special role.

For a discrete group Γ of isometries of $\mathbb{H}$ the most basic notion is that of *limit set* Λ: it is the set of cluster points of any of its orbits in the unit ball model. As the group is discrete, the limit set is a closed subset of the ideal boundary. The limit set is either composed of one or two points, or it is a perfect set, in particular infinite. The first two cases can only occur for elementary discrete groups, finite extensions of cyclic groups, and these are uninteresting examples. The group Γ acts on S too, and Λ is the unique closed set where this action is minimal, i.e. every orbit is dense. Moreover, a non-elementary Γ and any non-trivial normal subgroup $\Gamma' \subset \Gamma$ have the same limit set.

The *ordinary set*, or domain of discontinuity, is the part of the sphere at infinity $O = S - \Lambda$ where Γ acts properly discontinuously. Groups with non-empty ordinary set are also known as groups of the second kind. In these cases the limit set is nowhere dense and the ordinary set has full Liouville measure. The quotient O/Γ has a natural structure of a conformally flat manifold.

The *radial* limit set Λ_r is the set of ideal points such that any geodesic asymptotic to them intersects a neighborhood of an orbit of Γ infinitely times. It is a subset of the limit set.

The *convex hull* $\mathrm{Hull}(\Lambda)$ of the limit set is the smallest (hyperbolic) convex set in the closed unit ball which contains Λ. It contains the union of all the geodesics between couples of points in the limit set.

The group Γ is said *convex-cocompact* if it acts cocompactly on the quotient hull of its limit set, i.e. if $\mathrm{Hull}(\Lambda)/\Gamma$ is compact. This is a natural generalization of fundamental groups of compact hyperbolic manifolds.

A *cusp* of Γ is a conjugacy class of maximal parabolic subgroups. A representative of a cusp $P \subset \Gamma$ fixes a certain point ξ on the sphere at infinity and leaves invariant the family of horospheres centered at it. It contains a finite index abelian normal subgroup P' which acts on the family of horospheres as a rank $k \leq n$ discrete group of translations of the n-dimensional euclidean space. A cusp is said *bounded* (or parabolic) if the quotient of $\Lambda - \xi$ by P is compact. The *parabolic* limit set of Γ is the set $\Lambda_p \subset \Lambda$ of fixed points of the cusps.

We adopt the following definition: the discrete group Γ is said to be *geometrically finite* if it is possible to chose an equivariant system of disjoint open horoballs, based at the parabolic fixed points, such that *the quotient by Γ of the convex hull of the limit set with these balls removed is compact.* The usual notion of geometrically finiteness is that there exists a finite sided fundamental polyedron, and the two are known to be equivalent up to dimension three. Our definition is due to Epstein, Cannon, Holt, Levy, Paterson and Thurston [Ep], see also Ratcliffe [Ra] for an equivalent definition. We choose it since it is more natural in dealing with the construction of the geometric measure.

Amongst the consequences of the definition we quote: a geometrically finite Kleinian group Γ is finitely generated and contains a finite number of bounded cusps. The limit set of a geometrically finite group is made of radial limit points and the countable set of Γ-images of the cuspidal fixed points.

Since Γ is a finitely generated subgroup of some matrix group, by Selberg's lemma it contains a torsion-free normal subgroup of finite index. That is, up to finite covering we can assume that Γ does not contain elliptic elements. Hence Γ acts freely on $\mathbb{H} \cup O$.

In a note at the end of [NS] it is quoted an observation by G. Niblo, who gave a proof of the following statement. Any geometrically finite hyperbolic group Γ contains a finite index subgroup $\Gamma' \subset \Gamma$ such that:

- all parabolic subgroups of Γ' are abelian and torsion free;

- if P and P' are distinct cusps of Γ, even $P \cap \Gamma'$ and $P' \cap \Gamma'$ are non-conjugate in Γ'.

In view of this fact, in the following we also assume that any of the cusps of our group is isomorphic to some $\mathbb{Z}^k$.

To get a picture of the kind of quotients of $\mathbb{H}$ associated to a geometrically finite Γ we make the following definitions. We call

$$\begin{aligned} M_\Gamma &= \mathrm{Hull}(\Lambda)/\Gamma && \text{the convex hull quotient} \\ N_\Gamma &= \mathbb{H}/\Gamma && \text{the complete hyperbolic manifold without boundary} \\ \overline{N_\Gamma} &= (\mathbb{H} \cup O)/\Gamma && \text{the Kleinian manifold .} \end{aligned}$$

There are natural inclusions $M_\Gamma \subset N_\Gamma \subset \overline{N_\Gamma}$. The convex hull quotient M_Γ is a deformation retract of the complete manifold N_Γ. Also N_Γ is naturally identified with the interior of $\overline{N_\Gamma}$ and is a convex complete hyperbolic manifold (according to Thurston convex means that any path with fixed endpoints is homotopic to a geodesic arc). There is a compact manifold with boundary $\overline{N_\Gamma^0}$ inside $\overline{N_\Gamma}$ such that their difference is the disjoint union of a finite number of cusps.

The thick and thin decomposition of the manifold is reflected in the following naive picture: $\mathrm{Hull}(\Lambda)/\Gamma$ is made of a compact piece with boundary and a finite number of thin ends, one for each cusp.

An important observation from the point of view of the dynamics is the following. If we consider the geodesic flow on the unit tangent bundle of the (possibly infinite volume) hyperbolic manifold $\mathbb{H}/\Gamma$, we see that the nonwandering set is the quotient of the set of geodesics in $\mathbb{H}$ starting and ending at radial limit points.

Examples. The main example of a geometrically finite discrete group of hyperbolic isometries is a uniform lattice $\Gamma \subset \mathbb{M}^n$. Its limit set is the whole sphere at infinity and the quotient $\mathbb{H}/\Gamma$ is a compact hyperbolic manifold (in general an orbifold if Γ contains elliptic elements) of finite volume.

Lattices with cusps are called non-uniform. In this case $\mathbb{H}/\Gamma$ is an hyperbolic manifold of finite volume composed of a compact part and a finite number of ends, one for each cusp, diffeomorphic to a n-torus times $(0, \infty)$.

Classical examples of Kleinian groups with non-empty ordinary set are Schottky groups. Consider a collection of $2 \cdot r$ disjoint half-spaces in $\mathbb{B}^{n+1}$, i.e. regions bounded by spheres perpendicular to the boundary. Chose isometries $\gamma_1, ..., \gamma_r$ between pairs of the bounding hyperplanes and define Γ to be the Kleinian group generated by the γ_i's. The group Γ is isomorphic the free group with r generators, and if $r > 1$ the limit set is a Cantor set (perfect and totally disconnected). A little reflection on the construction shows that arranging the euclidean radii of the half-spaces we are able to produce limit sets of Hausdorff dimensions arbitrarly close to n.

Other example of Kleinian groups are quasi-Fuchsian representations of Fuchsian groups (see [Bo] for a discussion on the Hausdorff dimension of their limit sets).

In general, groups with infinite volume quotients arise naturally as fundamental groups of convex hyperbolic manifolds with boundary.

Examples of non convex-cocompact groups can be obtained by considering the normal subgroup corresponding to the fiber of the compact hyperbolic 3-manifolds that fiber over the circle constructed by Jörgensen and Thurston.

Interesting examples of Kleinian groups with cusps are certain subgroups of the Picard group, which consists of two-by-two matrices of Gaussian integers. The *Apollonius packing* residual set is the limit set of one of them.

Remark. It is well known that the main examples of finitely generated groups that are *hyperbolic in the sense of Gromov* are convex-cocompact discrete groups of hyperbolic isometries, e.g. fundamental groups of compact hyperbolic manifolds or Schottky groups. Groups with parabolic elements cannot be hyperbolic, but are the motivation for Gromov concept of *relative hyperbolicity* [Gr]. Automatic structures for them have been recently investigated by Neumann and Shapiro [NS].

1.3. *Disjoint balls.* Select a representative P of a cusp of Γ, and map its fixed point into the point at infinity in the upper half-space model for $\mathbb{H}$, with boundary $\mathbb{R}^n \cup \infty$. The subgroup P

acts on the boundary $\mathbb{R}^n$ as a free abelian group generated by k translations. It acts cocompactly on a certain k-plane $\mathbb{R}^k$ inside $\mathbb{R}^n$. As the quotient $(\Lambda - \infty)/P$ is compact, a fundamental domain for P in $\mathbb{R}^k$ can be lifted to give an n-dimensional closed parallelepiped $T \subset \mathbb{R}^n$ whose images by P have disjoint interiors and cover $\Lambda - \infty$.

The parabolic subgroup P leaves invariant the horizontal horospheres. If we take a sufficiently high horizontal horosphere H, its images under Γ are a P-invariant family of *disjoint* (euclidean) *spheres* resting on the boundary $\mathbb{R}^n$. We are free to take the basic horosphere H so high that its images have all diameter smaller than one.

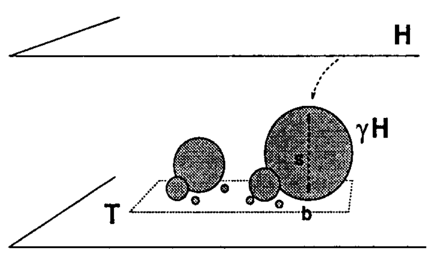
An image of the basic horoball B, the one bounded by H, is thus γB and can be parametrized by

$$\begin{array}{rl} \textit{resting point} & b \in \mathbb{R}^n \\ \textit{size}\ \text{(euclidean diameter)} & s \in (0,1) \end{array}$$

Resting points are the Γ-images of the cuspidal fixed point, $\gamma\infty$. The *size* is a natural parameter, in the sense that its logarithm is a "distance" from the basic horosphere at unit height H_1: indeed, if $s = s_\gamma = \mathrm{e}^{-t}$, we see that

$$t \;=\; \min_{y\in\gamma H} |\rho_o(y,\infty)| \;=\; \min_{y\in\gamma H}\,(y, H_1)$$

where o is the origin in $\mathbb{B}^{n+1}$ (hence the point e_{n+1} in $\mathbb{H}^{n+1}$).



As explained above the collection of disjoint balls is left invariant by the group of translations P. A set of representatives for them can be chosen such that their resting points all lie on the parallelepiped T. We define the quotient collection as

$$\mathcal{B} \;=\; \{\gamma B \;|\; \gamma \in P\backslash\Gamma/P\}\;.$$

The resting points of the balls in $\mathcal{B}$ describe a $(P\backslash\Gamma/P)$-orbit of the parabolic fixed point ∞.

We have as well a collection of balls in $\mathbb{R}^n$, the shadows of the disjoint balls projected by the vertical geodesics. Resting points become the centers of the balls, and diameters are still the sizes. The collection passes to the quotient $\mathbb{R}^n/P \;\simeq T^k \times \mathbb{R}^{n-k}$ and the centers of the balls are contained in the compact part coming from T. We will keep the same notation for balls and their shadows and identify a $(P\backslash\Gamma/P)$-orbit of the basic horoball B with the collection of euclidean balls $\mathcal{B} = \{B(b,s)\}$ in $\mathbb{R}^n$. Also, for any function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ we will denote by $B(b, \phi(s)\cdot s)$ the ball resting at b and diameter $\phi(s)\cdot s$, i.e. the same ball of $\mathcal{B}$ squeezed by the factor $\phi(s)$.

Consider the unit tangent bundle of $\mathbb{H}$. The Riemannian metric on $\mathbb{H}$ defines a canonical isomorphism between the latter and the bundle of cooriented contact elements of $\mathbb{H}$, and both have a natural contact structure. The set of unit vectors based at the basic horosphere H and pointing downward (u.h.s. model) forms a Legendrian submanifold of the unit tangent bundle. Its front is the projection on $\mathbb{H}$, i.e. the horosphere H. The geodesic flow acts on the Legendrian submanifold dilatating the volume of its front, i.e. contracting its height as

$$\mathrm{height}(g_t H) = e^{-t} \cdot \mathrm{height}(H) \ .$$

The $(P\backslash\Gamma/P)$-images of the dilated front read $\mathcal{B}_t = \{B(b, e^t s)\}$.

§2. Patterson-Sullivan measure

In this section we review some of the results obtained by Sullivan in [Su1,2,3,4]. From now on we denote by Γ a geometrically finite discrete subgroup of the Möbius group $\mathbb{M}^n$ with cusps, that we assume torsion-free and such that all its cusps are isomorphic to some $\mathbb{Z}^k$.

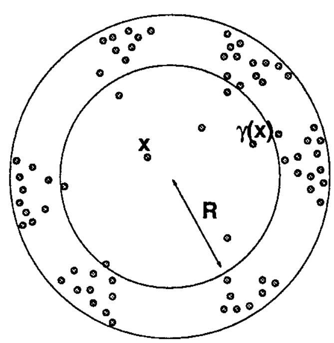
2.1. *The conformal measure.* Let $n(R)$ be the cardinality of the points of a given orbit Γy contained inside a ball of hyperbolic radius R centered at x. We define the *critical exponent* δ of Γ as the exponential growth rate of $n(R)$

$$\delta(\Gamma) = \limsup \frac{1}{R} \log n(R) \ .$$

It does not depend on the base points x and y. The critical exponent δ also governs the divergence of the *Poincaré series*

$$g_z(x,y) = \sum_{\gamma\in\Gamma} e^{-z(x,\gamma y)} \ .$$

The series converges for $z > \delta$, and diverges for $z < \delta$.



Generalizing a clever construction by Patterson, Sullivan constructed a remarkable collection of finite measures on the limit set of any geometrically finite Kleinian group. The spirit of the construction is best understood with the following definitions.

Let X be a smooth manifold, and let be given a collection of Riemannian metrics ρ's on X that are conformally equivalent. This means that for any two metrics ρ and ρ' their ratio ρ/ρ' is a well defined smooth function on X. A *conformal density of dimension s* is a map that assigns a finite measure μ^ρ on X to any metric in the collection

$$\{\text{Riemann metrics in a conformal class}\} \longrightarrow \{\text{finite measures}\}$$

with the prescriptions that all the measures belong to the same measure class (are absolutely continuous w.r.t. each other) and that the Radon-Nikodym derivative between two measures is the ratio of the respective metrics raised to the power s, i.e.

$$\frac{d\mu^{\rho}}{d\mu^{\rho'}} = (\rho/\rho')^{s} .$$

The restriction of the Hausdorff s-measure on a Borel subset of X of finite Hausdorff s-measure is a natural example.

The ideal boundary S of the hyperbolic space has a natural conformal structure. For any point $x \in \mathbb{H}$ a metric ρ on S is naturally defined as the visual metric from x. This means that the forward visual map sends the unit tangent sphere at x onto the boundary S, and ρ is the metric on S such that this map is an isometry. A Möbius transformation γ realizes an isometry between ρ and the metric $\rho' = |\gamma'| \cdot \rho$, the visual metric on the sphere as seen from γx.

We say a conformal density on S is *invariant* by the discrete group Γ if for any $\gamma \in \Gamma$

$$\gamma^{*} \mu^{\rho} = \mu^{|\gamma'|\cdot\rho}$$

where * denotes the pullback operation on measures, i.e. $\gamma^*\mu(A) = \mu(\gamma A)$ for any measurable A.

It is clear that a conformal density is uniquely determined by any of its representative measure. This justifies the following definition. We say a finite measure μ on S is *geometric* (or automorphic) for Γ if it is supported on the limit set Λ and satisfies $\gamma^*\mu = |\gamma'|^D\mu$ for any element γ of Γ, where D is the Hausdorff dimension of the limit set.

Theorem 2.1. *(Patterson* [Pa2], *Sullivan* [Su1]*) For a geometrically finite Kleinian group Γ the critical exponent δ is equal to the Hausdorff dimension D of the limit set and there exists a unique Γ-invariant conformal density of exponent δ supported on the limit set.*

Existence is given by Patterson's construction via a weak limit of a sum of Dirac measures supported on an orbit of the group weighted with the terms of the Poincaré series

$$\mu_z^x = \frac{1}{g_z(y,y)} \sum_{\Gamma} e^{-z(x,\gamma y)} \delta_{\gamma y} .$$

This family of measures is weakly bounded for $z > \delta$, hence we can take a weak limit as z goes to δ, say μ^x. We know "a posteriori" that *the Poincaré series diverges at the critical exponent*, thus the limit measure is supported on the limit set (otherwise we should slightly modify the terms of the series to get divergence, as originally done by Patterson). To see that the limit measure is geometric of dimension δ, we take y fixed and compare two such measures based at x and x'. The ratio $e^{-(x',\gamma y)}/e^{-(x,\gamma y)}$ goes to $e^{-\rho_{\bullet}(x',\xi)}$ as γy converges towards a point ξ on the limit set. Thus the Radon-Nikodym ratio has the expression

$$\frac{d\mu^{x'}}{d\mu^{x}}(\xi) = e^{-\delta\rho_{\bullet}(x',\xi)}$$

and the right hand side is the ratio of the visual metrics on S as seen from x' and x raised to the power δ. From the definition of the pullback of a measure we also see that for any $\gamma \in \Gamma$ we have $\gamma^*\mu^x = \mu^{\gamma x}$. These two facts together say that the collection of measures $\{\mu^x\}_{x\in\mathbb{H}}$ is a Γ-invariant conformal density of exponent δ supported on the limit set (μ^x is the measure associated to the visual metric on S from x).

We note that if we take for x the origin o in the unit ball model, and call $\mu = \mu^o$, we get

$$\gamma^*\mu = |\gamma'|^{\delta} \mu$$

for any $\gamma \in \Gamma$ (remember that the standard spherical metric on S is the visual metric from o).

Uniqueness of the invariant density is equivalent to the ergodicity of the action of Γ on the limit set, relative to the measure class determined by μ. This is also equivalent to the statement: there exists a unique probability measure μ supported on the limit set such that for any $\gamma \in \Gamma$ it transforms as $\gamma^*\mu = |\gamma'|^\delta \mu$.

The critical exponent is the Hausdorff dimension of the limit set and also of the radial limit set, since parabolic points are countable. Also, the measure μ has no atoms along parabolic points (we assume that the group is not elementary, i.e. that the limit set contains more than two points) due to the fact that if k is the rank of any of the cusps of Γ we have the bound $2\delta > k$.

Remark. We note that the critical exponent δ of a Kleinian group Γ is not an invariant of Γ as an abstract group. For example, it is possible to construct isomorphic Schottky groups (i.e. free groups) with limit sets of arbitrary Hausdorff dimension $0 < \delta < n$. The critical exponent depends on the representation of Γ in some Möbius group $\mathbb{M}^n$.

2.2. *The Bowen-Margulis measure.* The limit set does not support any Γ-invariant probability measure. To get invariance we must look at the diagonal action of Γ on $\Lambda \times \Lambda$. From a geometric measure μ, Sullivan defines the measure [Su1]

$$dm = \frac{d\mu \times d\mu}{|\xi - \eta|^{2\delta}} \times d(\text{arc length})$$

on $(\Lambda \times \Lambda - \text{diag}) \times \mathbb{R}$, which is Γ as well as time invariant, and descends to a finite measure m on the quotient. Following Yue we will call it the *Bowen-Margulis* measure, due to the similarity of the two constructions. The picture we get is the following: the set $(\Lambda \times \Lambda - \text{diag}) \times \mathbb{R}$, behaves like the unit tangent bundle of an hyperbolic space of dimension $\delta + 1$ with respect to the action of Γ as a lattice of isometries.

Theorem 2.2. *(Sullivan* [Su1,4]*) For a geometrically finite Kleinian group Γ*

- the counting function $n(R)$ is within a bounded ratio of $e^{\delta R}$ (see also [SV]*),*

- m gives full measure to the non-wandering set for the geodesic flow $(\Lambda_r \times \Lambda_r - diag) \times \mathbb{R}$,

- the measure m is recurrent, hence ergodic, for the geodesic flow (the geodesic flow is even Bernoulli [Ru]*),*

- the diagonal action of Γ is ergodic on $\Lambda \times \Lambda$ (this is sometimes called weak mixing) w.r.t. *$(\mu \times \mu)/|\xi - \eta|^{2\delta}$ as an invariant finite measure,*

- the conditional measures on the stable (unstable) leaves contract (expand) exponentially with exponent δ under the geodesic flow,

- δ is the measure theoretic entropy of the geodesic flow (with respect to m) and as well the topological entropy if there are no cusps (in formal agreement with Pesin-Margulis theory).

Since it will useful in the following we explain more carefully the definition of m. The unit tangent bundle of $\mathbb{H}$ is foliated in three ways: geodesics, expanding (unstable) horospheres, and contracting (stable) horospheres. The geodesic flow is the typical example of an Anosov flow [An]. Given a vector $v \in S\mathbb{H}$ the forward and backward visual maps give us two points on the sphere at infinity, the endpoints of the geodesic through v that we denote by v_∞ and $v_{-\infty}$. Thus, the forward visual map sends the unstable horosphere through v, say H^u_v (by definition the set of vectors w s.t. $w_{-\infty} = v_{-\infty}$ and $\rho_o(w, v_{-\infty}) = \rho_o(v, v_{-\infty})$), onto the sphere at infinity minus $v_{-\infty}$

$$p_u \ : \ H^u_v \ \rightarrow \ S - \{v_{-\infty}\}$$

and the same does the backward visual map $p_s : H^s_v \rightarrow S - \{v_\infty\}$ (in the following, we will use the notation $p_{u,v}$ or $p_{s,v}$ when it will be important to remind the point fixing the horosphere). Fix once for all a conformal measure μ, say the one based at the origin $o \in \mathbb{B}^{n+1}$. We construct a measure on each unstable leaf H^u_v by the prescription

$$d\mu^u(w) = e^{-\delta\rho_o(w,w_\infty)}\, d(p_u^*\mu)(w)$$

and analogously we define the measures $d\mu^s$ on stable leaves. It easily follows that the geodesic flow expands and contracts these measures with the wanted exponent, namely $g_t^*\mu^u = e^{\delta t}\mu^u$ and $g_t^*\mu^s = e^{-\delta t}\mu^s$. We define the measure m locally as

$$dm = d\mu^s \cdot d\mu^u \cdot dt \ .$$

It happens that the definition is invariant if we exchange the order of integration, that both $d\mu^u \cdot dt$ and $d\mu^u \cdot dt$ are holonomy invariant, and the measure dm is time as well Γ-invariant so that descends to the quotient SN_Γ. We get the expression

$$dm(v) = e^{-\delta\beta_o(v_\infty, v_{-\infty})}\, d\mu(v_{-\infty}) \cdot d\mu(v_\infty) \cdot dt \ .$$

As we can think at the unit tangent bundle of the hyperbolic space as $(S \times S - \text{diag}) \times \mathbb{R}$ and noting that the exponential of the Busemann cocycle, for the choice of the origin o, is proportional to $|v_\infty - v_{-\infty}|^{-2\delta}$ in the unit ball model, we recover the definition we have given at the beginning.

We note that by construction the conditionals of the Bowen-Margulis measure on stable and unstable horospheres are the measures $d\mu^s$ and $d\mu^u$ respectively.

2.3. *The density function of μ.* Now we discuss the metric structure of the geometric measure on the limit set, as presented in [Su4]. Let μ the Patterson-Sullivan measure on the limit set of the Kleinian group Γ. The aim is to have some control on the quantity

$$\mu(\xi, r) = \mu \text{ (ball of radius } r \text{ centered at } \xi)$$

for ξ in the limit set and the standard spherical metric on the boundary of the unit ball model.

For the class of convex cocompact groups (i.e. no cusps, and such that Γ acts cocompactly on the convex hull of the limit set), this is uniformly within a bounded ratio of r^δ, just like in the case of fundamental groups of compact hyperbolic manifolds where the geometric measure in the standard volume on the boundary sphere.

In presence of cusps the density function has oscillations which are due to the way a random geodesic wanders into the cuspidal neighborhoods. More precisely it is governed by the distribution of a certain point $v(\xi, r)$ inside the convex hull of the limit set.

Notation. Here and in the following the symbol $\asymp$ stands for "is within a bounded ratio of" for some constants that only depend on the group Γ, i.e. $f(r) \asymp g(r)$ means: there exists a positive K such that for any $r > 0$ we have $K^{-1} < f(r)/g(r) < K$. "Asymptotic equivalence" will be indicated by the symbol $\sim$, i.e. $f(r) \sim g(r)$ for r big means $f(r)/g(r) \to 1$ for $r \to \infty$.

We take as reference point the origin o in the unit ball, and make the assumption that it belongs to the convex hull. We define $v(\xi, r)$ to be the endpoint of the geodesic segment from the origin heading towards ξ and having length $t = \log 1/r$.

Construct an eigenfunction of the Laplace-Beltrami operator with eigenvalue $\delta(\delta - n)$, averaging the basic eigenfunctions with the measure μ. A basic eigenfunction is the Poisson kernel, or *height function*

$$h(x, \xi) = \text{height of } x \text{ when } \xi \text{ has been put at } \infty \text{ in the u.h.s. model}$$

i.e. $\exp(-\rho_o(x, \xi))$, and is normalized to have value one at $o \in \mathbb{H}$. It can also be described as $|\sigma'|(\xi)$ if σ is an hyperbolic isometry sending x to the origin. The eigenfunction is

$$\phi_\mu(x) = \int |\sigma'|^\delta d\mu \ .$$

By definition ϕ_μ is Γ-invariant. Sullivan proved an estimate for it when x is the point a hyperbolic distance t from the origin inside a rank k cuspidal neighborhood (i.e. x is at horospheric distance t from o with respect to a parabolic point of rank k in the limit set).

Theorem 2.3. *(Sullivan [Su4]) There exists a positive $\bar{t}$ such that:*

- if the point $x = v(\xi, r)$ is at horospheric distance $t > \bar{t}$ from the orbit $\Gamma \cdot o$ with respect to a rank-k cuspidal fixed point then $\phi_\mu(x)$ is within a bounded ratio of $e^{t(k-\delta)}$.

- if the point $x = v(\xi, r)$ is at horospheric distance less than $\bar{t}$ from the orbit $\Gamma \cdot o$ with respect to all cuspidal fixed points then $\phi_\mu(x)$ is within two positive constants.

The metric structure of μ is described by the following

Theorem 2.4. *(Sullivan [Su4]) The density function of the Patterson-Sullivan measure on the limit set of a geometrically finite Kleinian group satisfies*

$$\mu(\xi, r) \asymp r^\delta \cdot \phi_\mu(v(\xi, r))$$

uniformly for $\xi \in \Lambda$.

Thus $\mu(\xi, r)$ as function of the radius depends on the region where $v(\xi, r)$ actually lies. It goes as r^δ if $v(\xi, r)$ projects inside a fixed compact region of the convex hull quotient (the $\bar{t}$-ball around o in M_Γ), and as $r^{2\delta - k}$ if it projects inside a rank k cusp neighborhood. From the above we see that:

- balls around radial limit points are weighted essentially by r^δ, as geodesics heading toward such points reenters a compact region infinitely often by ergodicity of the geodesic flow (and indeed δ is the Hausdorff dimension of the radial limit set);
- balls centered at rank k parabolic fixed points are all weighted by $r^{2\delta-k}$, since any geodesic asymptotic to them eventually lies in a cusp neighborhood.

It is of importance to have a uniform estimate of the μ mass of balls around parabolic fixed points representing a cusp neighborhood. Indeed we have the simple corollary of the above theorem

Corollary 2.1. *Let b the resting points of the disjoint balls $\mathcal{B}$ as in section 1.3 in a rank k cusp orbit and s their sizes. For any smooth function $a : (0,1) \to (0,1)$*

$$\mu(b, a(s) \cdot s) \asymp a(s)^{2\delta - k} \cdot s^\delta$$

uniformly on those balls resting on a compact part of $\Lambda - \infty$, in particular on T.

Proof. We just observe that to the sizes of such balls correspond points at the boundary in the chosen compact part of the manifold referred in the Theorem 2.3, and shrinking the radius introduces a decay with exponent $2\delta - k$. More formally, given two distinct balls $B(b, s)$ and $B(b', s')$ we know that the ratio of their μ-masses is

$$\frac{\mu(b, s)}{\mu(b', s')} \asymp (s/s')^\delta$$

since the sizes represent points in (the boundary of) the chosen compact part. But we also know that if the function a takes values smaller than one

$$\frac{\mu(b, a(s)s)}{\mu(b, s)} \asymp (a(s)s/s)^{2\delta-k} = a(s)^{2\delta-k}$$

as we are inside the cusp neighborhood. Changing from the metric of the unit disk model to the metric of the u.h.s. model, as we are inside a compact part of the boundary $\mathbb{R}^n$, introduces only bounded ratio distorsions. The corollary follows by collecting these two facts together. □

Remark. Here we note that an uniform estimate as above holds if one consider the sizes as the diameters of the balls in the euclidean metric of the unit ball. Indeed, all of this discussion can be translated in this different language.

§3. Metric structure of the Patterson-Sullivan measure.

3.1. *Measures on metric spaces.* Let Λ be a subset of $\mathbb{R}^n$. The *Hausdorff s-dimensional measure* of Λ is defined taking coverings of Λ by balls of radii $r_i \leq \varepsilon$. It is the limit as $\varepsilon \to 0$ of the infimum over such coverings of the sums $\sum r_i^s$. If we only allow covering by balls centered at Λ itself we get the *covering measure* $\mathcal{C}^s(\Lambda)$ introduced by Tricot [Tr] (actually some technicalities are needed to make the measure countably subadditive).

In his investigation on limit sets of Kleinian groups Sullivan has been led to a dual construction based on packings rather than coverings [Su4] rediscovering Tricot's *packing measure* [Tr]. It is defined considering packings of disjoint balls centered at Λ and radii $r_i \leq \varepsilon$. The measure $\mathcal{P}^s(\Lambda)$ is the limit as $\varepsilon \to 0$ of the supremum over such packings of the sums $\sum r_i^s$ (for a rigorous definition we refer to the original papers or to the recent book by Mattila [Mat]).

The inequality $\mathcal{C}^s(\Lambda) \leq \mathcal{P}^s(\Lambda)$ holds. This inequality is generally sharp for fractal-like sets since equality forces s to be an integer and the set Λ to be rectifiable, as follows from combining the works of Preiss [Pr] and Saint Raymond and Tricot [SRT].

3.2. *Convex-cocompact Kleinian groups.* Now we consider the sphere S^n, and its group of conformal automorphisms $\mathbb{M}^n$. Let Γ be a discrete subgroup of $\mathbb{M}^n$ and recall the definition of conformal density given in section 2.1. The following is a tautology: *a Γ-invariant conformal density μ is unique if and only if the action of Γ is ergodic with respect to the measure class defined by μ* [Su1]. Uniqueness means that any two such conformal densities are proportional, their Radon-Nikodym derivative is a constant almost everywhere, and does not depends on the metric (in the fixed conformal class) with respect to which they are computed.

Let $\Gamma_0 \subset PSL(2,\mathbb{C})$ be a finitely generated discrete Kleinian group. It is called *structurally stable* if all sufficiently near representations into $PSL(2,\mathbb{C})$ are injective. It is called *non-rigid* if it admits arbitrarily close injective representations which are not conjugate in $PSL(2,\mathbb{C})$. Sullivan proved that the class of structurally stable non-rigid Kleinian groups coincides with the class of *geometrically finite* Kleinian groups without cusps. The latter is as well the class of *convex-cocompact* discrete subgroups of $PSL(2,\mathbb{C})$, Kleinian groups which acts cocompactly on the quotient hull of their limit sets.

Let Γ be a convex-cocompact Kleinian group and Λ its limit set inside the Riemann sphere. Call $\delta(\Gamma) = \delta$ the Hausdorff dimension of Λ. The Patterson-Sullivan measure μ is the unique conformal density of exponent δ supported on Λ and such that the action of Γ is ergodic with respect to it. It has be shown by Sullivan that the density of any representative of μ satisfies the uniform bounds

$$c \cdot r^\delta \;\leq\; \mu(\xi, r) \;\leq\; C \cdot r^\delta$$

for μ-almost all points $\xi \in \Lambda$, all sufficiently small r, and some constants c and C (this follows from Theorem 2.4). There follows from Besicovitch covering (and packing) theorem that μ is equivalent to both the covering and the packing measures of dimension δ restricted to Λ [Su4], in particular both these measures are finite.

We call *modulus* of the Kleinian group Γ the ratio between the packing and the covering measures of its limit set

$$\beta(\Gamma) \;=\; \frac{\mathcal{P}^\delta(\Lambda)}{\mathcal{C}^\delta(\Lambda)}$$

computed for example with respect to the standard spherical metric. It does not change under conjugacies in $PSL(2,\mathbb{C})$. As for its meaning, we have the following rigidity improvement of Sullivan's bounded ratio result (which, as follows from [SRT], implies that the Patterson-Sullivan measure does not have a density unless the group Γ is a lattice).

Theorem 3.1. *Let Γ be a convex-cocompact Kleinian group with critical exponent δ. Let μ be the Patterson-Sullivan measure (e.g. normalized) on the limit set Λ of Γ. Then for μ-almost all points $\xi \in \Lambda$ the following holds*

$$\frac{\limsup_{r\downarrow 0}\ \mu(\xi,r)/r^\delta}{\liminf_{r\downarrow 0}\ \mu(\xi,r)/r^\delta} = \beta(\Gamma)\ .$$

Proof. Denote by μ_c and μ_p the covering and the packing δ-dimensional measures restricted on the limit set. By the observations above they are locally finite measures supported on Λ, hence they are Radon measures, and they are representatives of two equivalent conformal densities (this follows from their definitions). By definition the Radon-Nikodym derivative $d\mu_c/d\mu_p$ at the point $\xi \in \Lambda$ is the limit of the ratio $\mu_c(\xi,r)\ /\ \mu_p(\xi,r)$ for r decreasing to zero, which exists μ_p-almost everywhere. Moreover, for a set of finite covering and packing measure $\liminf_{r\downarrow 0}\mu_p(\xi,r)/r^\delta = 1$ and $\limsup_{r\downarrow 0}\mu_c(\xi,r)/r^\delta = 1$ respectively μ_p and μ_c-almost everywhere ([SRT] Corollaries 7.1 and 7.2.). This implies that $\beta(\Gamma) = \limsup_{r\downarrow 0}\mu_p(\xi,r)/r^\delta$, and $\beta(\Gamma)^{-1} = \liminf_{r\downarrow 0}\mu_c(\xi,r)/r^\delta$. The theorem follows from uniqueness of the conformal density. □

3.3. *Kleinian groups with cusps.* Geometrically finite Kleinian groups with cusps are not convex-cocompact, and the density function $\mu(\xi,r)$ of the Patterson-Sullivan measure does not satisfies uniform bounds from above and below. From Theorem 2.4 and his estimates on the eigenfunction ϕ_μ Sullivan deduced that [Su4]:

- if all the cusps of Γ have ranks $k \geq \delta$ then there are positive constants c and C such that for almost any $\xi \in \Lambda$

$$c \leq \frac{\mu(\xi,r)}{r^\delta} \leq_{\text{i.o.}} C$$

- if all the cusps of Γ have ranks $k \leq \delta$ then there are positive constants c and C such that for almost any $\xi \in \Lambda$

$$c \leq_{\text{i.o.}} \frac{\mu(\xi,r)}{r^\delta} \leq C$$

where the notation "i.o." (infinitely often) means that the inequality is satisfied for a sequence of radii $r_i \to 0$ depending on the point ξ. There follows from Besicovitch covering theorem that the Patterson-Sullivan measure is equivalent to the packing measure in the first case, and to the covering measure in the second case. It there are cusps with rank $k > \delta$ (or with $k < \delta$) then the covering measure of the limit set is zero (respectively the packing measure is infinite). In particular, groups with cusps have $\beta(\Gamma) = \infty$, unless δ is an integer and all the cusps have ranks equal to δ (e.g. non-uniform lattices).

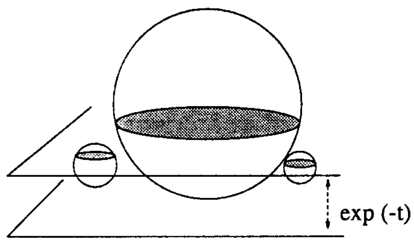
Remark. It seems interesting to study $\beta(\Gamma)$ as a function on a neighborhood of the variety of representations (up to conjugation) of a convex-cocompact Kleinian group Γ_0. By Sullivan [Su5] any such neighborhood is non-singular and consists of quasi-conformal conjugates of Γ_0. Compare the characterization of β in Theorem 3.1 with the maximal distorsion K of a quasi-conformal homeomorphism f between two metric spaces X and Y: for any point $x \in X$ the following holds

$$\limsup_{r\downarrow 0} \frac{\sup\{\text{dist}(fx,fy) \mid \text{dist}(x,y) = r\}}{\inf\{\text{dist}(fx,fy) \mid \text{dist}(x,y) = r\}} \leq K\ .$$

The ratio $\beta(\Gamma)$ is meaningful even in higher dimensions, for representations of an abstract group Γ_0 in the Möbius group $\mathbb{M}^n$ of the n-sphere. It does not change for the trivial inclusion $\Gamma \subset \mathbb{M}^n \subset \mathbb{M}^m$ with $m \geq n$ and under conjugation in $\mathbb{M}^m$. As follows from a recent paper by Yue [Yu] the set $\beta^{-1}(1)$ is somehow the set of "Fuchsian representations" (i.e. conjugated to lattices in some lower dimensional Möbius group) of Γ_0, and from the discussion above we see that $\beta^{-1}(\infty)$ is the set of non-Fuchsian representations with cusps. We plan to investigate on that in the future.

3.4. *Packings and coverings from the disjoint balls.* Let Γ be a geometrically finite Kleinian group with cusps, and P one of its cusps. As in section 1.3 we put the fixed point of P at ∞ in the u.h.s. model, and look at the collection of balls representing a cusp neighborhood $\mathcal{B} = \{B(b,s)\}$. The first thing we notice is the following. Imagine to cut these balls by means of an horizontal hyperplane at height e^{-t}, and observe the shadows projected on the boundary $\mathbb{R}^n$ by the vertical geodesics. We get a packing of the parallelepiped T by those balls which sizes s exceed e^{-t} of radii $\sqrt{se^{-t}}\sqrt{1-e^{-t}/s}$. Actually, to avoid balls of zero radius we could consider still those balls with $s > e^{-t}$ but cut them at half this height. The packing we get is be made of balls with radii (apart some constants)

$$e^{-t} < r < e^{-t/2} \ .$$



In the case of a non-uniform lattice, where the limit set is the entire ideal sphere (and the geometric measure projected to the boundary $\mathbb{R}^n$ is proportional to Lebesgue measure), we actually get a packing of the limit set. In general, we can only talk of a family of e^{-t}-separated points in the limit set. We record this observation in the following form

Packing lemma 3.1. *There exists a positive constant c such that the resting points b's of those balls of $\mathcal{B}$ with sizes $s > e^{-t}$ for all t sufficiently big form a set of $(c \cdot e^{-t})$-separated points inside the compact $\Lambda \cap T$.*

Less obvious is that with the same balls, just opportunely enlarging the radii but still keeping them decreasing to zero as t increases, one can construct a family of coverings.

In the theory of Diophantine approximation this is the content of a very classical theorem due, among others, to Dirichlet (a pigeon-hole proof, see [HW]): for any real number x and positive integer N there are couples of relatively prime integers p and q such that $p < N$ and

$$|px - q| < 1/N \ .$$

An analogous theorem by Swan holds for an imaginary quadratic number field, and says that any complex number can be approximated as above by a sequence of relatively prime integers. In such a case this amounts to saying that disjoint balls from one single cusp (the number of cusps is the class number of the field) suffices to produce a covering of the limit set. See section 4.3 for the connection between number theory and Kleinian groups.

Recall that for groups with finite volume quotients one has a well known *generalized Dirichlet theorem* which deals with the possibility of covering the sphere at infinity with the shadows of the horoballs of the cusps orbits [Pa1]. If we consider the limit set instead of the sphere at infinity the result still holds, as it only uses the compactness of a fundamental domain once the cusps neighborhoods are removed. This observation is due to Stratmann and Velani and implies certain rough estimates on some counting function [SV]. As we are going to improve the estimates in the following sections here we limit ourselves to state it (actually in an equivalent form, since the original result involves sizes measured in the euclidean metric of the unit ball model) for the simplest case of a group with one single cusp.

Covering lemma 3.2. *(Stratmann and Velani* [SV]*) There exist constants c and M such that for all sufficiently big times t the set $\Lambda - \{\infty\}$ is covered with multiplicity less than M by*

$$\bigcup B\left(b, c\sqrt{se^{-t}}\right)$$

where the balls are those with sizes $s > \mathrm{e}^{-t}$.

Sketch of the proof. We assume that Γ has only one parabolic cusp. By definition of geometric finiteness we know that the intersection of a fundamental domain D for Γ on $\mathbb{H}$ with the convex hull of the limit set, with the basic horoball from the cusp removed, $(\mathrm{Hull}(\Lambda) \cap D) - B$, is compact. Hence, if we enlarge sufficiently the basic horoball we can cover this set. Thus we can cover the convex hull of the limit set with the Γ-images of the enlarged basic ball. This is equivalent to the existence of a constant $c > 1$ such that

$$\mathrm{Hull}(\Lambda) \subset \bigcup_{B} B(b, cs) \ .$$

Place the cusp at ∞ in the upper-half space model. Consider the horizontal horosphere at height $c\mathrm{e}^{-t}$, which is a copy of the boundary $\mathbb{R}^n$, and observe that it intersects the family of enlarged balls in a collection of n-balls which (once projected down to the boundary) cover the limit set $\Lambda - \{\infty\}$ (indeed the trace left by the vertical geodesics from the points of Λ in the horosphere belongs to the convex hull). The sizes of the balls involved clearly verify $s > \mathrm{e}^{-t}$. Plain euclidean geometry says that the horizontal plane at height h in $\mathbb{R}^n \times \mathbb{R}^+$ cuts a ball of diameter d resting on the boundary $\mathbb{R}^n$ in an n-ball of radius $\sqrt{dh}\sqrt{1 - h/d} < \sqrt{dh}$. In particular, we get a covering of the limit set by the n-balls $B(b, c\sqrt{s\mathrm{e}^{-t}})$. The original balls are disjoint in $\mathbb{R}^n \times \mathbb{R}^+$, hence the multiplicity of the cover is finite uniformly on t. □

Remark. The case of more cusps is analogous and the covering is obtained from the images of all the cusps simultaneously. From the argument that the different cusps boundaries are a bounded distance apart in the convex hull quotient it can be seen that for any ball $B'(b', s')$ representing a cusp not equivalent to the one at infinity, there exist a ball $B(b, s)$ from the orbit of infinity s.t. the size s is within a bounded ratio of s', and b' belong to the shadow of $B(b, cs)$ for some constant c. This implies that any asymptotic result involving the cardinalities of the balls representing all the cusps of Γ also gives the same asymptotics when restricted to a single cusp orbit (see [SV] for details).

§4. Bounded ratio estimates

In this section we collect rough estimates on some natural counting functions which follow only from the existence of a geometric measure and the divergence of the Poincaré series. Most of the results of this section are implicitly contained in [Su3,4] and are essentially equivalent to the ones in [SV] once observed that the comparison amounts to the passage from the Poincaré series to an Eisenstein series relative to any of the cusps. The geometrically finite case is analogous to the better known situation of finite volume: the two series are comparable.

4.1. *Poincaré series divergence and the abundance of balls.* Sullivan observed that the divergence of the Poincaré series implies that for the orbit of any cusp, *the sum of the diameters of the disjoint horoballs (in the disk model) raised to the δ power is infinity.* This comes out from the fact that the points of any Γ-orbit are placed on these horospheres. Any such horosphere contains a lattice $\mathbb{Z}^k$ of points. Their contribution of each horosphere to the Poincaré series is comparable to the δ-power of the euclidean diameter in the unit ball model. But for a bounded region on the sphere at infinity these diameters are comparable to the ones computed in the u.h.s. model with the given cusp placed at infinity, hence for any measurable set Q in $\mathbb{R}^n$ of positive μ measure *the sum of the diameters of the balls resting at Q raised to the δ-power is infinity.* This observation, namely the appropriate abundance of balls of any size, is at the basis of his application of the Borel-Cantelli lemma to a geometrical proof of Khintchine approximation theorem and to the logarithm law for geodesic excursion.

Here we recall Sullivan's argument to prove the expected abundance of disjoint balls from comparison with the Poincaré series, since along the discussion we will prove a couple of useful lemmas.

We borrow the following three elementary lemmas from [NS], whose proof is a simple exercise in hyperbolic geometry:

Lemma 4.1. *If d' and d'' denote the hyperbolic lengths of two geodesic segments joining the boundary of an horoball B to an exterior point x, then they are within a bounded distance for some universal constant, namely*

$$|d' - d''| \leq \sinh^{-1}(1)/2 \ .$$

Lemma 4.2. *If d denotes the hyperbolic distance between two points lying on an horosphere ∂B, and d_ϵ is their euclidean distance computed along the horosphere, we have the exact formula*

$$d_\epsilon = 2\sinh(d/2)$$

which also says that d is within a bounded distance of $2\log d_\epsilon$ (it is indeed asymptotic to it).

Lemma 4.3. *Let x be outside the horoball B, y and y' be the endpoints of any two geodesic segments from x to ∂B. The euclidean distance between y and y' along the horosphere is less than 2.*

We want to split the Poincaré series

$$g_z(x,y) = \sum_{\gamma\in\Gamma} e^{-z(x,\gamma y)}$$

in such a way as to isolate the contribution of the portion of orbit placed along each horosphere in one cusp orbit.

As usual, we refer to the situation described in section 1.3 with the chosen cusp at infinity in the u.h.s. model and the parabolic subgroup P being isomorphic to $\mathbb{Z}^k$. We choose a fundamental parallelepiped T for the action of $\mathbb{Z}^k$ on the boundary $\Lambda \cap \mathbb{R}^n$, take the reference point x at height h somewhere above it, and the point y placed on the horizontal horosphere H and above T.

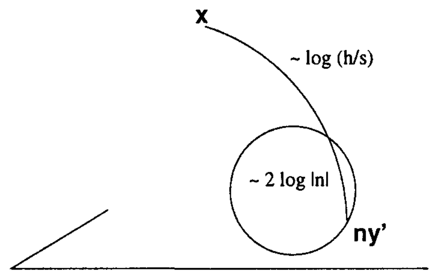
We limit ourselves to the balls resting on T. Consider one of the disjoint balls $\gamma B = B(b,s) \in \mathcal{B}$. Its boundary $\partial B(b,s)$ contains a full $\gamma P \gamma^{-1} \simeq \mathbb{Z}^k$ orbit of y. Moreover, the highest point of the ball, the one having height equal to the size, is at bounded distance from this orbit, say from a point $y' = \gamma y$. The orbit can be described by the collection ny', with $n \in \mathbb{Z}^k$ and the euclidean distance along $\partial B(b,s)$ between y' and ny' is proportional to $|n|$.

Lemma 4.4. *The contribution of the point ny' in the Poincaré series is within a bounded ratio of*

$$h^{-z} \cdot s^z \cdot |n|^{-2z} \ .$$

Proof. By the above lemmas 4.1-3 the hyperbolic distance between x and ny' is at bounded distance from the sum of the length of any geodesic segment joining x to $\partial B(b,s)$ and $2\log|n|$.

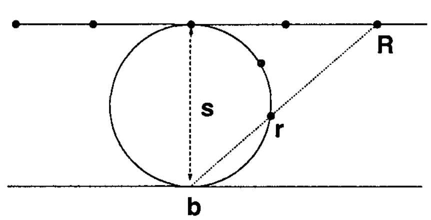
The above geodesic segment, by the same reasoning considering the horizontal horosphere through x and the fact that the balls we are considering are based at the compact T, is at bounded distance to the one representing the horospheric distance between x and $B(b,s)$ relative to ∞, and this is the logarithm of $h/(\text{size})$ (see the picture). $\square$



The same argument, plain euclidean lattice point counting (simply use Gauss formula on an horizontal horosphere and map it to an horosphere resting on the boundary of the u.h.s. model) also shows

Lemma 4.5. *Given any ball $B(b,s) \in \mathcal{B}$ of size s greater than e^{-t}, the cardinality of Γ images of y lying on its boundary and with heights greater than e^{-t} is (eventually) within a bounded ratio of $(e^t \cdot s)^{k/2}$.*

Proof. The inversion along the sphere centered at b with radius s maps the horosphere $\partial B(b,s)$ onto the horizontal horosphere H_s at height s. By definition, if r denotes the euclidean distance between a point on $\partial B(b,s)$ and the resting point b, and R the euclidean distance between its image on H_s and b, their product is constant: $r \cdot R = s^2$ (see the picture below). We assume the point y' is the heighest point of the horosphere $\partial B(b,s)$, which is true up to bounded distance. It is also the tangency point between the two horospheres. The points ny' are mapped by the inversion into points at euclidean distance n from y' lying over the horizontal horosphere. Gauss formula says that inside the ball of euclidean radius N and center y' on H_s there are $\asymp N^k$ points of the form ny' with $n \in \mathbb{Z}^k$. Projecting down to the horosphere $\partial B(b,s)$ we get N^k orbit points at height bigger than s/N^2 asymptotically for big N, as a simple computation in euclidean geometry shows. Calling $e^{-t} = s/N^s$ we are done. $\square$



Remark. May be it is worth to remark here that the proper abundance of balls is a corollary of these lemmas. Indeed the full Poincaré series is made of:

- a sum over balls resting on T

$$\Sigma = h^{-z} \sum_{\text{balls of } \mathcal{B}} s^z \sum_{\mathbb{Z}^k} |n|^{-2z} \ ,$$

- the $\mathbb{Z}^k$ shifts of these balls, which contribute like

$$\asymp h^k \sum_{\mathbb{Z}^k} |n|^{-2z} \cdot \Sigma \quad ,$$

- and an uninfluential sum due to the portion of the orbit on the horizontal horosphere.

For z going to δ the condition $k < 2\delta$ gives a convergent sum over the lattice $\mathbb{Z}^k$ (i.e. there is no atom deposited on parabolic fixed points in the weak limit giving the geometrical measure), so that for fixed h we see

Proposition 4.1. *The sum of the sizes of the disjoint balls $\mathcal{B}$ raised to the z power is comparable to the Poincaré series.*

In particular, the divergence of the Poincaré series is equivalent to the divergence of the sum of the sizes of those balls resting on T raised to the δ power.

Also, we can compare two Poincaré series with two different reference points, say with x' at height h times the height of x: *they are within a bounded ratio of $h^{k-\delta}$*, and this is indeed the estimate of the eigenfunction ϕ_μ referred in 2.3.

4.2. *Horoballs counting function.* In this section $\mathcal{B}$ is the collection of disjoint balls representing a fixed cusp at infinity, labeled by their sizes and resting points, sitting in the parallelepiped T. We recall that the whole family of balls is invariant by the parabolic subgroup and T is a fundamental domain for its action on $\Lambda - \infty$. Hence those balls of $\mathcal{B}$ are the images of the fundamental horoball by a choice of representatives of the two-sided coset $P\backslash\Gamma/P$.

We define the *horoballs counting function* as

$$\begin{aligned} N(t) &= \text{card}\left\{\text{balls of } \mathcal{B} \text{ s.t. } s > \mathrm{e}^{-t}\right\} \\ &= \text{card}\left\{\gamma \in P\backslash\Gamma/P \text{ s.t. } \text{size}(\gamma B) > \mathrm{e}^{-t}\right\} \quad . \end{aligned}$$

For an open set Q in the limit set, we will call $N(Q,t)$ the function which counts all the balls with sizes bigger than e^{-t} and resting points b belonging to Q.

Stratmann and Velani introduced a different notion of horoball counting function, measuring sizes in the unit ball model, and deriving estimates which are essentially equivalent to the ones to be obtained in this section [SV]. As we will see later our definition has the advantage to have a dynamical meaning in the contest of the Bowen-Margulis measure, and this will allow us to use mixing of the geodesic flow to sharpen the asymptotic behavior.

Suppose for simplicity that the one considered is the only cusp (otherwise see the remark at the end of section 3.2). The Covering lemma 3.2 says that for any time t sufficiently big we can cover with bounded multiplicity the region $\Lambda \cap T$ in the following way: take those balls with sizes $s > \mathrm{e}^{-t}$ and enlarge them forming the balls with diameters $c\sqrt{s\mathrm{e}^{-t}}$. Taking the μ-masses of these balls this says that

$$\mu(\Lambda \cap T) \; < \; \sum_{s > \mathrm{e}^{-t}} \mu(b, c\sqrt{s\mathrm{e}^{-t}}) \; < \; M\, \mu(\Lambda \cap T) \quad .$$

Using Sullivan's estimate for the μ-masses of such balls, i.e. taking in Corollary 2.1 the function $a(s) = \sqrt{\mathrm{e}^{-t}/s}$ (the presence of the constant c only gives uniformly bounded corrections), we get the existence of some constants c and C such that

$$c\,\mathrm{e}^{\delta t} \; < \; \sum_{s > \mathrm{e}^{-t}} (s\mathrm{e}^{t})^{k/2} \; < \; C\,\mathrm{e}^{\delta t} \quad .$$

Above, the terms inside the sum are comparable, by lemma 4.5, with the cardinalities of those points in the Γ-orbit with height bigger than e^{-t} sitting on each horosphere, hence

Theorem 4.1. *The cardinality of those Γ-images of y with height bigger than e^{-t} belonging to the balls resting on T (i.e. the images of y under the coset $P\backslash\Gamma$) is within a bounded ratio of $e^{\delta t}$.*

The point, now, is that $k/2$ is less than δ. Hence, once an orbit point belonging to one ball is counted for the first time, its further contributions to the above sum grow like $e^{(k/2)t}$ (remember that sizes are bounded from above), which is negligible compared to the $e^{\delta t}$ growth of the whole sum in Theorem 4.1. This is essentially the same argument which in the analytical theory of numbers shows that the two sums $\sum_{p\leq x}\log(p)$ and $\sum_n \sum_{p^n\leq x}\log(p)$, where p's are the primes, are asymptotically comparable. Hence

Theorem 4.2. *The horoball counting function $N(t)$ is within a bounded ratio of $e^{\delta t}$. The same holds for the counting function relative to any measurable open set Q of positive μ measure.*

Slightly changing the notation we can as well state this growth estimate as

Corollary 4.1. *There exists a constant $\lambda > 1$ such that for all sufficiently small $\varepsilon > 0$ the cardinality of those balls of $\mathcal{B}$ with sizes belonging to $(\varepsilon, \lambda\varepsilon)$ is within a bounded ratio of $\varepsilon^{-\delta}$.*

4.3. *Arithmetical examples.* The main example is the discrete group $PSL(2,\mathbb{Z})$, acting on the upper half plane $\mathbb{H}^2 = \{\mathrm{Im}(z) > 0\} \subset \mathbb{C}$ as fractional linear transformations. The quotient space $\mathbb{H}^2/PSL(2,\mathbb{Z})$ is called the *modular orbifold.* It has one cusp, the parabolic subgroup fixing the point at infinity and generated by the translation $z \mapsto z+1$. The images of the horizontal line at unit height bound a well known family of disjoint horoballs: the *Ford disks.* Resting points are the elements of the *Farey sequence*, placed inside the unit interval of the bounding real line (a fundamental domain for the parabolic subgroup). Sizes are the numbers $1/n^2$ for integers n, and the cardinality of balls having size $1/n^2$ is $\varphi(n)$, Euler totient function, which counts the cardinality of numbers $p < n$ relatively prime to n.

Mertens formula gives the asymptotic growth of the sum of the $\varphi(n)$ (see for example [HW]) and in the language of the counting function is

$$\mathrm{card}\left\{\text{balls of } \mathcal{B} \text{ with sizes } s > e^{-t}\right\} = \frac{3}{\pi^2}\, e^{t} + \mathcal{O}(t\, e^{t/2}) \ .$$

Above we recognize in the leading term the (inverse of the) volume of the unit tangent bundle of the modular orbifold. This formula also holds for those balls resting on any subinterval of $(0,1)$, and the constant in the leading term is multiplied by the length of the interval. This observation is the basis of the elementary proof given by Verjovsky of the vague convergence of probability measures supported on the closed horospheres to the Haar measure on the unit tangent bundle [Ve1]. An interesting observation by Zagier says that the Riemann hypothesis can be formulated as a statement about the error term in such asymptotic for mean values of smooth compact supported functions on the unit tangent bundle of the modular orbifold [Za1,2].

The collection of measures $e^{-t}\sum_{s>e^{-t}} \delta_b$ converges weakly to $3/\pi^2$ times Lebesgue measure on $(0,1)$. We could normalize the measures in the family and still obtain weak convergence to the probability Lebesgue measure. The error for mean values of characteristic functions (for good but generic sets) is estimated of order $t\, e^{-t/2}$. It is an immediate consequence of the classical Franel-Landau theorem that the Riemann hypothesis would imply an error $\sim e^{-(\frac{3}{4}-\varepsilon)t}$ for any strictly positive ε for mean values of Lipschitz test functions. See Verjovsky [Ve2] for a similar "if and only if" theorem about the Riemann hypothesis.

Here, the critical exponent of the Poincaré series is one, and indeed Mertens formula says that the counting function grows as $e^{\delta t}$. The connection between Mertens formula and the Riemann zeta function comes from the classical identity

$$\frac{\zeta(z-1)}{\zeta(z)} = \sum \frac{\varphi(n)}{n^z}$$

where $\zeta(z) = \sum_{n\geq 1} 1/n^z$ is the Riemann zeta function, and the right hand side is nothing but

$$\sum_{\text{balls}} (\text{size})^{z/2} \ .$$

For z equal to two we see on the right Sullivan's result (see Proposition 4.1) on the divergence of the sum of the sizes of the balls raised to the δ power, and on the left the pole of the Riemann zeta function at one, divided by its value in 2, related to the volume of the orbifold.

It would be nice for a generic Kleinian group with cusps to find, if any, the "qualitative" analogue of a couple of theorems by Landau in number theory. Euler totient function $\varphi(n)$ counts the cardinality of disjoint balls representing the only cusp of the modular orbifold of size $1/n^2$. The infinitude of prime numbers says that

$$\limsup \frac{\varphi(n)}{n} = 1$$

and Landau proved a result about the lim inf, namely

$$\liminf \frac{\varphi(n)}{n} \log\log n = \mathrm{e}^{-C}$$

where C is the Euler constant. We are not able to estimate something as fine as $\varphi(n)$, but only a sufficiently large sum of its values as in corollary 4.1.

In dimension three we have similar examples taking the groups of two by two complex matrices of determinant one with entries the integers of an imaginary quadratic number field. They are fundamental groups of finite volume hyperbolic orbifolds with cusps. For the cusp at infinity, corresponding to the lattice of algebraic integers, the analogue of Mertens formula reads

$$\text{card}\,\{\text{balls with sizes } \ s > \mathrm{e}^{-t}\} = \frac{\pi}{|d|^{1/2}\zeta_k(2)} \cdot \mathrm{e}^{2t} + \mathcal{O}(\mathrm{e}^{\frac{3}{2}t})$$

where d is the discriminant of the field, and ζ_k its Dedekind zeta function. Again it is a particular case of a more general statement: the sum of Dirac masses placed at the resting points of the horospheres with height bigger than e^{-t}, normalized by e^{-2t}, converges weakly to the Lebesgue measure on the boundary real plane (the image of the geometric measure under the stereographic projection). Even in this case the relative Riemann-hypothesis is related to the speed at which the measures supported on the closed horospheres from the cusps become asymptotically equidistributed (see section 6).

4.4. *The Poincaré series and the eigenfunction ϕ_μ.* As already noted, the same asymptotic estimates for the horoball counting function can be obtained measuring sizes in the unit ball model and as well measuring the distance of each horoball from a fixed point, e.g. the origin of the unit ball model for $\mathbb{H}$. In particular there follows that the same *radial* counting function $n(R)$ is within a bounded ratio of $\mathrm{e}^{-\delta R}$, and the critical exponent is indeed a limit and not only a lim sup.

Also, one could regard the Poincaré series $g_z(x,y)$ and estimate its dependence on z near the critical value. In other word by similar estimates we can prove, even in this generic case

Proposition 4.2. *(Stratmann and Velani [SV]) The counting function $n(R)$ of a geometrically finite Kleinian group is within a bounded ratio of $\mathrm{e}^{-\delta R}$.*

and its immediate corollary

Proposition 4.3. *There are strictly positive functions $c(x,y)$ and $C(x,y)$ of two variables in the unit ball such that the Poincaré series above the critical exponent verifies*

$$\frac{c(x,y)}{z-\delta} \le g_z(x,y) \le \frac{C(x,y)}{z-\delta} \ .$$

This kind of inequality was obtained by Patterson by analytic means in the case $\delta > n/2$ [Pa2] and by Sullivan in the case of convex-cocompact groups [Su1]. It roughly says that for fixed references points the Poincaré series behaves as a function having a pole at the critical exponent.

Remember that $\phi_\mu(x)$ was obtained averaging the Poisson kernel $P(x,\xi)$ with the geometric measure μ on the sphere. The Poisson kernel can be viewed as the dilatation of an isometry sending x to the origin, as well as the exponential of the horospheric distance between the origin and x w.r.t. ξ. Thus

$$\phi_\mu(x) = \int e^{\rho_o(x,\xi)} \, d\mu(\xi) \ .$$

Reconsidering μ as a weak limit of Dirac masses for some sequence $z_j \to \delta$, as in [Pa2] we can see that

$$\phi_\mu(x) = \lim_{z_j \to \delta} \frac{g_{z_j}(x,y)}{g_{z_j}(o,y)}$$

for some fixed y. By the above Proposition 4.3 we get the result that, apart for a constant, we can recover the eigenfunction as

$$\phi_\mu(x) = \lim_{z_j \to \delta} (z_j - \delta)\, g_{z_j}(x,y) \ .$$

Also, from the expression of ϕ_μ as a limit of the ratio of two Poincaré series we recover the estimate of the eigenfunction near a cuspidal end: if x is at height h w.r.t. a rank k cusp placed at ∞, $\phi_\mu(x)$ is within a bounded ratio of $h^{k-\delta}$.

It has been observed by Sullivan that ϕ_μ is always square integrable in a unit neighborhood of the convex hull of the limit set modulo Γ (with respect to the smooth hyperbolic volume measure), but a sufficient condition for this to hold in the whole infinite volume quotient manifold is indeed $\delta > n/2$. This explains the reason Patterson's analytic argument to prove the propositions above does not hold in the most general case.

4.5. *Eisenstein series.* We already know the relation between the orbital counting function and the Poincarè series. If $n(R)$ counts the cardinality of the orbit Γy inside a ball of radius R around x, in the series any orbit point contributes as the exponential of minus z times its distance from x. The limsup of $\log n(R)$ over R is also the critical exponent of the series. An horoball is a limit of an hyperbolic ball which center approaches an ideal point in the boundary sphere. Both the counting function and the series generalize in a natural way if the center converges towards a parabolic fixed point of the group.

Put as usual the chosen cusp at ∞ in the u.h.s. model. Horospheres centered at it are parametrized by the Busemann function, e.g. normalized to be zero at the origin, which exponential is the height function. As the parabolic subgroup would give a divergent number of orbit points for any height, we are forced to define the counting function as

$$M(t) = \operatorname{card}\left\{\gamma \in P\backslash\Gamma \text{ s.t. } h(\gamma x) > e^{-t}\right\}$$

where h denotes the height function $h(\cdot) = h(\cdot,\infty)$, and x is a point in $\mathbb{H}$. It counts orbit points higher than e^{-t} with shadow in T. Here we note that the previously defined horoballs counting function roughly counts points of the two-sided cosets $P\backslash\Gamma/P$, and they are comparable thanks to the condition $\operatorname{rank}(P) = k < 2\delta$.

The associated series is

$$E(x;z) = \sum_{P\backslash\Gamma} h(\gamma x)^z .$$

Note that if we naively take the Poincaré series $g_z(x', x)$ and try to send $x' \to \infty$ we realize that its terms are asymptotic to $h(\gamma x)/h(x')$ to the power z, hence, to get something not trivial, we must renormalize forgetting the divergent denominator $h(x')$ and counting only once any orbit of the cuspidal group.

$E(x;z)$ is called an *Eisenstein series* relative to the cusp. It is made summing the images of the height function, already invariant by any parabolic transformation fixing ∞, by the left cosets $P\backslash\Gamma$. The result is an automorphic function on the upper half-space. Moreover, as $h(x)^z$ is an eigenfunction of the Laplace-Beltrami operator on $\mathbb{H}^{n+1}$ with eigenvalue $-z(n-z)$, what we get is a eigenfunction on the quotient manifold $\mathbb{H}/\Gamma$, though not square integrable.

In Theorem 4.1 we actually proved an estimate on $M(t)$, it is within a bounded ratio of $e^{\delta t}$, and so of the terms in the Eisenstein series. This implies

Proposition 4.4. *The Eisenstein series of a geometrically finite Kleinian group with respect to any of its cusps is comparable to the Poincaré series.*

In particular, it has the same critical exponent of the Poincaré series and diverges like $(z-\delta)^{-1}$ as in Proposition 4.3, i.e. it has a simple pole-like divergence.

Remark. Actually, we can mimic most of the elementary classical estimates of the Eisenstein series of a non-uniform lattice of $PSL(2,\mathbb{R})$ (a classical reference is the book by Kubota [Ku]). The very same argument giving the asymptotic of two Poincaré series near a cuspidal end, and hence the behavior of the eigenfunction ϕ_μ, gives

Proposition 4.5. *Let $E(x,z)$ be the Eisenstein series of a geometrically finite Kleinian group with respect to a rank k cusp. For fixed z above the critical exponent and $h(x)$ big enough*

$$E(x;z) = h(x)^z + \mathcal{O}\left(h(x)^{k-z}\right) .$$

§5. Equidistribution of parabolic points

In this section we derive finer estimates on the horoballs counting function using the strong ergodic properties of the geodesic flow. In particular, we will make use of the results by Rudolph, who proved that the flow is mixing, and even Bernoulli, with respect to the Bowen-Margulis measure [Ru]. We use the notation introduced in section 1.3 for the disjoint balls representing a cusp P of the Kleinian group Γ.

5.1. *Replacing an orbit by parabolic points in the boundary.* In the Patterson construction the limit measure is obtained from the family of measures μ_z (here the reference point has been fixed to be the origin o of $\mathbb{B}^{n+1}$)

$$\mu_z = \frac{1}{g_z(y,y)} \sum_{\Gamma} e^{-z(o,\gamma y)} \delta_{\gamma y} .$$

We want to compare this family of measures with the family obtained replacing the Dirac masses belonging to each one of the disjoint horospheres by Dirac masses on their basis b's. We choose the point y belonging to the horizontal horosphere $H = \partial B$. The points of the orbit $\Gamma \cdot y$ are placed along the horospheres $\Gamma \cdot H$. We can split the sum over Γ as we have done in section 4.1. For any γH we pick an image $y' = \gamma y \in \gamma H$ in such a way that y' is the nearest point to the highest point of the sphere between all the $\gamma P \gamma^{-1} \cdot y \simeq \mathbb{Z}^k \cdot y'$ (the nearest point may not be unique, we just take one of them).

We write, with abuse of notation,

$$\mu_z = \frac{1}{g_z(y,y)} \sum_{\text{balls}} \left(\sum_{\mathbb{Z}^k} e^{-z(o,ny')} \delta_{ny'} \right)$$

and we are going to compare it with

$$\tilde{\mu}_z = \frac{1}{g_z(y,y)} \sum_{\text{balls}} \left(\sum_{\mathbb{Z}^k} e^{-z(o,ny')} \right) \delta_b \ .$$

We will show that the weak limits of μ_{z_i} and $\tilde{\mu}_{z_i}$ agree if $z_i \to \delta$ is a sequence such that the μ_{z_i} converge. In particular, put $y = o$ in the Poincaré series defining the measures μ_z and $\tilde{\mu}_z$ above.

Proposition 5.1. *The geometric probability measure is the weak limit of $\tilde{\mu}_z$ as $z \to \delta$ from above.*

Proof. Let f be a Hölder function on the sphere at infinity of Hölder exponent α. We assume that the support of f, when seen in the boundary of the u.h.s. model, is contained in a compact portion of $\mathbb{R}^n$ (this is always the case for any f which is zero in an open subset of the limit set, since we can take the u.h.s. model with the infinity corresponding to a parabolic point inside this subset). For example, we can assume that the support of f lies inside the parallelepiped T. We extend f to a Hölder function on a portion of the u.h.s., say the one below the horizontal horosphere at unit height, putting constant value on the vertical geodesics.

We need an elementary computation in the u.h.s. model, that we illustrate in dimension 2, where $\mathbb{H}^2 = \{z \in \mathbb{C} \text{ s.t. } \mathrm{Im}(z) > 0\}$. The isometry $z \mapsto -1/z$ maps the horizontal horosphere at unit height to the sphere resting on the origin and diameter 1. The point $i + n$ is sent to the point $(i - n)/(n^2 + 1)$, thus the height (the imaginary part) is asymptotic to $1/n^2$ (for big n) and the distance from the vertical geodesic passing through the base of the ball (the real part) is asymptotic to $1/n$. In general, referring to the notation we have introduced above, if ny' is the orbit point on the horosphere of size s and based at b, its height is within a bounded ratio of

$$\text{height}(ny') \asymp \frac{s}{n^2}$$

and the euclidean distance with the vertical line ℓ_b through b is within a bounded ratio of

$$\text{dist}(ny', \ell_b) \asymp \frac{s}{n} \ .$$

In particular, the exponential of minus the hyperbolic distance (o, ny') is within a bounded ratio ofs/n^2 (this is the estimate in section 4.1) uniformly on balls resting on a compact part of $\mathbb{R}^n$.

Taking the mean value of the test function f extended as above we get the comparison (we treat the constants of the bounded ratio estimates as one)

$$\begin{aligned}
|\mu_z(f) - \tilde{\mu}_z(f)| &\leq \frac{1}{g_z(o,o)} \sum_{\text{balls}} \sum_{\mathbb{Z}^k} e^{-z(o,ny')} |f(ny') - f(b)| \\
&\leq \frac{1}{g_z(o,o)} \sum_{\text{balls}} \sum_{\mathbb{Z}^k} \left(\frac{s}{n^2}\right)^z \|f\|_\alpha \left(\frac{s}{n}\right)^\alpha \\
&\leq \|f\|_\alpha \frac{1}{g_z(o,o)} \sum_{\text{balls}} s^{z+\alpha} \left(\sum_{\mathbb{Z}^k} \frac{1}{|n|^{2z+\alpha}} \right) \\
&\leq \|f\|_\alpha \frac{1}{g_z(o,o)} \sum_{\text{balls}} s^{z+\alpha} \ .
\end{aligned}$$

We used again the fact that the sums over $\mathbb{Z}^k$ are finite whenever $z \geq \delta$ by the condition $2\delta > k = \mathrm{rank}(P)$. For z converging to δ from above and α positive, the Poincaré series $g_z(o,o)$ is diverging, since it is bigger than some constant divided $z - \delta$ (see Proposition 4.2), while the sum over the balls is convergent since we are above the critical exponent. Hence

$$|\mu_z(f) - \tilde{\mu}_z(f)| \rightarrow 0$$

and z goes to δ. For $z > \delta$ the collection μ_z is a family of probability measures supported on the limit set, tight, and such that any weak limit for $z_i \rightarrow \delta$ is the probability geometric measure μ on Λ. □

This shows that the geometric measure can also be derived as a weak limit of measures supported on the ideal boundary, formed by Dirac masses on the orbit of a parabolic fixed point, each one weighted with the partial Poincaré series restricted to the ball. It is actually a kind of equidistribution result for parabolic fixed points, paraphrasing the theory of series we could say "in the sense of Abel".

5.2. *Equidistribution of parabolic fixed points.* We construct a family of (infinite) measures on $\Lambda - \infty$ by placing Dirac masses on the base points of those balls with height bigger than e^{-t} and normalizing with the appropriate power of this height, namely

$$\nu_t = \mathrm{e}^{-\delta t} \sum_{s > \mathrm{e}^{-t}} \delta_b .$$

The ν_t-mass of any compact subset of $\mathbb{R}^n$ is bounded from below and from above uniformly in t by the results in 4.2. As the family of balls is invariant by the cuspidal subgroup $P \simeq \mathbb{Z}^k$, we can regard ν_t as a family of (uniformly) finite measures on the cylinder $T^k \times \mathbb{R}^{n-k} = \mathbb{R}^n/\mathbb{Z}^k$, supported on the compact part coming from the fundamental parallelepiped T. Hence, the family converges weakly by subsequences as time goes to infinity. The natural guess is that they converge to the image of the geometrical measure on the sphere under stereographic projection and with a convenient scaling factor. In other words, the limit should play a rôle analogue to the rôle of Lebesgue measure in the boundary $\mathbb{R}^n$ in the finite volume case.

We identify the fixed point ∞ of the cusp P with the north-pole of the unit sphere $S^n \simeq S$. Let $p_\infty : S - \infty \rightarrow \mathbb{R}^n$ denote the inversion along the sphere centered at ∞ and with radius $\sqrt{2}$ (unit ball model). It is the variation of the stereographic projection described in 1.1. The product of the (euclidean) distances of a point $\xi \neq \infty$ on the sphere and its image $p_\infty(\xi)$ from ∞ is constant. Indeed p_∞ is a conformal map, and the conformal factor is a constant times $|\infty - \xi|^2$ (euclidean norm in the disk model). The standard volume on the unit n-sphere is the pullback of the euclidean volume on $\mathbb{R}^n$ times the Radon-Nikodym ratio $|\infty - \xi|^{2n}$.

In our case the Hausdorff dimension of the limit set is δ, hence the natural "geometric measure" on $\Lambda - \infty \subset S - \infty \simeq \mathbb{R}^n$ is

$$d\mu^\infty(\xi) = |\infty - \xi|^{-2\delta} \, d(p_\infty^* \mu)(\xi) .$$

This measure is finite on compact sets and δ-conformal.

We note that if we identify $S - \infty$ with an unstable horosphere H_x centered at ∞ for some point $x \in \mathbb{H}$ by means of the forward visual map, $d\mu^\infty$ is proportional to $d\mu^u$, the measure used in the construction of the Bowen-Margulis measure (see 2.2). We expect the vague limit of ν_t for t going to infinity to be proportional to μ^∞.

The other way around, we can pull back our family ν_t on the sphere and put the required Radon-Nikodym ratio (we add to the family a Dirac mass at infinity to have positive mass for any time t)

$$\mu_t = \delta_\infty + \mathrm{e}^{-\delta t} \sum_{s > \mathrm{e}^{-t}} |\infty - b|^{2\delta} \, \delta_b .$$

A weak limit of such a family of measures, if it exists, is geometric (δ-conformal and Γ-invariant) almost by definition (certainly any limit is in the same measure class by the uniform bounds on the horoball counting function in 4.2). To see this take any $\gamma \in \Gamma$ and try to measure a sufficiently small open Q on the sphere, let's say not containing ∞ and such that γQ also does not contain ∞, so small that the dilatation coefficient (in spherical metric) between Q and γQ is essentially a constant $|\gamma'|$. Any ball $B(b,s)$ resting on Q has an image $B(b',s') = \gamma B(b,s)$ resting on γQ. Asymptotically for small sizes s' is equal to s times the linear distorsion in euclidean $\mathbb{R}^n$, which is $r = |\gamma'||\infty - b|^2/|\infty - b'|^2$ (Leibnitz rule). If we treat the ratio above as constant too we get

$$
\begin{aligned}
\gamma^*\mu_t(Q) &= \mu_t(\gamma Q) \\
&= e^{-\delta t} \sum_{s'>e^{-t}} |\infty - b'|^{2\delta}\delta_{b'}(\gamma Q) \\
&\simeq e^{-\delta t} \sum_{rs>e^{-t}} |\infty - b'|^{2\delta}\delta_{b}(Q) \\
&\simeq e^{-\delta t} \sum_{s>e^{-(t+\log r)}} \frac{|\infty - b'|^{2\delta}}{|\infty - b|^{2\delta}} |\infty - b|^{2\delta}\delta_{b}(Q) \\
&\simeq e^{-\delta t}\, |\gamma'|^{\delta}\, r^{-\delta} \sum_{s>e^{-(t+\log r)}} |\infty - b|^{2\delta}\delta_{b}(Q) \\
&\simeq e^{-\delta(t+\log r)} |\gamma'|^{\delta} \sum_{s>e^{-(t+\log r)}} |\infty - b|^{2\delta}\delta_{b}(Q) \\
&\simeq |\gamma'|^{\delta}\, \mu_{t+\log r}(Q)
\end{aligned}
$$

and the error factor goes uniformly to one for small sizes, which are the only ones that count as t goes to infinity. Thus, if the whole family μ_t had a weak limit by Sullivan uniqueness theorem this would be proportional the geometric measure μ.

We have tried for some time to prove existence of the above weak limit only using the information we have collected so far and disjointness of the balls (the argument which gives the quasi-independence of the shadows of the disjoint balls and enables Sullivan to use a mild version of Borel-Cantelli lemma). As we will see in the next section this is a consequence of the mixing property of the geodesic flow. Namely, we postpone to the following section 5.5 the proof of

Theorem 5.1. *The measures μ_t converge weakly to (a constant times) the geometric measure μ on the limit set. The measures ν_t converge vaguely to (a constant times) μ^∞ on the limit set minus the reference parabolic point.*

Note that the two assertions are equivalent. Also, if we consider the balls whose resting points describe a $P\backslash\Gamma/P$-orbit of the parabolic fixed point ∞ and count those with height bigger than e^{-t}, the above theorem implies the following sharpening of the rough results in section 4.

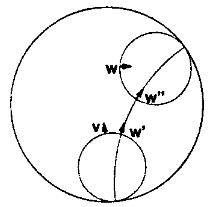
Theorem 5.2. *There exits a constant C such that asymptotically*

$$\textit{card}\left\{\textit{balls of } \mathcal{B} \textit{ s.t. } \ s > e^{-t}\right\} \sim C \cdot e^{\delta t} \ .$$

Remark. In view of the discussion in section 4.3 we refer to the above theorem as a *generalized Mertens formula.* The collections of parabolic fixed points entering in the definition of the measures ν_t, restricted to a fundamental domain for the cuspidal subgroup on the limit set, form what we may call a *generalized Farey sequence* and their asymptotic equidistribution is the content of the sharpening of Mertens formula quoted in 4.3.

5.3. *Cells.* We start recalling a basic tool in the theory of Anosov flows, *cells*, i.e. foliated charts for the stable (or unstable) foliation.

Fixed a vector $v \in S\mathbb{H}$, any unit vector $w \in S\mathbb{H}$ is uniquely obtained from the following series of movements. Start from v and move along the unstable leaf to get a vector $w' \in H^u_v$. Then act with the geodesic flow to get a vector $w'' = g_t w'$ belonging to the weakly unstable leaf through v. Finally move along the stable leaf through w'' to get w.



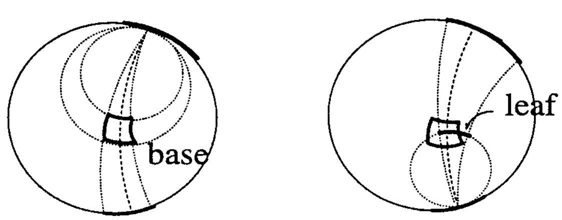
Now we define a stable cell. Let v a vector in the unit tangent bundle of $\mathbb{H}$, and D a neighborhood of v in the unstable horosphere through v, $v \in D \subset H^u_v$. We also think at D as the inverse image of a domain $D_\infty \subset S - \{v_{-\infty}\}$ containing v_∞ under the forward visual map $p_u : H^u_v \to S - \{v_{-\infty}\}$. Then act on D with the geodesic flow for a time up to r, i.e. set

$$B = \{g_t w' \text{ for } w' \in D \text{ and } t \in (0,r)\} .$$

This will be called the *base* of the cell, and is a domain in the weakly unstable leaf through v (i.e. for any $w'' \in B$, $w''_{-\infty} = v_{-\infty}$). Finally, given a domain $D_{-\infty} \subset S - \{v_\infty\}$ containing the point $v_{-\infty}$, we attach to any vector w'' of the base B a piece of stable leaf, the inverse image of $D_{-\infty}$ under the backward visual map $p_{s,w''} : H^s_{w''} \to S - \{v_\infty\}$. The cell is

$$\Delta^s_v = \Delta^s_v(D_\infty, (0,r), D_{-\infty}) = \bigcup_{w'' \in B} p^{-1}_{s,w''}(D_{-\infty}) .$$

It is the union of portions of stable leaves through the points w'' in the base.



Analogously we define a cell Δ^u_v which is a foliated chart for the unstable foliation: attaching pieces of unstable leaves to points in the base, a domain in the weakly stable leaf through v. The discussion below has an analogue for unstable cells.

The following lemma is a consequence of the fact that the group Γ acts discretely on $\mathbb{H}$ (we assume it contains no elliptic elements).

Lemma 5.1. *For any vector $v \in S\mathbb{H}$, sufficiently small cells around v project bijectively to the quotient $S(\mathbb{H}/\Gamma)$.*

The "sufficiently small" above is uniform as long as v is based in the thick part of the manifold.

We explain how to integrate functions on the cells with respect to the Bowen-Margulis measure. The Bowen-Margulis measure is defined locally as (see section 2.2)

$$dm = d\mu^s \cdot d\mu^u \cdot dt .$$

(with this notation we mean that we must start integrating from the left). Using the fact that $g_t^* d\mu^u = e^{\delta t} d\mu^u$ we exchange the order of integration and obtain the local expression

$$\begin{aligned} dm(w) &= d\mu^s_{w''}(w) \cdot e^{\delta t} dt \cdot d\mu^u_v(w') \\ &= e^{-\delta \rho_o(w, w_{-\infty})} d(p^*_{s,w''}\mu)(w) \cdot e^{\delta t} dt \cdot d\mu^u_v(w') \ . \end{aligned}$$

We describe a construction which will be used in the following section. Let Δ^s_v a stable cell as above. Let h be a function on the unstable horosphere through v supported on D which is integrable with respect to the measure $d\mu^u$. Let h' be a function on the real line with support inside $(0, r)$ and such that

$$\int_{\mathbb{R}} h'(t)\, e^{\delta t} dt \ = \ \int_0^r h'(t)\, e^{\delta t} dt \ = \ 1 \ .$$

Let h'' a function on the ideal sphere supported on $D_{-\infty}$ and such that

$$\int_S h''\, d\mu \ = \ \int_{D_{-\infty}} h''\, d\mu \ = \ 1 \ .$$

We construct a function h''' on $S\mathbb{H}$ supported on the cell Δ^s_v multiplying h, h', h'' and the factor

$$\alpha(w) \ = \ e^{\delta \rho_o(w, w_{-\infty})}$$

in the following way. If the vector $w \in \Delta^s_v$ is obtained from v from the points $w' \in H^u_v$, $w'' = g_t w'$ then

$$h'''(w) \ = \ \alpha(w) \cdot (h'' \circ p_{s,w''})(w) \cdot h'(t) \cdot h(w') \ .$$

The function h''' is defined to be zero outside the cell.

Lemma 5.2. *In the notation above the following holds*

$$\int_{S\mathbb{H}} h'''\, dm \ = \ \int_{H^u_v} h\, d\mu^u \ .$$

Proof. The proof consists in the following computation, using the local definition of the Bowen-Margulis given above.

$$\begin{aligned} \int_{S\mathbb{H}} h'''(w)\, dm(w) &= \int_D \int_0^r \left(\int_{H^s_{w''}} \alpha(w)\, e^{-\delta \rho_o(w, w_{-\infty})}\, (h'' \circ p_{s,w''})(w)\, d(p^*_{s,w''}\mu)(w) \right) \cdot \\ &\quad \cdot h'(t)\, e^{\delta t} dt \cdot h(w')\, d\mu^u(w') \\ &= \int_D \int_0^r \left(\int_S h''\, d\mu \right) h'(t)\, e^{\delta t} dt\, h(w')\, d\mu^u(w') \\ &= \int_D \left(\int_0^r h'(t)\, e^{\delta t} dt \right) h(w')\, d\mu^u(w') \\ &= \int_D h(w')\, d\mu^u(w') \ . \end{aligned}$$

□

5.4. *Asymptotic equidistribution of horospheres.* The following argument is not new, it dates back to Margulis' pioneering thesis in the late sixties. It has been used by Bowen and by Marcus to get unique ergodicity results for horospheric flows and foliations in axiom-A dynamical systems, or more recently by Eskin and McMullen in a finite volume hyperbolic manifold situation [EM], and by Kleinbock and Margulis in a more general lattice in semisimple Lie groups setting [KM]. Our point is that the reasoning in Proposition 2.2.1. of the last reference can be adapted to the infinite volume case, and provides an asymptotic equidistribution result for any unstable leaf when expanded by the geodesic flow. The ingredient required is the mixing property of the geodesic flow, and it is provided by the work by Rudolph, who actually proved that the geodesic flow is Bernoulli w.r.t. the Bowen-Margulis measure on the nonwandering set [Ru].

Fix a point $v \in SN_\Gamma$ and take any of its lifts in the covering, still denoted by $v \in S\mathbb{H}$. Then consider the unstable horosphere H_v^u through v, equipped with the measure $d\mu^u$, and a compact support function $h \in L^2(H_v^u, d\mu^u)$. Since h has compact support, we can decompose it in a finite sum of functions whose supports inject on the quotient SN_Γ, so we just assume that $\mathrm{supp}(h)$ injects on SN_Γ. Also, there is no loss of generality in assuming that the support of h contains the initial point v.

Consider a function $f \in L^2(SN_\Gamma, dm)$, and define the collection of functions $f_t = f \circ g_t \circ i$ on H_v^u, where i is the injection $\mathrm{supp}(h) \hookrightarrow SM$, and g_t is the geodesic flow on the unit tangent bundle of the manifold. If f is continuous, the integral

$$\int_{H_v^u} h\ f_t\ d\mu^u$$

is well defined for any t. In the following we identify the function f with its pullback $f \circ \pi^{-1}$ on $S\mathbb{H}$, where π is the projection $\pi : S\mathbb{H} \to SN_\Gamma$ (i.e. f is a Γ-invariant function on $S\mathbb{H}$). If we define the functions $\hat{h} = h \circ p_{u,v}$ and $\hat{f}_t = f_t \circ p_{u,v}$ on $S - v_{-\infty}$ (indeed the support of $\hat{h}$ does not contain $v_{-\infty}$), we get the expression

$$\begin{aligned}\int_{H_v^u} h\ f_t\ d\mu^u &= \int_{H_v^u} h(w')\ f(g_t w')\ \mathrm{e}^{-\delta\rho_o(w',w'_\infty)}\ d(p^*_{u,v}\mu)(w') \\ &= \int_{S-v_{-\infty}} \hat{h}(w'_\infty)\ \hat{f}_t(w'_\infty)\ \mathrm{e}^{-\delta\rho_o(w',w'_\infty)}\ d\mu(w'_\infty)\end{aligned}$$

where $p_{u,v}$ is the forward visual map from H_v^u onto $S - \{v_{-\infty}\}$, and w' is the intersection point between the geodesic from $v_{-\infty}$ and w_∞, and the horosphere H_v^u. The following lemma is a straightforward consequence of the fact that $g_t^* d\mu^u = \mathrm{e}^{\delta t} d\mu^u$.

Lemma 5.3. *In the notation above the following holds*

$$\int_{H_v^u} h\ f_t\ d\mu^u\ =\ \mathrm{e}^{-\delta t} \int_{H_{g_t v}^u} (h \circ g_{-t})\ f\ d\mu^u\ .$$

We are ready to prove

Theorem 5.3. *For any point $v \in SN_\Gamma$, any function $h \in L^2(H_v^u, d\mu^u)$ with compact support and any $f \in L^2(SN_\Gamma, dm)$ uniformly continuous, as t tends to infinity*

$$\int_{H_v^u} h(w)\ f(g_t w)\ d\mu^u(w) \longrightarrow \left(\int_{H_v^u} h\ d\mu^u\right) \cdot \left(\int_{SN_\Gamma} f\ dm\right)\ .$$

Proof. Fixed a small ε, we pick two non-negative functions $h' \in L^2(\mathbb{R}, \mathrm{e}^{\delta r'} dr')$ and $h'' \in L^2(S, d\mu)$, both of mean value one, and such that $\mathrm{supp}(h') \subset (0, r)$ and $\mathrm{supp}(h'') \subset B(v_{-\infty}, r)$ (the latter means the ball of radius r centered at $v_{-\infty}$ in the metric of the sphere) for some small r to be determined. As we can assume that supp(h) injects in SN_Γ, by Lemma 5.1 we can take r so small that the stable cell

$$\Delta_v^s \;=\; \Delta_v^s(p_{u,v}(\mathrm{supp}(h)), r, \mathrm{supp}(h''))$$

inside $S\mathbb{H}$ embeds in the quotient SN_Γ. We recall the construction (see section 5.3): a point $w \in \Delta_v^s$ is uniquely obtained from a point $w' \in H_v^u$ acting with the geodesic flow to get a point $w'' = g_{r'}w'$ belonging to the same stable horosphere passing through w.

By construction

$$\int_{H_v^u} h(w')\, f(g_t w') d\mu^u(w') \;=\; \int_S h''\, d\mu \;\cdot\; \int_0^r h'(r')\, e^{\delta r'} dr' \;\cdot\; \int_{H_v^u} h(w')\, f(g_t w')\, d\mu^u(w')$$

since both h' and h'' have mean value one.

We define a function h''' on SN_Γ supported on the cell Δ_v^s, multiplying our h, h' and h'' and the factor $\alpha(w) = e^{\delta\rho_o(w,w_{-\infty})}$ as in Lemma 5.2

$$h'''(w) \;=\; \alpha(w)\cdot(h''\circ p_{s,w''})(w)\cdot h'(r')\cdot h(w')\ .$$

A computation similar to Lemma 5.2 gives

$$\int_{H_v^u} h(w')\, f(g_t w') d\mu^u(w') \;=\; \int_{SN_\Gamma} h'''(w)\, f(g_t w')\, dm(w)$$

where of course the integral is restricted to Δ^s, and the function $f(g_t w')$ is constant along the w'-fibers of the cell. The stable leaves are contracted exponentially by the geodesic flow and the r'-leaves are not expanded. Let $\ell_{w'}$ denote the weakly stable w'-fiber of the cell. Then the maximal distance between $g_t w'$ and the points of $g_t\ell_{w'}$ does not increase as t grows. From uniform continuity of f we see that we can choose r so small that the value $f(g_t w')$ is within an arbitrarily small constant of $f(g_t w)$ for any other vector $w \in \ell_{w'}$. There follows that for r sufficiently small the above integral is within $\varepsilon/2$ of

$$\int_{SN_\Gamma} h'''\, g_t^* f\, dm$$

for all positive t. By the mixing property of the geodesic flow [Ru], we can find a positive $\bar{t}$ such that for $t > \bar{t}$ the above integral is within $\varepsilon/2$ of

$$\left(\int_{SN_\Gamma} h'''\, dm\right)\cdot\left(\int_{SN_\Gamma} f\, dm\right)\ .$$

Since by Lemma 5.2

$$\int_{SN_\Gamma} h'''\, dm \;=\; \int_{H_v^u} h\, d\mu^u$$

our proposition follows from triangular inequality. □

Remark. Following the same reasoning in [KM], we see that we can relax the requirement "uniformly continuous" by *almost* uniformly continuous in the statement of Theorem 5.3 (it means that the function f can be approximated by two monotone sequences of uniformly continuous functions $f_i' \leq f \leq f_i''$ a.e., both converging to f a.e., see the reference for details). Examples of almost uniformly continuous functions are characteristic functions of sets with zero measure boundary with respect to a Borel measure. While we do not have any control of the μ-mass of the boundary of a ball in the limit set, we do have the conditional of dm on geodesics which is Lebesgue measure. From Fubini theorem there follows that the Theorem 5.3 still holds for a function f which resembles a characteristic function along the time direction. This observation will be used in the following section to prove equidistribution of parabolic fixed points.

This theorem can be interpreted in the following way. For any vector v in the unit tangent bundle of N_Γ, consider any (small) ball in the unstable leaf through v, and all its g_t-images, equipped with the measures $d\mu^u$. After renormalization by a constant times the factor $\exp(-\delta t)$, this is a family of probability measures on SN_Γ weakly convergent to the Bowen-Margulis probability measure.

In particular, if $v_{-\infty}$ is a rank-k cuspidal fixed point, we can take for h the constant function of value one. Indeed, in the quotient SN_Γ the leaf H_v^u is not compact, it is a k-torus times $\mathbb{R}^{n-k}$, but its intersection with the nonwandering set $(\Lambda \times \Lambda - \text{diag}) \times \mathbb{R}$ modulo Γ is compact, and this is where $d\mu^u$ is supported. Thus, the expression $\int_{H_v^u} h f_t d\mu^u$ has the interpretation of the mean value of f w.r.t. a family of finite measures supported on the unstable fronts $g_t H_v^u$ from the cusp (the conditionals of the Bowen-Margulis measure, with constant mass for any t).

Theorem 5.4. *The conditionals of the Bowen-Margulis measure on a family of unstable fronts from a cusp, normalized by the factor $e^{-\delta t}$ to have all the same mass, converge weakly to the Bowen-Margulis measure as time goes to infinity.*

5.5. *Proof of equidistribution of parabolic fixed points.* The aim is to prove vague convergence of the measures ν_t defined on the boundary $\mathbb{R}^n$ (the cusp in question is at ∞ in the u.h.s. model), i.e. we must pick a compact support continuous test function ψ on $\mathbb{R}^n$ and look at $\nu_t(\psi)$. The statement in Theorem 5.1 is

$$e^{-\delta t} \sum_{s > e^{-t}} \psi(b) \;\to\; C \cdot \mu^\infty(\psi)$$

for t going to infinity and some constant C.

We will prove something more, in the same spirit of Bowen's equidistribution of closed geodesics for Anosov geodesic flows. In the notation of section 5.2., for an arbitrary strictly positive ε we consider the measures

$$\nu_t^\varepsilon \;=\; e^{-\delta t} \sum_{e^{-t} < s < e^{-(t-\varepsilon)}} \delta_b \;.$$

Theorem 5.5. *For any $\varepsilon > 0$, there exists a positive constant $C(\varepsilon)$ such that the measures ν_t^ε converge vaguely to $C(\varepsilon) \cdot \mu^\infty$ for t going to infinity.*

Proof. This proof is inspired by the ideas in [Vel]. As the measures ν_t^ε are invariant by the parabolic subgroup P, and are supported on the limit set, it will be sufficient to consider test functions ψ supported inside the parallelepiped T (a fundamental domain for the action of P on $\Lambda - \infty$). We will take ψ uniformly continuous.

Let H be the horizontal horosphere from which we constructed the disjoint balls in section 1.3. We fix a unit vector v based on H such that $v_\infty = \infty$ (i.e. $H = H_v^s$), $v_{-\infty} = \gamma'\infty$ for some $\gamma' \in \Gamma$, and $v_{-\infty}$ is contained in the support of ψ. We assume the horizontal horosphere H has euclidean height one, to avoid unimportant constants along the proof.

We pull back the test function ψ with the backward visual map $p_{s,v} : H_v^s \to S - \{\infty\} \simeq \mathbb{R}^n$ and define the function $f = \psi \circ p_{s,v}$ on the stable horosphere H_v^s.

Then, as in the proof of Theorem 5.4, we pick two more functions: f' = characteristic function of $(0, \varepsilon)$, and $f'' \in L^2(S, d\mu)$ uniformly continuous, with mean value one, and sufficiently small support around $\infty \in S$.

We construct the unstable cell Δ_v^u (a foliated chart for the unstable foliation) as follows. The base of the cell is

$$B \;=\; \{g_r w' \text{ for } w' \in \text{supp}(f) \text{ and } r \in (0, \varepsilon)\}$$

To any point w'' in the base B we attach the images of $\mathrm{supp}(f'')$ by $p_{u,w''} : H^u_{w''} \to S - w''_{-\infty}$ and define

$$\Delta^u_v \;=\; \Delta^u_v(\mathrm{supp}(\psi),(0,\varepsilon),\mathrm{supp}(f'')) \;=\; \bigcup_{w''\in B} p^{-1}_{u,w''}(\mathrm{supp}(f'')) \;.$$

We recall that any vector w in the cell uniquely defines the two vectors $w' \in H^s_v$ and $w'' = g_r w' \in H^u_w$ as in the definitions of B and Δ^u_v. If the supports of f' and f'' are sufficiently small this cell embeds in the quotient SN_Γ. As in the proof of Theorem 5.3, we define a function f''' on $S\mathbb{H}$. supported on the cell Δ^u_v, multiplying f, the characteristic function f', and f'' times the factor $\alpha(w) = \mathrm{e}^{\delta\rho_o(w,w_\infty)}$ as

$$f'''(w) \;=\; \alpha(w)\cdot(f''\circ p_{u,w''})(w)\cdot f'(r)\cdot f(w') \;.$$

Since the cell is embedded in SN_Γ, the function $\overline{f'''} = f'''\circ\pi^{-1}$ is well defined on SN_Γ, where π is the projection $\pi : S\mathbb{H} \to SN_\Gamma$. By the previous remark such $\overline{f'''}$ is almost uniformly continuous. as the only discontinuity is in the time-variable and there it is a characteristic function. Also, it is trivially square integrable since it is bounded and has compact support. As α is exactly the inverse of the Radon-Nikodym factor with which the measures μ^u are constructed by pullback of the geometric measure μ (same reasoning as in Lemma 5.2), we see that on any strictly unstable w''-leaf of the cell

$$\int_{H^u_{w''}} (f''\circ p_{u,w''})\,\alpha(w'')\,d\mu^u(w'') \;=\; \int_S f''\,d\mu \;=\; 1 \;.$$

We also recall that the conditional measures of dm on stable leaves are contracted by the geodesic flow as $g^*_r d\mu^s = \mathrm{e}^{-\delta r} d\mu^s$. We get for the mean value of the function $\overline{f'''}$ the expression

$$\begin{aligned}
\int_{SN_\Gamma} \overline{f'''}\,dm &= \int_{\Delta^u_v} f'''\,dm \\
&= \int_0^\varepsilon \mathrm{e}^{-\delta r} dr \cdot \int_{H^s_v} f\,d\mu^s \\
&= \frac{1-\mathrm{e}^{-\delta\varepsilon}}{\delta}\cdot\int_{H^s_v} f\,d\mu^s \;.
\end{aligned}$$

We are in position to apply Theorem 5.4 to the function $\overline{f'''} \in L^2(SN_\Gamma, dm)$. Let v' be the vector based at the same point of v but such that $v'_\infty = v_{-\infty} = \gamma'\infty$ for some $\gamma' \in \Gamma$. The parabolic subgroup P preserves the unstable (horizontal) horosphere $H^u_{v'} \simeq H$. There exists a compact region $T \subset H^u_{v'}$ such that the images $P\cdot T$ cover the intersection of $H^u_{v'}$ with the convex hull of the limit set and have disjoint interiors (see section 3.1). We define h to be the characteristic function of T. Hence the Γ-orbit of $T = \mathrm{supp}(h)$ cover the intersection between the disjoint spheres $\Gamma\cdot H$ and the convex hull of the limit set. From this observation and using Lemma 5.3 we get

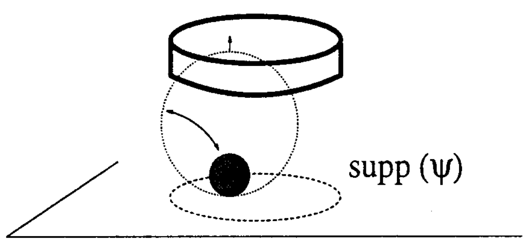
$$\int_{H^u_{v'}} h(w)\,\overline{f'''}(g_t w)\,d\mu^u(w) \;=\; \mathrm{e}^{-\delta t}\sum_{\gamma\in\Gamma}\int_{\gamma g_t H^u_{v'}} f'''(w)\,d\mu^u(w) \;.$$

The right-hand side is a sum of contributions of the intersections between the Γ-images of $H^u_{g_t v'}$ and the support of f'''. But $\mathrm{supp}(f''') = \Delta^u_v$, and its intersections with the Γ-images of $H^u_{g_t v'}$ are some of the the unstable leaves of the cell. The contribution of each unstable leaf in the integral is the value of f at the point w' of $\mathrm{supp}(f) \subset H^u_v$ corresponding to the leaf.

To see which leaves do contribute in the integral above, we identify a fundamental domain for Γ on the u.h.s. model for $\mathbb{H}$ with a full stable leaf coming from ∞. The base of the cell Δ_v^u is a "parallelepiped" in $\mathbb{H}^{n+1}$ embedded in a fundamental domain as in the picture below. Any point of it should be thought as having a unit vector pointing upward. By contrast, the Γ-images of $H^u_{g_t v'}$ are the boundaries of the disjoint balls $\mathcal{B}$ dilatated by the factor $\exp(t)$. Let $\partial B(b, s) = \gamma H$ for some $\gamma \in \Gamma$. The point where the unit normal vector of the dilatated sphere $\partial B(b, e^t s)$ points upward is the highest point $(b, e^t s) \in \mathbb{H}^{n+1}$. This says that in the quotient SN_Γ the horosphere $H^u_{g_t v'}$ intersects the cell Δ_v^u along the unstable leaves passing through those points $(b, e^t s)$ which belong to the base B. Each such horosphere contributes as

$$\begin{aligned}
\int_{\gamma H^u_{g_t v'}} f'''(w)\, d\mu^u &= \int_{\gamma H^u_{g_t v'} \cap \Delta^u_v} \alpha(w) \cdot (f'' \circ p_{u,w''})(w) \cdot f'(r) \cdot f(w')\, d\mu^u_{w''}(w) \\
&= \left(\int_S h'' d\mu \right) \cdot f'(r) \cdot f(w') \\
&= f(w') = f(b, e^t s) = \psi(b)
\end{aligned}$$

if the point $(b, e^t s)$ belongs to the base B of the cell, and zero otherwise (see the picture below).



Collecting these facts together we get

$$\int_{H^u_{v'}} h(w')\, \overline{f'''}(g_t w')\, d\mu^u(w') = e^{-\delta t} \sum_{e^{-t} < s < e^{-(t-\epsilon)}} \psi(b) \ .$$

According to Theorem 5.4, for t going to infinity we have

$$\int_{H^u_{v'}} h(w)\, \overline{f'''}(g_t w)\, d\mu^u(w) \to \left(\int_{H^u_{v'}} h\, d\mu^u \right) \cdot \left(\int_{SN_\Gamma} \overline{f'''}\, dm \right)$$

and the right-hand side is a constant times

$$\int_{H^s_v} f\, d\mu^s \ .$$

The theorem follows since the measure $d\mu^s$ on the stable leaf through v is proportional, after identification by the backward visual map, to the measure $d\mu^\infty$ on the boundary $S - \infty \simeq \mathbb{R}^n$ as defined in section 5.2. In other words

$$\int_{H^s_v} f'\, d\mu^s = \text{constant} \cdot \int_{\mathbb{R}^n} \psi\, d\mu^\infty \ .$$

□

Remark. This theorem shows in particular that the constant $\lambda > 1$ in Corollary 4.1 can be chosen arbitrarly near to one.

Proof of Theorem 5.1. Consider the collection of constants $C(r)$ in Theorem 5.5. They are not decreasing as function of r, since for $r \geq r'$ and any bounded domain $U \subset \mathbb{R}^n$ we have $\nu_t^r(U) \geq \nu_t^{r'}(U)$ for any time $t > 0$. Also, for any bounded domain U the masses $\nu_t^r(U)$ are all bounded by a finite constant K (depending on U) as follows from Theorem 4.2. Hence there exists and is finite the limit $C(r) \to C$ for $r \to \infty$ (actually, from the proof of Theorem 5.5 we see that $C(r)$ is equal to a constant times $(1 - \mathrm{e}^{-\delta r})$). We also see that the each measure ν_t in section 5.2 is equal to ν_t^{t+r} for any $r > 0$ since the sizes of the balls are bounded by one (see section 1.3).

Fix an $\varepsilon > 0$. Take a continuous test function ψ on $\mathbb{R}^n$ with bounded support U. There exists a finite r' such that for any $r > r'$

$$\|\psi\|_\infty \cdot K \cdot \mathrm{e}^{\delta(1-r)} \; < \; \varepsilon/3 \; .$$

By the observation above we find a finite r'' such that for any $r > r''$

$$|C(r)\mu^\infty(\psi) - C\mu^\infty(\psi)| \; \leq \; |C(r) - C| \cdot |\mu^\infty(\psi)| \; < \; \varepsilon/3 \; .$$

Now take an $r > \max\{r', r''\}$. By Theorem 5.5 there exists a $\bar{t}$ such that for any $t > \bar{t}$

$$|\nu_t^r(f) - \mu^\infty(f)| \; < \; \varepsilon/3 \; .$$

Finally let $t > \bar{t}$ and use triangular inequality

$$\begin{aligned}
|\nu_t(\psi) - C\mu^\infty(\psi)| \; &= \; |\nu_t^{t+1}(\psi) - C\mu^\infty(\psi)| \\
&\leq \; |\nu_t^{t+1}(\psi) - \nu_t^r(\psi)| \; + \; |\nu_t^r(\psi) - C(r)\mu^\infty(\psi)| \; + \; |C(r)\mu^\infty(\psi) - C\mu^\infty(\psi)| \\
&\leq \; (\mathrm{e}^{-\delta t} \sum_{s > \mathrm{e}^{-(t+1-r)}} |\psi(b)|) \; + \; \varepsilon/3 \; + \; \varepsilon/3 \\
&\leq \; \|\psi\|_\infty \cdot K \cdot \mathrm{e}^{\delta(1-r)} \; + \; \varepsilon/3 \; + \; \varepsilon/3 \qquad \text{(from Theorem 4.2)} \\
&\leq \; \varepsilon \; .
\end{aligned}$$

□

Remark. From the above proof we see that Theorem 5.2 can be improved as

Theorem 5.6. *For any $\varepsilon > 0$ there is a constant $C(\varepsilon)$ such that asymptotically*

$$\textit{card}\,\{\textit{balls of } \mathcal{B} \textit{ s.t. } \; \mathrm{e}^{-t} < s < \mathrm{e}^{-(t-\varepsilon)}\} \; \sim \; C(\varepsilon) \cdot \mathrm{e}^{\delta t} \; .$$

§6. Arithmetical examples

6.0. *Introduction.* Equidistribution of wave fronts in finite volume hyperbolic manifolds is a well known consequence of the mixing property of the geodesic flow. In particular, spheres (fronts emanating from a point in the manifold) and closed horospheres (fronts from an ideal point, a cuspidal end) are asymptotically equidistributed in the unit tangent bundle with respect to the Liouville measure, and hence in the manifold itself.

In this section we will be concerned with equidistribution of the closed horospheres from the cuspidal ends of certain arithmetic hyperbolic 3-manifolds, and in the same spirit of Verjovsky [Ve1] we will try to say something about error terms for mean values of characteristic functions. We consider the quotients of $\mathbb{H}^3$ by the groups of two by two complex matrices with unit determinant, and entries the integers of an imaginary quadratic number field, for example the Gaussian integers $\mathbb{Z}[i]$.

The point is that equidistribution of closed horospheres is closely related to the fact that a certain orbit, in the covering, of a parabolic fix point simulates the geometric measure on the ideal boundary, which in the finite volume case is Lebesgue measure. The natural orbit is the analogue of the Farey series (originally introduced by Ford and more recently used for example by Sullivan to get a geometrical proof of Khintchine theorem in the theory of Diophantine approximations), that is the collection of resting points, inside a fundamental domain for the cuspidal subgroup. of the disjoint horoballs representing one cuspidal end of the hyperbolic manifold. Counting them, for the arithmetic manifolds we are interested in, is the content of a well known formula in number theory, due to Mertens for the integers, and we generalize it to give an equidistribution result.

We also use these disjoint balls to construct a family of measures on a fundamental domain for the group in the upper half-space model for $\mathbb{H}^3$, which gives asymptotic equidistribution of closed horospheres both in the unit tangent bundle and in the frame bundle, together with an estimate of the error for certain characteristic functions.

Finally, we explain the relation between the asymptotics of the measures supported on the closed horospheres and the spectral theory on the hyperbolic manifold, by means of Eisenstein series. as first observed by Sarnak in a two dimensional situation. In the same spirit of Zagier [Za1,2], who observed that the Riemann hypothesis is related to the speed of convergence of analogous measures in the modular surface, we use the Rankin-Selberg unfolding trick to derive a similar "if and only if" theorem about the Riemann hypothesis for imaginary quadratic number fields.

6.1. *Imaginary quadratic number fields.* We briefly recall some basic number theory, and introduce the "geometrical" picture of ideals in an imaginary quadratic number field. As general references we quote the very classical book by Hecke [He] and [Kn] for the analytic theory of numbers.

Given a *field* k, let $k[x]$ denote the ring of polynomials in x with coefficients in k, and $k(x)$ the quotient field, ratios of polynomials in $k[x]$ whose denominators are non-zero. Also, we denote by $\mathbb{Q}$ the field of rational numbers, and by $\mathbb{Z}$ the integers, or rational integers.

An *algebraic number* is a root of an equation $f(x) = 0$ where $f(x)$ is a non-zero polynomial in $\mathbb{Q}[x]$. Given an algebraic number θ, there exists a polynomial p in $\mathbb{Q}[x]$ of smallest degree and leading coefficient one which has θ as a root. The degree n of p is called the *degree* of θ (w.r.t. the field $\mathbb{Q}$), and the n roots of p, $\theta^{(1)}, ..., \theta^{(n)}$, are called the *conjugates* of θ.

An *algebraic number field* is an algebraic extension of $\mathbb{Q}$, i.e. is of the form $k = \mathbb{Q}(\theta)$ for some algebraic number θ. The degree of $\mathbb{Q}(\theta)$ (over $\mathbb{Q}$) is the minimum degree of a polynomial defining θ, hence the degree of θ, and has the interpretation of a dimension, $[\mathbb{Q}(\theta) : \mathbb{Q}]$. Every number α of $\mathbb{Q}(\theta)$ is obtained exactly once in the form

$$\alpha \;=\; g(\theta) \;=\; c_0 + c_1\theta + ... + c_{n-1}\theta^{n-1}$$

where the coefficients c_j's are in $\mathbb{Q}$, and the *conjugates* of α are defined to be the numbers $\alpha^{(i)} = g(\theta^{(i)})$, for $i = 1, ..., n$.

The *algebraic integers* ϑ of an algebraic number field k are the zeros of monic polynomials in $\mathbb{Z}[x]$, i.e. polynomials of the form

$$f(x) \;=\; x^n + a_{n-1}x^{n-1} + ... + a_0$$

with coefficients a_j rational integers, that are contained in k. They form a ring, and are the natural generalization of the rational integers.

Example. An *imaginary quadratic* number field is of the form $k = \mathbb{Q}(\sqrt{-D})$ for D a positive square-free rational integer (with no repeated prime factors), and its integers are

$$\vartheta = \mathbb{Z}[i\sqrt{D}] = \{a + b\sqrt{D}i \mid a, b \in \mathbb{Z}\} \quad \text{if } D \neq 1 \bmod 4$$
$$\vartheta = \mathbb{Z}[(1+i\sqrt{D})/2] = \{a + b(1+\sqrt{D}i)/2 \mid a, b \in \mathbb{Z}\} \quad \text{if } D = 1 \bmod 4 \ .$$

There is no unique factorization in a generic algebraic number field, and restoring this property is the reason for the introduction of divisor theory. A *(fractional) divisor* is an object $\mathfrak{a} = (\alpha_1, ..., \alpha_r)$ generated by some $\alpha_j \in k$, and is called *principal* when it is generated by a single number, (α). We say that $\mathfrak{a} = (\alpha_1, ..., \alpha_r)$ divides $\mathfrak{b}$, i.e. $\mathfrak{a}|\mathfrak{b}$, if any of the β_j's that generate $\mathfrak{b}$ is of the form

$$\beta_j = \gamma_1\alpha_1 + ... + \gamma_r\alpha_r$$

with γ_i's algebraic integers. Together with the definition of multiplication

$$(\alpha_1, ..., \alpha_r) \cdot (\beta_1, ..., \beta_s) = (\alpha_1\beta_1, \alpha_1\beta_2, ..., \alpha_r\beta_s)$$

this gives to the set $\mathfrak{M}$ of non-zero fractional divisors of k a group structure, the identity being the principal divisor $\mathfrak{o} = (1)$. Principal divisors $\mathfrak{P}$ also form a group, the quotient $\mathfrak{J} = \mathfrak{M}/\mathfrak{P}$ is called the *class group* of k, and its order the *class number.* Also, *integral divisors* are those generated by algebraic integers. The *units* $\vartheta^\times$ of the field are the $\mathfrak{u}$ such that $\mathfrak{u}$ and $\mathfrak{u}^{-1}$ are integral.

In $\mathfrak{M}$ there is a well defined concept of *greatest common divisor*, namely $(\mathfrak{a}, \mathfrak{b})$ is the divisor generated by the generators of $\mathfrak{a}$ and those of $\mathfrak{b}$, and this gives unique factorization of any (integral) divisor into *prime* (integral) divisors. The *norm* $||\cdot||$ is a multiplicative function from $\mathfrak{M}$ to the divisors of $\mathbb{Q}$. On principal divisors, it is defined as the product of all the associates of (any) a generator: $||(\alpha)|| = \alpha^{(1)}\alpha^{(2)}...\alpha^{(n)}$.

The "physical" model for divisors are *ideals* (or non-zero finitely generated ϑ-submodules of k), and we show the correspondence in the case of an imaginary quadratic number field $k = \mathbb{Q}(\sqrt{-D})$, where any divisor is generated by at most two numbers. To a divisor $\mathfrak{a} = (\alpha_1, \alpha_2)$ there corresponds the ideal $\mathfrak{a} = \vartheta \cdot \alpha_1 + \vartheta \cdot \alpha_2 \subset k$. In particular, to an integral divisor there corresponds an integral ideal inside ϑ, and $\mathfrak{o}$ is identified with the integers ϑ. As we can think at ϑ as a lattice inside $\mathbb{C}$ generated by 1 and a complex number τ, integral ideals are sublattices of it. The geometrical picture of divisibility is: the integral ideal $\mathfrak{a}$ divides $\mathfrak{b}$ if the lattice $\mathfrak{b}$ is a sublattice of $\mathfrak{a}$. The norm of an integral ideal $\mathfrak{a}$ has the geometrical interpretation of the cardinality of fundamental domains of ϑ contained in a fundamental domain for α (or, better, the ratio of their areas). In particular, the norm of a principal integral ideal (α) is the square modulus of α (as a complex number).

The object of the analytical theory of numbers is the study of *arithmetical functions*, functions on the (arithmetic) semigroup of integral ideals in an algebraic number field. The set $\mathcal{A}$ of non-zero arithmetical functions inherits the structure of a (Dirichlet) ring from the notion of divisibility of integral ideals, the product being defined by

$$f * g\,(\mathfrak{a}) = \sum_{\mathfrak{b}|\mathfrak{a}} f(\mathfrak{b})\, g(\mathfrak{a}/\mathfrak{b}) \ .$$

The unit in $\mathcal{A}$ is the function e equal to one on $\mathfrak{o}$ and zero otherwise. A special role is played by the constant function o (equal to one for any ideal) and its Dirichlet inverse, the Möbius function μ (i.e. $\mu * o = o * \mu = e$). The *Dirichlet series* of an arithmetical function f is the formal series

$$\hat{f}(z) = \sum_{\mathfrak{a}} \frac{f(\mathfrak{a})}{||\mathfrak{a}||^z}$$

and the map {arithmetical functions} $\to$ {Dirichlet series} is a homeomorphism of rings (usual multiplication between formal series). In particular, the Dirichlet series of the constant function

$$\zeta_k(z) = \sum_{\mathfrak{a}} \frac{1}{\|\mathfrak{a}\|^z}$$

is called the *(Riemann, Dedekind) zeta function* of the number field k. Unique factorization allows to get a multiplicative expression for the zeta function over the set of prime integral ideals. The series is absolutely convergent for $\mathrm{Re}(z) > 1$, has an analytic continuation in the whole complex plane, a pole at $z = 1$ (responsible for the prime number theorem), and a functional equation which relates its values at couples of points obtained by reflection around the line $\mathrm{Re}(z) = 1/2$. The *Riemann hypothesis* for the field k is the same as for the rationals: all the nontrivial zeros of the zeta function have real part $1/2$.

6.2. *Arithmetic three-manifolds.* Take an imaginary quadratic number field, say $k = \mathbb{Q}(\sqrt{-D})$ where D is a square free positive integer. Let ϑ be its ring of algebraic integers, and form the group of matrices $\Gamma_k = PSL(2, \vartheta)$, i.e. two by two complex matrices whose entries are integers of the field, and whose rows and columns are made of relatively prime integers.

The u.h.s. model for the hyperbolic three space is $\mathbb{H}^3 = \{(z,r) \mid z \in \mathbb{C},\ r \in \mathbb{R}_+\}$, and it is standard to identify its points with a subspace of quaternions as

$$(z,r) \to p = z + rj = x + yi + rj$$

$(1, i, j, k$ is the usual base of the skew-field of quaternions). The group of orientation preserving isometries of $\mathbb{H}^3$ is $PSL(2,\mathbb{C})$. The action of a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is given by

$$p \mapsto Ap = (ap + b)(cp + d)^{-1}$$

so that the induced action on the ideal boundary $\mathbb{C} \cup \infty$ is by fractional linear transformations.

The group Γ_k is a discrete subgroup of $PSL(2,\mathbb{C})$, and the quotient $M_k = \mathbb{H}^3/\Gamma_k$ is an hyperbolic manifold (actually an orbifold because of the presence of torsion elements) with a number of cusps equal to the class number h of k. It has finite volume, which has been computed several times starting from Humbert in 1919, given by

$$|d|^{3/2}\zeta_k(2)/4\pi^2$$

where d is the discriminant of k over $\mathbb{Q}$, and ζ_k is the Dedekind zeta function of the field. From now on, we shall assume that there are no more units in the field other than ± 1, that is, we shall assume $D \neq 1, 3$, in order to simplify the notation in the arguments below.

Cusps are in one-to-one correspondence with the elements of the ideal class group, they are the orbits of the action of Γ_k on $\mathbb{P}(k)$ (thought inside the ideal boundary $\mathbb{C} \cup \infty$). To the identity in the ideal class group of k, that is ϑ itself, there corresponds the cusp at infinity. The relative parabolic subgroup P preserves the family of horizontal horospheres, each one homeomorphic to (ideal boundary) $-$ (cusp) $\simeq \mathbb{C}$, acting on them as translations by the integers of the field. Indeed. we can think at ϑ as the lattice $\mathbb{Z} + \mathbb{Z}\tau$ in the complex plane, its modulus being

$$\tau = \begin{cases} \frac{1}{2}(1 + i\sqrt{D}) & \text{if } D = 3 \pmod 4 \\ i\sqrt{D} & \text{if } D \neq 3 \pmod 4 \end{cases}.$$

We identify the quotient torus $T = \mathbb{C}/P$ with the horizontal horosphere H at unit height modulo P, and it is immediate to compute its area (using the fact that the discriminant of the field is $-D$ if $D = 3 \pmod 4$ and $-4D$ otherwise).

Lemma 6.1. *area*$(T) = Im(\tau) = |d|^{1/2}/2$.

The images of the horizontal horosphere H at unit height under Γ_k are disjoint euclidean spheres resting on the bounding complex plane. The family is left invariant by the parabolic subgroup, and can be described as

Lemma 6.2. *The disjoint spheres representing the cusp associated to* ϑ *have resting points* $r = a/b$ *and sizes* $s = 1/|b|^2$, *where* a *and* b *run between the couples of integers such that ideal*$(a, b) = \vartheta$.

In particular, a set of balls representing the images of the horizontal horosphere by the two-sided quotient $P\backslash\Gamma_k/P$ are those where the a's in the above lemma are taken modulo b (the integral ideal generated by b).

Remark. In general, the correspondence between cusps of M_k and elements of $\mathfrak{J}$, the ideal class group of the field, is as follows. Let $\mathfrak{a}$ be a representative of an element of $\mathfrak{J}$. It is an ideal of the form (s, t) with $s, t \in \vartheta$. Let A be the following matrix of $SL(2, k)$

$$A = \begin{pmatrix} s & u \\ t & v \end{pmatrix}$$

for some $u, v \in \mathfrak{a}^{-1}$. Then A^{-1} sends the point in the boundary sphere s/t to the point at infinity. In the u.h.s. model it is immediate to see that the stabilizer of ∞ is the parabolic subgroup of upper triangular matrices of unit determinant. There follows that the stabilizer of s/t is

$$P_{s/t} = \left\{ A \begin{pmatrix} w & z \\ 0 & w^{-1} \end{pmatrix} A^{-1} \text{ with } w \in \vartheta^{\times} \text{ and } z \in \mathfrak{a}^{-2} \right\} .$$

To this cusp of M_k there corresponds a family of closed horospheres, tori (elliptic curves) of the form $\mathbb{C}/\mathfrak{a}^{-2}$. The area of the one at unit height in the u.h.s. model, call it $T_{\mathfrak{a}}$, can be written in terms of $\mathrm{Im}(\tau)$ and the norm of $\mathfrak{a}$. Indeed, by the very definition of norm of a fractional ideal, the area of this torus divided by the area of T is the norm of $\mathfrak{a}^{-2}$, hence

$$\mathrm{area}(T_{\mathfrak{a}}) = \mathrm{area}(T)/\|\mathfrak{a}\|^2 .$$

The family of disjoint spheres can be described in terms of couples of points and some arithmetic condition.

6.3. *Horoballs counting function.* Consider a cusp $P \subset \Gamma_k$, map its fixed point at infinity in the u.h.s. model, and consider the collection of disjoint balls which are the images of the horoball B bounded by the horizontal hyperplane H at unit height. We denote by T the euclidean torus obtained quotienting the basic horosphere H by the parabolic subgroup P, as well as a fundamental domain for P on the boundary $\mathbb{C}$. We denote by $\mathcal{B}$ the collection of $P\backslash\Gamma_k/P$-images of the horizontal horoball B resting on T. They are parametrized by the couples $(b, s) = (\text{resting point}, \text{diameter})$, and we also think of them as points in the u.h.s. model $\mathbb{H}^3$, namely the highest points of the balls.

Let Q be a small domain in $\mathbb{H}^3$ contained in a fundamental domain for the action of P, e.g. the cone $T \times \mathbb{R}_+$ above T. We define the *horoballs counting function* relative to Q as

$$n(Q, t) = \mathrm{card}\left\{\text{balls of } \mathcal{B} \text{ s.t. } (b, e^t s) \in Q\right\} .$$

If we think at the starting spheres as fronts emanating from (the images) of the cusp, and act with the geodesic flow until time t, $n(Q, t)$ counts the cardinality of those images of the fronts which point upward (to the cusp at ∞) and are inside Q.

In particular, if we take the *cone* C composed of all the geodesics from the point at infinity and the portion of the horizontal horosphere at unit height with shadow S in the boundary $\mathbb{C}$, then $n(C, t)$ counts the cardinality of those balls resting on S with heights bigger than e^{-t}.

The scope of this section is to find an asymptotics for the counting function which holds for generic well behaved domains Q. We call a domain $Q \subset \mathbb{H}^3$ *good* if it has a piecewise smooth boundary.

Theorem 6.1. *For sufficiently big times t the counting function relative to a good Q satisfies the asymptotic formula*

$$n(Q,t) \;=\; \mathrm{area}(T)\frac{vol(Q)}{\pi \cdot vol(M_k)} \cdot e^{2t} \;+\; \mathcal{O}(e^{\frac{3}{2}t}) \;.$$

The idea, which is based on an elementary argument in number theory, is that the highest points of the disjoint balls, in this arithmetical case, can be viewed as a subset of positive density of a lattice in an appropriate euclidean space. The euclidean space turns out to be $\mathbb{R}^4$, and $n(Q.t)$ is counting points in a domain dilated by the factor $e^{t/2}$. The theorem shows the "volume + perimeter" behavior of euclidean lattice point counting (Gauss formula). Indeed, the counting function $n(Q,t)$, taking as Q the cone over a full fundamental domain T of the cuspidal subgroup, is the sum of the values of the Euler totient function, and its asymptotic is what in the case of integers is known as Mertens formula.

The main geometrical tool in the proof is the following simple observation (an analogous map in dimension two has been used in [Vel]).

Lemma 6.3. *There exists a fibration ψ from $\mathbb{C}\times\mathbb{C}^*$ onto $\mathbb{H}^3$, with fibers S^1, which is volume preserving in the sense that for any Borel set $Q\subset\mathbb{H}^3$ the euclidean volume of $\psi^{-1}(Q)$ is equal to π times the hyperbolic volume of Q.*

Proof of Theorem 6.1. Here we deal about the counting function for the disjoint balls associated to the cusp at infinity, the one corresponding to the identity in the ideal class group. Moreover, we work the proof in the case of a principal ideals field, i.e. the one considered is the only cusp. The general situation does not contain any additional difficulty, apart for a more involved notation. We remind, again, that for the cone with shadow T *this is nothing but the proof of a Mertens-type theorem for imaginary quadratic number fields*, and indeed the cardinality of all the balls sitting on T and with sizes $\geq 1/x$ is the sum

$$\sum_{\|\mathfrak{a}\|\leq x} \varphi_k(\mathfrak{a})$$

where φ_k is the Euler-type function on k that counts the cardinality of residue classes modulo $\mathfrak{a}$ of integral ideals $\mathfrak{b}$ relatively prime to $\mathfrak{a}$.

We want to generalize the classical asymptotic formula for this sum to an equidistribution result, hence we look at the geometrical meaning of the proof. Consider the fibration $\psi:\mathbb{C}\times\mathbb{C}^*\to\mathbb{H}^3$ given, using coordinates in the u.h.s. model for $\mathbb{H}^3$, by

$$(z,z') \;\mapsto\; (z/z', |z'|^{-2}) \;.$$

Equip $\mathbb{C}\times\mathbb{C}^*$ with the standard euclidean volume, and take the inverse image of the domain in the hyperbolic space $\tilde{Q}=\psi^{-1}Q$. The above map is volume preserving, in the sense that for any $Q\subset\mathbb{H}^3$ (we use the subscript ϵ to mean euclidean volume)

$$\mathrm{vol}_\epsilon(\tilde{Q}) \;=\; \pi\cdot\mathrm{vol}(Q)$$

as a computation shows.

Lemma 6.2 implies (taking into account that we must divide by the order of the group of units) that

$$n(Q,t) \;=\; \frac{1}{2}\,\mathrm{card}\left\{(a,b)\in\vartheta\times\vartheta \;\mid\; \mathrm{ideal}(a,b)=\vartheta\;,\;\; (a,b)\in e^{t/2}\tilde{Q}\right\}$$

where the notation $\lambda\tilde{Q}$ stands for dilatation of the euclidean domain $\tilde{Q}$ by the factor λ.

Rename the independent variable as $x=e^t$, and call $N(x)$ the function $2\cdot n(Q,t(x))$ (from here on we have fixed Q, hence we will simplify our notation). At this point $N(x)$ is estimated in a very classical way, relating it, with the aid of the Möbius function, with the function

$$M(x) \;=\; \mathrm{card}\left\{(a,b)\in\vartheta\times\vartheta \;\mid\;\; (a,b)\in e^{t/2}\tilde{Q}\right\}$$

that counts the full number of points of the lattice $\vartheta \times \vartheta$ in the domain $\sqrt{x}\tilde{Q}$ not caring whether their coordinates are relatively prime or not.

Indeed let (a, b) be a lattice point belonging to $\sqrt{x}\tilde{C}$ such that ideal$(a, b) = \vartheta$, i.e. it contributes to the counting function $N(x)$. If (d) is any principal ideal then clearly the lattice point (ad, bd) belongs to $\sqrt{x}|d|\tilde{Q}$, and it is counted by the function $M(x|d|^2)$. Since we have assumed that in the field we are working with every ideal is principal in this way we obtain all the lattice points. We recall that if $\mathfrak{d}$ denotes the ideal generated by d its norm as an ideal is $\|\mathfrak{d}\| = |d|^2$. From the discussion above there follows that

$$M(x) = \sum_{\|\mathfrak{d}\| \leq x} N(x/\|\mathfrak{d}\|)$$

and by the Möbius inversion formula (see [Kn]) this implies

$$N(x) = \sum_{\|\mathfrak{d}\| \leq x} \mu(\mathfrak{d})\, M(x/\|\mathfrak{d}\|) \ .$$

Above, μ is the Möbius function, a function on the arithmetical semigroup of the nonzero integral ideals which is the Dirichlet inverse of the constant function. What is important for us is the following couple of facts: it is a bounded function (indeed takes the values ± 1), and its Dirichlet series gives the inverse of the zeta function of the field.

Now $M(x)$ is easy to estimate since it counts the number of lattice points in a domain of euclidean $\mathbb{R}^4$. The lattice is $\vartheta \times \vartheta$ and has a fundamental domain with volume $\mathrm{Im}(\tau)^2$. The number of lattice points inside a certain region is given by

$$\frac{1}{\mathrm{Im}(\tau)^2}\mathrm{Volume} \ + \ (\mathrm{bounded}) \cdot \mathrm{Perimeter} \ .$$

This is true if $\tilde{Q}$ is a good domain (and Q good implies $\tilde{Q}$ good). From the above definitions we see that the volume of $\sqrt{x}\tilde{Q}$ goes as x^2 and its perimeter as $x^{3/2}$. Also we take care of the boundary effect calling $e(x)$ the bounded function above, namely

$$M(x) = x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} + e(x) \cdot x^{3/2} \cdot \mathrm{per}_\epsilon(\tilde{Q})$$

and

$$N(x) = \sum_{\|\mathfrak{d}\| \leq x} \mu(\mathfrak{d}) \left(x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} \frac{1}{\|\mathfrak{d}\|^2} + e(x) \cdot x^{3/2} \cdot \mathrm{per}_\epsilon(\tilde{Q}) \frac{1}{\|\mathfrak{d}\|^{3/2}} \right) \ .$$

The leading term of the r.h.s. above is the coefficient of the volume. We substitute to it the whole sum over all ideals

$$x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} \left(\sum_{\mathfrak{d}} \frac{\mu(\mathfrak{d})}{\|\mathfrak{d}\|^2} \right) = x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} \frac{1}{\zeta_k(2)}$$

and get

$$\begin{aligned} N(x) &= x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} \frac{1}{\zeta_k(2)} + \\ &\quad + x^2 \cdot \frac{\mathrm{vol}_\epsilon(\tilde{Q})}{\mathrm{Im}(\tau)^2} \sum_{\|\mathfrak{d}\| > x} \frac{\mu(\mathfrak{d})}{\|\mathfrak{d}\|^2} + \\ &\quad + e(x) \cdot x^{3/2} \cdot \mathrm{per}_\epsilon(\tilde{Q}) \sum_{\|\mathfrak{d}\| \leq x} \frac{\mu(\mathfrak{d})}{\|\mathfrak{d}\|^{3/2}} \ . \end{aligned}$$

We estimate the last two terms of the r.h.s. above. The norm of the second one is bounded by a constant times

$$\begin{aligned} x^2 \cdot \left| \sum_{\|\mathfrak{d}\|>x} \frac{\mu(\mathfrak{d})}{\|\mathfrak{d}\|^2} \right| &\le x^2 \cdot \sum_{n>x} \frac{\text{card}\,\{\text{ideals with norm } n\}}{n^2} \\ &\le x^2 \cdot \sum_{n>x} \frac{1}{n^{3/2}} \\ &\le \mathcal{O}(x^{3/2}) \end{aligned}$$

(since μ is bounded by one). Actually, a more accurate estimate is valid, since a general theorem tells us that the cardinality of integral ideals of norm n, for any algebraic number field, is $\mathcal{O}(n^\varepsilon)$ for any small positive ε [Kn], so that the term is $\mathcal{O}(x^{1+\varepsilon})$. The norm of the third one is bounded by a constant times

$$\begin{aligned} |e(x)| \cdot x^{3/2} \cdot \left| \sum_{\|\mathfrak{d}\|\le x} \frac{\mu(\mathfrak{d})}{\|\mathfrak{d}\|^{3/2}} \right| &\le |e(x)| \cdot x^{3/2} \cdot \sum_{\|\mathfrak{d}\|\le x} \frac{1}{\|\mathfrak{d}\|^{3/2}} \\ &\le |e(x)| \cdot x^{3/2} \cdot \left(\zeta_k(3/2) + \mathcal{O}(x^{-1/2})\right) \\ &\le |e(x)| \cdot \mathcal{O}(x^{3/2}) \end{aligned}$$

(see [Kn] ch. 4, Proposition 2.6, but indeed we just needed the sum to be bounded, and this is not an arithmetical information but is equivalent to the observation that the Poincaré series of the group Γ_k is convergent as we are above the critical exponent).

Thus, the whole error term is bounded by a constant times $x^{3/2}$ for x sufficiently large. We have proven more, indeed, that

$$N(x) = x^2 \cdot \frac{\text{vol}_\epsilon(\tilde{Q})}{\text{Im}(\tau)^2 \zeta_k(2)} + |e(x)| \cdot \mathcal{O}(x^{3/2}) + \mathcal{O}(x^{1+\epsilon}) \ .$$

In particular, making use of our expressions for the area of T (Lemma 6.1) and for the volume of the quotient manifold we obtain

$$N(x) = 2\,\text{area}(T) \frac{\text{vol}(Q)}{\pi \cdot \text{vol}(M_k)} \cdot x^2 + (\text{bounded}) \cdot x^{3/2}$$

which gives the result. □

Remark. The lifted domains $\tilde{Q} \subset \mathbb{C} \times \mathbb{C}^*$ are S^1-invariant, hence we could expect to get better estimates for the function $e(x)$ above than its boundedness. In particular, for suitable Q we could expect to estimate $e(x) = \mathcal{O}(x^{-1/2+\epsilon})$, so to get the full error in Theorem 6.1 of order $x^{1+\epsilon}$.

6.4. *Arithmetic and dynamics in the frame bundle.* The group $PSL(2,\mathbb{C})$ is the group of orientation preserving isometries of $\mathbb{H}^3$, hence the orthonormal frame bundle $\mathcal{F}M_k$ of our manifolds is naturally identified with the homogeneous space $\Gamma_k \backslash PSL(2,\mathbb{C})$ and it fibers onto the unit tangent bundle with fiber $U(1)$.

Let g, u, v denote the standard base of the Lie algebra $\mathfrak{psl}(2,\mathbb{C})$ over the complex, so that $g_\mathbb{R}$ generates the one parameter subgroup giving the geodesic flow on the unit tangent bundle. The foliation induced by the complex vector fields g and v is the weakly stable foliation of the "holomorphic Anosov action" of $\mathbb{C}^*$ defined by g (we refer to [Gh2] for a definition). The strictly unstable foliation is generated by u, and induces the $\mathbb{C}$-action by right translation with upper diagonal matrices.

Consider the cusp at infinity, and take an orthonormal frame based at the point $(0,1) \in \mathbb{H}^3$ with the first vector pointing down (u.h.s. model, this corresponds to the identity in $PSL(2,\mathbb{C})$ under a natural identification $PSL(2,\mathbb{C}) \simeq \mathcal{F}\mathbb{H}^3$). Its u orbit describes the horizontal horosphere, with parallel displacement of the framing, and descends to a closed torus in $\mathcal{F}M_k$, and again to a closed horosphere in the manifold. The g action on such torus gives a $\mathbb{C}^*$-family of closed tori immersed in the frame bundle, call them T_w with $w \in \mathbb{C}^*$, which project to the closed tori in the manifold, (a $U(1)$-family for each one). Other closed holomorphic curves immersed in the frame bundle are given by the closed orbit of the g action, corresponding to closed geodesics.

Consider the weakly stable leaf L from the cusp, i.e. the leaf of the stable foliation passing from the frame based at $(0,1)$ and pointing upward. This is a copy of the complex affine group, and there exists a volume preserving map from the euclidean $\mathbb{C} \times \mathbb{C}^*$ to it, which composed with the projection onto $\mathbb{H}^3$ gives the fibration ψ described in lemma 6.3, and used in the proof of Theorem 6.1.

In the quotient $\mathcal{F}M_k$ the intersection points between the torus T_w and the weakly stable leaf L, read in euclidean coordinates of the latter, are of the form $(w^{-1/4}a, w^{-1/4}b)$ where a and b run between the couples of relatively prime integers of k, and $|w| \cdot \text{area}(T_1)$ is the area of the torus. That is, the "relatively prime" lattice appear naturally as describing the intersection points of the closed fronts from the cusp at infinity and the weakly stable leaf from the cusp, a manifestation of the fascinating beauty of such arithmetic hyperbolic three-manifolds, as claimed by Thurston [Th2]:

> *Find topological and geometric properties of quotients spaces of arithmetic subgroups of $PSL(2,\mathbb{C})$. These manifolds often seems to have special beauty.*

This explains what we have really done in the previous section. We have fixed a (small) domain in the stable leaf, counted the cardinality of its intersections with the tori T_w, and proved its asymptotic for $w \to \infty$. Up to Möbius inversion, this is plain lattice point counting in the euclidean space of dimension four. Also, this tells us that the perimeter-like error is proper of counting in the frame bundle, and should be improved while quotienting by S^1 first, projecting to the unit tangent bundle, and by S^2 down to the manifold, as in the circle problem.

6.5. *Weak limits.* The first natural way to interpret Theorem 6.1 is in term of the geometric (Patterson-Sullivan) measure on the sphere at infinity, which in the finite volume case with cusps is the standard spherical measure. Recall that the geometric measure on the limit set of a Kleinian group is obtained via a weak limit of a family of measures on the unit disc, defined by Dirac masses placed on one orbit of the group, as the parameter of the Poincaré series decreases to the critical value (the Hausdorff dimension δ of the limit set, in our case 2). In our case, the limit set is the whole sphere, and the geometric measure is its standard spherical volume. Once fixed the cusp, we project the limit set on the boundary $\mathbb{C} = (\text{ideal boundary}) - (\text{cusp})$ via stereographic projection, and construct a measure on it putting Radon-Nikodym derivative $|\text{cusp} - \xi|^{-2\delta}$ (euclidean distance) to the push forward of the geometric measure. The result is the standard Lebesgue measure ν on the complex plane. From the disjoint spheres we define a family of measures on $\mathbb{C}$ as

$$\nu_t = e^{-2t} \sum_{\text{size} > e^{-t}} \delta_b \ .$$

From Theorem 6.1 there follows

Corollary 6.1. *The measures ν_t converges weakly to (a constant times) ν, as time goes to infinity, with an error, for mean values of characteristic functions of balls, which is at least of order $e^{-(1/2)t}$.*

Clearly, a similar statement holds if we replace ν by the geometric measure on the sphere, and the family ν_t by the same Dirac masses weighted as $|\mathrm{cusp} - b|^{2\delta}$. This statement is to be compared with Theorem 5.1.

Also, the disjoint spheres represent the closed horospheres, fronts coming out from the cusps, and we get exponentially fast equidistribution for them, too. We recall that the geodesic flow on a compact hyperbolic manifold is the typical example of an Anosov flow. The unit tangent bundle is foliated in three ways: geodesics, the expanding horospherical foliation and the contracting horospherical foliation. A *standard box* $\triangle$ is a foliated chart for the strictly unstable foliation made of equal area leaves starting from a base B, which is identified with a domain in the weakly stable leaf asymptotic to the cusp. More precisely, start with a vector v of $S\mathbb{H}^3$, the center, pointing to the chosen cusp. Take a small ball B in the weakly stable leaf through v, and attach to any such vector small balls of the strictly unstable leaf of equal area ℓ. The box is

$$\triangle = \bigcup_{q\in B} \ell_q .$$

If the box is sufficiently small, it embeds in the quotient SM_k. We identify the base B with a domain in a fundamental domain for Γ_k in the hyperbolic space. The volume of the standard box, volume in the unit tangent bundle, is by construction

$$\mathrm{Vol}(\triangle) = \ell \cdot \mathrm{vol}(B)$$

by the same reasoning as in the proof of Theorem 5.5. Now, let T be a closed leaf of the unstable foliation in the unit tangent bundle of M_k coming from the cusp. Its intersection with the standard box is made of a certain number of equal area pieces of T. The geodesic flow moves the torus T enlarging its area (the area of its front) as $\mathrm{area}(g_tT) = e^{2t}\cdot\mathrm{area}(T)$. We denote by p_t the probability measure concentrated on g_tT. By the some reasoning as in Theorem 5.5 we see that $p_t(\triangle)$ is ℓ times $n(B,t)$ over the area of g_tT. But Theorem 6.1 gives the asymptotic value of the cardinality as

$$n(B,t) = \mathrm{area}(T)\,\frac{\mathrm{vol}(B)}{\pi\cdot\mathrm{vol}(M_k)}\,e^{2t} + (\text{bounded})\,e^{\frac{3}{2}t}$$

where we recognize in the denominator the volume of the unit tangent bundle of M_k. To sum up, calling p the Haar probability measure on M_k, we have

Corollary 6.2. *For any (sufficiently small) standard box there exist a constant c such that for sufficiently big times t*

$$|p_t(\triangle) - p(\triangle)| < c\cdot e^{-t/2} .$$

Remark. The same reasoning with standard boxes can be made in the frame bundle, and actually the same result holds, the exponent one-half in the error being optimal in this case.

Moreover, it follows from the proof that the above bound is uniform for standard boxes inside a compact region of the unit tangent bundle. The leaves of the weakly unstable foliation are dense (this is a consequence of the fact that the orbit of any cusp is dense in the limit set), hence we can approximate any bounded region of SM_k by standard boxes. There follows

Corollary 6.3. *The probability measures p_t converge vaguely to p as times go to infinity.*

It should be remarked that this last corollary is well known to hold for a generic finite volume hyperbolic manifold with cusps. As clearly explained in [EM], this is simply a consequence of mixing of the geodesic flow, combined with the effect of strictly negative curvature (actually the fact that an appropriate fattened ε-neighborhood of a closed expanding front remain an ε-neighborhood of the new front under the geodesic flow, uniformly in time). Nevertheless, an exponential error cannot be deduced from mixing, as decay of matrix coefficients only allows to get estimates for smooth square summable functions (for analogous results from the point of view of subgroup actions on homogeneous spaces see [KM]).

The last observation is a rephrasing of the above in the language of number theory. The ring of integers ϑ of an imaginary quadratic number field is a lattice $\mathbb{Z} + \mathbb{Z}\tau$ in $\mathbb{C}$. Consider the quotient elliptic curve T. We refer to the resting points of those disjoint balls (the images of the cusp at infinity, associated to the identity in the ideal class group, under the two-sided cosets $P\backslash\Gamma_k/P$) with sizes bigger than x^{-1} as the *generalized Farey sequence* of order x. They are torsion points of T of order smaller than x, though not all of them. The asymptotics of their cardinality for $x \to \infty$ is well known, and it is the asymptotics of the sum of an Euler-type function (see the proof of Theorem 6.1). We have been unable to find any reference to the following

Corollary 6.4. *The generalized Farey sequence is asymptotically equidistributed in the torus T with respect the to Lebesgue measure, with an error for characteristic functions of good domains which is $\mathcal{O}(x^{-1/2})$.*

Note, again, that a purely geometrical proof of the above comes from the results in [EM] combined with our discussion, but says nothing about error terms.

6.6. *Rankin-Selberg tranfsorm and the Riemann hypothesis.* It has been shown by Sarnak that the asymptotic behavior of the measures supported on the closed horospheres from the cusps is governed by the analytical properties of the relative Eisenstein series [Sa1]. His work treats finite area surfaces with cusps, but clearly generalizes to higher dimensions. The general philosophy is that such asymptotic is controlled by the part of the discrete spectrum of the Laplace-Beltrami operator orthogonal to cusp forms, and for the arithmetic manifold we are concerned with there is no such eigenvalue outside the continuous spectrum (actually Sarnak also showed that the first eigenvalue satisfies $\lambda_1 \geq 3/4$, and he conjectured $\lambda_1 \geq 1$, an analogue of the $1/4$ conjecture for the modular group [Sa2]). Zagier observed that the speed of convergence of analogous measures in the modular orbifold is related to the Riemann hypothesis [Za1,2]. In this section we show how to derive a similar statement in the case of our arithmetic hyperbolic three manifolds, immediate consequence of the Rankin-Selberg unfolding trick.

Let f be a function on the manifold, i.e. a Γ-invariant function on the covering $\mathbb{H}^3$. We denote by $x = (z, y) \in \mathbb{C} \times \mathbb{R}^+$ the coordinates in the u.h.s. model, with the chosen cusp P fixing ∞. The function f is invariant by the parabolic subgroup P, hence it has a Fourier expansion as

$$f(z, y) \;=\; c_0(y) \;+\; \text{non-zero coefficients} \;\;.$$

We note here that the constant coefficient is nothing but the mean value of the function with respect to a measure m_y supported on the closed horosphere at height y, that is the projection of our measures on the hyperbolic manifold (the correspondence reads $y = \mathrm{e}^{-t}$) apart from the normalizing factor area(T).

The *Rankin-Selberg transform* of f is defined to be

$$R(f; s) \;=\; \int_0^\infty c_0(y)\, y^{s-2}\, dy$$

and it is formally the Mellin transform of the function $c_0(y)/y$. Writing down the integral defining c_0 we see an integral over a fundamental domain for the parabolic subgroup, and the "unfolding trick" consists in replacing it by a $P\backslash\Gamma$ orbit of a fundamental domain for Γ, then identified with the manifold $\mathbb{H}/\Gamma$. The fact that f is automorphic leads to the identity, holding in the region of the complex plane where the integrals have sense,

$$R(f; s) \;=\; \int_M f(x)\; E(x; s)\; d\mathrm{vol}_{\mathbb{H}^3}$$

where $E(x;s)$ is the Eisenstein series relative to the cusp, i.e. the sum of the height function y to the power $1+s$ of the $P\backslash\Gamma$ orbit of the point x

$$E(x;s) = \sum_{\gamma\in P\backslash\Gamma} y(\gamma x)^{1+s} .$$

We know quite a lot about these Eisenstein series, thank to the seminal work by Selberg [Se]. We start by recalling some basic facts which hold for any finite volume hyperbolic three-manifold with cusps. A classical reference is the book by Kubota [Ku].

The series converges uniformly, hence is holomorphic, in the region $\mathrm{Re}(s) > 1$, since it is comparable with the Poincaré series and this value gives its critical exponent. It is automorphic, as it is made summing up the images of the height function y under the left cosets of the group modulo the parabolic subgroup, and it is an eigenfunction of the Laplace-Beltrami operator Δ with eigenvalue $\lambda(s) = (1-s^2)$ (here we consider Δ as positive definite). By the general theory of Selberg, it admits a meromorphic continuation in the whole complex plane. The Laplace-Beltrami operator admits a unique self-adjoint extension to $L^2(M, d\mathrm{vol}_{\mathbb{H}})$, and this space splits as the direct sum of three Δ-invariant orthogonal subspaces.

- Θ describes the continuous spectrum, one series for each cusp, and is given by the meromorphic continuation of the Eisenstein series in the line $\mathrm{Re}(s) = 0$. In λ-space this gives the continuous spectrum $[1,\infty)$.

- Θ_0 is the portion of the discrete spectrum coming from the poles of the Eisenstein series. The pole at $s = 1$ has constant residue, given by the constant eigenfunction with eigenvalue $\lambda_0 = 0$. Other poles are a finite number of real $s_i \in (0,1)$, their residues are eigenfunction u_i with eigenvalues $\lambda_i = 1 - s_i^2$.

- $\mathcal{H}_0$ is the part of the discrete spectrum orthogonal to Θ_0, and is given by cusp forms (cusp forms are in some sense orthogonal to the closed horospheres from the cusps, having zero integral along them).

In our arithmetical case we have a more precise information, and we borrow it from [EGM]. For simplicity, we deal with the case of class number one. Consider the function

$$E^*(x;s) = \frac{1}{|\vartheta^\times|}\,\zeta_k(1+s)\,E(x;s)$$

where ζ_k is the Dedekind zeta function of the field k and $|\vartheta^\times|$ is the order of the group of units.

Theorem 6.2. *(Elstrodt, Grunewald and Mennike* [EGM]*) $E^*(x;s)$ has a meromorphic continuation with only a simple pole at $s = 1$, with residue $4\pi^2/|d|$, and zeros at the points $-n$, for $n \geq 2$, and verifies a functional equation relating its values between s and $-s$.*

There follows that $E(x;s)$ is holomorphic in the region $\mathrm{Re}(s) > 0$, apart for the pole in $s = 1$, with residue

$$\frac{4\pi^2}{|d|\zeta_k(2)}$$

and have other poles for $s = \rho - 1$, where ρ runs between the nontrivial zeros of the Dedekind zeta function ζ_k.

This same analytic behavior is inherited by the Rankin-Selberg transform of f, and we can draw some consequences.

Assume first of all that $f(z,y)$ decreases faster than any power of y as y goes to infinity, e.g. it has compact support. This says that the part of integral $\int_1^\infty c_0(y)\, y^{s-2} dy$ is everywhere analytic. Assume also that for small y the zeroth coefficient admits an expansion as

$$c_0(y) = a_0 + a_1\, y^{\alpha_1} + a_2\, y^{\alpha_2} + \ldots$$

Then, looking at the integral $\int_0^1 c_0(y)\, y^{s-2}$, it is a standard fact of Mellin transform theory that $R(f;s)$ has a meromorphic continuation with poles only at the points $s = 1 - \alpha_i$, and residues a_i. In particular, an estimate of the error as

$$c_0(y) - a_0 \;=\; \mathcal{O}(y^{3/2-\varepsilon})$$

for any small positive ε would imply that $R(f;s)$ has no poles in the region $\mathrm{Re}(s) > -1/2$, and hence that the Dedekind zeta function ζ_k has no zeros with real part bigger than 1/2.

For the converse we need some remarks. First, we know that $R(f;s)$ has no pole in the region $\mathrm{Re}(s) > 0$ other than the one at $s = 1$, whose residue gives the right normalization needed to reconstruct the mean value of f with respect to the volume probability measure on the manifold. If we knew that $R(f;\varepsilon + it)$ is a summable function in t, we could shift the integral in the Mellin inversion formula to get

$$c_0(y) - a_0 \;=\; \mathcal{O}(y^{1-\varepsilon}) \;.$$

Second, we know that the Eisenstein series is an eigenfunction of the Laplace-Beltrami operator even after the analytic continuation. In particular we can substitute $\lambda(s)^{-n}\Delta^n E(x;s)$ to $E(x;s)$ in the formula for the Rankin-Selberg transform of f, integrate by parts, and observe that a sufficient degree of differentiability of f ensures the summability needed above. If, moreover, the zeta function had no zeros of real part greater than 1/2, we could shift the integral until any line $\mathrm{Re}(s) = -1/2 + \varepsilon$ and get an error $\mathcal{O}(y^{3/2-\varepsilon})$ for any positive ε.

We are in position to summarize the results. Denote by m_t the probability measures uniformly supported on the closed tori of area $\mathrm{e}^{2t}\mathrm{area}(T)$, and by m the probability measure given by the volume form in the hyperbolic manifold $M_k = \mathbb{H}^3/\Gamma_k$.

Theorem 6.3. *For any compactly supported smooth function f on $\mathbb{H}^3/\Gamma_k$ and any strictly positive ε the following holds*

$$|m_t(f) - m(f)| \;=\; \mathcal{O}(e^{-(1-\varepsilon)t}) \;.$$

Remark. Note that an error $\mathcal{O}\left(\mathrm{e}^{-(1-\varepsilon)t}\right)$ is consistent with the estimates given in the previous paragraphs, where the measures lived in the unit tangent bundle of the manifold. Indeed, an estimate $e(x) = \mathcal{O}(x^{-1/2+\varepsilon})$ in the proof of the Theorem 6.1 would give the same asymptotics.

Theorem 6.4. *The Riemann hypothesis for the Dedekind zeta function ζ_k of an imaginary quadratic number field k is true if and only if for any smooth function f with compact support on $\mathbb{H}^3/\Gamma_k$ and any strictly positive ε*

$$|m_t(f) - m(f)| \;=\; \mathcal{O}(e^{-(3/2-\varepsilon)t}) \;.$$

References

[An] D. Anosov, *Geodesic flows on closed riemannian manifolds with negative sectional curvature*, Proc. Steklov Inst. Math. vol. 90, American Math. Soc., Providence (1967) (English translation).

[BP] R. Benedetti and C. Petronio, *Lectures on hyperbolic geometry*, Universitext, Springer-Verlag (1992).

[Bo] R. Bowen, Hausdorff dimension of quasi-circles, *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 11-25.

[EGM] J. Elstrodt, F. Grunewald and J. Mennicke, Eisenstein series on three-dimensional hyperbolic space and imaginary quadratic number fields, *J. reine angew. Math.* **360** (1985), 160-213.

[EM] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups, *Duke Math. J.* **71** (1993), 181-209.

[Ep] D.B.A. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Paterson and W.P. Thurston, *Word processing in groups*, Jones and Bartlett Publishers, Boston (1992).

[Gh1] E. Ghys, Dynamique des flots unipotents sur les espaces homogènes, *Séminaire Bourbaki, exposé n° 747*, Astérisque **206** (1992), 93-136.

[Gh2] E. Ghys, Holomorphic Anosov flows, *Invent. Math.* **119** (1995), 585-614.

[Gr] M. Gromov, Hyperbolic groups, in *Essays in Group Theory*, S.M. Gersten (Ed), Mathematical Sciences Research Institute Publications **8**, Springer-Verlag (1987), 75-263.

[He] E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen,* Chelsea, New York (1948) , translated as *Lectures on the theory of algebraic numbers*, Graduate Texts in Math. **77**, Springer-Verlag (1981).

[HW] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, Oxford Univ. Press, London (1975).

[KM] D.Y. Kleinbock and G.A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, *Amer. Math. Soc. Transl. (2)* **171** (1996), 141-172.

[Kn] J. Knopfmacher, *Abstract Analytic Number Theory*, North-Holland (1975).

[Ku] T. Kubota, *Elementary theory of Eisenstein series*, Halsted Press (1973).

[La] S.P. Lalley, Renewal theorems in symbolic dynamics, with applications to geodesic flows, noneuclidean tessellations and their fractal limits, *Acta Math.* **163** (1989), 1-55.

[Ma] B. Maskit, *Kleinian groups*, Grundleheren der mathematischen Wissenschaften **287**, Springer-Verlag (1988).

[Mat] P. Mattila, *Geometry of sets and measures in Euclidean spaces, fractals and rectifiability*, Cambridge University Press (1995).

[Mi] J. Milnor, Hyperbolic geometry: the first 150 years, in *The Mathematical Heritage of Henri Poincaré*, Proceedings of Symposia in Pure Mathematics, vol. **39**, American Mathematical Society (1983).

[Ni] J. Nicholls, *The ergodic theory of discrete groups*, Cambridge Univ. Press (1989).

[NS] W.D. Neumann and M. Shapiro, Automatic structures, rational growth, and geometrically finite hyperbolic groups, *Invent. Math.* **120** (1995), 259-287.

[Pa1] S.J. Patterson, Diophantine approximation in Fuchsian groups, *Phil. Trans. Royal Soc. London* A **262** (1976), 527-563.

[Pa2] S.J. Patterson, The limit set of a Fuchsian group, *Acta Math.* **136** (1976), 241-273.

[Pr] D. Preiss, Geometry of measures in $\mathbb{R}^n$: distribution, rectifiability, and densities, *Annals of Math.* **125** (1987), 537-643.

[Ra] J. Ratcliffe, *The foundations of hyperbolic manifolds*, Graduate Texts in Math. **149**, Springer-Verlag (1994).

[Ru] D.J. Rudolph, Ergodic behaviour of Sullivan's geometric measure on a geometrically finite hyperbolic manifold, *Ergod. Th. & Dynam. Sys.* **2** (1982), 491-512.

[Sa1] P. Sarnak, Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series, *Comm. on Pure and Applied Math.* **34** (1981), 719-739.

[Sa2] P. Sarnak, The arithmetic and geometry of some hyperbolic three manifolds, *Acta Math.* **151** (1983), 253-295.

[Se] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, *J. Indian Math. Soc.* **20** (1956), 47-87.

[SRT] X. Saint Raymond and C. Tricot, Packing regularity of sets in n-space, *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 133-145.

[SV] B. Stratmann and S.L. Velani, The Patterson measure for geometrically finite groups with parabolic elements, new and old, *Proc. London Math. Soc. (3)* **71** (1995), 197-220.

[Su1] D. Sullivan, The density at infinity of a discrete group of hyperbolic motions, *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 419-450.

[Su2] D. Sullivan, Discrete conformal groups and measurable dynamics, *Bull. Amer. Math. Soc.* **6** (1982), 57-73.

[Su3] D. Sullivan, Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics, *Acta Math.* **149** (1983), 215-237.

[Su4] D. Sullivan, Entropy, Hausdorff measures old and new, and limit set of geometrically finite Kleinian groups, *Acta Math.* **153** (1984), 259-277.

[Su5] D. Sullivan, Quasiconformal homeomorphisms and dynamics II: Structural stability implies hyperbolicity for Kleinian groups, *Acta Math.* **155** (1985), 243-260.

[Th1] W. Thurston, *Geometry and topology of 3-manifolds*, Princeton notes (1978).

[Th2] W. Thurston, Three dimensional manifolds, Kleinian groups and hyperbolic geometry, in *The Mathematical Heritage of Henri Poincaré*, Proceedings of Symposia in Pure Mathematics, vol. **39**, American Mathematical Society (1983).

[Tr] C. Tricot, Two definitions of fractional dimension, *Math. Proc. Cambridge Philos. Soc.* **91** (1982), 57-74.

[Ve1] A. Verjovsky, Arithmetic, geometry and dynamics in the unit tangent bundle of the modular orbifold, in *Dynamical systems*, Pitman Research Notes in Mathematics Series 285, Santiago 1990, Longman (1993).

[Ve2] A. Verjovsky, Discrete measures and the Riemann hypothesis, *Kodai Math. J.* **17** (1994), 596-608.

[Yu] C.B. Yue, Dimension and rigidity of quasi-Fuchsian representations, *Annals of Math.* **143** (1996), 331-355.

[Za1] D. Zagier, Eisenstein series and the Riemann zeta function, in *Automorphic forms, representation theory and arithmetic*, Tata Institute of Fundamental Research, Bombay 1979, Springer-Verlag (1981), 275-301.

[Za2] D. Zagier, Introduction to modular forms, in *From number theory to physics*, M. Waldschmidt, P. Moussa, J.-M. Luck and C. Itzykson (Eds.), Springer-Verlag (1992), 238-291.