On the bass note of a Schottky group

by

PETER G. DOYLE

AT&T Bell Labs
Murray Hill, NJ, U.S.A.
and
Princeton University
Princeton, NJ, U.S.A.

Introduction

1.1 Classical Schottky groups

Let $C_1, \dots, C_n$ be a collection of circles in the Riemann sphere that bound disjoint open disks $D_1, \dots, D_n$. Note that circles may be tangent, but otherwise they don't intersect. (See Figure 1.) Let F denote the complement of $D_1 \cup \dots \cup D_n$, that is, the closure of the common exterior of $C_1, \dots, C_n$. Suppose that n is even, and that for $i=1, \dots, n/2$ we have specified a Möbius transformation γ_i mapping the exterior of C_{2i-1} to the interior of C_{2i}. The group Γ of Möbius transformations generated by the γ_i's is a Kleinian group with fundamental domain F. It is called a *classical Schottky group*.

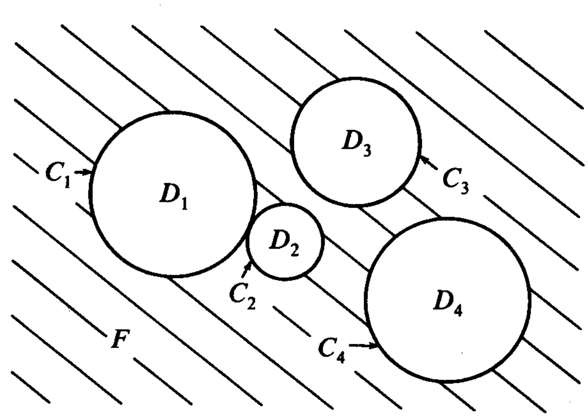


Fig. 1. Circles in the Riemann sphere.

1.2. A lower bound for the bass note

View the Riemann sphere as the boundary of hyperbolic 3-space $\mathbf{H}^3$, and view Γ as a group of isometries of $\mathbf{H}^3$. Consider $\lambda_0(\mathbf{H}^3/\Gamma)$, the bottom of the spectrum of the Laplacian $\Delta = -\operatorname{div}\operatorname{grad}$ on $\mathbf{H}^3/\Gamma$ (note the minus sign). Except in trivial cases, $\lambda_0(\mathbf{H}^3/\Gamma)$ is a bona fide eigenvalue; we call it the "lowest eigenvalue" of $\mathbf{H}^3/\Gamma$, or of Γ. Physically, it is the square of the frequency of the bass note of $\mathbf{H}^3/\Gamma$—the lowest note you would hear if you hit $\mathbf{H}^3/\Gamma$ with a mallet.

We will prove the conjecture of Phillips and Sarnak [13] that for any classical Schottky group Γ,

$$\lambda_0(\mathbf{H}^3/\Gamma) \geqslant L_2 > 0,$$

where L_2 is some universal constant.

1.3. Implications

Let $\delta(\Gamma)$ denote the exponent of convergence of the Poincaré series

$$\sum_{\gamma} \exp(-s \cdot \varrho(z, \gamma w)),$$

where z, w are fixed points of $\mathbf{H}^3$ and $\varrho(a, b)$ is the hyperbolic distance from a to b. Let $d(\Lambda(\Gamma))$ denote the Hausdorff dimension of the limit set $\Lambda(\Gamma)$. By work of Patterson [11, 12] and Sullivan [17, 18],

$$\delta(\Gamma) = d(\Lambda(\Gamma)),$$

and

$$\lambda_0(\mathbf{H}^3/\Gamma) = \delta(\Gamma)\,(2-\delta(\Gamma)),$$

as long as $\delta(\Gamma)>1$. Thus our universal lower bound for λ_0 implies a universal upper bound $U_2<2$ for $\delta(\Gamma)$ and $d(\Lambda(\Gamma))$.

1.4. Some history

The problem of finding an upper bound for $\delta(\Lambda)$ can be traced back to Schottky [16]. Burnside [5] conjectured $\delta \leqslant 1$; this was disproved by Myrberg [9, 10]. Akaza [1, 2] gave examples with $\delta = 1.5$. Sarnak [15] and Phillips proved the existence of examples with

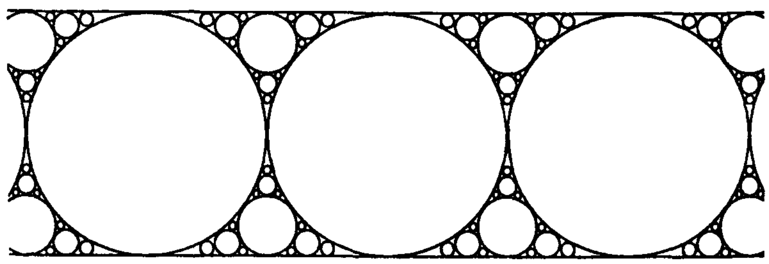


Fig. 2. The Apollonian packing.

$\delta \geqslant 1.75$. Phillips and Sarnak [13] conjectured the existence of a universal upper bound $U_2 < 2$, and proved the analogous result in higher dimensions. Brooks [3, 4] proved the conjecture for the special class of groups for which the disks $D_1, \ldots, D_n$ are a subset of the disks of the Apollonian packing. (See Figure 2.) Phillips, Sarnak, and Brooks have suggested that the supremum of $\delta(\Gamma)$ over all such "Appolonian" Schottky groups should equal the supremum over all classical Schottky groups, but this is not known.

1.5. Rayleigh's cutting method

To get a lower bound for the bass note of $\mathbf{H}^3/\Gamma$, we will apply a classical method from physics called Rayleigh's cutting method. This method was introduced by Rayleigh [14] as a way of estimating the bass note of a Helmholtz resonator. The idea is to cut the system into pieces whose lowest eigenvalue can be estimated, and then observe that if the lowest eigenvalue of each of the pieces is $\geqslant c$, then the lowest eigenvalue of the original system is $\geqslant c$.

1.6. Infinitely skinny tubes that grow more or less exponentially

In applying Rayleigh's method, our approach will be to cut our manifold into an infinite number of infinitely skinny tubes. A crude estimate shows that we can get a lower bound for the lowest eigenvalue of a tube as long as its cross-section grows more or less exponentially. Thus to get a lower bound for the lowest eigenvalue of a manifold it suffices to show that it can be cut into infinitely skinny tubes in such a way that the cross-section of each and (almost) every tube grow more or less exponentially. This result complements the known fact that the rate of exponential growth of a manifold gives an upper bound for the bass note. Here we have a sort of converse: A definite rate of exponential growth gives a lower bound for the bass note, provided that the growth can be "correlated" by means of tubes.

1.7. Plan

In section 2, we will make precise the notion of cutting a manifold into tubes, and show how a cutting into tubes gives a lower bound for the bass note. In section 3, we will show how to cut $\mathbf{H}^3/\Gamma$ into tubes, so as to prove the existence of a universal lower bound for $\lambda_0(\mathbf{H}^3/\Gamma)$.

2. Cutting a manifold into tubes

2.1. The lowest eigenvalue λ_0

Our goal is a method for getting lower bounds for the lowest eigenvalue λ_0 of a manifold M. Among several equivalent definitions for λ_0, the following will be most convenient to work with:

Definition. Let M be a non-compact complete n-dimensional Riemannian manifold with boundary. Let $TF(M)$ ("test functions") denote the set of C^∞ functions $u: M \to [0, \infty)$ that have compact support and do not vanish identically. We define $\lambda_0(M)$ as the infimum over $TF(M)$ of the Rayleigh quotient

$$\frac{\displaystyle\int_M |\operatorname{grad} u|^2}{\displaystyle\int_M u^2}$$

2.2 The cutting method

To estimate λ_0 we will use *Rayleigh's cutting method* (see Rayleigh [14], Maxwell [8]). This method belongs to a class of related methods known collectively as *Rayleigh's short-cut method*. For a general discussion of the short-cut method see Doyle and Snell [6]. (In the references given here, Rayleigh's method is applied to conductance problems; the generalization from conductance problems to bass-note problems is straightforward.)

The cutting method is based on *Rayleigh's cutting law,* one form of which is the following:

Proposition. *If M is obtained by gluing together parts of the boundary of another manifold M_{cut}, then*

$$\lambda_0(M) \geq \lambda_0(M_{\text{cut}}).$$

Proof. This follows from the definition of λ_0 that we have adopted. □

2.3. Cutting into tubes

Our method for getting lower bounds for λ_0 is based on cutting M up into infinitely skinny tubes. In other applications it will be most convenient to consider tubes modeled on $[0, \infty)$ that begin inside the manifold and run out to infinity in one direction. Here, we will consider only tubes modeled on $\mathbf{R}$ that run out to infinity in both directions. There are two excuses for this. The first is that these doubly-infinite tubes are best for the specific application we have in mind. The second is that you can always get a singly-infinite tube by folding a doubly-infinite tube in half.

Definition. A *cutting of M into tubes* consists of a measure space T (to index the tubes) and measurable maps

$$f: T \times \mathbf{R} \to M,$$
$$\sigma: T \times \mathbf{R} \to [0, \infty)$$

(to show how they run, and tell their cross-section) such that

(i) f pushes the measure $\sigma(\tau, x) d\tau dx$ on $T \times \mathbf{R}$ over onto the Riemannian volume measure on M, that is, for any integrable function u on M, we have

$$\int_M u = \int d\tau \int u(f(\tau, x))\, \sigma(\tau, x)\, dx.$$

(This makes precise the notion that σ tells the cross-section of the tube.)

(ii) for almost all τ, the curve $f(\tau, \cdot)$ is piecewise C^1, parametrized by arc length, and proper. (The tubes may zig-zag a little, but must run out to the boundary.)

(iii) for almost all τ we have $0 < \int_a^b \sigma(\tau, x)\, dx < \infty$ whenever $-\infty < a < b < \infty$. (The tubes are neither too fat nor too thin.)

2.4. Inhomogeneous strings

Our purpose in cutting M into tubes is to reduce our n-dimensional eigenvalue problem to a bunch of 1-dimensional inhomogeneous string problems (Sturm-Liouville problems). An inhomogeneous string is described by two functions σ and ϱ, telling its cross-section and its density as a function of length. When we come to consider the strings that arise from our tubes, we will want to set $\varrho=\sigma$, to indicate that the mass per unit length is proportional to the cross-section. For the moment we will distinguish σ from ϱ, not so much for the sake of generality as to make clearer their differing roles.

Definition. If $\sigma:\mathbf{R}\to[0,\infty)$ is measurable, and

$$0<\int_a^b \sigma(\tau,x)\,dx<\infty \quad\text{whenever}\quad -\infty<a<b<\infty,$$

we will say that σ is *neither too big nor too small.* Thus (iii) above states that for almost all τ, $\sigma(\tau,\cdot)$ is neither too big nor too small.

Definition. Let

$$\sigma,\varrho:\mathbf{R}\to[0,\infty)$$

be measurable, with σ and ϱ neither too big nor too small. Then we define $\lambda_0(\sigma,\varrho)$ as the infimum over $TF(\mathbf{R})$ of the Rayleigh quotient

$$\frac{\displaystyle\int \sigma\left(\frac{d\varphi}{dx}\right)^2 dx}{\displaystyle\int \varrho\varphi^2\,dx}.$$

2.5. Using piecewise differentiable test functions

Our method from converting information about the tubes into information about M will involve pulling back a test function on M to each of the tubes. The functions on the tubes that we will obtain in this way will not necessarily be smooth, because we are allowing the tubes to zig-zag. It would be possible to work only with cuttings into smooth tubes, but we prefer to allow ourselves the extra latitude in cutting, and pay the price by smoothing the pulled-back test functions, rather than the tubes themselves.

LEMMA. *If $\varphi:\mathbf{R}\to[0,\infty)$ is piecewise C^1, has compact support, and doesn't vanish identically, then*

$$\frac{\displaystyle\int \sigma\left(\frac{d\varphi}{dx}\right)^2 dx}{\displaystyle\int \varrho\varphi^2\,dx}\geqslant\lambda_0(\sigma,\varrho).$$

Proof. Choose $b:\mathbf{R}\to[0,\infty)$ to be smooth, with support in the interval $[-1,1]$ and $\int_{-1}^{1} b(x)\,dx=1$, and convolve φ with $(1/\varepsilon)\,b(x/\varepsilon)$. The result is a member of $TF(\mathbf{R})$ whose Rayleigh quotient approaches that of φ as $\varepsilon\to 0$. □

2.6. What the tubes tell us about the original space

We now show how to turn information about the lowest eigenvalues of the tubes into information about M. Note that in treating the tubes as inhomogeneous strings, we set $\varrho=\sigma$, as promised.

PROPOSITION. *Suppose M can be cut into tubes is such a way that for almost all τ,*

$$\lambda_0(\sigma(\tau,\cdot),\sigma(\tau,\cdot))\geqslant\lambda.$$

Then $\lambda_0(M)\geqslant\lambda$.

Proof. Let u be an element of $TF(M)$. Then

$$\begin{aligned}\int_M|\operatorname{grad} u|^2 &= \int d\tau\int \sigma(\tau,x)\,|\operatorname{grad} u(f(\tau,x))|^2\,dx\\ &\geqslant \int d\tau\int \sigma(\tau,x)\left(\frac{\partial(u\circ f)}{\partial x}\right)^2 dx\end{aligned}$$

(note the use of parametrization by arc length)

$$\geqslant \lambda\int d\tau\int \sigma(\tau,x)\,((u\circ f)(\tau,x))^2\,dx$$

(note the need for properness of the tubes)

$$=\lambda\int_M u^2.$$

□

2.7. Twiddling σ and ϱ

In estimating λ_0 for our tubes, we will want to know what effect altering σ and ϱ by some bounded factor will have on $\lambda_0(\sigma,\varrho)$. There are two reasons for this. The first is that when all we are after is a very conservative lower bound for λ_0, we may want to hack the space into tubes in such a way that we have only very gross information about the growth of the cross-section of the tubes. The second reason is that to get a lower bound for $\lambda_0(\sigma, \varrho)$, it may be convenient to assume that σ and ϱ are pretty smooth.

LEMMA. *For any K $(0<K<\infty)$,*

$$\lambda_0(K\sigma, \varrho) = \lambda_0\left(\sigma, \frac{1}{K}\varrho\right) = K\lambda_0(\sigma, \varrho). \qquad \square$$

LEMMA. *Let $\tilde\sigma, \tilde\varrho: \mathbf{R} \to [0, \infty)$ be measurable, with $\tilde\sigma$ and $\tilde\varrho$ neither too big nor too small. If*

$$\sigma \geq \tilde\sigma$$

and

$$\varrho \leq \tilde\varrho,$$

then

$$\lambda_0(\sigma, \varrho) \geq \lambda_0(\tilde\sigma, \tilde\varrho). \qquad \square$$

LEMMA. *Let $\tilde\sigma, \tilde\varrho: \mathbf{R} \to [0, \infty)$ be measurable, with $\tilde\sigma$ and $\tilde\varrho$ neither too big nor too small. Suppose that for constants K_1, K_2 $(0<K_1, K_2<\infty)$, we have*

$$\sigma \geq \frac{1}{K_1}\tilde\sigma$$

and

$$\varrho \leq K_2 \tilde\varrho.$$

Then

$$\lambda_0(\sigma, \varrho) \geq \frac{1}{K_1 K_2}\lambda_0(\tilde\sigma, \tilde\varrho). \qquad \square$$

2.8. Estimating λ_0 for a string

The great thing about a string is that it is easy to get lower bounds for the lowest eigenvalue by exhibiting an appropriate superharmonic function. Of course the same thing works in higher dimensions, but it is easier to cook up a superharmonic function on the line than in n-space.

Rather than appeal to established theory, we will find it simplest to concoct our own proof of how to get a lower bound for λ_0 out of a suitable superharmonic function. This proof, which is based on some ideas of Holland [7], may seem a little mysterious. Its advantage is that it works directly with the definition of λ_0 that we have been using, rather than the definition of λ_0 as some kind of eigenvalue.

LEMMA. *If σ is piecewise C^1, and if there is a C^2 function $\varphi_0\colon \mathbf{R}\to(0,\infty)$ that satisfies*

$$-\frac{d}{dx}\left(\sigma\frac{d\varphi_0}{dx}\right)\geqslant\lambda\varrho\varphi_0$$

then $\lambda_0(\sigma,\varrho)\geqslant\lambda$.

Proof. Setting

$$\varphi_0=e^{-\psi_0},$$

we find that

$$\frac{d}{dx}\left(\sigma\frac{d\psi_0}{dx}\right)-\sigma\left(\frac{d\psi_0}{dx}\right)^2\geqslant\lambda\varrho.$$

Let φ be in $TF(\mathbf{R})$. Then

$$\int\sigma\left(\frac{d\varphi}{dx}\right)^2dx=\int\left[\frac{d}{dx}\left(\sigma\frac{d\psi_0}{dx}\right)\varphi^2+\sigma\frac{d\psi_0}{dx}\frac{d}{dx}(\varphi_2)+\sigma\left(\frac{d\varphi}{dx}\right)^2\right]dx$$

(integration by parts)

$$=\int\left[\frac{d}{dx}\left(\sigma\frac{d\psi_0}{dx}\right)+2\sigma\frac{d\psi_0}{dx}\left(\frac{1}{\varphi}\frac{d\varphi}{dx}\right)+\sigma\left(\frac{1}{\varphi}\frac{d\varphi}{dx}\right)^2\right]\varphi^2\,dx$$

$$=\int\left[\frac{d}{dx}\left(\sigma\frac{d\psi_0}{dx}\right)+\sigma\left(\frac{d\psi_0}{dx}+\left(\frac{1}{\varphi}\frac{d\varphi}{dx}\right)\right)^2-\sigma\left(\frac{d\psi_0}{dx}\right)^2\right]\varphi^2\,dx$$

(completing the square)

$$\geqslant\int\left[\frac{d}{dx}\left(\sigma\frac{d\psi_0}{dx}\right)-\sigma\left(\frac{d\psi_0}{dx}\right)^2\right]\varphi^2\,dx$$

$$\geqslant\lambda\int\varrho\varphi^2\,dx. \qquad\square$$

2.9. Tubes that grow more or less exponentially

If the cross-section σ of a tube does not grow exponentially in one direction or the other, there is no hope that $\lambda_0(\sigma, \sigma)$ will be positive. The reason is that $\lambda_0(\sigma, \sigma)$ measures the exponential rate of decay 36the heat kernel for the tube, and since the flow of heat along the tube is slow and no heat is lost, the temperature can't decay exponentially unless there is an exponential amount of material over which to distribute the heat.

On the other hand, if σ is growing exponentially with a certain minimum rate, and if it doesn't waver too much, then we can get a positive lower bound for $\lambda_0(\sigma, \sigma)$. The bigger the minimum growth rate and the smaller the amount of wavering, the better the lower bound will be.

LEMMA. *Suppose that*

$$\sigma(x) = e^{A(x)},$$

where $A: \mathbf{R} \to \mathbf{R}$ *is piecewise* C^1 *and*

$$\frac{dA}{dx} \geqslant a > 0.$$

Then

$$\lambda_0(\sigma, \sigma) \geqslant \frac{a^2}{4}.$$

Proof. Let

$$\varphi_0(x) = e^{-\frac{a}{2}x}.$$

Then

$$-\frac{d}{dx}\left(\sigma \frac{d\varphi_0}{dx}\right) = \frac{a}{2}\left(\frac{dA}{dx} - \frac{a}{2}\right)\sigma\varphi_0 \geqslant \frac{a^2}{4}\sigma\varphi_0,$$

so

$$\lambda_0(\sigma, \sigma) \geqslant \frac{a^2}{4}. \qquad \square$$

Definition. If $f: \mathbf{R} \to (0, \infty)$ is measurable and satisfies

$$\frac{1}{K} e^{A(x)} \leqslant \sigma(x) \leqslant K e^{A(x)}, \quad 0 < K < \infty$$

for some piecewise C^1 function $A: \mathbf{R} \to \mathbf{R}$ with

$$\frac{dA}{dx} \geqslant a > 0$$

then we say that f grows *more or less exponentially*, with *rate* a and *factor* K.

PROPOSITION. *Suppose for a cutting of M into tubes we can find constants a, K such that for almost all τ, $\sigma(\tau, \cdot)$ grows more or less exponentially, with rate a and factor K. Then*

$$\lambda_0(M) \geqslant \frac{a^2}{4K^2}. \qquad \square$$

2.10. Growth criteria

It will be handy to have some simple criteria to tell when a function is growing more or less exponentially.

PROPOSITION. *Let $f: \mathbf{R} \to (0, \infty)$ be measurable, and suppose that for some sequence*

$$-\infty \leftarrow \ldots < x_{-1} < x_0 < x_1 < \ldots \to \infty$$

we have

$$f(x_i) \leqslant f(x) \leqslant f(x_{i+1}), \quad x \in [x_i, x_{i+1}].$$

Let

$$a = \inf_i \frac{\log f(x_{i+1}) - \log f(x_i)}{x_{i+1} - x_i}$$

and

$$K = \sup_i \frac{f(x_{i+1})}{f(x_i)}.$$

If $a>0$ and $K<\infty$ then f grows more or less exponentially, with rate a and factor K. $\square$

COROLLARY. *In particular, if*

$$|x_{i+1} - x_i| \leqslant L < \infty$$

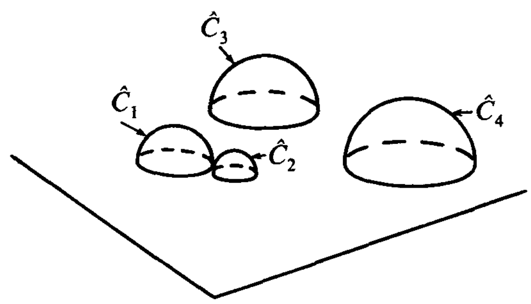


Fig. 3. Planes in hyperbolic 3-space.

and

$$1 < r \leqslant \frac{f(x_{i+1})}{f(x_i)} \leqslant R < \infty$$

then f grows more or less exponentially, with rate $(\log r)/L$ *and factor R.* □

3. Cutting up $\mathbf{H}^3/\Gamma$

3.1. The geometrical problem

Thanks to Rayleigh's method, we have now reduced our problem to showing that $\mathbf{H}^3/\Gamma$ can be cut into tubes whose cross-section grows more or less exponentially. This is a problem of pure geometry. The solution is straightforward and elementary, though a bit complicated. It relies mainly on a crude estimate of how densely you can pack circles on the Riemann sphere.

3.2. Cutting up the fundamental domain

There is an obvious fundamental domain $\hat{F}$ for the action of Γ on $\mathbf{H}^3$. Let $\hat{C}_1, \dots, \hat{C}_n$ be the planes in $\mathbf{H}^3$ that meet the sphere at infinity in the circles $C_1, \dots, C_n$, let $\hat{D}_1, \dots, \hat{D}_n$ be the corresponding open half-spaces that meet the sphere at infinity in the disks $D_1, \dots, D_n$, and let $\hat{F}$ be the complement of $\hat{D}_1 \cup \dots \cup \hat{D}_n$ in $\mathbf{H}^3$. (See Figure 3.) Note that whereas we described F as the *exterior of* $C_1, \dots, C_n$, we can best describe $\hat{F}$ as the *interior* of $\hat{C}_1, \dots, \hat{C}_n$.

The manifold $\mathbf{H}^3/\Gamma$ is obtained from $\hat{F}$ by gluing the boundary components $\hat{C}_1, \dots, \hat{C}_n$ together in pairs. If we cut $\mathbf{H}^3/\Gamma$ apart along the $n/2$ surfaces along which it was glued, we recover $\hat{F}$. By the cutting law, any lower bound for $\lambda_0(\hat{F})$ is also a lower bound for $\lambda_0(\mathbf{H}^3/\Gamma)$. Hence instead of cutting up $\mathbf{H}^3/\Gamma$ we will work on cutting up $\hat{F}$. We will forget all about the pairings, and drop the assumption that n is even. Here's what we will prove:

THEOREM. *Let $\hat{F}$ be the manifold (with boundary) corresponding to the finite collection of circles $C_1, \ldots, C_n$ in the Riemann sphere. Then*

$$\lambda_0(\hat{F}) \geq L_2 > 0$$

for some universal constant L_2.

3.3. The picture in the upper half-space model of $\mathbf{H}^3$

To describe how to cut up $\hat{F}$, it is convenient to use the upper half-space model for $\mathbf{H}^3$. In this model, we represent $\mathbf{H}^3$ as the upper half-space

$$\{(x, y, z) | z > 0\},$$

outfitted with the metric

$$ds = \frac{1}{z}\sqrt{(dx^2+dy^2+dz^2)}.$$

The sphere at infinity appears as the (x, y)-plane, together with a point ∞ at infinity.

In this model, which was already tacitly used in Figure 3, the circles $C_1, \ldots, C_n$ appear as circles, possibly degenerating into straight lines, in the (x, y)-plane. The planes $\hat{C}_1, \ldots, \hat{C}_n$ appear as hemispheres, possibly degenerating into planes, that are perpendicular to the (x, y)-plane.

3.4. The case of no circles

Suppose first that there are no circles at all ($n=0$). Then $\hat{F}$ is all of $\mathbf{H}_3$, and we can make all of our tubes vertical (parallel to the z-axis). The cross-section of the tubes grows exponentially as a function of the length along the tube, thanks to the factor of $1/z$ in the metric: Let $dx\,dy$ denote the cross-section of a tube at height 1, that is, where it passes through the surface $z=1$. Then the cross-section at height h is

$$\frac{1}{h^2}\,dx\,dy.$$

But moving from height 1 to height h corresponds to going a distance

$$d = -\log h$$

along the tube, so the cross-section as a function of distance along the tube is

$$e^{2d}\, dx\, dy.$$

Applying the final proposition of section 2, we conclude that

$$\lambda_0(\mathbf{H}^3) \geqslant 1.$$

In fact,

$$\lambda_0(\mathbf{H}^3) = 1,$$

so by cutting we haven't thrown anything away.

3.5. The case of one circle

Suppose that there is only one circle. Moving a point of the circle to ∞ and normalizing, we may assume that the circle coincides with the x-axis, and that $\hat{F}$ is the domain

$$\{(x, y, z) | y \geqslant 0, z > 0\}.$$

Again, we can cut $\hat{F}$ into vertical tubes; again, we conclude that

$$\lambda_0(\hat{F}) \geqslant 1;$$

again, the correct answer is

$$\lambda_0(\hat{F}) \geqslant 1.$$

3.6. The case of two tangent circles

Suppose that there are two circles that are tangent. Moving the point of tangency to ∞ and normalizing, we may assume that the circles coincide with the lines $y=0$ and $y=1$, and that $\hat{F}$ is the domain

$$\{(x, y, z) | 0 \leqslant y \leqslant 1, z > 0\}.$$

Once again, we can cut into vertical tubes, etc.

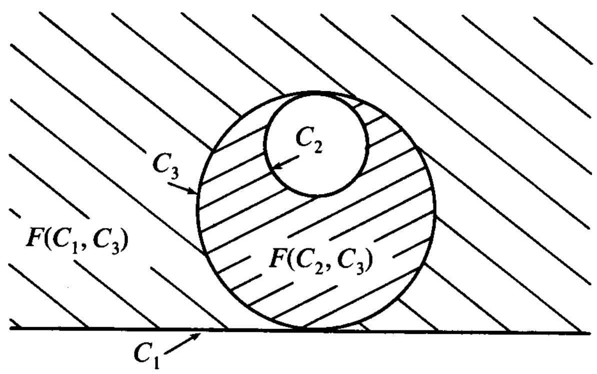


Fig. 4. Introducing a third circle.

3.7. The case of two non-tangent circles

It may seem that we have gone as far as we can go with verticla tubes, but this isn't quite true. Suppose that there are two circles C_1 and C_2 that aren't tangent. Pick a point on C_1, move it to ∞, and normalize as in the one-circle case above. C_2 is a bona fide circle in the half-plane

$$\{(x, y) | y \geqslant 0\}.$$

Now construct a circle C_3 between C_1 and C_2 that is tangent to C_1 and C_2. (See Figure 4.) Decompose $\hat{F}$ into the part $\hat{F}(C_1, C_3)$ between $\hat{C}_1$ and $\hat{C}_3$, and the part $\hat{F}(C_2, C_3)$ between $\hat{C}_2$ and $\hat{C}_3$. If we move the point of tangency of C_1 and C_3 to ∞, we can cut $\hat{F}(C_1, C_3)$ into vertical tubes. Similarly, if we move the point of tangency of C_2 and C_3 to ∞, we can cut $\hat{F}(C_2, C_3)$ into vertical tubes. As for the boundary between $\hat{F}(C_1, C_3)$ and $\hat{F}(C_2, C_3)$ we can simply ignore it, since it has measure 0. In this way, we get a cutting of $\hat{F}$ into tubes, some of which are vertical from one point of view and some from another. Once again, we conclude that

$$\lambda_0(\hat{F}) \geqslant 1,$$

whereas in fact

$$\lambda_0(\hat{F}) = 1.$$

Taking a second look at the argument we have just gone through, we see that it can be simplified as follows: Cut $\hat{F}$ into the two pieces $\hat{F}(C_1, C_3)$ and $\hat{F}(C_2, C_3)$. We already know that

$$\lambda_0(\hat{F}(C_1, C_3)) \geqslant 1$$

and

$$\lambda_0(\hat{F}(C_2, C_3)) \geqslant 1.$$

By the cutting law,

$$\begin{aligned}\lambda_0(\hat{F}) &\geqslant \lambda_0(\hat{F}(C_1, C_3) \underset{\text{disjoint}}{\cup} \hat{F}(C_2, C_3)) \\ &= \min(\lambda_0(\hat{F}(C_1, C_3)), \lambda_0(\hat{F}(C_2, C_3))) \\ &\geqslant 1.\end{aligned}$$

3.8. The case of three or more circles

From now on we will assume that there are three or more circles. We can assume that some pair of circles are tangent: If not, pick one of the circles (say C_1) and enlarge it until it hits another of the circles. Call the enlarged circle C_1' and cut $\hat{F}$ into the two pieces

$$\hat{F}(C_1, C_1')$$

and

$$\hat{F}(C_1', C_2, \ldots, C_n).$$

By the cutting law,

$$\begin{aligned}\lambda_0(\hat{F}) &\geqslant \min(\lambda_0(\hat{F}(C_1, C_1')), \lambda_0(\hat{F}(C_1', C_2, \ldots, C_n))) \\ &\geqslant \min(1, \lambda_0(\hat{F}(C_1', C_2, \ldots, C_n))).\end{aligned}$$

Hence it suffices to find a lower bound for

$$\lambda_0(\hat{F}(C_1', C_2, \ldots, C_n)).$$

The argument we just went through shows that without loss of generality, we may assume that C_1 and C_2 are tangent. Moving the point of tangency to ∞ and normalizing, we may assume that C_1 and C_2 are the lines $y=0$ and $y=1$. (See Figure 5.)

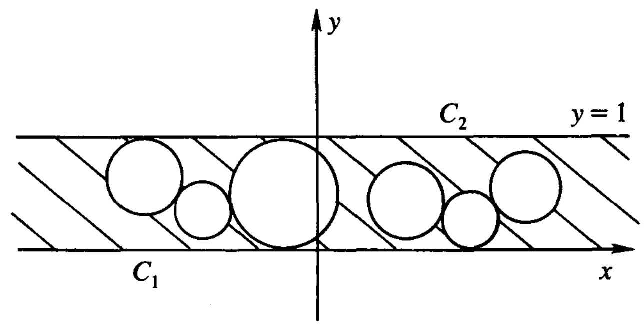


Fig. 5. Normalized configuration.

3.9. The obvious strategy doesn't work

Take a look at the domain $\hat{F}$. It lies above the (x, y)-plane, between the planes $y=0$ and $y=1$, and on or above the hemispheres $\hat{C}_3, \ldots, \hat{C}_n$. (From now on, all geometrical terms used will be Euclidean by default.)

How shall we cut this space into tubes? The obvious thing is to start by cutting the space apart along the cylinders that lie above the circles $C_3, \ldots, C_n$. This yields $n-2$ domains that are congruent from the hyperbolic point of view, along with a residual piece. The residual piece can be cut into vertical tubes, so all we have to do is show how to cut the hyperbolically congruent domains into tubes that grow more or less exponentially. Unfortunately, this can't be done; the bottom of the spectrum of one of these domains is 0. The problem is that the intersection of one of these domains with the plane $z=\varepsilon$ has Euclidean area on the order of ε^2, hence hyperbolic area on the order of $\varepsilon^2/\varepsilon^2=1$. (See Figure 6.) Since there isn't an exponential room at infinity, there is no way to cut the domain into tubes that grow exponentially.

Notice that in higher dimensions, the diffculty we have just encountered disappears. In the upper half-space model of $\mathbf{H}^4$, the Euclidean measure of the intersection of an analogous domain with the hyperplane at height ε is still on the order of ε^2, only now the hyperbolic measure is $\varepsilon^2/\varepsilon^3=1/\varepsilon$. Thus there is plenty of room at infinity, and it is a simple matter to construct an appropriate cutting into tubes.

Back in $\mathbf{H}^3$, we recognize that because of the difficulty we have just discussed, we must allow at least some of the tubes that begin over a given hemisphere to spread out beyond the corresponding circle. In so doing, they will most likely stray out over other hemispheres, and confusion is liable to result. Our task will be to avoid this confusion.

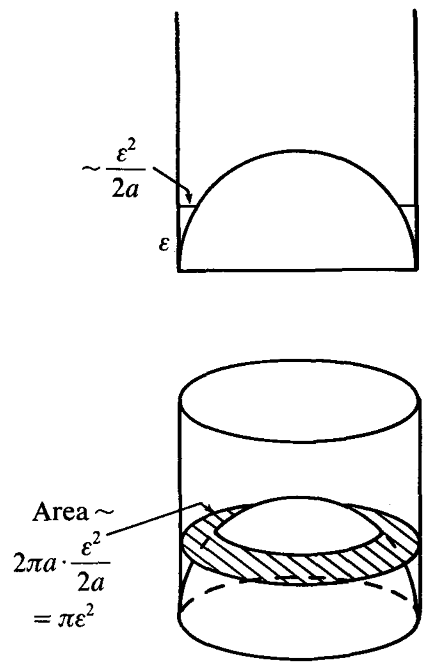


Fig. 6. No room at infinity.

3.10. Keeping the tubes more or less vertical

Our strategy for laying out the tubes will be to make them all drop down more or less vertically toward the (x, y)-plane. To get started, we will make all of the tubes vertical above height $z=1$. (See Figure 7.) This is okay because all of the circles have radius $\leq 1/2$, so no hemisphere protrudes above $z=1/2$. Below $z=1$, we will be forced to bend the tubes. In so doing, we would like to make sure that the tubes stay more or less vertical, in the sense that the angle that the tubes make with the vertical stays bounded above by some universal constant $<\pi/2$. This way, the true cross-section of a tube will be more or less the same as its horizontal cross-section, that is, the area of its intersection with the plane z=const. Similarly, the length (either hyperbolic or Euclidean) of a section of tube will be more or less the same as that of its projection onto the z-axis. Thus, instead of worrying about the growth of the cross-section as a function of length along the tube, we can worry about the growth of the horizontal cross-section as a function of height above the (x, y)-plane. In particular, if we can arrange things so that the tubes stay more or less vertical, and so that the horizontal cross-section of every tube never decreases, and increases by a definite factor every time the vertical distance

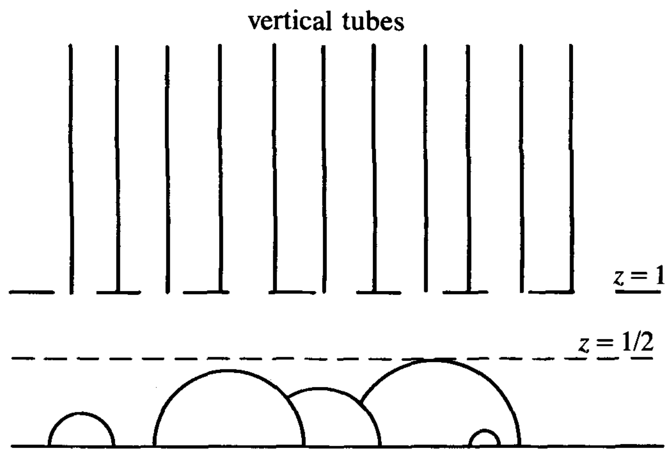


Fig. 7. Cutting vertically down to height 1.

to the (x, y)-plane drops by a factor of 2 (or 8, or whatever), then the tubes will be growing more or less exponentially.

Of course near the tops of the hemispheres there is no way to keep the tubes more or less vertical. This is a real nuisance. To get around this difficulty, we will outfit each hemisphere with a duncecap, as shown in Figure 8. A hemisphere of radius a gets a cap whose apex is at height $z=2a$ and whose brim rests along the parallel at height $z=a/2$. Call the region lying on or above the spruced-up hemispheres G. If we can cut G up into tubes that grow more or less exponentially, then the same goes for $\hat{F}$. (See Figure 9.) One way to see this is to consider that the obvious "squash the hats" map from G onto $\hat{F}$ taking (x, y, z) to $(x, y, z-f(x, y))$ is a quasi-isometry from the point of view of the hyperbolic metric. So from now on we will concentrate on cutting up G, rather than $\hat{F}$.

3.11. Working our way down

In extending the tubes down toward the (x, y)-plane, we will proceed one step at a time. On the first step, we will go from height 1 to height 1/2, on the second step we will go

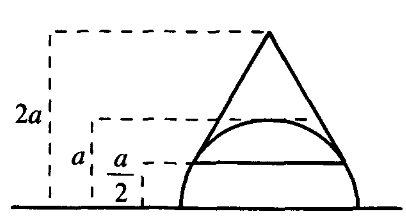


Fig. 8. Duncecap dimensions.

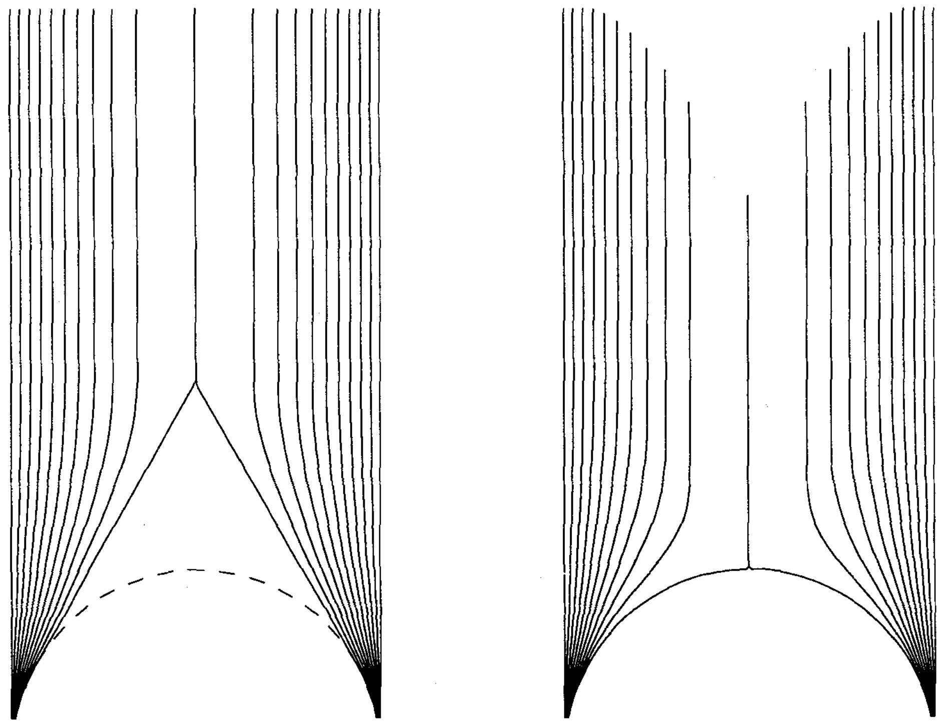


Fig. 9. Duncecaps.

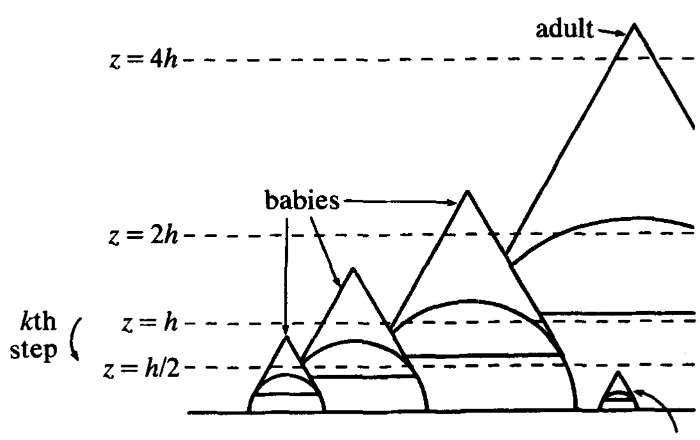


Fig. 10. Relevant hemispheres.

from height 1/2 to height 1/4, etc. On the kth step we go from height $h=1/2^{k-1}$ down to height $h/2=1/2^k$. As we do this, we will only need to consider hemispheres of radius $>h/4$, since these are the only ones that protrude into the region we are cutting up. So let us define the *relevant hemispheres* to be those hemispheres (other then C_1 and C_2) whose radius is $>h/4$. (See Figure 10.) Note that we do not consider a hemisphere relevant if its radius is exactly $h/4$, so that the apex of its hat is at height exactly $h/2$. Of course which hemispheres are relevant depends on which step we're working on.

It is precisely because we only have to consider big hemispheres that this one-step-at-a-time approach will allow us to avoid the confusion that we anticipated from sending tubes starting above one hemisphere out over other hemispheres. There may well be other hemispheres beneath where we send our tubes, but they're irrelevant because they're small and don't get in the way. When we're down so that our distance to the plane is comparable to their radii, *then* we'll worry about those little hemispheres.

3.12. Associating pieces of the plane to relevant hemispheres

In going from height h down to height $h/2$, we proceed by dividing the (x, y)-plane into pieces, one for each relevant hemisphere, together with a residual piece. We consider each piece separately: The tubes that begin over a piece remain over that same piece throughout this stage of their descent. This is precisely what we tried to do before, only now the piece we associate to each hemisphere extends beyond the base of the hemisphere, and the hemispheres we have to consider and the pieces we associate to them change at each step. As before, we will make the tubes vertical over the residual piece. What we have to figure out is how to associate a piece to each relevant hemisphere, and how to cut up the space above it. Of course these two problems are really the same: What piece we associate to a hemisphere depends on how we plan to cut up the space above it.

3.13. Taking care of the babies

Among the relevant hemispheres, we will distinguish those whose hats intersect the region we are trying to cut up, i.e. whose radius is $\leqslant 2h$, and refer to them as *babies*. Each hemisphere spends three steps as a baby: If the radius is a, the hemisphere will be a baby when $h/4<a\leqslant h/2$, when $h/2<a\leqslant h$, and when $h<a\leqslant 2h$. During each of these steps, the piece of the (x, y)-plane that we associate to the hemisphere will consist of the base of the hemisphere, and no more. By deflecting the tubes radially, we can arrange

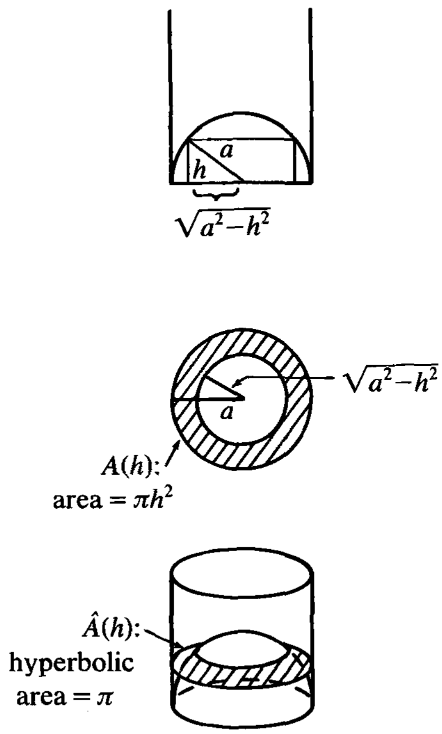


Fig. 11. No room at infinity (reprise).

things so that the tubes stay more or less vertical, and so that the horizontal cross-section of every tube never decreases, and increases by a definite factor between the beginning of the first step and the end of the third. (Look back at Figure 9.) This takes care of the babies.

3.14. Room to grow

Consider a hemisphere $\hat{C}$, associated to the circle C and the disk D. Let the radius of C be a. We assume that $\hat{C}$ is relevant, and not a baby, that is, that $a>2h$. At height h, the tubes trapped over $\hat{C}$ cover an annulus $A(h)$ in the plane of outer radius a and inner radius $\sqrt{a^2-h^2}$. (See Figure 11.) The Euclidean area of $A(h)$ is

$$\pi a^2-\pi(a^2-h^2)=\pi h^2.$$

This annulus is the image under projection into the plane of the annulus $\hat{A}(h)$ consisting the intersection of the plane $z=h$ with the locus of points above $\hat{C}$ and inside the cylinder over C. Earlier we determined that the hyperbolic area of $\hat{A}(h)$ was on the order of 1. Now we see that it is exactly π.

In order to insure that the tubes over $\hat{C}$ have enough room to grow, we will need to annex around the base of the hemisphere a region B having Euclidean area on the order of that of $A(h)$, that is, we must have

$$\mathrm{Area}(B) \geqslant K \cdot \mathrm{Area}(A(h)) = K \cdot \pi h^2$$

for some universal constant K. To see that this is the right condition, let $\hat{B}(h)$ be the portion of the plane $z=h$ that lies over B. The hyperbolic area of $\hat{A}(h) \cup \hat{B}(h)$ is

$$\frac{1}{h^2}(\pi h^2 + \mathrm{Area}(B)) = \pi + \frac{\mathrm{Area}(B)}{h^2},$$

while the hyperbolic area of $\hat{A}(h/2) \cup \hat{B}(h/2)$ is

$$\frac{1}{(h/2)^2}(\pi (h/2)^2 + \mathrm{Area}(B)) = \pi + 4\frac{\mathrm{Area}(B)}{h^2}.$$

Our assumption on the area of B guarantees that the ratio of the latter quantity to the former will always be $\geqslant (1+4K)/(1+K)$. Thus in going from height h down to height $h/2$ the hyperbolic area available to the tubes increases by a definite factor, which is the kind of condition we need.

3.15. Annexing annuli

Of course this condition on the area of the annexed region B is not in itself enough: It matters how the annexed territory is distributed around C, because all of the tubes have to keep on expanding more or less exponentially, and it might not be possible to avoid pinching tubes over one part of the hemisphere even though there is plenty of extra room somewhere around on the far side of the hemisphere.

To make sure that the territory we annex is nicely distributed, the ideal thing would be to annex an annulus, since then we could just divert our tubes radially away from the center of C, and all would be well. Unfortunately, if C is tangent or nearly tangent to the base of another relevant hemisphere, or to C_1 or C_2, hostilities will arise when we try to annex territory belonging to or coveted by the other hemisphere. (See Figure 12.)

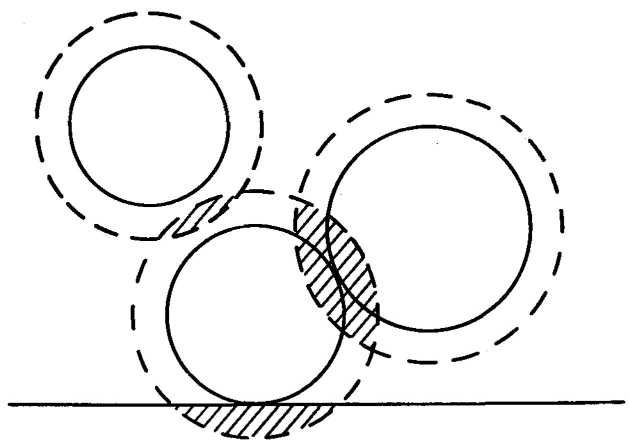


Fig. 12. Disputed territory.

3.16. Avoiding conflicts between neighboring hemispheres

To avoid this kind of conflict, we will prescribe a policy of universal appeasement. If the annulus claimed by $\hat{C}$ overlaps that claimed by another hemisphere, we instruct $\hat{C}$ to renounce its claim to any territory beyond the line tangent to C and perpendicular to the line joining the center of C to the center of the base of the competing hemisphere. (See Figure 13.) A similar rule applies if the annulus claimed by $\hat{C}$ intersects D_1 or D_2.

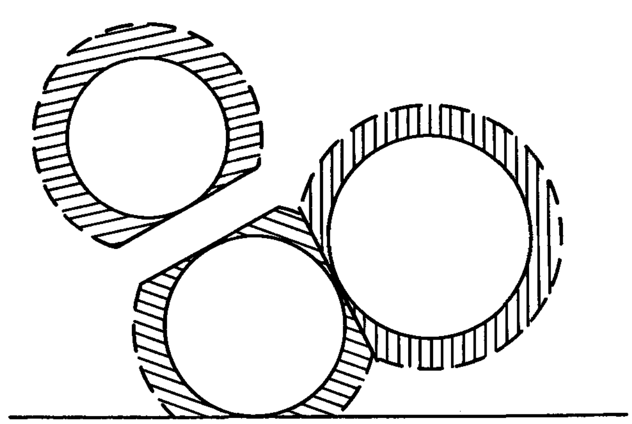


Fig. 13. Universal appeasement.

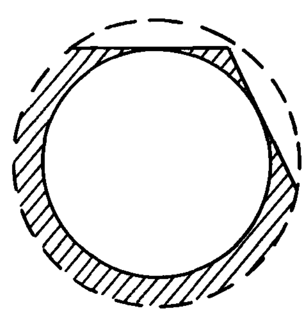


Fig. 14. Overlap.

3.17. Overlap of renounced territory

The effect of this policy of appeasement is that the domain associated to each circle is no longer an annulus, but an annulus with certain pieces snipped away. The question arises, whether two of the snipped-away pieces can overlap, as shown in Figure 14. This will depend on how large the annuli are that we originally tried to associate to the hemispheres.

This question of the possible overlap of snipped-away parts is important, because now that we don't have a complete annulus to work in, we can no longer simply spread our tubes out radially away from the center of the hemisphere. If we do this, the tubes near the middle of the snipped-away part will not be growing. If these tubes are to grow, we must allow them to expand laterally, which will squeeze tubes out away from the middle of the snipped-away part. (Peek ahead at Figure 24.) But if two snipped-away parts overlap, there will be nowhere for the tubes to spread out into. We want to arrange things so that snipped-away parts can never overlap.

3.18. Avoiding overlap

It turns out that to insure that snipped-away parts don't overlap, it is sufficent to make sure that the annuli that the hemispheres try to annex are sufficiently small, though of course we must still make them large enough to be useful. Specifically, say that to the hemisphere $\hat{C}$, with notation as above, we associate an annulus of outer radius

$$\sqrt{a^2+Kh^2} \sim a+Kh^2/2a,$$

where K is some universal constant yet to be specified. The claim is that if K is chosen small enough, then no two snipped-away parts will ever overlap. The truth of this

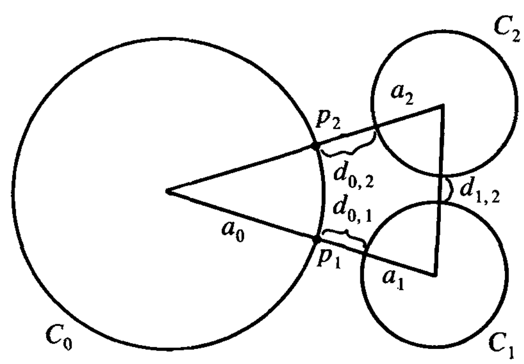


Fig. 15. Notation for the elementary geometric lemma.

statement depends on only the grossest of estimates of how densely one can pack circles in the plane.

3.19. An elementary geometric lemma

LEMMA. *Let C_0, C_1, C_2 be three circles bounding disjoint open disks in the plane. (See Figure 15.) Let the radii of these circles of a_0, a_1, a_2. Assume that*

$$a_0, a_1, a_2 \geqslant 1.$$

Let p_1 and p_2 denote the points on C_0 closest to C_1 and C_2. Let

$$d_{0,1} = \operatorname{dist}(C_0, C_1) = \operatorname{dist}(p_1, C_1)$$

and

$$d_{0,2} = \operatorname{dist}(C_0, C_2) = \operatorname{dist}(p_2, C_2).$$

Then for universal constants K_1, K_2, if

$$d_{0,1}, d_{0,2} \leqslant K_1$$

then

$$\operatorname{dist}(p_1, p_2) \geqslant K_2.$$

This conclusion remains true in the limit when one or both of the circles C_1 and C_2 are allowed to degenerate into lines (circles of infinite radius).

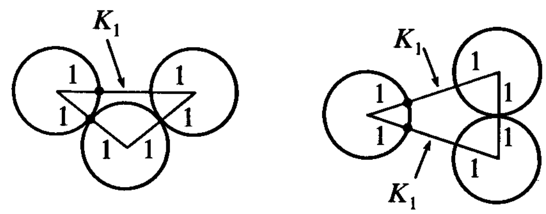


Fig. 16. Special configurations.

Proof. Choose any K_1 between 0 and 1. Call a configuration of circles a K_1-configuration if

$$d_{0,1}, d_{0,2} \leqslant K_1.$$

Clearly $\mathrm{dist}(p_1, p_2)$ is a continuous positive function on the space of K_1-configurations. The only way for $\mathrm{dist}(p_1, p_2)$ to approach 0 is for something bad to happen as one of the radii a_0, a_1, a_2 approaches 0, but it's easy to see that no such bad thing happens. Hence there is some appropriate choice for K_2.

More concretely, consider the two special K_1-configurations for which $a_0 = a_1 = a_2 = 1, d_{0,1} = K_1, d_{1,2} = 0$, and either $d_{0,2} = 0$ or $d_{0,2} = K_1$. (See Figure 16.) An arbitrary K_1-configuration can be transformed into one of these two special configurations by a sequence of steps that do not increase $\mathrm{dist}(p_1, p_2)$. (Without going into detail, the steps are: make $a_1 = a_2 = 1$; assume $d_{0,1} \geqslant d_{0,2}$; make $d_{0,1} = K_1$; make $d_{1,2} = 0$; either make $d_{0,2} = 0$ or make $d_{0,2} = K_1$; make $a_0 = 1$.) Setting

$$K_1 = 2\sqrt{2} - 2 = 0.8284271247 \ldots$$

and computing $\mathrm{dist}(p_1, p_2)$ for the two special configurations, we find (see Figure 17)

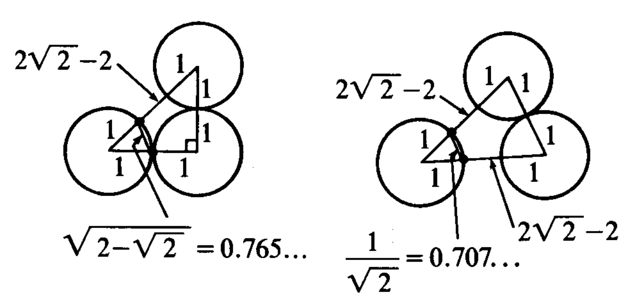
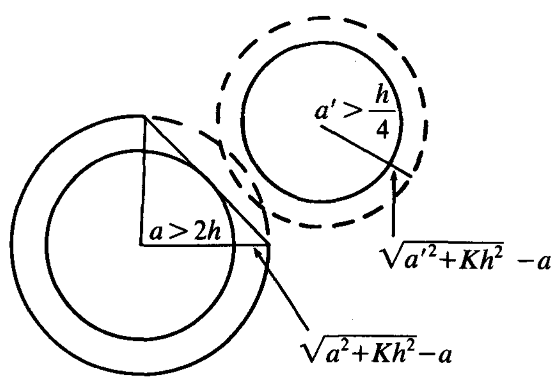


Fig. 17. The case where $K_1 = 2\sqrt{2} - 2$.



$$
\begin{aligned}
\operatorname{dist}(C, C') &\leqslant (\sqrt{a^2+Kh^2}-a)+(\sqrt{a'^2+Kh^2}-a') \\
&\leqslant K\frac{h^2}{2a}+K\frac{h^2}{2a'} \\
&= 2\left(\frac{h}{a}+\frac{h}{a'}\right)\cdot K\frac{h}{4} \\
&< 2\left(\frac{1}{2}+4\right)\cdot K\frac{h}{4} \\
&= 9\cdot K\frac{h}{4} \\
&\leqslant K_1\frac{h}{4} \quad \text{if} \quad K \leqslant \frac{K_1}{9}
\end{aligned}
$$

Fig. 18. Making sure conflicting hemispheres are close.

that we can choose

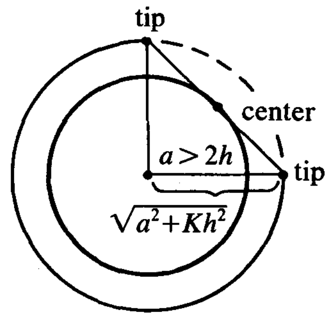
$$K_2 = \min\left(\sqrt{2-\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} = 0.7071067811 \ldots \qquad \square$$

3.20. How little area to ask for

The elementary geometric lemma tells us how small to make K in order to make sure that no two snipped-away parts overlap. By choosing

$$K \leqslant \frac{K_1}{9},$$

we can guarantee that any relevant hemisphere with which C conflicts is within a distance $K_1 \cdot h/4$ of C. (See Figure 18.) Thus by the elementary geometric lemma if two



$$\begin{aligned}\text{dist(center, tip)} &= \sqrt{(a^2+Kh^2)-a^2}\\ &= \sqrt{K}h\\ &\leqslant \frac{K_2}{2}\cdot\frac{h}{4} \quad \text{if} \quad \sqrt{K}\leqslant\frac{K_2}{8},\\ \text{i.e. if} \quad & K\leqslant\frac{K_2^2}{64}\end{aligned}$$

Fig. 19. Making sure snips are short.

hemispheres both conflict with C the distance between the centers of the corresponding snips must be $\geqslant K_2\cdot h/4$. (Recall that we only have to worry about relevant hemispheres, which have height $>h/4$.) But by choosing

$$K\leqslant\frac{K_2^2}{64},$$

we can guarantee that the distance from center to tip of either snip is $\leqslant K_2/2\cdot h/4$. (See Figure 19.) Thus if the snips overlap, the distance between their centers must be $<K_2\cdot h/4$, a contradiction.

For the values $K_1=2\sqrt{2}-2$ and $K_2=1/\sqrt{2}$, we find that these conditions on K will be satisfied for $K=1/128$. With this value of K, each hemisphere is asking for growth room of 1/128th the area that its tubes cover at height h.

3.21. Laying out the tubes

Now we're in great shape. We have annexed our annulus, and although parts have been snipped away, no two snipped-away parts overlap. Within the sector defined by one of the snipped-away parts, the annexed area is on the order of the original area. (See

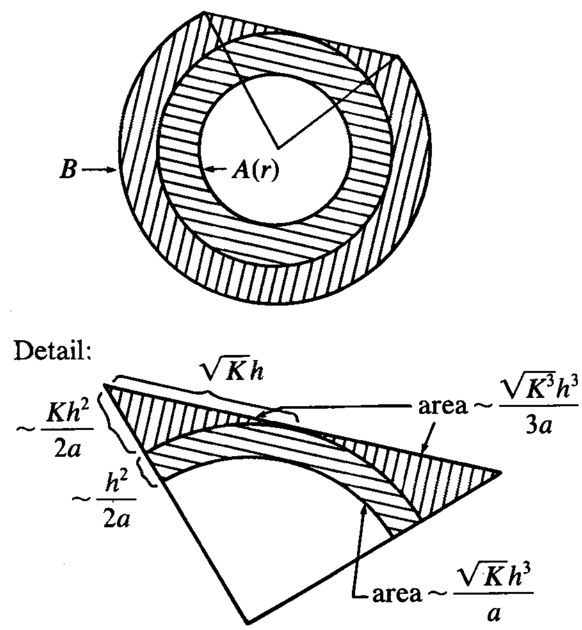


Fig. 20. Comparing areas within a snipped sector.

Figure 20.) Thus we hope and expect that there is enough room for the tubes to grow. To verify this, let us specify precisely how the tubes are to run.

The region we are supposed to be cutting up is the portion of G between height h and height $h/2$ and lying above $D \cup B$, the union of the base of the hemisphere and the annexed territory. Just as we divided the plane up into pieces, we will divide $D \cup B$ into pieces, and treat each such piece separately. We begin by cutting $D \cup B$ along the radii that extend to the tips and centers of the snips. (See Figure 21.) This yields pieces of

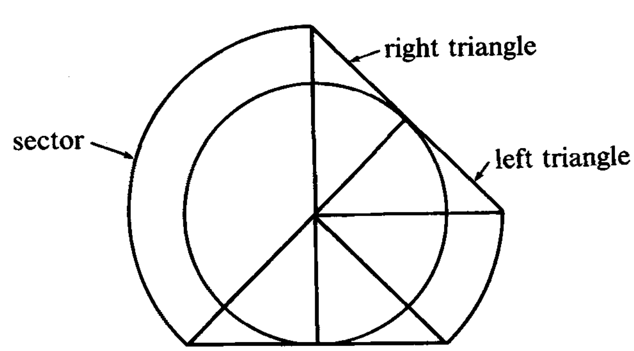


Fig. 21. Dividing up $D \cup B$.

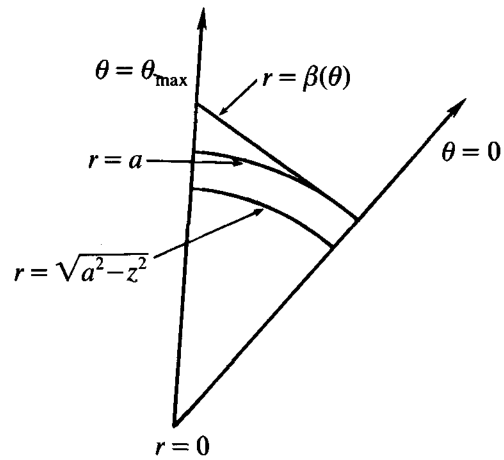


Fig. 22. Polar coordinates.

three kinds: sectors, right triangles, and "left triangles". In discussing how to lay out the tubes we will ignore left triangles, since they are just like right triangles.

For a given piece S, either a sector or a triangle, let $\hat{S}(z)$ be the portion of the plane at height z that lies above S, and let

$$H = \bigcup_{z \in [h/2, h]} \hat{S}(z).$$

Our task is to extend the tubes that have run into $\hat{S}(h)$ down through the region H until they run into $\hat{S}(h/2)$. Since the tubes are to remain more or less vertical, this amounts to specifying a correspondence φ_z between $\hat{S}(h)$ and $\hat{S}(z)$ for each $z \in [h/2, h]$. Our strategy is to choose these correspondences so that the available horizontal cross-section is shared equally among the tubes, that is, so that

$$\frac{\text{Area}(\varphi_z(A))}{\text{Area}(\hat{S}(z))} = \frac{\text{Area}(A)}{\text{Area}(\hat{S}(h))}$$

for all $A \subseteq \hat{S}(h)$.

This condition doesn't determine the φ_z's, so we require in addition that the φ_z's be nice and smooth, and take "radii" of $\hat{S}(h)$ to "radii" of $S(z)$. To see what this entails, introduce polar coordinates (r, θ) in the (x, y)-plane so that the point $r=0$ is the center of the disk D and the ray $\theta=0$ lies just to the right of S. (See Figure 22.) Then H is determined by the inequalities

$$\frac{h}{2} \leqslant z \leqslant h,$$

$$0 \leqslant \theta \leqslant \theta_{\max},$$

$$\sqrt{a^2-z^2} \leqslant r \leqslant \beta(\theta),$$

where

$$\beta(\theta) = \sqrt{a^2+Kh^2}$$

when S is a sector, and

$$\beta(\theta) = a \sec \theta,$$

$$a \sec \theta_{\max} = \sqrt{a^2+Kh^2}$$

when S is a right triangle. Set

$$\alpha = \sqrt{a^2-x^2},$$

and define

$$u = \frac{\displaystyle\int_0^\theta \frac{\beta^2-\alpha^2}{2}}{\displaystyle\int_0^{\theta_{\max}} \frac{\beta^2-\alpha^2}{2}} = \frac{\displaystyle\int_0^\theta \beta^2-\alpha^2}{\displaystyle\int_0^{\theta_{\max}} \beta^2-\alpha^2},$$

$$v = \frac{\left(\dfrac{r^2-\alpha^2}{2}\right)}{\left(\dfrac{\beta^2-\alpha^2}{2}\right)} = \frac{(r^2-\alpha^2)}{(\beta^2-\alpha^2)}.$$

The function $u = u(\theta, z)$ tells what fraction of the area of $\hat{S}(z)$ has θ-coordinate between 0 and θ. The function $v = v(r, \theta, z)$ tells what fraction of the area of the infinitesimal strip between angles θ and $\theta + d\theta$ has r-coordinate between α and r. We make the tubes run along the curves u=const, v=const. The correspondences φ_z associate points in $\hat{S}(h)$ and $\hat{S}(z)$ that have equal values of u and v. These correspondences are illustrated in Figures 23 and 24. Figure 23 shows a sector; here the tubes are being deflected radially. Figure 24 shows a right and left triangle together; here in addition to being deflected radially the tubes are being squeezed from the center of the snip out towards the tips.

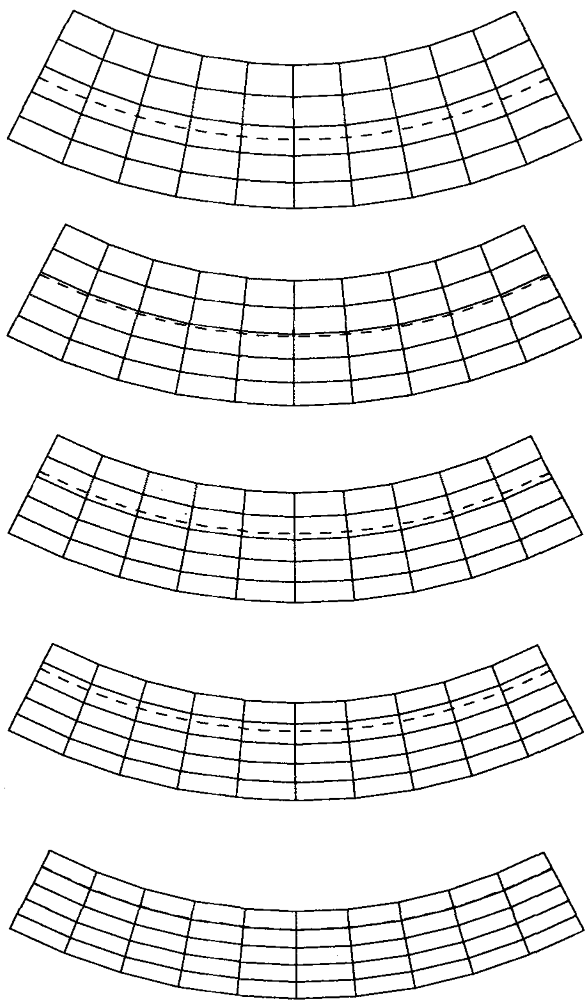


Fig. 23. Tubes above a sector.

Have we succeeded in keeping the tubes more or less vertical? This comes down to checking whether there is a universal upper bound for

$$\left(\frac{\partial r}{\partial z}\right)_{u,v}^2 + r^2 \left(\frac{\partial \theta}{\partial z}\right)_{u,v}^2,$$

where the subscript u, v indicates that u and v are to be held constant. For a sector, this fact is geometrically obvious. For a triangle, this fact is not quite so obvious, but nonetheless true. To see why, consider that the only free parameter is the ratio h/a, and

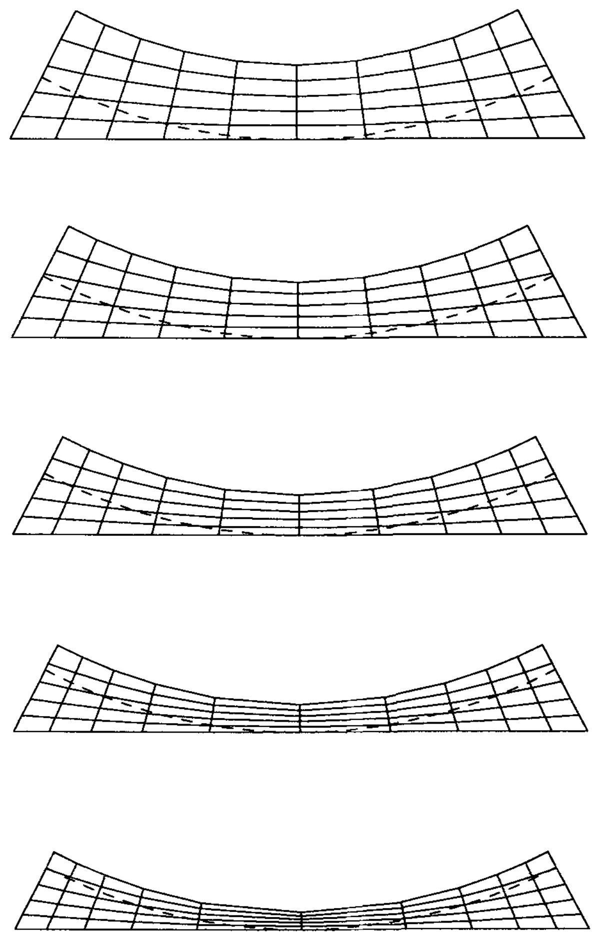


Fig. 24. Tubes above a pair of triangles.

the only way the expression above can fail to be bounded is if something bad happens as h/a goes to 0. So you just have to convince yourself that nothing bad happens as h/a goes to 0. As a last resort, this can be verified by computation.

3.22. Taking stock

Our task was to cut G into tubes that grow more or less exponentially. To make sure that we have accomplished this, let us follow the course of a representative tube, as it threads its way down toward the plane. To start off, it drops straight down until it reaches height 1. Below height 1, its journey is divided into steps of three kinds.

Steps spent over the residual piece of the plane. In the course of one of these steps, the tube remains vertical, and its hyperbolic horizontal cross-section grows exponentially.

Steps spent over a sector or a triangle. In the course of one of these steps, the tube remains more or less vertical; its hyperbolic horizontal cross-section does not decrease, and increases by a definite factor from the beginning to the end of the step.

Steps spent taking care of a baby hemisphere. These steps come in sets of three. In the course of one of these sets of three steps, the tubes remain more or less vertical; the hyperbolic cross-section does not decrease, and increases by a definite factor between the beginning of the first step and the end of the third step.

Taken together, these facts imply that we have indeed succeeded in cutting G into tubes that grow more or less exponentially, with rate bounded below and factor bounded above by universal constants. But as we remarked before, this implies that we can do the same for the fundamental region $\hat{F}$, so there is a universal lower bound for $\lambda_0(\hat{F})$, and hence for $\lambda_0(\mathbf{H}^3/\Gamma)$. □

3.23. What is the constant?

In the proof we have just gone through, we neglected to determine precise values for various "universal constants". As a result, while we now know that there is a universal lower bound L_2 for $\lambda_0(\hat{F})$, and hence a universal upper bound U_2 for $\delta(\Gamma)$, we do not know concrete values for these bounds. With sufficient patience, we could work out precise bounds. However, the resulting bounds are liable to be exceedingly poor. To get good bounds by the cutting method, a keener knife and a steadier hand will be needed.

Acknowledgements

I would like to thank Ralph Phillips, Peter Sarnak and Bob Brooks for telling me about this problem, and their ideas on how to solve it. They also made helpful comments on earlier versions of this paper, as did Curt McMullen. The referee filled me in on the history of the problem and supplied some of the references cited in the Introduction. Most of all, I would like to thank Dennis Sullivan for advice and encouragement.

References

[1] Akaza, T., Poincaré theta-series and singular sets of Schottky groups. *Nagoya Math. J.*, 24 (1964), 43–65.

[2] — (3/2)-dimensional measure of singular sets of some Kleinian groups, *J. Math. Soc. Japan*, 24 (1972), 448–464.
[3] Brooks, R., The spectral geometry of the Apollonian packing. *Comm. Pure Appl. Math.*, 38 (1985), 357–366.
[4] — On the deformation theory of classical Schottky groups. *Duke Math. J.*, 52 (1985), 1009–1024.
[5] Burnside, W., On a class of automorphic functions. *Proc. London Math. Soc.*, 23 (1891), 49–88.
[6] Doyle, P. G. & Snell, J. L., *Random walks and electric networks*. MAA Carus Monographs, 1984.
[7] Holland, C. J., A minimum principle for the principal eigenvalue for second-order linear elliptic equations with natural boundary conditions. *Comm. Pure. Appl. Math.*, 31 (1978), 509–519.
[8] Maxwell, J. C., *Treatise on Electricity and Magnetism*. 3d ed., Clarendon, Oxford, 1891.
[9] Myrberg, P. J., Zur Theorie der Konvergenz der Poincaréschen Reihen. *Ann. Acad. Sci. Fenn. Ser. A*, 9 (1916), 1–75.
[10] — Zur Theorie der Konvergenz der Poincaréschen Reihen (zweite Abh.). *Ann. Acad. Sci. Fenn. Ser. A*, 11 (1917), 1–29.
[11] Patterson, S. J., The limit set of a Fuchsian group. *Acta Math.*, 136 (1976), 241–273.
[12] — Some examples of Fuchsian groups. *Proc. London Math. Soc.*, 39 (1979), 276–298.
[13] Phillips, R. & Sarnak, P., The Laplacian for domains in hyperbolic space and limit sets of Kleinian groups. *Acta Math.*, 155 (1985), 173–241.
[14] Rayleigh, J. W. S., On the theory of resonance, in *Collected scientific papers*, vol. 1, pp. 33–75. Cambridge Univ. Press, Cambridge, 1899.
[15] Sarnak, P., Domains in hyperbolic space and limit sets of Kleinian groups, in *Differential Equations*, ed. I. W. Knowles and R. T. Lewis, pp. 485–500. North-Holland, Amsterdam, 1984.
[16] Schottky, F., Über eine specielle Function, welche bei einer bestimmten linearen Transformation ihres Argumentes unverändert bleibt. *Crelle's J.*, 101 (1887), 227–272.
[17] Sullivan, D., Discrete conformal groups and measurable dynamics. *Bull. Amer. Math. Soc.*, 6 (1982), 57–73.
[18] — Entropy, Hausdorff measures old and new, and the limit sets of geometrically finite Kleinian groups. *Acta Math.*, 153 (1983), 259–277.

Received March 3, 1986
Received in revised form July 8, 1987