Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Signature:

_____ 4/22/2009

Jacob K. Geerlings Date

Limit sets of Kleinian groups: properties, parameters, and pictures

by

Jacob K. Geerlings

Adviser

David Borthwick

Department of Mathematics and Computer Science

_____
Dr. David Borthwick
Adviser

_____
Dr. Emily Hamilton
Committee Member

_____
Dr. Keith Berland
Committee Member

_____
4/22/09

Limit sets of Kleinian groups: properties, parameters, and pictures

by

Jacob Geerlings

Adviser
David Borthwick

An abstract of
A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Mathematics and Computer Science

2009

Abstract

Limit sets of Kleinian groups: properties, parameters, and pictures

By Jacob K. Geerlings

This thesis explores fractal images that are the limit sets of Kleinian groups. Properties and classification of Möbius transformations lead to groups of such transformations, and to the classification of some important group types. Möbius transformations are represented by matrices in SL(2,ℂ). The purpose of the early material is to foster an intuitive grasp of what happens in simple groups, so that more complex groups may be correctly pictured. The accumulation points of group's actions on points in the upper half space are on the plane and are called the limit set. The limit set of a Kleinian group Γ can have 0, 1, 2, or uncountably infinitely many points. A discrete group will have a limit set that is a proper subset of the plane. The limit set is the smallest nonempty, closed, Γ-invariant subset of the plane. Besides properties of the limit set, the thesis explores and utilizes a depth-first search to plot pictures of connected limit sets. The family used for the algorithm has connected limit sets that are origin-symmetric and pass through 1 and -1. Each group in the family has a parabolic commutator. The pictures are used to explore different features of the groups in the family—ones generated by parabolic, hyperbolic, and loxodromic transformations. The images, generated by Mathematica, by self-symmetry and spirals, serve to remind readers of facts about complex arithmetic and Möbius transformations. The final chapter also explores some limitations of the algorithm.

Limit sets of Kleinian groups: properties, parameters, and pictures

by

Jacob K. Geerlings

Adviser
David Borthwick

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Mathematics and Computer Science

2009

Table of Contents

Introduction 1

Chapter 1: Möbius transformations 4

Chapter 2: Groups of transformations 12

Chapter 3: Limit sets 23

Chapter 4: Graphing the limit set 31

References 43

Appendix: the code 44

Introduction: What we are looking for

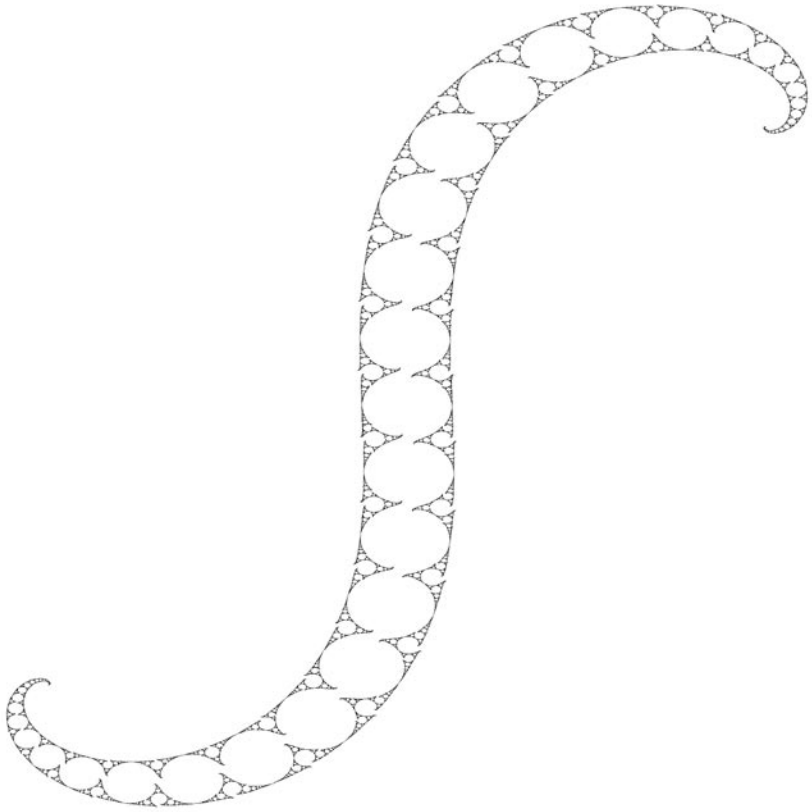
This thesis explores the limit sets of Kleinian groups—finitely generated, discrete groups of Möbius transformations. Then, the depth-first search algorithm as described by Mumford, Series, and Wright in *Indra's Pearls* [2] is applied to their special family of groups with parabolic commutator. Their normalization puts the important structures of the limit set all pretty close to, and rotation-symmetric, around the origin.

In plainer English, this thesis gives ways of repeatedly transforming the plane in ways that make fractals. Fractals, infinitely mirrored, repetitively branching forms, will be pictured and discussed. The *limit set*, which is a fractal in some cases, will be our tool for understanding these transformations of the plane. These geometric objects help us understand the transformations that generate them and the underlying complex number system.

The thesis was constructed with two main sources—the graduate reference text *Foundations of Hyperbolic Manifolds*, by John G. Ratcliffe [3], and the book *Indra's Pearls: The Vision of Felix Klein*, which is meant to be accessible to laymen. Ratcliffe's treatment of the relevant subject material (one chapter out of 13) is rigorous and dry, while Mumford, Series, and Wright skim over proofs, focusing on graphical intuition and readability. The thesis places the harder material from Ratcliffe in the middle, surrounded by the intuitive material based on *Indra's Pearls*.

The first chapter presents some preliminary information concerning complex numbers, Möbius transformations, and conjugation, leaning more toward intuition. The second chapter discusses groups and discreteness. Torsion-free Kleinian groups enter play there, as do Schottky and Fuchsian groups. This chapter moves the reader into what

is the heaviest chapter. Chapter 3 proves important properties of the limit set, drawing heavily from Ratcliffe, but adapted for three dimensions. Chapter 4 continues to investigate the limit set, but with parameters and an algorithm from Mumford, Series and Wright. The Mathematica code used to generate the fractal images is at the very end.



Chapter 1: Möbius transformations

In order to build a vocabulary with which we may describe and generate fractal patterns, an intuitive understanding of Möbius transformations is a start. But first, let us review some facts about the complex plane. The number i is defined as a solution to the equation $x^2 + 1 = 0$ (the other is $-i$). There is no solution in the real numbers, so i is called "imaginary." The complex plane is a plane of two axes: real numbers and imaginary. The imaginary axis moves through real multiples of i. A complex number is one with both imaginary and real parts, like $4 + 3i$, or in general $a + bi$ for a, b real. Each complex number is a point on the plane, with $4 + 3i$ situated four units right and three units up from the origin.

Addition of complex numbers works just like regular addition.

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

Multiplication is a bit more complicated, though, because $i^2 = -1$. So,

$$\begin{aligned}(a + bi)(c + di) &= ac + adi + bci + bdi^2 \\ &= (ac - bd) + (ad+bc)i\end{aligned}$$

Multiplication is more easily understood in a different coordinate representation, though. Instead of representing x and y positions, complex numbers can be represented by a distance from zero and a degree (in radians) from the positive real ray. Thus, i gets radius 1 and angle $\pi/2$. This can be written using Euler's formula as $i = e^{i\pi/2}$. From the exponential notation for complex numbers, $re^{i\theta}$, where r is $|a + bi|$, the distance from the origin, and θ is the angle from the positive real numbers, also called the "argument" of the number. By this representation, it is easy to see that two numbers, when multiplied, have radius equal to the product of radii, and argument equal to sum of the factors'

arguments. That is,

$$re^{i\theta}\, se^{i\varphi} = rs(e^{i\theta}\, e^{i\varphi}) = rse^{i\theta + i\varphi} = rse^{i(\theta + \varphi)}$$

The next item on the list is infinity. The number infinity (∞) does not need to be feared and hated (unless perhaps your computer program is stumped by it); it has a proper place in the sphere of mathematics. At the top. This is the value of stereographic projection. Stereographic projection maps the set of complex numbers and infinity to the unit sphere. Consider the unit sphere sitting in the center of the complex plane. From the top of the sphere, the point (0, 0, 1), make a line to a point in the plane. That line will intersect the unit sphere at one point other than (0, 0, 1). Associate those two points. Each point in the plane is thus mapped to a point on the sphere. Any line tangent to the sphere at (0, 0, 1) will be parallel to the plane, so it will not be associated with any point of the plane. However, as slopes of lines approach horizontal, the points of the plane get farther from 0 and the points on the sphere get closer to the pole. Hence, the North Pole is associated with ∞. This isomorphism maps circles in the plane to circles on the sphere, and lines in the plane become great circles through the North Pole. This inclusion of infinity with the set of complex numbers is called the extended plane, or $\mathbb{C}_\infty$. The unit sphere when associated with this mapping to the extended plane is called the Riemann sphere.

Definition: A Möbius transformation is a function M with the form

$$M(z) = \frac{az + b}{cz + d}, \text{ with } a, b, c, d \text{ complex and } ad - bc \text{ nonzero.}$$

Why is $ad - bc$ nonzero? If it were zero, then M could give the unhelpful value 0/0. $M\left(-\frac{d}{c}\right) = \frac{-\frac{ad}{c} + b}{-d + d} = \frac{ad - bc}{0}$. Also, M could be something like $\frac{z+1}{z+1}$, a constant function, which is not invertible. With our nascent conception of infinity, let us define

some intuitive facts like $1/0 = \infty$, $1/\infty = 0$, $\infty + a = a\infty = \infty/a = \infty$, and leave other things indeterminate, like 0/0 and 0∞. Now we have Möbius transformations which act on the entire extended plane. These relate to simpler transformations like translation ($z \mapsto z+a$), scaling ($z \mapsto az$), and the inversion ($z \mapsto 1/z$) as follows. The notation "$\mapsto$" means "is mapped to," so the first transformation is the transformation mapping z to $z + a$ for each complex number z.

Theorem 1.1: Every Möbius transformation is the composition of translations, scalings, and the inversion.

Proof: Beginning with the identity transformation, perform translations, scalings, and inversion to show that one can generate any Möbius transformation of the form $\frac{az+b}{cz+d}$.

$z \mapsto z$	identity
$z \mapsto c(cz + d)$	scale by c, translate by d, scale by c
$z \mapsto \frac{bc-ad}{c(cz+d)}$	invert, scale by $bc - ad$
$z \mapsto \frac{a}{c} + \frac{bc-ad}{c(cz+d)}$	translate by a/c.

$$= \frac{a(cz+d)+bc-ad}{c(cz+d)} = \frac{acz+bc}{c(cz+d)} = \frac{az+b}{cz+d}.$$

So, any Möbius transformation is a composition of translation, scaling, and inversion. □

This means that Möbius transformations preserve circles (including lines—circles through infinity), the angles between them, and orientation, as translations, scalings, and the inversion also do.

Theorem 1.2: A Möbius transformation other than the identity fixes at most two points ("fixes" means maps to itself, i.e. $f(z) = z$).

Proof: Set $\frac{az+b}{cz+d} = z$. Then, $cz^2 + (d-a)z - b = 0$. This quadratic equation has

solutions

$$\frac{a-d\pm\sqrt{d^2-2ad+a^2-\ 4bc}}{2c}=\frac{a-d\pm\sqrt{(a+d)^2-4}}{2c}$$

(assuming $ad - bc = 1$, which is easily arranged). The only trouble comes when $c = 0$. Then, the transformation fixes infinity and some other point, which point is given by the solution to $(d-a)z-b=0$ (if this equation has no solution, infinity is the only fixed point).

When $c = b = d - a = 0$, the equation is true for all z. This is only the case when the transformation is the identity. □

Hidden in this theorem (assuming Möbius maps are invertible, which is discussed later) is the fact that two Möbius maps that are not the same or the identity may only have the same action on two points. If a, b, and c are mapped to x, y, and z by both Möbius transformations h and g, the Möbius transformation $h^{-1}g$ has three fixed points. Thus, $h^{-1}g$ is the identity so $h = g$.

So we have found some points of interest to our transformations—fixed points. Now we need to be able to move them around without changing how interesting they are.

Definition: a map M is *conjugate* to a map N if there is an invertible map C such that

$$M = C \circ N \circ C^{-1}.$$

Conjugation can be thought of like coordinate change. Let's say N has a structure, perhaps a fixed point at $4 + 2i$ that would be more conveniently observed at the origin. Then, to move that structure to the origin, simply make C any map such that $C(4+2i) = 0$. An easy choice is $C(z) = z - 4 - 2i$. Then, $C^{-1}(z) = z + 4 + 2i$. So,

$$M(0) = C \circ N \circ C^{-1}(0) = C \circ N(4+2i) = C(4+2i) = 0$$

$M(0) = 0$, so the conjugation does indeed move the fixed point as desired.

When conjugating by Möbius maps, the result is another Möbius map, as Möbius transformations are closed under composition. To see this, let's compose two of them.

$$\frac{a\dfrac{(ez+f)}{gz+h}+b}{c\dfrac{ez+f}{gz+h}+d}=\frac{aez+af+bgz+bh}{cez+cf+dgz+dh}=\frac{(ae+bg)z+af+bh}{(ce+dg)z+cf+dh}$$

The result is another Möbius transformation. And, it looks like something else rather handy. Observe the result of a matrix multiplication of two matrices with entries the same as the coefficients above.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix}=\begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

This shows that if one represents Möbius transformations by matrices of their coefficients, then composition is represented by matrix multiplication. This is quite useful when keeping track of transformations in the computer. There is not one-for-one correspondence between these matrices and Möbius transformations. The same transformation is represented by multiplying each matrix entry by a constant. So, since it does not change the transformation, divide by $\sqrt{ad-bc}$, the root of the determinant. Then, each Möbius transformation is represented by a unique matrix with determinant 1. This set of matrices is called the *special linear group*, SL(2,ℂ) as it does not have just any matrix in it. Each matrix has determinant 1, and so is invertible and corresponds to a unique Möbius transformation.

The formula for the inverse matrix becomes a formula for inverse transformations as well, since they are unaffected by multiplication by a constant. Observe that the coefficients in the identity matrix give the identity transformation.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Conjugating with Möbius maps can produce some important results. Besides translation, which we have seen, the map $z \mapsto \frac{1}{z}$ is its own inverse, and conjugating by it switches the interior of the unit circle with the exterior, including swapping 0 and ∞ and i and $-i$. This conjugation is just like projecting back to the Riemann sphere, flipping its south and north poles, and projecting back to the plane from the new north pole. Another useful map is one that pairs the unit circle and the real line. The classic map for this is the Cayley map, $\frac{z-i}{z+i}$. Of course, since ∞ is our friend now, the image of the unit circle includes ∞, otherwise the circle that is the real line would be missing a point.

Now comes a very important classification of Möbius maps. These terms will pop up for the rest of the paper. When I say "conjugate" here, I mean by Möbius transformations.

Definition: A Möbius map is called *parabolic* if it has only one fixed point.

Looking back at facts about fixed points, this occurs when $\sqrt{(a+d)^2-4} = 0$ (so $a + d = \pm 2$), or if c is zero and $d = a$. These transformations are always conjugate to translations—conjugate the fixed point to ∞.

Definition: A Möbius transformation is called *elliptic* if it is conjugate to a map $z \mapsto kz$ for k on the unit circle, that is, $|k| = 1$.

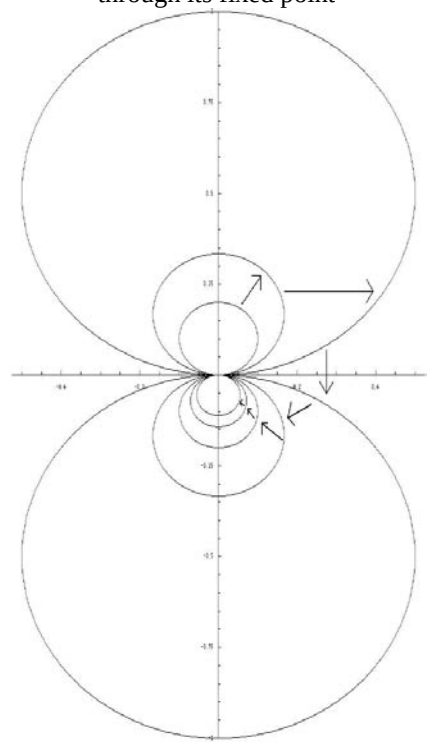
It is obvious that 0 and ∞ are the fixed points of the conjugate transformation. Remember, complex multiplication multiplies distances from the origin, and k is 1 away from the origin. Thus, an elliptic transformation is conjugate to a transformation that moves points in circles around the origin, effectively spinning the Riemann sphere.

Definition: A Möbius transformation M is *loxodromic* if it is conjugate to a map $z \mapsto kz$ with k not on the unit circle, $|k| \neq 1$. If k is real, M is also *hyperbolic*.

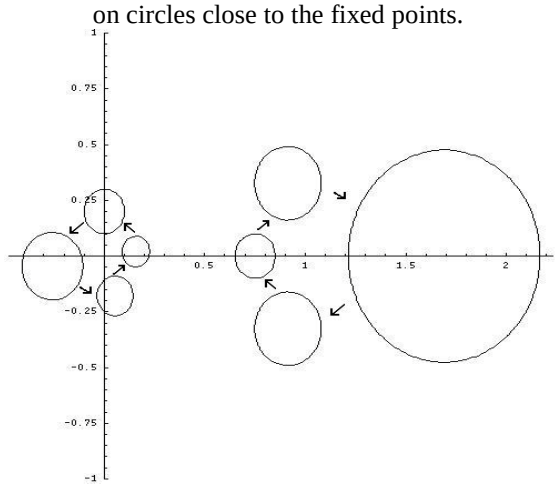
This kind of transformation, then, maps points in spirals towards or away from its fixed points (0 and ∞ if conjugated correctly). If it is hyperbolic, the spirals are quite straight.

Let us see some pictures. All the pictures shown have fixed points at 0. The parabolic and elliptic ones have fixed points at 0 and 1, to show the entire spiral. The parabolic transformation used was $\begin{pmatrix} 1 & 0 \\ 2i & 1 \end{pmatrix}$. Points move toward the fixed point from one side. Points on the other side travel away first in circles that take them finally toward the fixed point. The elliptic transformation was $z \mapsto iz$

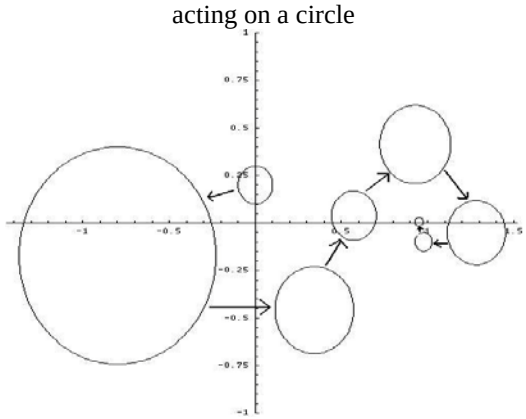
A parabolic transformation acting on a circle through its fixed point



An elliptic transformation acting repeatedly on circles close to the fixed points.



A loxodromic transformation acting on a circle



conjugated so it would be $z \mapsto \frac{iz}{(i-1)z+1}$. The loxodrome used was $z \mapsto 2iz$, conjugated to be $\frac{2iz}{(2i-1)z+1}$. Each transformation acted on a circles close to (or in the parabolic case, including) a fixed point. Notice in the elliptical case that the conjugate map of $z \mapsto iz$, which applied four times is the identity, maps the same circles to themselves after four iterations. It retains the property of the original map. And, by definition, the maps are of the same type under conjugation—a parabolic map cannot be conjugated to a hyperbolic one by Möbius transformations.

These are the important features of the three kinds of transformations: parabolic moves things in circles to the fixed point, elliptic moves in circles around the fixed points, and loxodromic moves in spirals from one fixed point to the other. Picture the inverses of the transformations—the parabolic one moves everything the other direction to the same fixed point. The inverse of the elliptic transformation simply rotates points in the opposite direction around the fixed points, and the inverse of the loxodromic map sends points to the other fixed point.

One more key feature of these maps is essential to our investigations. The quantity $(a+d)$, the sum of the first and last entry in the matrix representation of a transformation, tells us what type of transformation it is. We have already seen that this quantity, called the trace, denoted $\mathrm{Tr}(M)$ for a transformation M, must be ± 2 for parabolic transformations. If the trace is real and between -2 and 2, the transformation is elliptic. Otherwise, the transformation is loxodromic. To see this, conjugate fixed points to 0 and ∞. Then, the element is a scaling of the form $z \mapsto kz$. This has matrix representation $M = \frac{1}{\sqrt{k}}\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$, so $\mathrm{Tr}(M)$ is

$$\left(\frac{1}{\sqrt{k}}\right)(k+1) = \sqrt{k} + \frac{1}{\sqrt{k}}$$

If $|k| = 1$, and $k = a+bi$, then $\sqrt{k}$ is the number on the unit circle halfway anglewise to the positive reals and for any $a+bi$ on the unit circle $(a^2+b^2) = 1$,

$$a + bi + \frac{1}{a+bi} = \frac{(a+bi)^2 + 1}{a+bi} = \frac{(a+bi)(a^2+b^2) + (a-bi)}{a^2+b^2} = 2a$$

$\sqrt{k} + \frac{1}{\sqrt{k}}$ is real, then, and it cannot get any bigger than 2 or less than -2. Thus, if M is elliptic, trace is real and between 2 and -2. This leaves loxodromic transformations with the rest of the traces.

Chapter 2: Groups of transformations

Definition: A group is a set G and an operation on that set such that

1) $ab \in G \ \forall \ a,b \in G$. (closure)

2) $a(bc) = (ab)c$ (associativity)

3) There exists an e such that $ea = ae = a$. (identity)

4) $\forall \ a \in \ G$, there is $a^{-1} \in G$ such that $aa^{-1} \ = \ a^{-1}a \ = \ e$. (inverses)

So, the integers are a group under addition (addition is associative, identity is 0, inverses are opposites), but not multiplication—multiplicative inverses of integers are not always integers.

An easy way to make groups is with *generators*. Pick a few starting elements, then do the operation arbitrarily many times. With the integers and addition, this makes multiples—with two as a generator and addition as the operation, one can produce $\mathbb{Z}2$, or the set of positive and negative even integers (and zero) {..., -4, -2, 0, 2, 4, ...}

Definition: A set of elements $g_1, g_2, \ldots$ are generators of a group G if every element of G can be written as the product of g_i's.

So, 2 is a generator for $\mathbb{Z}2$, from before.

Example: The set of integers 1 through 6 under multiplication mod 7 is a group. Is two a generator? $2^2 = 2*2 = 4$, $2^3 = 2*4 = 1 \pmod 7$. No. 2 only generates the set {1, 2, 4}, a subgroup of the whole group (a *subgroup* is a subset that is a group by the operation). Three is a generator. Its powers are $3^0 = 1$, $3^1 = 3$, $3^2 = 2$, $3^3 = 6$, $3^4 = 4$, $3^5 = 5$, $3^6 = 1$, and the cycle restarts. Check these answers in the table below. Such a group, generated by a single element, is called *cyclic*.

Mod 7 multiplication table (a Cayley table)

(mod 7)	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

Now, returning to the topic of Möbius transformations, one can make groups of them just as easily as with integers. The operation desired is composition.

Example: The set of all Möbius transformations is a group under composition. It is closed under composition, composition is associative, the identity transformation is an element, and each element has an inverse.

This is a huge group, and the groups we seek are all subgroups of this one. The rest of this thesis will discuss properties of groups that are generated by a finite number of Möbius transformations.

Example: One can make a simple cyclic group by taking a Möbius transformation as a generator g. Elements would then be of the form g^m, where m is any integer, positive, negative or zero. We have already seen what happens in this case—parabolic transformations are conjugate to translations, so the group generated would be conjugate to a bunch of translations of varying sizes. A loxodromic transformation would compose with itself to spiral things into or out of the generator's sink and source. Its inverse would carry things to the other fixed point, and all elements of the group would fix the same two points. Elliptic transformations would carry points in circles around the fixed points. So, there is one point in the parabolic case and two points in the loxodromic and elliptic cases that is fixed by the longest strings of compositions, as long as the groups are cyclic.

Since composition of Möbius transformations works like matrix multiplication, with each transformation having a matrix representation in SL(2,$\mathbb{C}$), SL(2,$\mathbb{C}$) contains a group.

We will need a notion of "how far apart" two Möbius transformations are for the next definition. So here is a *metric* on the group of Möbius transformations of $\mathbb{C}_\infty$, based on the matrix representation of the transformations.

Definition: $D(g, h) = |g - h|$ for $g, h \in$ SL(2,$\mathbb{C}$).

This means subtract the two matrices entry by entry, then take the Euclidean distance (root of sum of the squares of each coordinate), pretending the resulting matrix is a point of $\mathbb{C}^4$. This metric determines a *topology* on groups of Möbius transformations, making groups of transformations *topological groups*. The topological term "open" refers to a set that does not contain its boundary, or (alternately defined) to a set whose points are all interior points (there is a neighborhood around them that is a subset of the surrounding space).

Proposition: For S an open set and M a Möbius transformation, $M(S)$ is open.

Proof: Let $m \in M(S)$. Then, $M^{-1}(m) \in S$. So, there is a neighborhood N of $M^{-1}(m)$ that is a subset of S. Since Möbius transformations (as compositions of inversion, translation, and scaling) preserve circles' interiors and exteriors, $M(N) \subset M(S)$, so m is an interior point of $M(S)$. This is true for any m, so $M(S)$ is open. □

Definition: a *discrete group* is a group all of whose points are open.

The points of a group of Möbius transformations are the images of points in the plane under the transformations in the group. So, each set of images of a point must be discrete (that is, every point in the set is isolated) for the group to be discrete (an isolated

point has a disk around it of some positive radius such that no other set member is in the disk). But it is unnecessary to check every single element.

Theorem 2.1: a group of Möbius transformations Γ is discrete if and only if $\{I\}$ is open in the group. (I refers to the identity)

Proof: By definition, if Γ is discrete, $\{I\}$ is open. Now, assume $\{I\}$ is open. For $g \in \Gamma$, the proposition gives that left multiplication by g preserves openness of sets. Thus, $g\{I\} = \{g\}$ is open. Each member of Γ is open. □

Example: The group of Möbius transformations generated by the parabolic map $z \mapsto z + c$ is discrete. Each transformation in the group is a certain distance from the next. Check the identity: The "closest" one can get to the identity transformation is to only do the generator or its inverse once. $\{I\}$ is isolated, so the group is discrete.

Theorem 2.2: A group X is discrete if and only if every convergent sequence $\{x_n\}$ in X is constant after some $m > 0$.

Proof: First, assume X is discrete with $x_n \to x$. Then, since X is discrete, there is a disk such that x is the only member of X in the disk. But since the sequence converges to x, the sequence at some point enters the disk not to leave again (this is what convergence means). The only option is that after this point the sequence is $x, x, x, \ldots$ So discreteness implies sequences are eventually constant.

Now, assume every convergent sequence in X is constant after some point. Assume the contrary; assume X is not discrete. Then, there is a point x of X that is not isolated. Make a sequence by taking disks of radius $1/n$. Then, since x is not isolated, there is a point of X in the disk at each level that is different from x. Set x_n to that point for each disk of radius $1/n$. Then, $x_n \to x$, but is not eventually constant, contradicting the

assumption. Therefore X must be discrete whenever every convergent sequence in X is eventually constant. □

Now we can use this to prove other useful facts. For example, we will be generating groups with parabolic and loxodromic elements. What if they have the same fixed point? Is this a group we seek? It turns out that this group is not discrete. (Theorem from Ratcliffe [3], proof is similar, but notably easier in three dimensions)

Theorem 2.3: If a group of Möbius transformations Γ is generated by a loxodromic element h and one other element g, and h and g have exactly one fixed point in common, Γ is not discrete.

Proof: Conjugation allows the common fixed point to be ∞, and the other fixed point of h to be 0. This makes $h = kz$ for some $k \in \mathbb{C}$. If necessary, swap h for h^{-1} so that $|k|<1$. Since g fixes ∞, g can be written $g = ax + b$ for some $a, b \in \mathbb{C}$.

Any transformation of the form $h^m g h^{-m}$ is in Γ, so iterating m one makes a sequence in Γ.

$$h^m\, g\, h^{-m}\,(z) = h^m\, g\,(k^{-m}) = k^m(ak^{-m}\, z + b)$$

$$= az + k^m b.$$

$|k| < 1$, so the sequence $\{h^m\, g\, h^{-m}\}$converges to az. However, the sequence is never eventually constant as each $k^m b$ is distinct. So Γ is not discrete. □

Definition: The set $\Gamma x = \{g(x) : g \in \Gamma\}$ is called the *orbit* of x.

In English, the orbit of a point is the set of points it is mapped to by the transformations in Γ.

We will need an extension of the idea of Möbius transformations to continue. To initiate an analogy, let us consider the group of Möbius transformations that have only

real coefficients. Since Möbius maps preserve circles and angles, interiors and exteriors of circles are preserved as well—if two points are separated by a circle, their images will be separated by the image of the circle. To see this, make a circle including the two points, and map the two circles by the transformation. The image of the figure will be two circles intersecting at the same angles. The points must come in order by continuity—the interior point, an intersection point, the exterior point, and then the other intersection point of the two circles, so one image point will be inside the image of the circle, and one outside. So, a Möbius transformation with real coefficients (which obviously preserve the real axis, a circle on the Riemann sphere) will also preserve which plane (upper or lower) is mapped to which. These Möbius transformations are members of SL(2,ℝ), the set of 2×2 matrices with real entries and determinant 1. For real coefficients,

$ad–bc=1$,

$$\frac{ai+b}{ci+d}=\frac{(ai+b)(-ci+d)}{d^2+c^2}=\frac{bd+ca-bci+adi}{d^2+c^2}$$

This quantity has imaginary part $(ad-bc)i=i$, so the upper half plane is mapped to itself. The upper half plane is given the name $\mathbb{H}^2$, as it is the upper half of 2-space. SL(2,ℝ) preserves the real line and the upper half plane, so can be said to act on the upper half plane. The transformations act on more points than that by extension, but the entire action of any transformation is determined by its action on the real line.

The analogy then, takes us to $\mathbb{H}^3$, which is the half of three-space above ℂ. What would a Möbius transformation look like in this space, though? Generalizing inversions, translations, and scaling, the Möbius transformations in three-space are compositions of reflections over spheres. $\mathbb{H}^3$ does not include ∞ or ℂ, merely complex numbers with a

third real, nonzero component. Thus, the boundary of $\mathbb{H}^3$ is $\mathbb{C}_\infty$ just as the boundary of the unit ball is the unit sphere.

The generalization of how these transformations act in the upper half space is controlled by the same four constants as in the plane, otherwise this would not be a useful extension. The action of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on a point (z, y) with $z \in \mathbb{C}, y > 0$ is

$$(z, y)| \to \left(\frac{(az+b)\left(\bar{c}\bar{z} + \bar{d}\right) + a\bar{c}y^2}{|cz+d|^2 + |c|^2 y^2}, \frac{y}{|cz+d|^2 + |c|^2 y^2} \right)$$

Note that as $y \to 0$, the height of the image also goes to 0. Also, the complex coordinate approaches $\frac{(az+b)(\bar{c}\bar{z}+\bar{d})}{|cz+d|^2} = \frac{az+b}{cz+d}$, with $cz + d \neq 0$, but $-\frac{d}{c}$ just maps to ∞, as usual. So this definition matches our current understanding and extends it continuously (and uniquely) into the upper half plane. One other feature of these will be quite important, and that is the hyperbolic metric.

Definition: the hyperbolic metric d on $\mathbb{H}^3$ is the function determined by

$$\cosh d((z, y), (z', y')) = \frac{|z - z'|^2 + y^2 + y'^2}{2yy'}$$

The reason we need this metric is that it is preserved under Möbius transformations. So, for two points a,b in $\mathbb{H}^3$ and any Möbius transformation M, $d(a,b) = d(M(a), M(b)$. For comparison, notice that a bit of algebra shows that

$$|M(z) - M(z')| = \frac{|z - z'|}{|cz + d|\,|cz' + d|}$$

Definition: a group of Möbius transformations Γ acts *discontinuously* on a subset X of $\overline{\mathbb{H}^3}$ if and only if for each compact subset K of X, the set $K \cap gK$ is nonempty for only finitely many g in Γ.

Example: A group Γ generated by a single parabolic transformation $z \mapsto z + 1$ (an element of the upper half plane group) acts discontinuously on the upper half plane. For any compact (i.e. closed and bounded) region of the plane, the group eventually moves that region away from itself; only finitely many members of the group map it to a set overlapping the original. The group is not discontinuous at ∞—any set containing ∞ will overlap its image at ∞ under all transformations in the group, and there are infinitely many. So, Γ acts discontinuously on $\mathbb{H} \cup \mathbb{R}$. Considered as acting on the plane and the upper half space, Γ acts discontinuously on $\mathbb{H}^3 \cup \mathbb{C}$.

Example: A group Γ generated by a single elliptic transformation does not act discontinuously on the plane unless it is finite. If the group is finite, then Γ is discontinuous on $\overline{\mathbb{H}^3}$ by definition. But if the group is infinite, it is not even discrete—in the plane, points are rotated around the fixed points by any value whatsoever. Since the identity can be approximated arbitrarily well by iterating the generator, the group is not discontinuous on $\mathbb{H}^3$. A ball can be mapped arbitrarily close to itself, so it can intersect itself under infinitely many of the group's transformations.

By the same argument (circles return to their starting point), any group with an elliptic element does not act discontinuously on a compact region with some area.

Definition: a *Kleinian group* is a finitely generated discrete group of Möbius transformations.

However, we will concern ourselves for the remainder of the paper primarily with groups that are *torsion-free* Kleinian groups. This means that no element (but the identity) generates a finite subgroup. This saves the trouble of dealing with elliptic transformations, which have a tendency to mess up our algorithms later on, as we will

see.

One more definition before we see pictures again. This is an important subset of Kleinian groups that will aid in our intuition about Kleinian groups in general.

Definition: a *Schottky group* is a Kleinian group generated by transformations which pair distinct circles in the plane (This is a restricted definition of Schottky groups that will be sufficient for our purposes.

A loxodromic transformation pairs two distinct circles C1 and C2 (a parabolic transformation cannot meet these requirements unless the circles are tangent) if

- The transformation maps C1 to C2, mapping the inside of C1 to the outside of C2, and the outside of C1 to the inside of C2.
- The transformation's attractive fixed point is in C2, and the repelling fixed point is in C1.
- The image of C2 under repeated iterations of the transformation is smaller circles containing the attractive fixed point.

To see what this means, let us pair two circles, both of radius 1, with centers at 2 and -2. Let's find the transformation that maps the one at -2 to the one at 2. They are obviously mapped to each other by $z \mapsto z + 4$, but this does not map the outside of the first to the inside of the second. So, first map to 0. $z \mapsto z + 2$. Then invert with $z \mapsto 1/z$. Then map 0 to the next center with the first transformation again. So the final function becomes $1/(z+2)+2$. Some points are checked in the table. If inside the left circle, a point should move outside the right. If outside the

Z	1/(*z*+2)+2
0	5/2

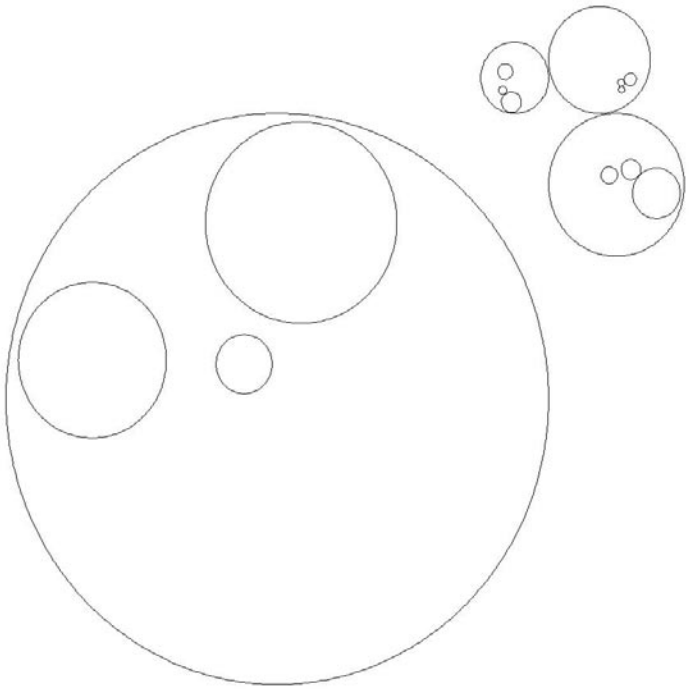
-2	∞
-5/2	0
-3	1
$\sqrt{5}$	$\sqrt{5}$

left, it should move inside the right.

For a more in depth example, let us pair the circles at -1 + *i* and 2 + *i* with radii ½ and 1, respectively. First, move -1 + *i* to the origin by adding 1 – *i*. Then, invert and change size—we have a circle of radius ½, we need one of radius one. $z \mapsto 1/z$ takes ½ to 2, so compensate by using $z \mapsto 1/(2z)$. Then translate again by adding 2+*i*. The final composition becomes

$$\frac{1}{2(z+1-i)} + 2 + i.$$

Below is a Schottky group on two generators. In this image, opposite circles are paired by loxodromic transformations, with an iteration of each transformation on each other circle. See how the image circles get smaller. Also, each interior of an original circle is an image of the plane outside the opposite circle—there are three circles with three circles in each one. There is a transformation in the group that pairs any two of these circles, so the three circles inside another will propagate deeper and smaller into the plane. This is our first taste of the fractal possibilities of Kleinian groups. Once a structure appears—like three circles in one circle—it is transported by the group elements to many other levels and locations by the composed Möbius transformations.



One other kind of group will be a helpful illustration.

Definition: A *Fuchsian group* is a Kleinian group that leaves invariant a circle.

"Leaves invariant" means maps to itself. That does not mean every point is fixed, it only means the map of, for example, the unit circle will be the unit circle. Remember: the only transformation that fixes more than two points is the identity. So every Fuchsian group has an associated circle that is the image of itself under each transformation in the group.

An especially symmetrical Fuchsian group is the one generated by

$$a = \begin{pmatrix} \sqrt{2} & i \\ -i & \sqrt{2} \end{pmatrix} \text{ and } b = \begin{pmatrix} \sqrt{2} - i & \sqrt{2} \\ \sqrt{2} & \sqrt{2} + i \end{pmatrix}$$

The special symmetry happens because the four Schottky circles are tangent at the intersections of the lines $y = \pm x$ with the unit circle. To check that the unit circle is preserved by each generator, one need only check three points—three points determine a circle.

z	$a(z)$	$\lvert a(z)\rvert$	$b(z)$	$\lvert b(z)\rvert$
1	$(1+2i\sqrt{2})/3$	1	$(7-4i\sqrt{2})/9$	1
i	1	1	$-i$	1
-1	$(1-2i\sqrt{2})/3$	1	1	1

Since in the complex plane, lines and circles are the same, any Fuchsian group is

conjugate to a group that fixes the real axis. This should remind you of the group that preserves the upper half plane, SL(2,$\mathbb{R}$). Furthermore, the image circles of the Schottky circles, because they cross the real axis twice at right angles, are centered on the real axis. Note: the inversion $1/z$, which maps $-i$ to i is not a member of SL(2,$\mathbb{C}$). Its determinant is -1. The slightly different transformation $-1/z$ does preserve the upper half plane. The set of all these transformations is called the upper half plane group. Any Fuchsian group is conjugate to a subgroup of the upper half plane group.

Chapter 3: Limit sets

In the last chapter, the hyperbolic metric on the upper half space $\mathbb{H}^3$ was given. This metric is only useful in the upper half space, not on the plane—notice that $y = 0$ gives infinity. So, the metric is useful on the interior of the space, but not on the boundary, for the boundary is infinitely far away. However, if one simply used the Euclidean metric, one could take distances and check convergence to the boundary. This would be given by

$$|z, y| = \sqrt{|z|^2 + y^2},$$

This metric encounters no such problems on the boundary. This is a metric on $\overline{\mathbb{H}^3}$. It is the metric by which points converge in the following definition.

Definition: A point x of $\mathbb{C}_\infty$ is a *limit point* of a group of Möbius transformations Γ if there is a sequence of group elements $\{g_i\}$ and a point $y \in \overline{\mathbb{H}^3}$ s.t. $\lim g_i(y) = x$, and this sequence is not eventually constant. The *limit set* $L(\Gamma)$ of a group is the set of all limit points of Γ.

Remember that the hyperbolic metric preserves distances under Möbius transformations, so convergence would not occur at all under that metric. This kind of limit point, a limit point of a topological group, is not to be confused with the limit point of a set or sequence. It is related, but not quite the same. The relation is that the limit set is the limit points of the orbits of Γ, which are called limit points of Γ directly for convenience. The definition is stated in terms of the upper half-space and the plane, but can be extended to all dimensions, including upper half space groups having limit points on the extended (including ∞) real line. The useful proposition below extends this a bit. (From Ratcliffe 12.1.2 [3])

Proposition: For a torsion-free Kleinian group Γ and any point x in $\mathbb{H}^3$,

$$L(\Gamma) = \overline{\Gamma x} \cap \mathbb{C}_\infty$$

Proof: The definition of limit point gives $(\overline{\Gamma x} \cap \mathbb{C}_\infty) \subset L(\Gamma)$.

Now, let a be a limit point of Γ. This means there is a sequence $\{g_i\}$ of elements of Γ and a point $y \in \mathbb{H}^3$ such that $\{g_i(y)\}$ converges to a.

Because of the properties of the hyperbolic metric,

$$d\big(g_i(x), g_i(y)\big) = d(x, y), \text{ for all } x \text{ in } \mathbb{H}^3, i \text{ positive integer.}$$

So, each element of the sequence takes x and y the same distance from each other. So $|g_i(x) - g_i(y)|$ gets smaller and smaller as $i \to \infty$. The sequences $\{g_i(x)\}$ and $\{g_i(y)\}$ have the same limit point, and a is a complex number, so $a \in \overline{\Gamma x} \cap \mathbb{C}_\infty$ and $\overline{\Gamma x} \cap \mathbb{C}_\infty \subset L(\Gamma)$.

□

This will prove very important when graphing. Instead of selecting a point in $\mathbb{H}^3$ and calculating where it is taken by the elements of the group, a point in $\mathbb{C}$ will suffice.

This next theorem comes easily from the idea of limit points.

Theorem 3.1: A fixed point of a parabolic or loxodromic element of a group Γ is a limit point of Γ.

Proof: Let $h \in \Gamma$ be parabolic or loxodromic, fixing a point a. h (or h^{-1}) has a for an attractive fixed point. Thus, for an $x \in \mathbb{C}$ close to a (or even not a fixed point of h), the sequence $\{h^n(x)\}$ (or $\{h^{-n}(x)\}$) converges to a. □

So, we have limit points of groups of Möbius transformations which are the accumulation points of the action of the group on any point that's not a fixed point. The following lemma is useful in the proof of the next theorem, modified from Ratcliffe's 12.1.3 [3].

The points fixed by every element in a group are very important to the structure of the limit set. To see this, let's observe what happens to the limit set if all group elements have the same fixed points. For simplicity's sake, we assume that there is not an elliptic transformation in the group.

Theorem 3.2: For a torsion-free Kleinian group, the following are equivalent (each implies the others).

(1) All elements of Γ fix the same point.

(2) $L(\Gamma)$ has 0, 1, or 2 points.

(3) $L(\Gamma)$ is finite.

Proof: First, show (1) $\Rightarrow$ (2). If more than two points are fixed by all elements of Γ, the only transformation in the group is the identity, and the only sequences given are $\{I^n(x)\}$. None of these sequences are distinct, so $L(\Gamma)$ has 0 points. If exactly one point is fixed by all elements of Γ, then that point is a limit point, as there is a parabolic or loxodromic element that fixes that point (4.1). If a parabolic and a loxodromic element share a fixed point, the group is not discrete (2.3). Thus, all elements of the group are parabolic elements fixing the same point. Conjugate this fixed point to ∞. Then, all elements are of the form $z \mapsto z + c$, with c complex, so any convergent sequences converge to ∞.

Now, if all elements of Γ fix the same two points $\{a, b\}$, then each element of Γ is a loxodromic (or hyperbolic) transformation fixing a and b. By 4.1 again, a and b are in the limit set, that is, $\{a, b\} \subset L(\Gamma)$. To show a subset in the other direction, use the proposition. For a point $x \in \mathbb{H}^3$, the limit set is a subset of its orbit's closure. Because all elements of Γ have the same fixed points, the only accumulation points for the orbit are the two fixed points. If there were a sequence of group elements taking x to another point

in the plane, it would be a fixed point of a transformation (if the group is discrete), or the group would fail to be discrete. Thus,

$$L(\Gamma) \subset \{a, b\}$$

So (1) ⇒ (2).

(2)⇒(3) obviously (0, 1, and 2 are finite).

Show (3)⇒(1):

Assume that $L(\Gamma)$ is finite. Let x be a point not in the limit set. Assume first that Γx is finite. As long as Γ is discontinuous at x, there are finitely many elements g of Γ such that $g(x) = x$. Thus, Γ is finite, and the only finite group without elliptic elements is the trivial group, which fixes all points. Assume now that Γx is infinite. Γx then has a limit point a in $\mathbb{C}_\infty$. This limit of the orbit is a limit point of the group, so since $L(\Gamma)$ is finite, and the limit set is invariant under the action of the group, Γa is finite. Since the only elements in the group are parabolic and loxodromic, a finite orbit is only achieved at a fixed point. a is fixed by all elements of the group. (3)⇒(1). □

So, if the limit set is finite at all, the group is very simple. And the limit set, if it has more than two points, has infinitely many points! We will investigate "how much" infinitely presently, because the limit set is uncountable—any mapping to the integers will not cover all points of the limit set. To show this, and a few other things about the limit set, we will delve briefly into the *hyperbolic convex hull*. This treatment in 3 dimensions is based on Ratcliffe's n-dimensional presentation [3].

Definition: A subset K of $\mathbb{H}^3$ is *hyperbolic convex* if any two points of C are connected by a hyperbolic line segment, ray, or line contained in C.

But these hyperbolic line segments are not quite the same as regular ones. In our

case, the set will be convex if we can travel to any two points by either straight lines or by straight lines on the surface of spheres centered on (orthogonal to) $\mathbb{C}$ (remember that straight lines are just circles through ∞).

Definition: The *hyperbolic convex hull* of a set $K \subset \overline{\mathbb{H}^3}$ is the intersection (denoted $C(K)$) of all hyperbolic convex subsets of $\overline{\mathbb{H}^3}$ that contain K.

So, $C(K)$ is the "smallest" hyperbolic convex set that contains K. It works like connecting the points of K by hyperbolic lines. But it does it in the most efficient way because of the intersection. Sure, the regular convex hull (connect everything with lines) of a set K would be hyperbolic convex, but there may be a smaller one. $C(K)$ is the smallest. This lemma will give the next very useful theorem.

Lemma: For a Kleinian group Γ and K a closed, Γ-invariant subset of $\mathbb{C}_\infty$, $C(K)$ is closed and Γ-invariant.

Proof: The proof requires a different model than the upper half space model. The *projective disk* model will be what we need, or D^3. In this model of the unit ball, the hyperbolic lines are straightened out, so the hyperbolic convex hull of K corresponds to the regular convex hull of the D^3 version of K. This straightening wreaks havoc on the metric, but we will not be needing it here. So, $C(K)$ is now just the convex hull of K in D^3. Its compactness makes it closed and bounded, so now we must check if it is Γ-invariant.

Let $g \in \Gamma$.

$gC(K) = g(\cap \{S : S \supset K$ and S is a convex subset of $\overline{D^3}\}$

$= \cap \{gS : S \subset K$ and S is a convex subset of $\overline{D^3}\}$

But, each S is a superset of K, which is invariant under Γ. Applying g to S will not

change the final intersection.

$gC(K) = \cap \{gS : gS \subset K$ and gS is a convex subset of $\overline{D^3}\}$

Then, since we pick an arbitrary convex gS a subset of $\overline{D^3}$, we can just call it S.

$gC(K) = \cap \{S : S \subset K$ and S is a convex subset of $\overline{D^3}\}$

$= C(K)$.

So, C(K) is Γ-invariant and closed. These remain true outside of the D^3 model, as Γ-invariant and closed do not depend on the metric of the space, but rather the topology.

□

Theorem 3.3: Let Γ be a Kleinian group with an infinite limit set. Then every nonempty, Γ-invariant, closed subset of $\mathbb{C}_\infty$ contains the limit set, $L(\Gamma)$.

Proof: Let $K \subset \mathbb{C}_\infty$ be nonempty, Γ-invariant, and closed. Then, K is infinite, as Γ has either a parabolic or loxodromic element to take a single point of K to infinite orbits toward a fixed point, and the group does not satisfy the conditions of Theorem 3.2. Where $C(K)$ is the hyperbolic convex hull of K, $C(K)$ is also Γ-invariant and closed. And, $C(K) \cap \mathbb{C}_\infty$ is just K—the "connecting lines" that make K hyperbolic convex will be hyperbolic. Now, pick a point in $C(K)$ not on $\mathbb{C}_\infty$. It satisfies this chapter's proposition with $L(\Gamma) = \overline{\Gamma x} \cap \mathbb{C}_\infty$. But Γx is in $C(K)$, as it is Γ-invariant, and its closure is in $C(K)$, as it is closed. So $L(\Gamma) \subset C(K)$. And $L(\Gamma) \subset \mathbb{C}_\infty$ by definition, so $L(\Gamma) \subset C(K) \cap \mathbb{C}_\infty = K$. K contains the limit set. □

So, any time K is a Γ-invariant set, its closure contains the limit set. The limit set is the "smallest" closed, Γ-invariant set, to put it another way. The following proofs are entirely my own.

Theorem 3.4: Let F denote the set of all fixed points of loxodromic elements of a

torsion-free Kleinian group Γ with an infinite limit set. Then, $\bar{F} = L(\Gamma)$.

Proof: First, notice that by 3.1 and the fact that $L(\Gamma)$ is closed, $\bar{F} \subset L(\Gamma)$.

We must show F is nonempty in order to show that it is eligible for Theorem 3.3. Assume a group has two parabolic elements (not the identity) with different fixed points (if the group has a single fixed point, the limit set is not infinite). Conjugate one fixed point to zero, the other to ∞. The transformations t_1, t_2 will have the form:

$$t_1 = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, t_2 = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

Then, $t_1\, t_2 = \begin{pmatrix} 1 + bc & b \\ c & 1 \end{pmatrix}$, and Tr($t_1\, t_2$) = 2+bc. If $t_1\, t_2$ is loxodromic, F is nonempty. For $t_1\, t_2$ to not be loxodromic (the group has no elliptic elements), Tr($t_1\, t_2$) must be 2 or -2. This occurs if bc = 0 or -4. bc = 0 implies t_1 or t_2 is the identity, so assume bc = -4.

$$t_1\, t_2 t_2 = \begin{pmatrix} 1 + 2bc & b \\ 2c & 1 \end{pmatrix}$$

Tr($t_1\, t_2 t_2$) = 2+2bc = -6, so there must be a loxodrome in Γ. F is nonempty.

Let $a \in F$ be a fixed point of $h \in \Gamma$, a loxodrome. For $g \in \Gamma$, $g(a)$ is also in F. It is the fixed point of $g\, h\, g^{-1}(z)$, which is loxodromic because it is conjugate to h. Thus, F is Γ-invariant.

$\bar{F} \subset \mathbb{C}_\infty$ is nonempty, closed, and Γ-invariant, so $L(\Gamma) \subset \bar{F}$.

$$\bar{F} = L(\Gamma) \qquad \square$$

Theorem 3.5: For a Kleinian group Γ, if $L(\Gamma)$ is infinite, $L(\Gamma)$ is perfect (has no isolated points).

Proof: Let a be an isolated point of $L(\Gamma)$. Then, a is an isolated point of F, and there is a loxodromic element $h \in \Gamma$ that fixes a. For $b \in F$ not fixed by h, but fixed by another

element g ($\in \Gamma$). Then, assume a is the attractive fixed point of h (so, swap h for its inverse if necessary). Repeated conjugation by h makes a sequence in F converging to a. That is,

$$h^m \circ g \circ h^{-m}(b) \to a \text{ as } m \to \infty.$$

So, a is not isolated in $L(\Gamma)$, contradicting the assumption, so no such point exists. □

It is well known in mathematics that any perfect set is also uncountable. So, we have seen so far that if a limit set has more than two points, it has uncountably infinitely many. Another important fact is that the limit set is not nearly all of $\mathbb{C}$ for a discrete group. It is *nowhere dense*, which means $L(\Gamma)$ is dense no neighborhood in $\mathbb{C}_\infty$. A set X is said to be *dense* in another set Y if every point in Y is in X or has a sequence in X converging to it. So, the rationals are dense in the real numbers. Any real number can be approximated with fractions. But, no neighborhood in the plane can be approximated by the limit set of a discrete group. Alternately, the interior of the closed set $L(\Gamma)$ is empty.

Definition: The ordinary set of a Kleinian group Γ is the set of points $O(\Gamma)$ in $\mathbb{C}_\infty$ but not in $L(\Gamma)$. It is just everything else in the plane but the limit set.

Theorem 3.6: The limit set of a torsion-free Kleinian group Γ is nowhere dense in $\mathbb{C}_\infty$.

Proof: If the limit set is finite, this is obvious. Assume the limit set is infinite. So long as the group is discrete, there is a point not in the limit set. Then, the nonempty set $O(\Gamma)$ is Γ-invariant because its complement is. Thus, its closure $\overline{O(\Gamma)}$ is also Γ-invariant. So, by the ever-useful Theorem 3.3, $L(\Gamma) \subset \overline{O(\Gamma)}$. Since $L(\Gamma)$ is entirely composed of limit points of its complement, it has no interior, and so is nowhere dense. □

Chapter 4: Graphing the limit set

As we consider the limit set, our experience with transformations in general may make us wish to make pictures of these closed, nowhere dense subsets of the plane. Some pictures may be difficult. In the case of Schottky groups, the limit sets are not always connected. As transformations map smaller and smaller circles, the accumulation points—the limit set—is a series of disconnected points. The set is perfect, as we discussed before, so no point is isolated.

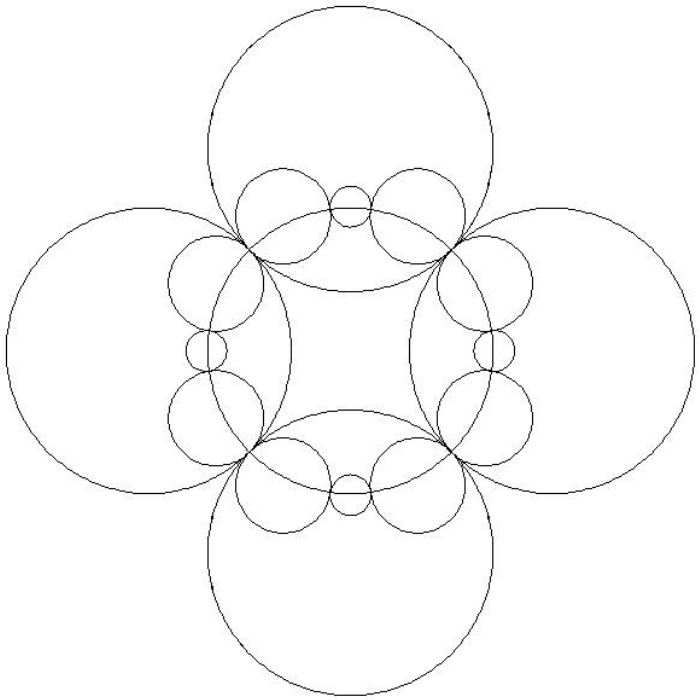
When limit points are not connected, the limit sets are collections of points that resemble the Cantor set, the classic example of a perfect disconnected set. Begin with the closed segment [0,1]. Remove the middle third. Left with $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, take out the middle third of each of those to get $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{4}{9}, \frac{2}{3}] \cup [\frac{8}{9}, 1]$. Continue taking out the middle third of each segment. The limit of this process is the Cantor set. Observe what happens at endpoints. 1 is first the side of the original segment, but as the process continues, a sequence is made converging to 1. This happens with every endpoint of the remaining segments, so no point in the Cantor set is isolated, and it is perfect.

Let us first explore a Fuchsian group, that generated by

$$a = \begin{pmatrix} \sqrt{2} & i \\ -i & \sqrt{2} \end{pmatrix} \text{ and } b = \begin{pmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{pmatrix}$$

We looked at a similar group in chapter 2, and this group also preserves the unit circle. The picture below draws two levels. The first is the paired circles of the generators, then each generator is applied to each circle. The disks become smaller with successive iterations, converging to points on the unit circle. The limit set of the group is the whole unit circle, not some jumble of points. One can ask, “What's different? Why

does this group have a connected limit set?" There are a few key differences, given as "necklace conditions" in chapter 6 of *Indra's Pearls* [2].



For the limit set to be connected, the tangency points must be mapped to each other. In this picture, if the bottom right point of tangency were mapped by *b* into the left circle at a point other than the bottom left point of tangency, the limit point would not be able to be continuous at that point. So that point must be mapped to itself by doing the transformation *b*, then *a*, then *B* then *A*, or by *ABab*. Likewise, the bottom left point must be a fixed point of *bABa*, the top left by *abAB*, and the top right by *BabA*. This will cause the points of tangency to be correctly mapped to each other. It will also be incredibly convenient if these transformations are parabolic, so points do not get dragged to some other fixed point. The transformation *abAB* is called the **commutator**, because if the operation commutes, then $abAB = aAbB = I$. If it does not, *abAB* is a quantity that suggests how close the two elements are to commuting with each other.

These four parabolic fixed point conditions are actually the same. Each form of the commutator above is a conjugation of the one before. $bABa = b(ABab)B$ (conjugating by b, B), $abAB = a(bABa)A$, $BabA = B(abAB)b$, and $ABab = A(BabA)a$. So if one is parabolic, all are parabolic, as they are conjugate to each other.

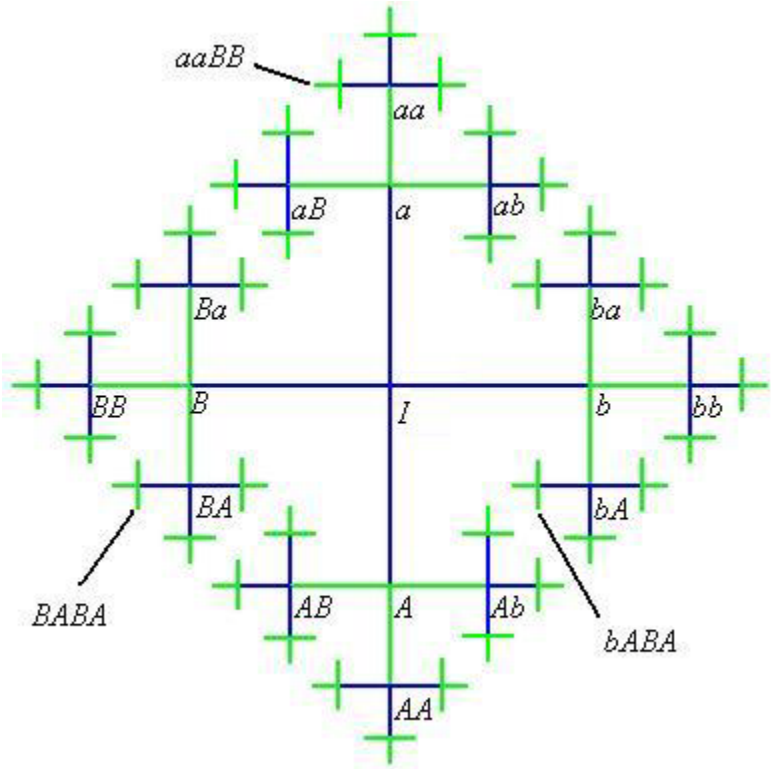
A final requirement, assuring that the generators do not share a fixed point, is that Tr($abAB$) = -2.

Before we see an algorithm, let us first explore the "tree" of compositions. An adaptation of the graphic favored by Mumford, Series, and Wright is below. This tree represents possible compositions, elements of a group generated by a and b. Beginning with the identity in the center, each step represents the action of a transformation, so that each location has a transformation named by the composition of the steps taken to get there. For example, notice the composition aabB is not on the tree—bB is redundant. Only “reduced” transformations are on the tree.

Our algorithm will need to be able to go deeper into the tree (farther from I), turn (if it has traversed enough of a branch to get a good picture of the limit set), and obviously plot limit points. Let us observe the some turns, though, to generate even more respect for the parabolic commutator. If one begins at a and goes as far right as possible, one gets the composition $\psi = abABabABabAB\ldots$ In the Fuchsian group above, the limit point this sequence converges to is the point of tangency of the initial Schottky circles for a and b. Also, if one begins at b and takes every possible left hand turn, the transformation made is $\phi = baBAbaBA\ldots$. Now, $\psi^{-1} = \phi$, and ψ is the commutator, so both are parabolic. ϕ and ψ have the same attractive fixed point, points are merely attracted from the “other side,” (consider $z \mapsto z+c$ and $z \mapsto z - c$). So, the limit point represented by the infinite sequence of transformations ϕ is the same as the limit point represented by ψ.

This coalescing happens anywhere one starts on the tree. From aB, turning all the way left gets one to $aBAbaBAb\ldots$, while starting at Ba and turning right gives

BabABabA…, again the inverse of the previous parabolic transformation—giving the same point in the limit set from both words. So, if one were to traverse the outermost edge of the tree (the "deepest" level), clockwise (or counterclockwise), the limit set would be the result in order. So, our algorithm must enter the tree, go deep enough so that if it skipped to the next branch, the skip would be visually almost indistinguishable, then skip to the next branch and plot the image of the tree element on a point. Here is another visualization to observe how the algorithm works. In this diagram, the generators are represented by directions. Traveling up applies *a*, and down applies *A*. Likewise, traveling right applies *b*, left, *B*. As before, observe that there are no redundancies in the diagram (*aA* is just represented by *I*). Each intersection of paths represents a word, and a few are labeled to show. The colors are alternated so the successive levels are clear.



The algorithm will enter a branch, say *a*. It will then take all available right turns until it is "deep enough" (this method is called the *depth first search*). Then it will spit out a limit point by applying the resulting group element to a point in the plane. Then it will go back one level and turn. The algorithm will traverse all relevant points by

following the right-hand wall in this way. For example, if it was told to provide the elements up to level three, it would give *a*, *ab*, *abA*, then go back to *ab* before going to *abb*, go back, turn, and give *aba*. Then, when it goes back, there are no available turns (a right turn would take it backward), so it goes back to *a* and does the same thing from the intersection *aa*, giving *aab*, *aaa*, and *aaB* before returning to *a* again, continuing to follow every available right turn.

Now, we have seen how to make the limit set a curve, and we have found what order the points of the set come in, and we have the basics of an algorithm for drawing that curve. All we need now are some groups. Mumford, Series, and Wright have provided a useful normalization of a family of groups based on two parameters—the trace of the generator matrices *a* and *b*. Here is what is meant by normalization. The limit sets are only interesting as far as conjugation. One need not graph carefully a conjugate limit set to one already studied—the conjugate contains all the information in a different view. A normalization, then, is a selection of which conjugation to use. Mumford, Series, and Wright explain the parameters they chose in this way. Two matrices have eight total parameters. Requiring determinant one reduces it to six. Conjugation halves that because any three points can be mapped to any other three by a Möbius transformation. Then, they apply the Markov identity

$$(\mathrm{Tr}\, a)^2 + (\mathrm{Tr}\, b)^2 + (\mathrm{Tr}\, ab)^2 = \mathrm{Tr}\, a\ \mathrm{Tr}\, b\ \mathrm{Tr}\, ab,$$

which guarantees that the commutators have trace -2. Selecting the three parameters to be the traces of *a*, *b*, and *ab*, the Markov identity reduces it to two necessary parameters.

Let us observe some things about Grandma's Recipe. First, it is origin symmetric. To see why, observe that *a* and *ab* have the same element in first and last positions. They are of the form

$$T = \begin{pmatrix} r & s \\ t & r \end{pmatrix}.$$

Grandma's Recipe for parabolic commutator groups:

For two complex numbers t_a and t_b, set $t_{ab} = t_a t_b - \frac{1}{2}\sqrt{(t_a t_b)^2 - 4(t_a^2 + t_b^2)}$ (to satisfy the Markov equation).

Compute $z_0 = \frac{(t_{ab} - 2)t_b}{t_b t_{ab} - 2t_a + 2i\, t_{ab}}$.

The generator matrices are then:

$$a = \begin{pmatrix} \frac{t_a}{2} & \frac{t_a t_{ab} - 2t_b + 4i}{(2t_{ab} + 4)z_0} \\ \frac{(t_a t_{ab} - 2t_b - 4i)z_0}{2t_{ab} - 4} & \frac{t_a}{2} \end{pmatrix}$$

$$b = \begin{pmatrix} \frac{t_b - 2i}{2} & \frac{t_b}{2} \\ \frac{t_b}{2} & \frac{t_b + 2i}{2} \end{pmatrix}$$

Thus, $$ab = \begin{pmatrix} \frac{t_{ab}}{2} & \frac{t_{ab} - 2}{2z_0} \\ \frac{(t_{ab} + 2)z_0}{2} & \frac{t_{ab}}{2} \end{pmatrix}$$ (Wright 229)

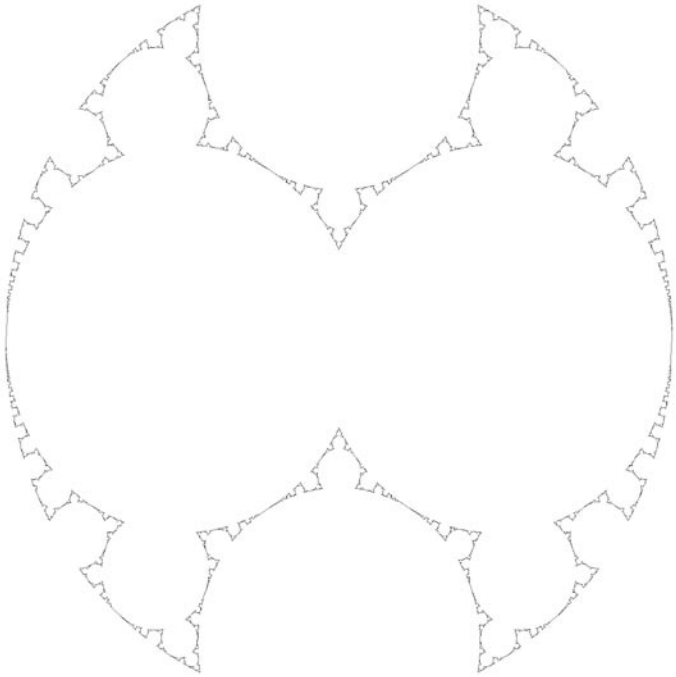
$$-T(-z) = -\frac{r(-z) + s}{t(-z) + r}$$

$$= \frac{rz - s}{-tz + r} = T^{-1}(z)$$

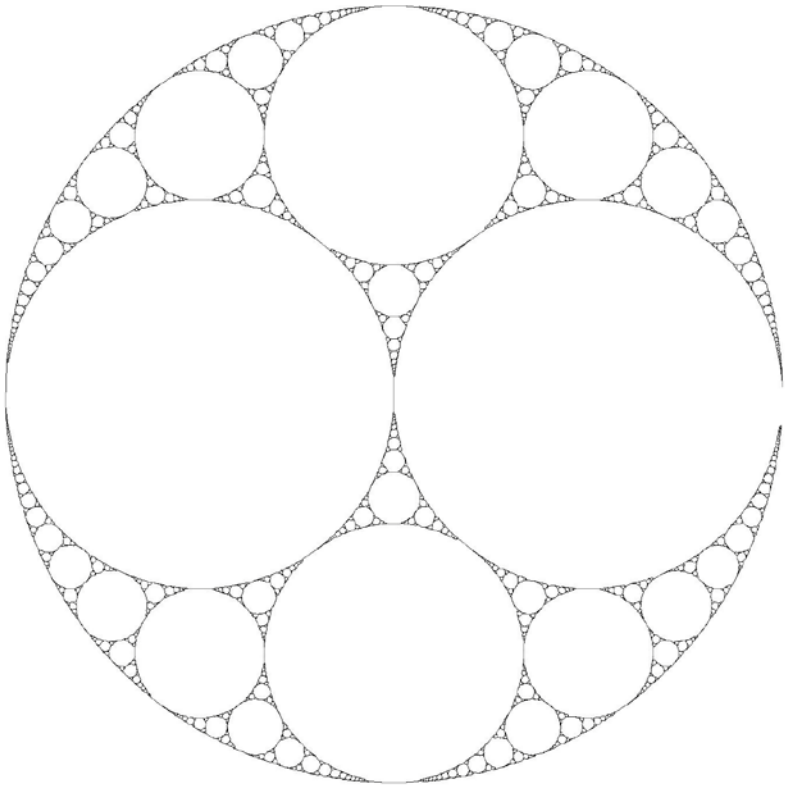
So *T* is not "odd," rather it is kind of "inverse-odd." Conjugating by a rotation of π/2 about the origin produces the inverse transformation. Thus, rotations of π/2 do not affect the limit set because $a(z) = -A(-z)$ and $ab(z) = -BA(-z)$, and *a* and *ab* generate the whole group.

According to the authors, the purpose of the quantity z_0 is to normalize to 1 the fixed point of *abAB*. It also serves to put the fixed point of *aBAb* at -1. So half the limit set is graphed as *a*'s branches are traversed, and one needs only go through that one main

$t_a = t_b = 2.2$



Parabolic generators: $t_a = t_b = 2.0$



branch to see what the whole set looks like (afterwards, rotate around the origin to finish).

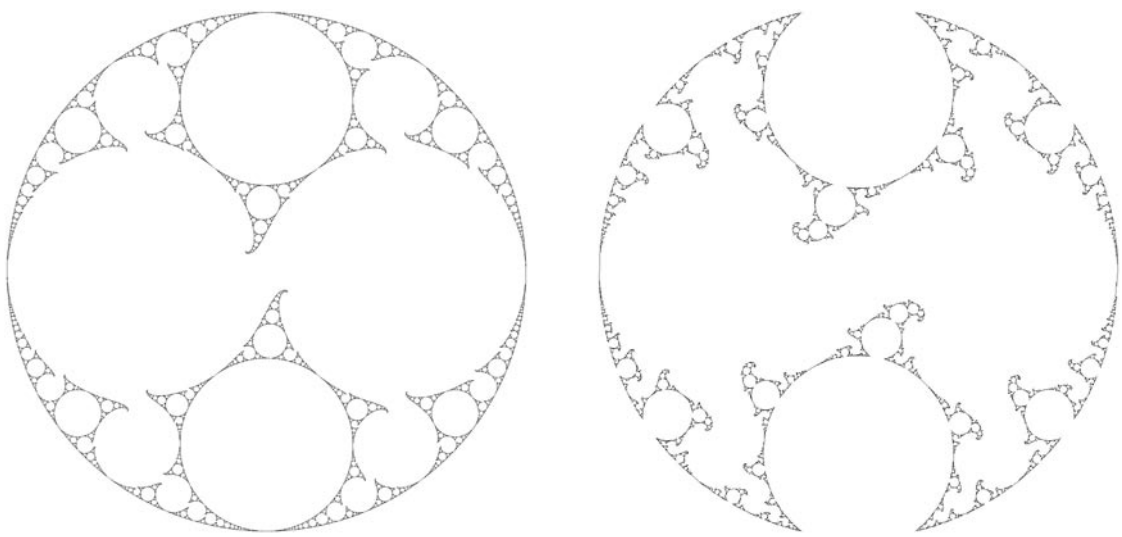
So, initial tries with $t_a = t_b = 3$ or $2\sqrt{2}$ produce limit sets that are circles (the unit circle, because of the normalization). These are both Fuchsian groups. The second, with $2\sqrt{2}$ for both traces, is the Fuchsian group seen in the last chapter. Things get more interesting as we move gradually to parabolic

generators, with trace 2. The limit set starts becoming our first true fractal curve.

Now, observe what happens when the generators move all the way to parabolic. Instead of meeting circles in infinitely many places, but not the whole circle, the parabolic group meets whole circles in an old fractal shape called the Appolonian gasket.

The limit set has become a continuous meandering curve tracing out infinitely many circles. There are Schottky circles that will generate this picture and the last. They are the real axis, two circles tangent to it and to each other, with points of tangency at 1, -1 with the real axis and $-i$ with each other. The fourth Schottky disk, which is paired to the whole upper half plane, is nestled in the ideal triangle between the three other circles, tangent to all three.

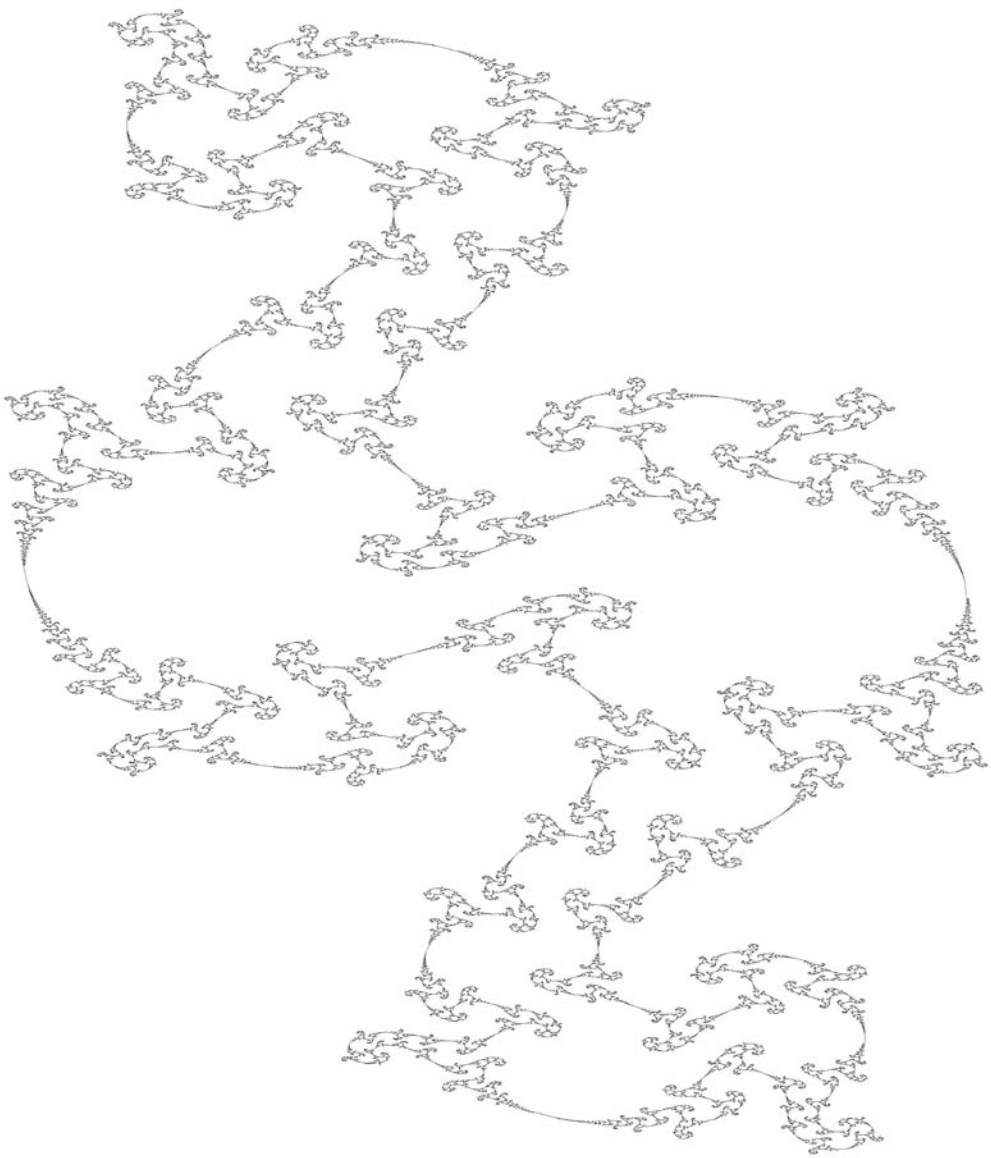
But these have been real traces. The real fun comes when there are imaginary parts to the traces. On the right is a picture with nearly the same traces as the last, $t_a =$



$1.9+.3i$, $t_b = 2.05$. It seems to make a valiant attempt at the gasket formation, but its generators are not quite parabolic. One is very close, and the other is pretty close, too, with some imaginary flavor. The imaginary part is what causes the reappearing spirals.

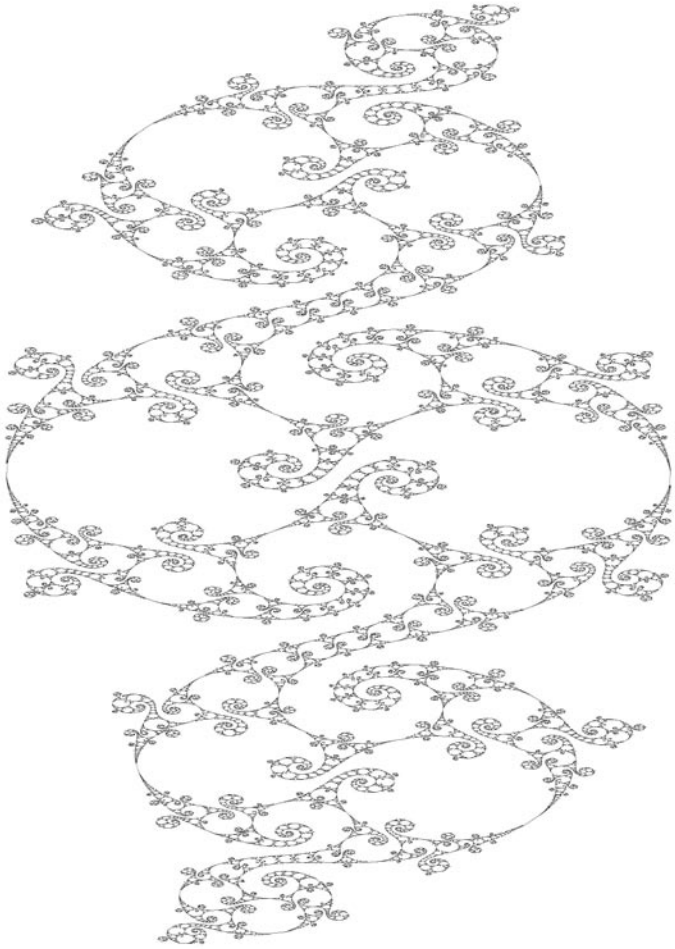
Remember that complex multiplication rotates points as it stretches them. The loxodromes of the group make spirals appear all over the place. The lefthand picture is a much better approximation, with $t_a = 2+.05i$, $t_b = 2$—one generator is still parabolic, but is still short of completing the gasket.

Here is one of my favorites. The generators have traces $1.8+.251i$ and $1.8-.251i$. There are a few prominent loxodromes—one carries points in a spiral between the two

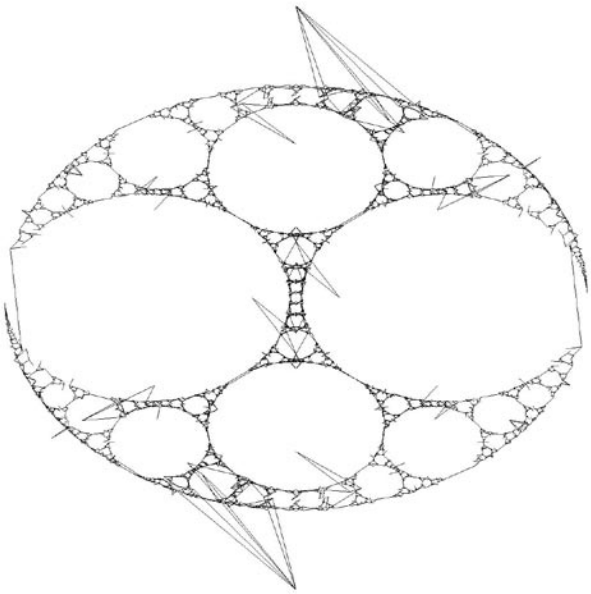


main circles, and others branch off in all directions. If you look closely at the beginning of the set (close to 1), you can see already there the twists and turns that repeat all over the limit set. The small spirals there are also axes of loxodromes. The fixed points of parabolic (or nearly parabolic) transformations are where the curve goes thin and straight, like at 1 and -1 (limit points of the commutator), and scattered throughout. Parabolic transformations can move things more slowly to the fixed point, so the curve through them marches regularly on with fewer twistings.

Twists and turns can be taken to a maximum; the next picture was generated by the nearly elliptical transformations $t_a = 2\cos(\pi/10) + .05i$, $t_b = 2\cos(\pi/10) - .05i$. The nearby elliptical group was set up by the authors of *Indra's Pearls* to show that the algorithm does not run correctly on some elliptic-generated groups, even though they are discrete. The limit sets keep retracing themselves in that case, as $a^{10} = b^{10} = I$ [2]. But, when the transformations become just barely loxodromic (because of complex trace), the algorithm works fine; the points pop out in order and can be connected by lines. Some of the features of the limit set are



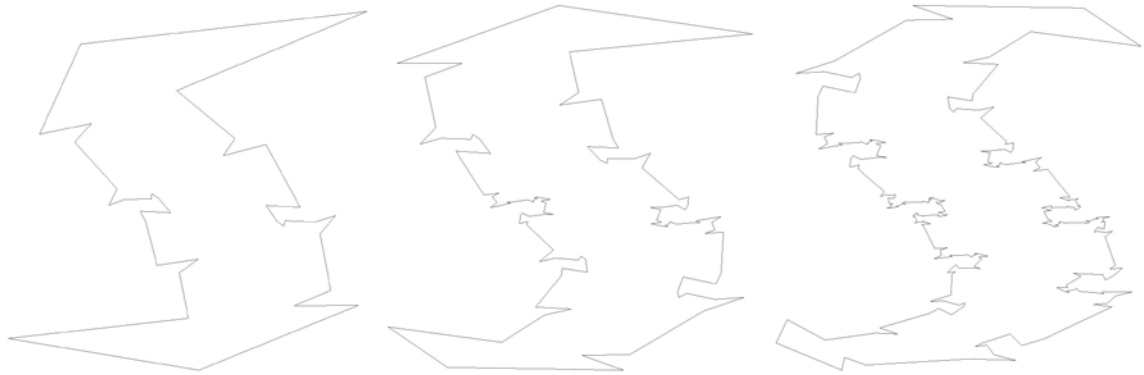
preserved. Compared with the picture in the book, there are 7 almost-circles (places where the limit set touches a circle in many but not all points) in between many of the sections, as in the book's picture of the elliptic case, where the quasicircles are less hidden by swirls.



Below is what happens when my version of the algorithm is run on the elliptic transformation without the additional imaginary part in the trace.

And so we bump into groups that this algorithm cannot handle. Groups that are not discrete are obviously problems. Since elements of the group get close to $\{I\}$, the limit set (if you wish to call it that) is every point in the plane! I do not want to spend hours of CPU time to make the computer draw me a black square.

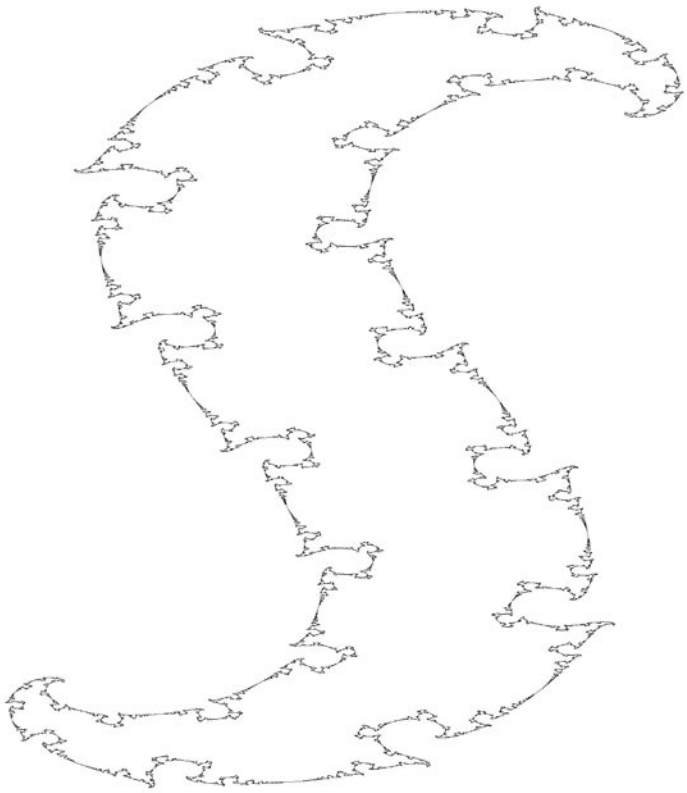
Other play with the algorithm will show how useful it is to delve deeply into the tree. These plots were made by limiting the depth the algorithm could traverse in the



Maximum level is 3 Maximum level is 4 Maximum level is 5

tree.

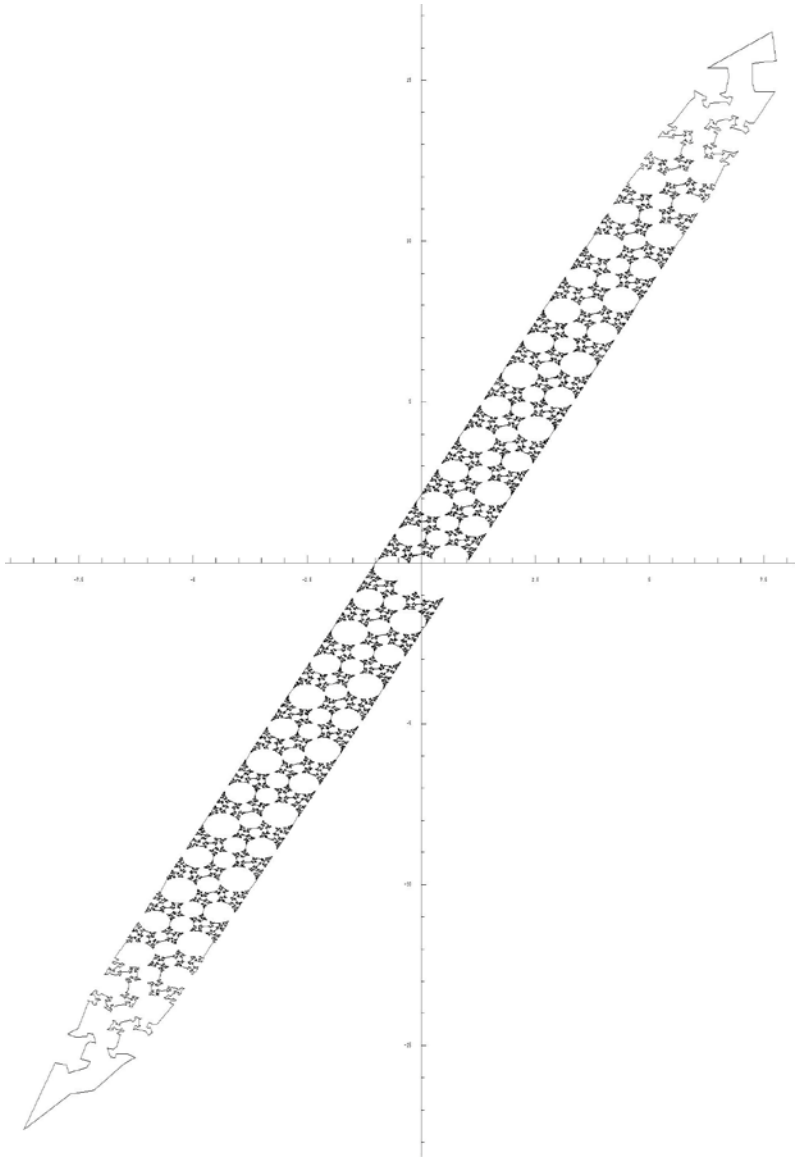
The final picture, when the level the algorithm explores in the tree is "big enough" (usually 100 is plenty, but some groups need more), is on the next page.



As we have added imaginary parts to the traces of our groups, so we have left behind proper Schottky groups. There are not circles that can be paired to generate these pictures. They are merely Kleinian groups, not Schottky or Fuchsian. Since the limit set has an interior and an exterior like a circle, they could still be called quasi-Fuchsian. But there is certainly no Möbius transformation conjugating those spirals to the real line.

One more example was an accidental bump into a group in which the transformation *bab* is, for purposes of computer approximation, a translation. For this group, if the level is allowed to go to 100, as is normal and necessary for other groups, the limit set stretches hundreds in the x- and y-directions. The approximately translational element takes the same limit set that is in between 0 and 2*i* and moves it up and right or down and left. Then, when the maximum level is reached, it gets jerky, just

like the previous poor approximations for an S-shape. So, the limit set stretches out toward ∞, repeating as it goes. The traces used were 1.64213876± 0.6658841*i*.



So play away if you wish. The Mathematica code I used to make the images is in an appendix if you should be curious and wish to try your own traces or alter the code to draw other families of groups. The program generates a list of points and then plots them, so be sure to set the variable named "countermax". An accidentally non-discrete group would run a loop for a very long time.

References

1. Jørgensen, T., On pairs of once-punctured tori, In: *Kleinian Groups and Hyperbolic 3-Manifolds,* edited by Y. Komori, V. Markovic, and C. Series, Cambridge University Press, Cambridge (2003), 183-207.
2. Mumford, D., Series, C., and Wright, David, *Indra's Pearls: The Vision of Felix Klein*, Cambridge University Press, Cambridge, (2002), 1-280.
3. Ratcliffe, J. G., *Foundations of Hyperbolic Manifolds*, Springer-Verlag, New York, (1994), 117-185, 573-585.

Appendix: the code

```
First cell, the helper functions:

levmax = 200;
epsilon = .1;
fixpt[mtrx_] := (mtrx[[1, 1]] - mtrx[[2, 2]] +
          Sqrt[((mtrx[[1, 1]] + mtrx[[2, 2]])^2 - 4 )])/(2*mtrx[[2, 1]]);
mobOnPt[mtrx_,
      z_] := (mtrx[[1, 1]]z + mtrx[[1, 2]])/(mtrx[[2, 1]]z + mtrx[[2, 2]]);
getXY[z_] := {Re[z], Im[z]};
initialize[tra_, trb_] :=
  {
    trab = (tra trb - Sqrt[(tra trb)^2 - 4(tra^2 + trb^2)])/2;
    z0 = (trab - 2)trb/(trb trab - 2tra + 2 I trab);
    a = {{tra/
            2, (tra*trab - 2trb +
                4I)/(z0(2trab + 4))}, {z0 (tra*trab - 2trb - 4I)/(2trab - 4),
          tra/2}};
    b = {{trb/2 - I, trb/2}, {trb/2, trb/2 + I}};
    A = Simplify[Inverse[a]];
    B = Inverse[b];
    gens = {a, b, A, B};
    fix = Map[fixpt, gens];
    word = Table[0*n, {n, 5000}];
    tags = Table[0*n, {n, 5000}];
    tags[[1]] = 1; lev = 1;
    word[[1]] = gens[[1]];
    newpt = 0;
    oldpt = 23;
    counter = 1;
    btflag = False;
    atflag = False;
    }
branchTermination[] :=
  {
    newpt = mobOnPt[word[[lev]], fix[[tags[[lev]]]]];
    If[Abs[newpt - oldpt] < epsilon \[Or] lev >= levmax,
      Sow[{oldpt, newpt}];
      oldpt = newpt;
      counter++;
      If[lev > levmax, Print[LOOKOUT!!]];
      btflag = True, btflag = False];
    }
goForward[] :=
  {
    lev = lev + 1;
    tags[[lev]] = Mod[tags[[lev - 1]] + 1, 4, 1];
    word[[lev]] = word[[lev - 1]].gens[[tags[[lev]]]];
```

```
    }
goBackward[] := {lev -= 1; btflag = False;}
availableTurn[] :=
  If[Mod[tags[[lev + 1]] + 1, 4, 1] == tags[[lev]], atflag = False,
    atflag = True]
turnAndGoForward[] :=
  {
    tags[[lev + 1]] = Mod[tags[[lev + 1]] - 1, 4, 1];
    If[lev == 0, word[[1]] = gens[[tags[[1]]]],
      word[[lev + 1]] = word[[lev]].gens[[tags[[lev + 1]]]]];
    lev = lev + 1;
    }
```

Next Cell, the tree-traversing loop—I left parameters for a group with ∞ in the limit set:

```
levmax = 25;
epsilon = .01;
countermax = 1000000;
initialize[1.64213876 - .76658841I, 1.64213876 + .76658841I];
segs = Reap[
      While[((lev != 0 \[Or] tags[[1]] == 1) \[And] counter < countermax),
          branchTermination[];
          While[! btflag, branchTermination[]; If [! btflag, goForward[]]];
          goBackward[];
          availableTurn[];

          While[lev != 0 \[And] ! atflag, availableTurn[];
            If[! atflag, goBackward[]]];
          If[tags[[1]] == 1 \[Or] lev != 0, turnAndGoForward[]]];];
Print["algorithm finished.  making ptlist"];
ptlist = Reap[
        Do[Sow[getXY[segs[[2, 1, i, 1]]]], {i, 2, Length[segs[[2, 1]]]}]][[2,
     ;    1]];
Show[Graphics[Line[ptlist]], Graphics[Line[-ptlist]], PlotRange -> All,
    Axes -> True, AspectRatio -> 1];
```

Similar code could be adapted for C or java. The key is the method of traversing the tree of words. One must follow right-hand walls until deep enough, then spit out approximations for limit points based on the current transformation (`word`).