KLEINIAN GROUPS AND FRACTALS - MATHCAMP 2021

CONTENTS

1. DAY ONE

Let's begin by looking at some nice pictures.

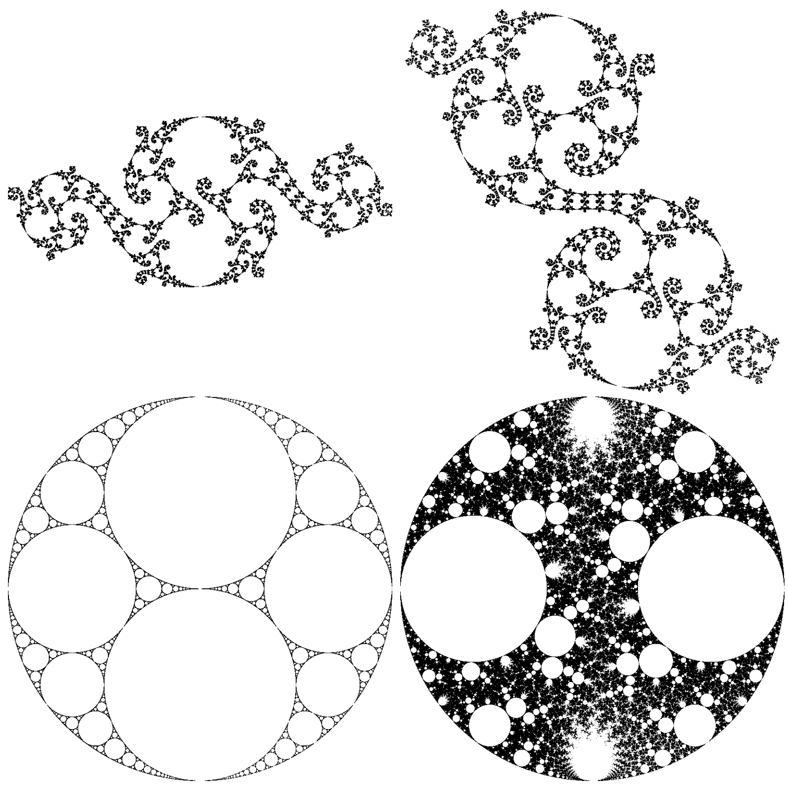

FIGURE 1. Some fractals.

These pictures were generated with the help of *Kleinian groups*.

To understand what a Kleinian group is, we will need to recall some geometric properties of complex numbers. Both translations and rotations in the plane have a nice description in terms of complex numbers. For example, consider the drawing by M. C. Escher shown in Figure 2. If we choose coordinates appropriately, and ignore colors, it is symmetric under the following transformations:

FIGURE 2. A symmetric tiling of the plane.

- The horizontal translation $z \mapsto z + 1$
- The vertical translation $z \mapsto z + i$
- The 90 degree counterclockwise rotation $z \mapsto iz$.
- More generally, any transformation of the form $z \mapsto az + b$ with $a \in \{1, i, -1, -i\}$, $b \in \mathbb{Z} + i\mathbb{Z}$.

Figure 3 shows another Escher drawing. It's also symmetric. We already know how to describe one of these symmetries using complex numbers. If we choose coordinates so that the boundary of the drawing is the unit circle $|z| = 1$, then the 90 degree rotation about the origin is given by $z \mapsto iz$. It turns out that the others can also be described efficiently in terms of complex numbers. The orientation-preserving transformations are of the form

$$z \mapsto \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{C}, ad - bc \neq 0\,.$$

For example, the 120 degree rotation whose center is just to the right of the origin has the parameters

$$a = (1 - e^{i\pi/4})^{-1}, \quad b = -2^{-1/4}i, \quad c = 2^{-1/4}i, \quad d = (1 - e^{-i\pi/4})^{-1}\,.$$

(See the homework for an explanation of where these numbers come from.)

Definition 1. A transformation of the form $z \mapsto \frac{az+b}{cz+d}$ is called a *Möbius transformation.*

A Möbius transformation is generally not a map $\mathbb{C} \to \mathbb{C}$ because the denominator $cz + d$ is zero when $z = -d/c$. Instead, Möbius transformations are maps from the *Riemann sphere* $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ to itself. We use the conventions

$$\frac{a(-d/c) + b}{c(-d/c) + d} = \infty, \qquad \frac{a\infty + b}{c\infty + d} = \frac{a}{c}\,.$$

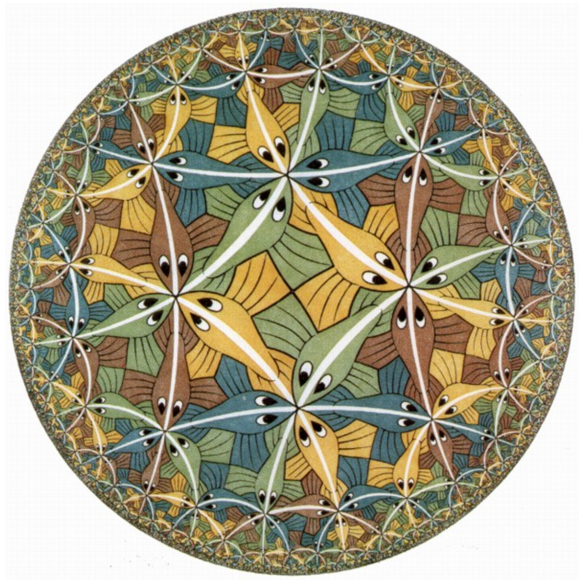

FIGURE 3. Another symmetric tiling.

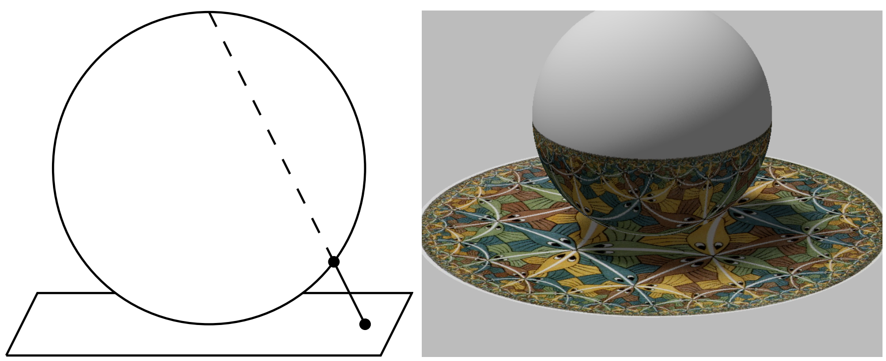

FIGURE 4. Left: Stereographic projection. To project a point P in the plane onto the sphere, draw a line from P to the north pole of the sphere. The other intersection of the line with the sphere is the stereographic projection of P. Right: A stereographic projection of Figure 3.

We call $\hat{\mathbb{C}}$ a "sphere" because there is a *stereographic projection* relating $\mathbb{C}$ to a sphere minus the north pole (see Figure 4). Adding ∞ fills in the missing point.

We leave it as an exercise to the reader to check that Möbius transformations form a group under composition. The (orientation-preserving) symmetry groups of the two Escher drawings are both subgroups of the group of Möbius transformations. They are both discrete in the sense that they do not contain arbitrarily small rotations or translations.

Definition 2. A *Kleinian group* is a discrete subgroup of the group of Möbius transformations.

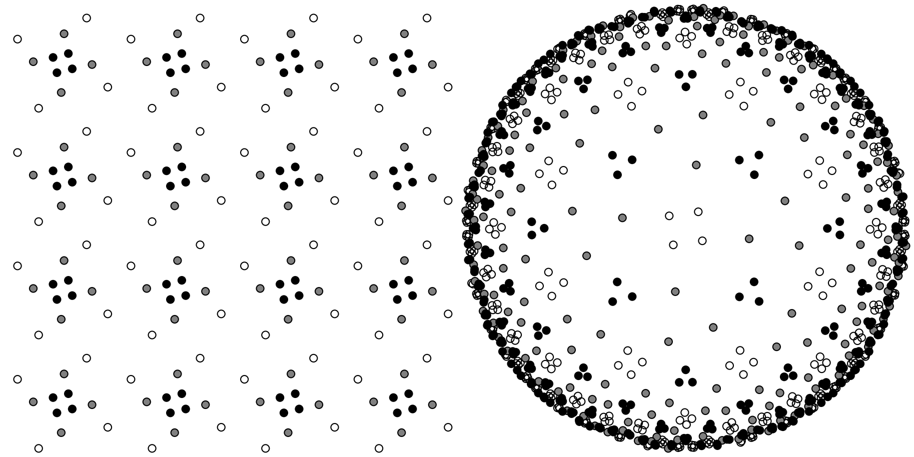

FIGURE 5. Some orbits of the symmetry groups Figures 2 and 3.

The fractals shown in Figure 1 were all generated by choosing a Kleinian group Γ and an element $z \in \hat{\mathbb{C}}$, and plotting $\gamma(z)$ for all $\gamma \in \Gamma$.

Definition 3. Let Γ be a Kleinian group. An *orbit* of Γ is a subset of $\hat{\mathbb{C}}$ of the form

$$\{\gamma \cdot z | \gamma \in \Gamma\}.$$

for some $z \in \hat{\mathbb{C}}$.

In other words, an orbit is the set of all points in $\hat{\mathbb{C}}$ that we can reach from z by applying elements of Γ.

We have just encountered some Kleinian groups. Some of their orbits are shown in Figure 5. Although these orbits are nice in their own right, somehow they are not quite as pretty as the fractals. To understand how to get fancier orbits, we will need to introduce the concept of *uniformization.* Although "uniformization" is a fancy word, you can just think of it as being about gift wrapping surfaces.

Let's look at an example of uniformization. Figure 6 shows another one of Escher's tilings. Let Γ be its group of symmetries. The plane is covered by copies of a fish tile and a butterfly tile. Any two fish tiles are related by an element of Γ, and likewise any two butterfly tiles are related by an element of Γ. We can also think of the plane as being covered by repetitions of a single tile, which consists of one fish and one butterfly. Such a tile is called a *fundamental domain* of Γ.

A *fundamental domain* for Γ is, essentially a way of choosing one point of each orbit of Γ. Any combination of one fish and one butterfly is a fundamental domain of Γ. If we want to be boring, a square is also a fundamental domain. Figure 7 illustrates three different fundamental domains.

Now let's consider the *quotient* $\mathbb{C}/\Gamma$. This amounts to taking opposite sides of the tile and gluing them together. This procedure is depicted in Figure 8. The end result is a torus.

Similarly, Figure 9 shows how a surface with two handles can be expressed as a quotient of the open unit disc.

The process of describing a surface as a quotient of a simpler space is called *uniformization.*

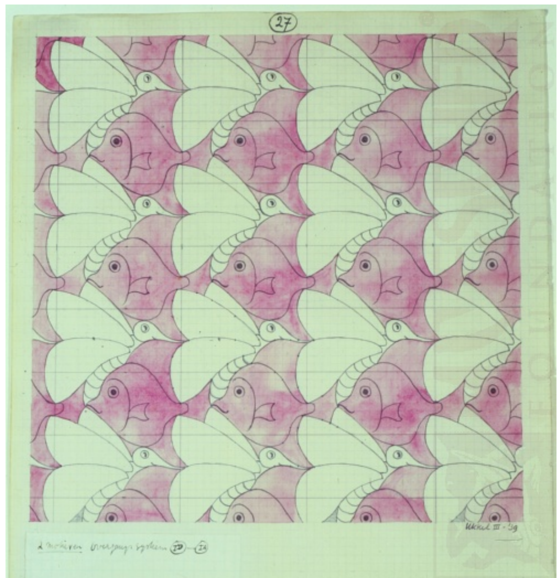

FIGURE 6. Another symmetric plane tiling.

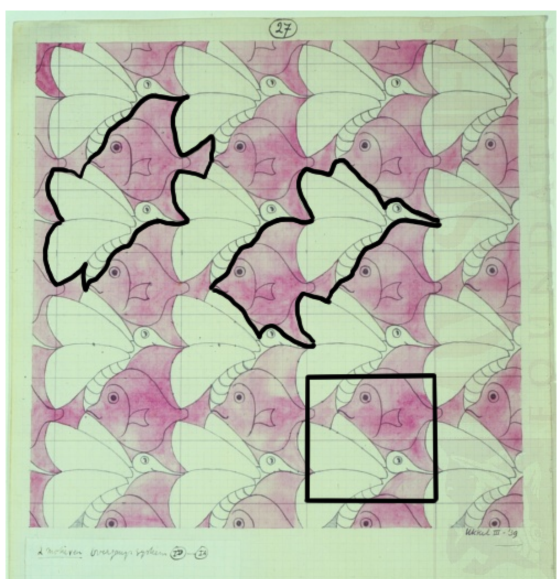

FIGURE 7. Some fundamental domains.

Theorem 4 (Uniformization theorem)**.** *Any orientable surface can be described as a quotient of either* $\hat{\mathbb{C}}$, $\mathbb{C}$ *or the open unit disc* $\mathbb{H}$ *by a Kleinian group.*

More specifically:

- A sphere is isomorphic to $\hat{\mathbb{C}}$.
- A plane, cylinder, and torus are quotients of $\mathbb{C}$.

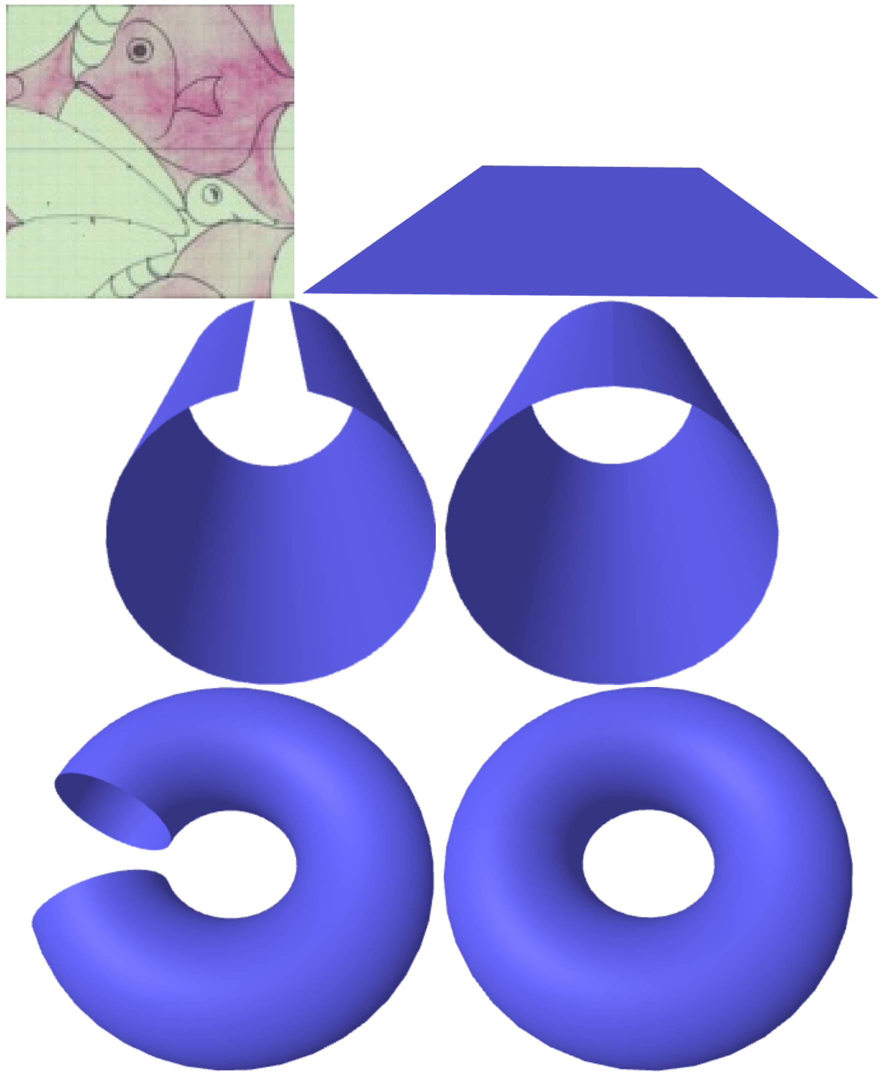

Figure 8. Top left: A fundamental domain of the drawing shown in Figures 6 and 7. Remaining: The quotient $\mathbb{C}/\Gamma$ is constructed by gluing matching edges together. In this case, the quotient is a torus.

- Any other orientable surface is a quotient of $\mathbb{H}$. The process of expressing a surface as a quotient of $\mathbb{H}$ is called *Fuchsian uniformization.*

This description is unique up to isomorphism. The group is called the *fundamental group* of the surface.

Remark 5. In 3 dimensions, there is a similar but more complicated result, known as Thurston's geometrization conjecture. Perelman proved this conjecture (and, as a consequence, also proved the Poincaré conjecture) in 2003.

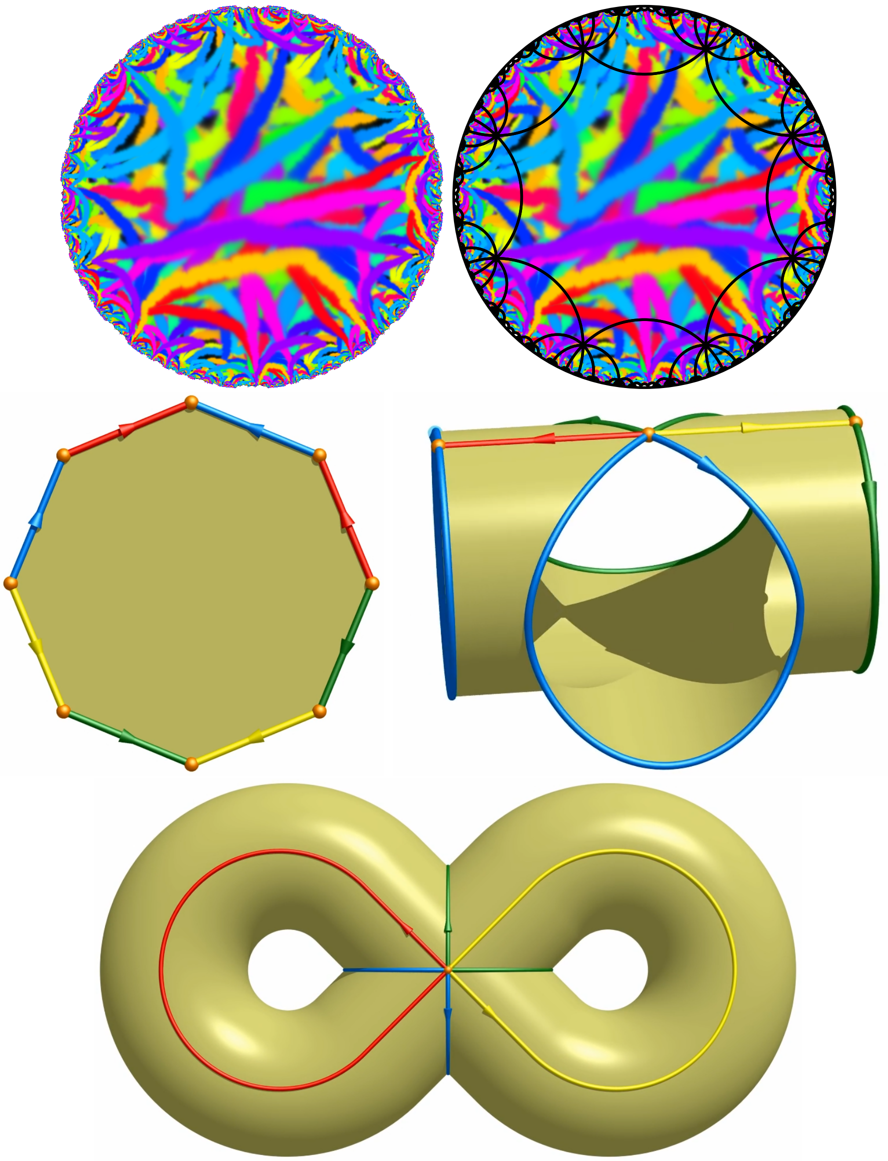

Figure 9. Top left: A symmetric design in the unit disc. Top right: A division of the design into tiles. Below: Gluing the edges of the tile together yields a two-handled surface. Images are taken from this video: `https://www.youtube.com/watch?v=G1yyfPShgqw`. See also `https://www.youtube.com/watch?v=I83K-on4X5A`.

Now let's consider a different type of uniformization, called Schottky uniformization. Figure 10 shows another symmetric tiling. In this case, the drawing is symmetric under scaling $z \mapsto 4z$ (and also under rotation, but we will ignore that). The plane minus the origin can be tiled with copies of the annulus $1 \leq |z| < 4$. If

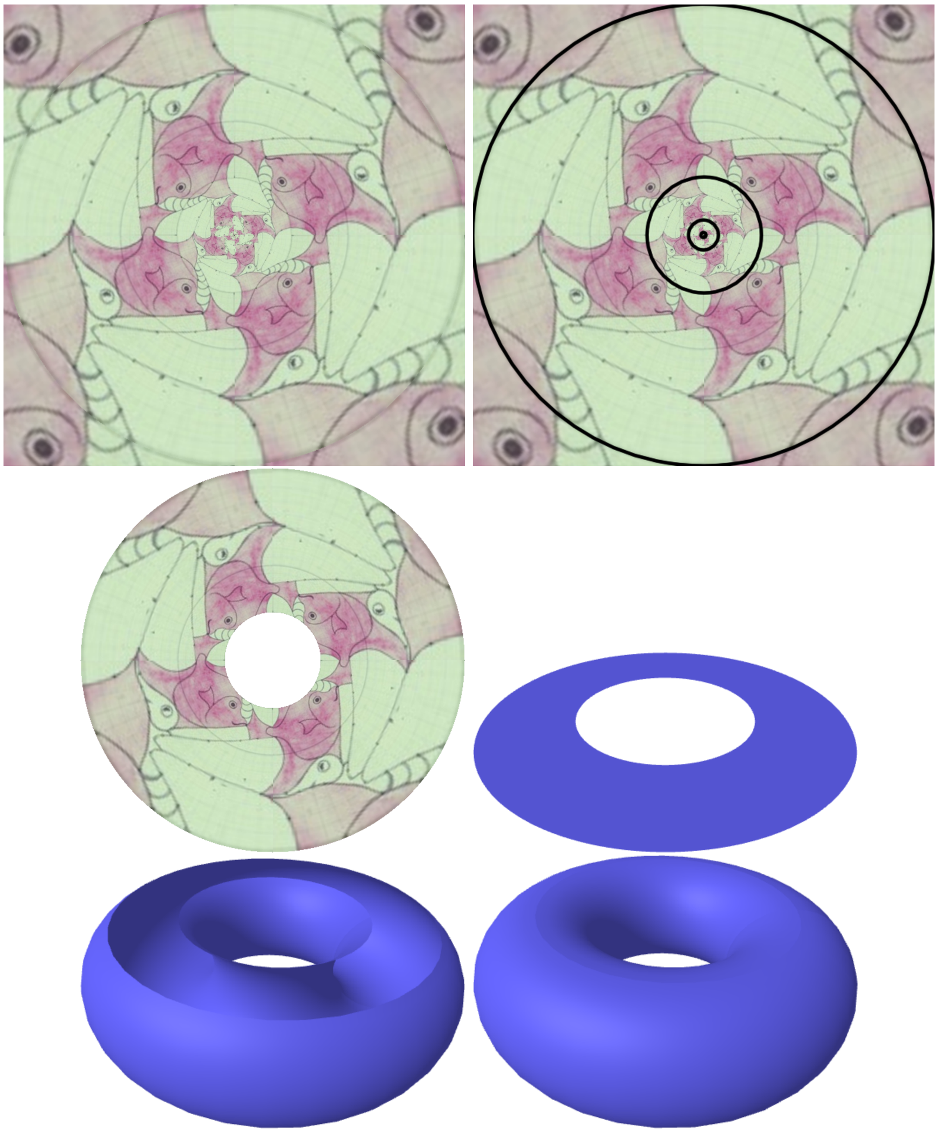

FIGURE 10. Top left: a tiling of the plane minus the origin. Top right: dividing the pattern into annulus shaped tiles. Below: the inner and outer boundaries of the tile can be glued to make a torus.

we glue the boundaries of the annulus together, we get a torus. So

$$(\hat{\mathbb{C}} \setminus \{0, \infty\})/\Gamma = \text{torus}\,.$$

An annulus is $\hat{\mathbb{C}}$ with two discs removed. If we instead remove four discs, we can glue two pairs of boundaries together to get a surface with two handles, as seen in Figure 11. We would therefore like to say that the surface with two handles is of the form U/Γ, where U is a subset of $\hat{\mathbb{C}}$ and Γ is a Kleinian group. We are left with the following questions:

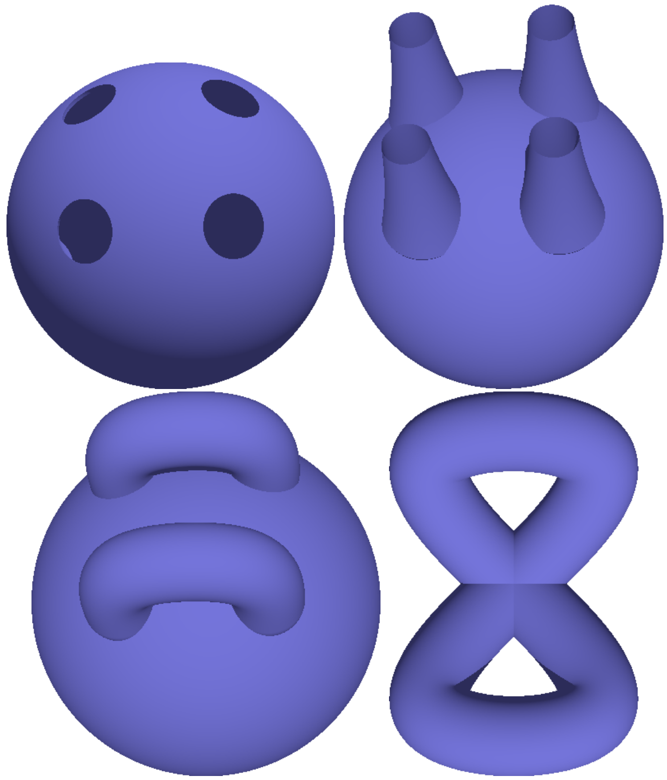

Figure 11. A tile that is shaped like a sphere with four holes can have its edges glued together to form a surface with two handles.

- How do we construct Γ?
- What does Γ look like?
- What does U look like?

To construct Γ, we will use the following ingredients:

- Four disjoint discs, labeled D_a, $D_{a^{-1}}$, D_b, $D_{b^{-1}}$.
- A Möbius transformation a that sends the complement of $D_{a^{-1}}$ to D_a, and a Möbius transformation b that sends the complement of $D_{b^{-1}}$ to D_b.

The group Γ will be the group of Möbius transformations generated by a and b. A Kleinian group constructed in this manner is called a *Schottky group*.

2. Day Two

To better understand the four-disc Schottky tiling, let's first look at the annulus tiling in more detail. Figure 12 gives an abstract description of the tiling shown in Figure 10.

Now let's look at the annulus tiling from a different perspective. The top half of Figure 13 shows the annulus tiling on both the complex plane and the Riemann sphere. In the bottom half of the figure, we have rotated the sphere by 90 degrees, and then projected the tiling back onto the plane. This is equivalent to applying the transformation $z \mapsto \frac{z+1}{z-1}$. In the bottom right image in Figure 13, the region

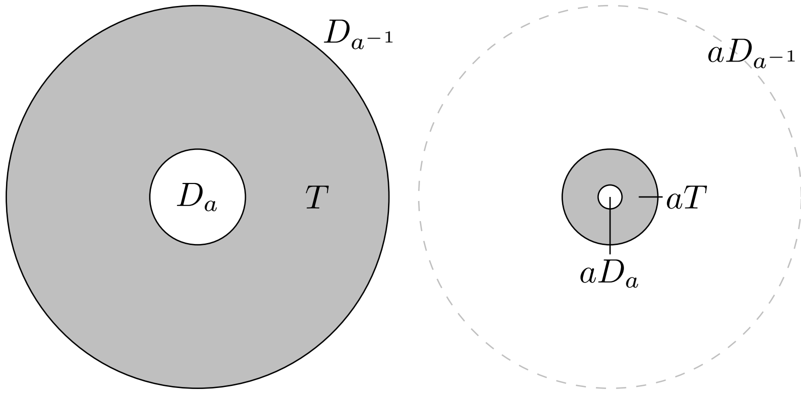

FIGURE 12. Left: An annulus-shaped tile T. Right: Applying a generator to get an adjacent tile.

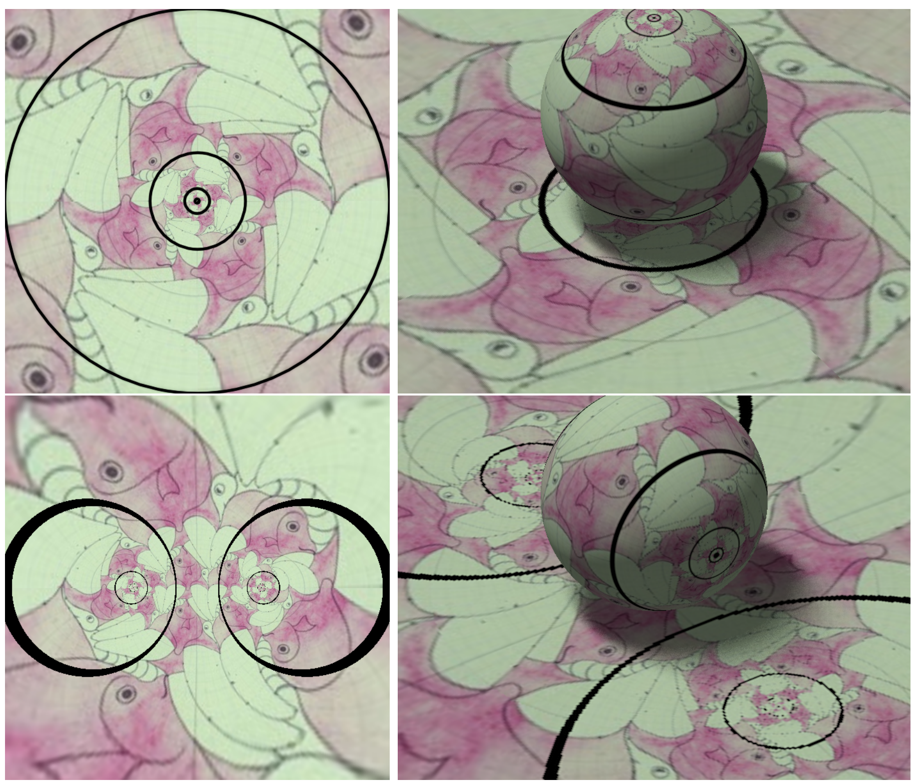

FIGURE 13. Top left: The tiling from Figure 10. Top right: The tiling projected onto the sphere. Bottom right: A 90 degree rotation of the sphere. Bottom left: The rotated tiling projected back onto the complex plane.

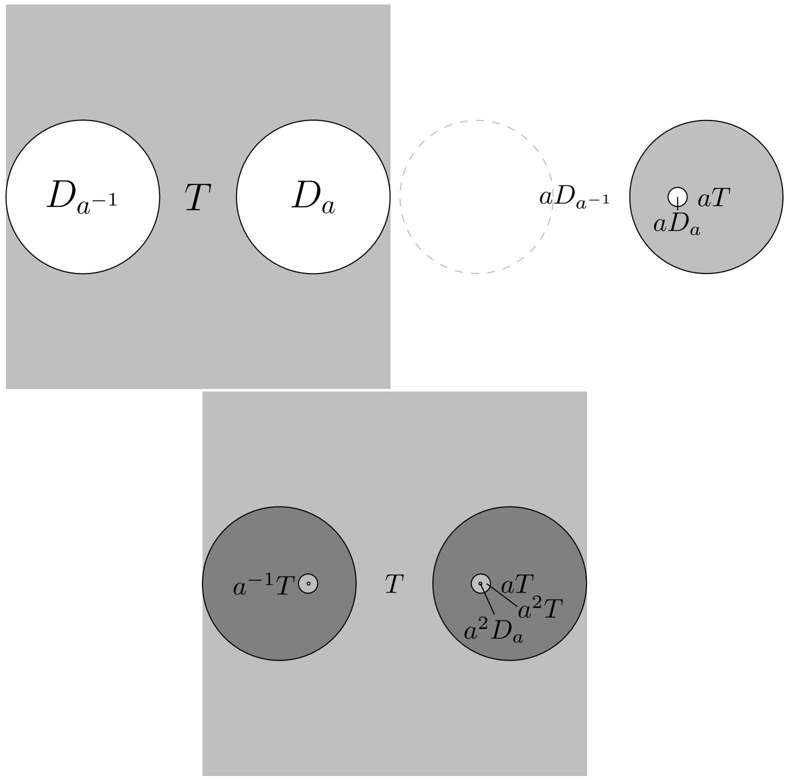

FIGURE 14. Top left: A Schottky tile consisting of $\hat{\mathbb{C}}$ minus two discs. To right: Applying a generator to get an adjacent tile. Bottom: A few more tiles.

outside of the two large circles, extending all the way to infinity, is a tile T. There is a Möbius transformation a that sends T to the annulus inside the right large circle and outside the next largest circle on the right. Similarly, a^{-1} sends T to the annulus inside the left large circle and outside the next largest circle on the left. The tiles $\ldots, a^{-2}T, a^{-1}T, T, aT, a^2T, \ldots$ tile $\hat{\mathbb{C}}$ minus two points. Figure 14 shows a more abstract version of this setup.

Now we are ready to look at the Schottky tiling with four discs. The left side of Figure 15 shows a single tile T, which is $\hat{\mathbb{C}}$ with four discs D_a, D_b, $D_{a^{-1}}$, $D_{b^{-1}}$ removed. The right side of the figure depicts the action of a Möbius transformation a, which sends T to the adjacent tile $aT \subset D_a$. Then a^{-1} automatically sends T to the adjacent tile $a^{-1}T \subset D_{a^{-1}}$. Similarly, we can choose a Möbius transformation b that sends T to an adjacent tile $bT \subset D_b$, and then b^{-1} will send T to an adjacent tile $b^{-1}T \subset D_{b^{-1}}$. Figure 16 shows the effects of the four transformations a, a^{-1}, b, b^{-1}. Some more tiles are shown in Figure 17.

I'd like to point out a key difference between the Schottky group Γ and the symmetry groups of the Escher tilings. The symmetry groups of the Escher tilings shown in Figure 18 are also generated by two elements, but there are some relations between them. For example, in the first drawing, abT and baT represent the same tile. In fact, ab and ba are the same symmetry. Relations between a and b come

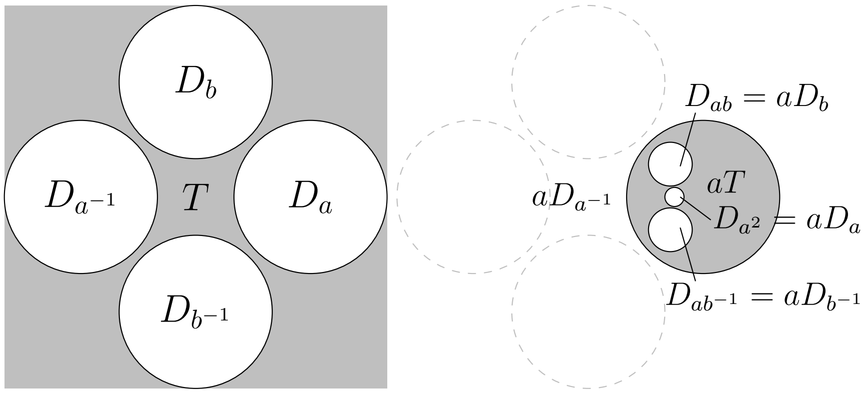

FIGURE 15. Left: A Schottky tile T consisting of $\hat{\mathbb{C}}$ minus four discs. Right: A transformation a, which sends T to an adjacent tile.

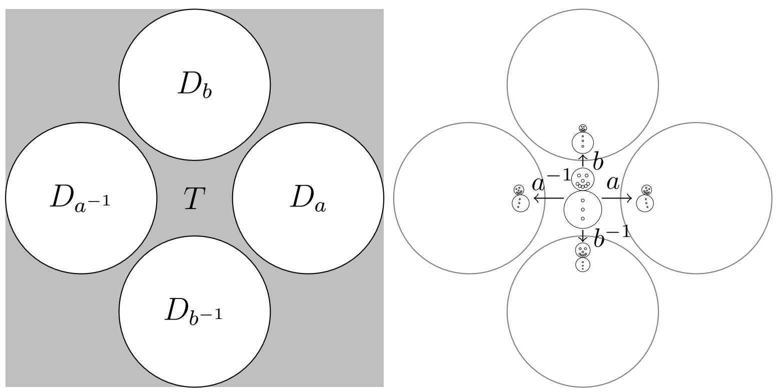

FIGURE 16. Left: A set of four disjoint Schottky discs D_a, $D_{a^{-1}}$, D_b, $D_{b^{-1}}$. The region outside of the discs is the fundamental tile T. Right: Transformations a, a^{-1}, b, b^{-1} such that a sends the complement of $D_{a^{-1}}$ to D_a and b sends the complement $D_{b^{-1}}$ to D_b.

from corners of the tile. But the tile for our Schottky group has no corners; hence the generators of the Schottky group do not satisfy any relations. Therefore, a Schottky group is a *free group*.

A more precise way of saying this is the following. Let Γ is a group generated by two elements a, b. Then any element $\gamma \in \Gamma$ can be written as a product

$$\gamma = g_1 g_2 \dots g_n$$

for some nonnegative integer n and $g_1, g_2, \dots, g_n \in \{a, a^{-1}, b, b^{-1}\}$, Such a product is called a *word*. If $g_i \neq g_{i+1}^{-1}$ for each i, then the word is said to be *reduced*. For the symmetry groups depicted in Figure 18, there is more than one reduced word representing each element of Γ. But for the Schottky group depicted in Figure 17, there is exactly one reduced word representing each element of Γ.

To sum up:

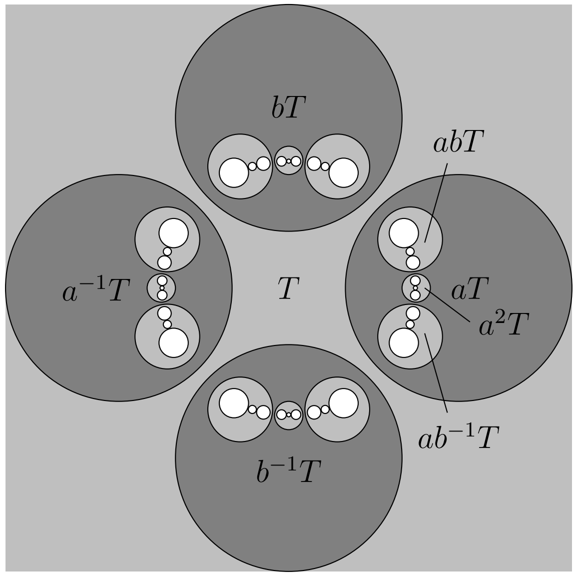

FIGURE 17. A tiling associated with a Schottky group.

	Fuchsian	Schottky
Tiles have. . .	corners but no holes	holes but no corners
Group has. . .	one relation per corner	no relations

The Fuchsian uniformization has simpler (hole-free) tiles, at a cost of having a more difficult to understand group.

We can prove rigorously that Schottky groups are free as follows.

Proposition 6. *Any reduced word starting with a maps T into D_a. Likewise, any reduced word starting with a^{-1}, b, or b^{-1} maps T into $D_{a^{-1}}$, D_b, or $D_{b^{-1}}$, respectively.*

Proof. We use induction on the length of the word. The case of length 1 follows from the definition of a and b. Now consider a reduced word $g_1 g_2 \dots g_n$, $n \geq 2$. By the induction hypothesis, $g_2 \dots g_n$ maps T into D_{g_2}. Since the word is reduced, $g_2 \neq g_1^{-1}$. So g_1 maps D_{g_2} into D_{g_1}. So $g_1 g_2 \dots g_n$ maps T into D_{g_1}. □

Corollary 7. *The group Γ is free.*

Proof. Suppose two different reduced words $g_1 \dots g_n$ and $h_1 \dots h_m$ represent the same element of Γ. Then the word $g_1 \dots g_n h_m^{-1} \dots h_1^{-1}$ represents the identity. This word is not necessarily reduced, but we can cancel aa^{-1}, $a^{-1}a$, bb^{-1}, $b^{-1}b$ pairs to get a nontrivial reduced word. But Proposition 6 implies that any nontrivial reduced word maps T into the complement of T, so it cannot represent the identity. □

Now let's try to understand the set U and its complement. Let $z \in \hat{\mathbb{C}}$ be a point not in U. Then in particular $z \notin T$, so $z \in D_{g_1}$ for some $g_1 \in \{a, a^{-1}, b, b^{-1}\}$. Likewise, $z \notin g_1 T$, so $z \in g_1 D_{g_2}$ for some $g_2 \neq g_1^{-1}$. By similar reasoning, there exists an infinite reduced word $g_1 g_2 g_3 \cdots$ so that z is contained in the infinite decreasing sequence of discs

$$D_{g_1} \supset g_1 D_{g_2} \supset g_1 g_2 D_{g_3} \supset \cdots .$$

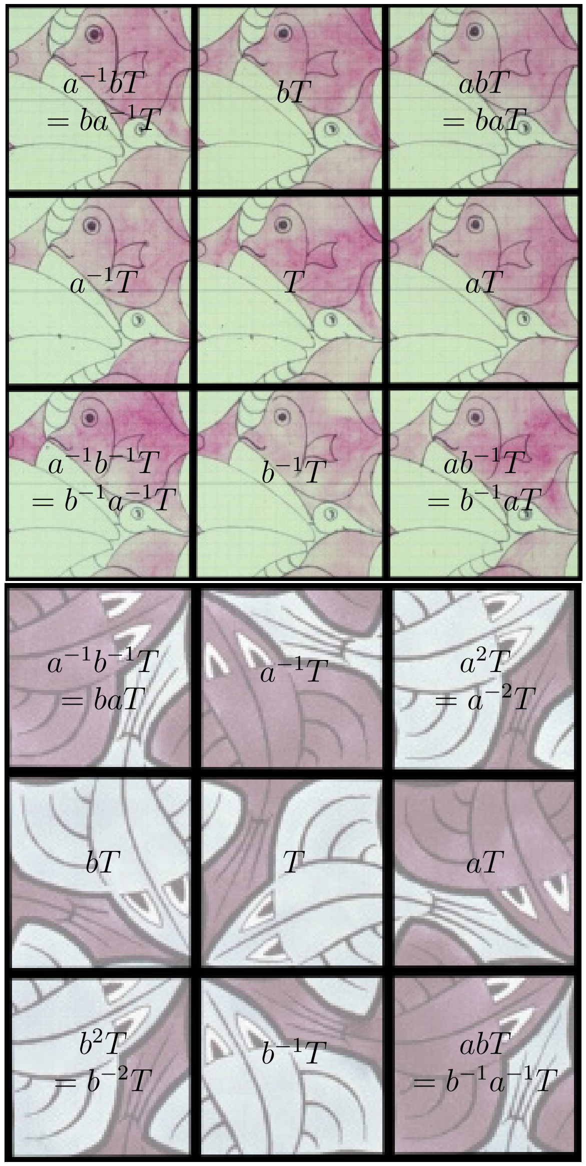

FIGURE 18. The generators of the symmetry groups of these Escher tilings satisfy some relations. Each relation can be attributed to a corner of the initial tile.

For convenience, we will write $D_{g_1 g_2 \ldots g_n}$ for $g_1 g_2 \ldots g_{n-1} D_{g_n}$. So the above sequence would be written as

$$D_{g_1} \supset D_{g_1 g_2} \supset D_{g_1 g_2 g_3} \supset \cdots . \tag{8}$$

Proposition 9. *The radii of the discs in the sequence* (8) *go to zero. In particular, the intersection of all of the discs is a single point.*

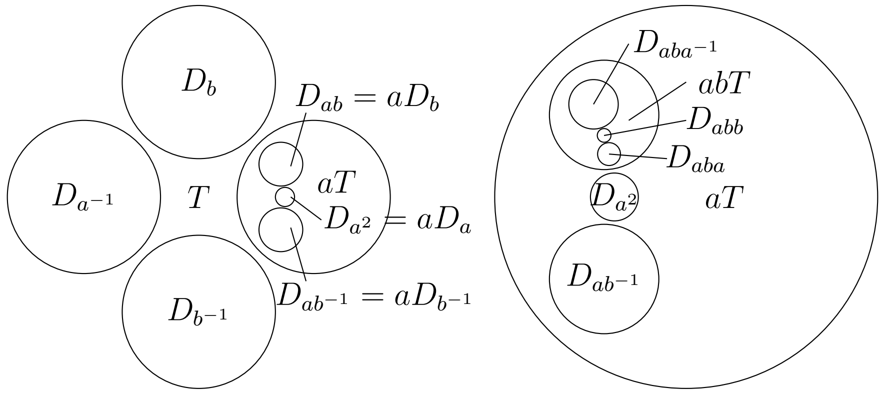

FIGURE 19. As tiles are added to the drawing, each disc is replaced with three smaller ones.

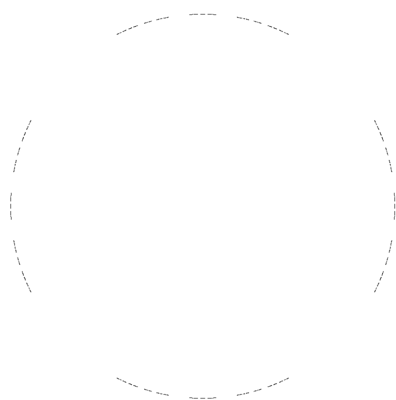

FIGURE 20. The limit set of the Schottky group depicted in Figures 16 and 17.

This will be proved in the homework.

Corollary 10. *The complement of* U *is in bijection with the set of infinite reduced words in* a*,* a^{-1}*,* b*,* b^{-1}*.*

Proposition 11. *Let* $g_1 g_2 \cdots$ *be an infinite reduced word, and let* $z \in U$*. Then the sequence* $z, g_1 z, g_1 g_2 z, \ldots$ *converges to the point corresponding to* $g_1 g_2 \cdots$*. In particular, the bijection of Corollary 10 depends only on the group* Γ *and not on the choice of circles.*

The group Γ acts in the way you would expect on infinite words. For example, $a \cdot (baba^{-1} \cdots) = ababa^{-1} \cdots$ and $b^{-1} \cdot (baba^{-1} \cdots) = aba^{-1} \cdots$.

Definition 12. A point $z \in \hat{\mathbb{C}}$ is a *limit point* of Γ if there exists $z' \in \hat{\mathbb{C}}$ and distinct $\gamma_1, \gamma_2, \ldots \in \Gamma$ such that the sequence $\gamma_1 z', \gamma_2 z', \ldots$ converges to z'.
The *limit set* of Γ is the set of limit points of Γ.

We see that the limit set of Γ is precisely the complement of U. Figure 20 depicts the limit set of a Schottky group. We note that the limit set is totally disconnected, is closed, and has no isolated points. A subset of $\hat{\mathbb{C}}$ having these properties is called a *Cantor dust*.

In conclusion:

- The group Γ is a free group with two generators.
- The limit set of Γ is a Cantor dust.
- Let U be the complement of the limit set of Γ. The quotient U/Γ is a surface with two handles.
- There is a fundamental domain for U that is constructed by removing four discs from $\hat{\mathbb{C}}$.

More generally, we can consider a similar situation with $2n$ pairs of circles rather than 2 pairs of circles. Then:

- The group Γ is a free group with n generators.
- If $n \geq 2$, then the limit set of Γ is a Cantor dust. For $n = 1$, the limit set consists of two points, and for $n = 0$, it is empty.
- Let U be the complement of the limit set of Γ. The quotient U/Γ is a surface with n handles.
- There is a fundamental domain for U that is constructed by removing $2n$ discs from $\hat{\mathbb{C}}$.

3. Day Three

There is a uniformization theorem for Schottky uniformization, as well.

Theorem 13 (Schottky uniformization)**.** *Any orientable closed surface has a Schottky uniformization.*

However, unlike Fuchsian uniformization, Schottky uniformization is not unique. Essentially, this is because Fuchsian uniformization unwinds all of the loops in our surface, while Schottky uniformization unwinds only half of them, and we can choose which half.

Anyway, now we are familiar with Fuchsian and Schottky uniformization. They look quite different—but occasionally they are actually the same! Consider the tiling in Figure 21. It is the limit of a Schottky tiling, where we allow the circles to touch. But it's also a pair of Fuchsian tilings.

As shown in Figure 22, the quotient U/Γ is a pair of punctured tori.

Proposition 14. *Let S be an orientable surface with at least one puncture. Assume S has a Fuchsian uniformization (i.e. S is not a plane or cylinder). Let $\overline{S}$ be the mirror image of S. Then the Fuchsian uniformizations of S and $\overline{S}$ combine to form a Schottky uniformization of $S \cup \overline{S}$.*

Corollary 15. *The fundamental group of a surface with at least one puncture is free.*

Now, what if we take a Schottky uniformization of $S \cup S'$ for some other surface S'? Amazingly, the limit set can look like the fractals in the top row of Figure 1! In the next part of the course, we'll learn how to generate these fractals.

Here are a few observations about Figure 21.

- The set of points not covered by any tile is a circle.
- The sequences of discs (8) still shrink to zero size. In particular, each infinite word determines a point on the circle, and every point on the circle is represented by some infinite word.

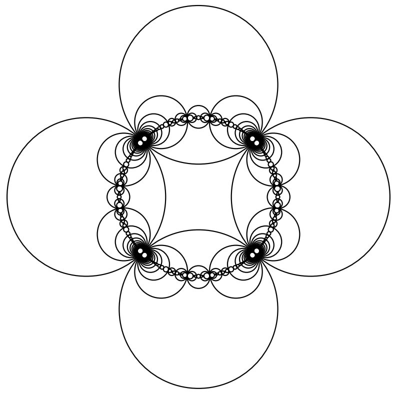

FIGURE 21. By allowing the Schottky circles to touch, we get a tiling that is both Schottky and Fuchsian.

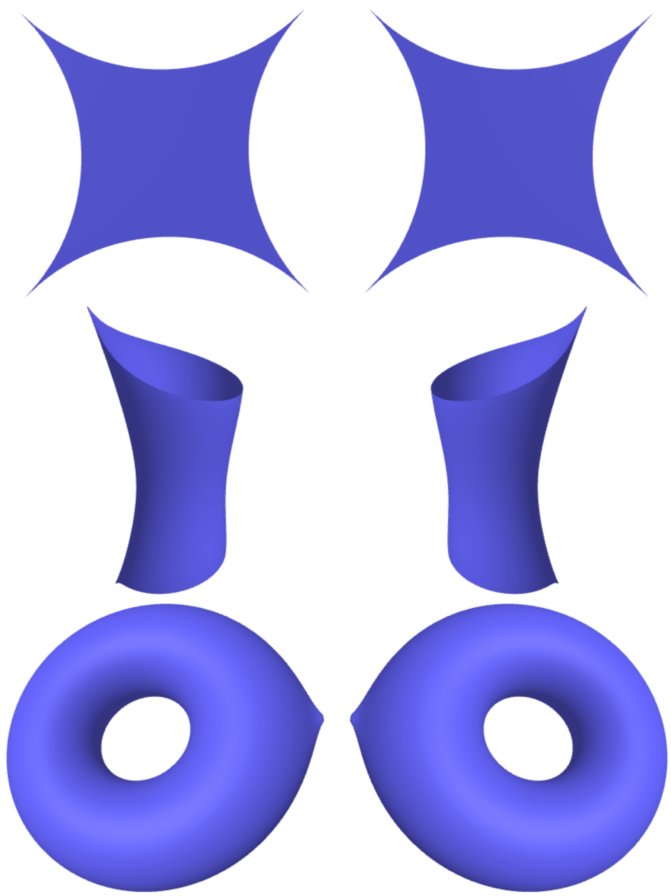

FIGURE 22. When U is disconnected, the quotient U/Γ becomes a pair of punctured tori.

- Points where two Schottky circles are tangent are represented by more than one word. For example, $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$ and $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$ both represent the point where D_a and D_b meet. This is reminiscent of how $0.999\dots$ and $1.000\dots$ represent the same real number.

These observations are related. The image of the set of infinite words can only be connected if some points are represented by more than one word.

Proposition 16. *Let f be a continuous map from the space of infinite words to $\hat{\mathbb{C}}$. If f is injective, then its image is totally disconnected.*

Proof. The proof of this fact involves some topology concepts that are beyond the scope of this course. But it is short, so we include it for the interested reader. Let V be a closed subset of the space of infinite words. The space of infinite words is compact, so V is also compact. Hence $f(V)$ is also compact. Since $\hat{\mathbb{C}}$ is Hausdorff, $f(V)$ must be closed. Therefore, f is a closed map. Since f is closed, continuous, and injective, it induces a homeomorphism from the space of infinite words to the image of f. In particular, the image must be totally disconnected. □

In Proposition 9, we asserted that when D_a and D_b are disjoint, the sequences of discs (8) shrink to zero size. When the circles D_a and D_b touch, the discs can actually fail to shrink to zero size; see the left half of Figure 24 for an example where they don't. However, when this happens, we can always perturb the circles so that they do not touch, as shown on the right side of the figure. So we can nonetheless assign a point to each infinite word.

Proposition 17. *The infinite words*

$$aba^{-1}b^{-1}aba^{-1}b^{-1}\dots, \qquad bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$$

represent fixed points of $aba^{-1}b^{-1}$, and $aba^{-1}b^{-1}$ has no other fixed points. In particular, if the infinite words $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$ and $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$ represent the same point, then $aba^{-1}b^{-1}$ has only one fixed point.

Proof. It is clear that $aba^{-1}b^{-1}$ fixes the images of $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$ and $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$. So we just need to show that $aba^{-1}b^{-1}$ does not fix any other points. Let z be a fixed point of $aba^{-1}b^{-1}$. We claim that $z \in D_a$ or $z \in D_b$. Indeed, if $z \notin D_b$, then $b^{-1}z \in D_{b^{-1}}$, $a^{-1}b^{-1}z \in D_{a^{-1}}$, $ba^{-1}b^{-1}z \in D_b$, and $z = aba^{-1}b^{-1}z \in D_a$. If $z \in D_a$, then we can show by induction that $z \in D_{(aba^{-1}b^{-1})^n a}$ for all z, implying that z is represented by $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$. Likewise, if $z \in D_b$, then z must be represented by $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$. □

Figures 23 and 24 illustrate the importance of $aba^{-1}b^{-1}$ having only one fixed point. In the left half of Figure 23, the Schottky circles are tangent to each other, but the transformations transformations a and b do not map tangency points to tangency points. So the point represented by the word $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$ is not on the boundary of D_b; hence it cannot be the same as the point represented by $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$. It is possible to move the Schottky circles so that they no longer touch, as seen in the right half of the figure. In the right half of Figure 24, the transformations a and b do map tangency points of the circles to tangency points of circles. So the tangency point of D_a and D_b is a fixed point of $aba^{-1}b^{-1}$. This point is represented by $bab^{-1}a^{-1}bab^{-1}a^{-1}\dots$. However, $aba^{-1}b^{-1}aba^{-1}b^{-1}\dots$ represents the other fixed point of $aba^{-1}b^{-1}$, which is in the interior of D_a.

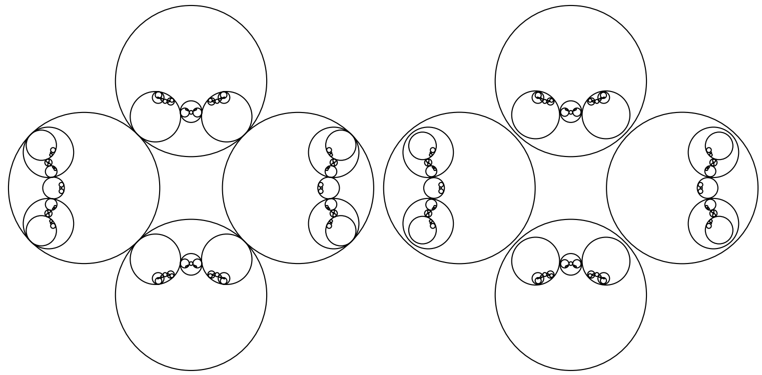

FIGURE 23. Left: Another tiling where the Schottky circles touch. Although the fundamental domain is disconnected into two pieces T_1 and T_2, aT_2 borders both T_1 and T_2. So we still end up getting a connected surface. Right: By moving the circles slightly, we get a connected fundamental domain for the same group.

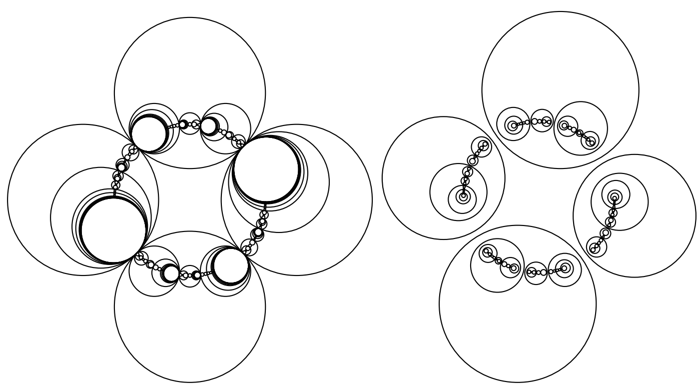

FIGURE 24. Left: a third tiling where the Schottky circles touch. In this case, there are circles that the tiling never fills. The transformation $aba^{-1}b^{-1}$ has two fixed points, one at the point where D_a and D_b meet, and one inside D_a. The discs D_a, D_{ab}, $D_{aba^{-1}}$, $D_{aba^{-1}b^{-1}}$, $D_{aba^{-1}b^{-1}a}$, ... must contain both points, so this sequence of discs cannot shrink to zero size. Right: A different set of Schottky circles for the same group. So this Schottky group corresponds to a closed surface; the original tiles just failed to cover the whole thing.

In fact, this condition about fixed points is the key ingredient behind the fractals in Figure 1: they were generated by choosing Möbius transformations a and b such that $aba^{-1}b^{-1}$ has only one fixed point.

So far, we have almost entirely avoided doing any calculations. But if we actually want to construct fractals, we will need to do some. Let's start by verifying that Möbius transformations form a group.

Proposition 18. *The set of Möbius transpositions forms a group under composition.*

Proof. The Möbius transformation $z \mapsto \frac{1z+0}{0z+1} = z$ is the identity. Next, we check that the composition of two Möbius transformations is again a Möbius transformation.

(19)
$$\frac{a_1 \frac{a_2 z + b_2}{c_2 z + d_2} + b_1}{c_1 \frac{a_2 z + b_2}{c_2 z + d_2} + d_1} = \frac{a_1(a_2 z + b_2) + b_1(c_2 z + d_2)}{c_1(a_2 z + b_2) + d_1(c_2 z + d_2)} = \frac{(a_1 a_2 + b_1 c_2)z + (a_1 b_2 + b_1 d_2)}{(c_1 a_2 + d_1 c_2)z + (c_1 b_2 + d_1 d_2)}$$

Composition of functions is always associative. Finally, we leave it as an exercise to the reader to check that the inverse of $z \mapsto \frac{az+b}{cz+d}$ is $z \mapsto \frac{dz-b}{-cz+a}$. □

You may notice that the formula (19) looks a lot like the formula for multiplication of 2×2 matrices

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{pmatrix}.$$

Definition 20. We will write $\mathrm{GL}_2(\mathbb{C})$ be the group of invertible 2×2 matrices under multiplication.

This is short for "the general linear group of degree 2 over $\mathbb{C}$".

Proposition 21. *There is a surjective homomorphism from* $\mathrm{GL}_2(\mathbb{C})$ *to the group of Möbius transformations that sends the matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *to the transformation* $z \mapsto \frac{az+b}{cz+d}$*. The kernel consists of multiples of the identity. Hence the group of Möbius transformations is isomorphic to the quotient* $\mathrm{GL}_2(\mathbb{C})/\mathbb{C}^\times$*.*

Definition 22. We will refer to the group of Möbius transformations as $\mathrm{PGL}_2(\mathbb{C})$.

(The P stands for "projective").

We would like to classify Möbius transformations. There is a classification of matrices up to conjugacy, called *Jordan normal form.*

Theorem 23 (Jordan normal form)**.** *Every element of* $\mathrm{GL}_2(\mathbb{C})$ *is conjugate to an element of the form*

$$\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \text{ or } \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$$

for some $x, y \in \mathbb{C}^\times$*.*

This means that every Möbius transformation is conjugate to one of the form $z \mapsto xz/y$ or $z \mapsto (xz+1)(x) = z + 1/x$. Note that transformations of the form $z \mapsto z + 1/x$ are all conjugate to $z \mapsto z + 1$ since $x((z/x) + 1/x) = z + 1$.

Corollary 24. *Every element of* $\mathrm{PGL}_2(\mathbb{C})$ *is conjugate to an element of the form* $z \mapsto \lambda z$ *for some* $\lambda \in \mathbb{C}^\times$*, or to* $z \mapsto z + 1$*.*

The transformation $z \mapsto z+1$ is a translation. If we write $\lambda = re^{i\theta}$ with r, θ real, then we see that $z \mapsto \lambda z$ is the composition of the scaling $z \mapsto rz$ and the rotation $z \mapsto e^{i\theta} z$.

To generate fractals, we saw that it was important to identify transformations with only one fixed point. So let's look at the fixed points of the transformations above.

The transformation $z \mapsto re^{i\theta}z$ has two fixed points: 0 and ∞. If $r < 1$, then 0 is an *attracting* fixed point in the sense that the transformation moves points closer to 0, and ∞ is a *repelling* fixed point. If $r > 1$, then ∞ is repelling and 0 is attracting. If $r = 1$, then the fixed points are neither attracting or repelling; the points just move around in a fixed circle.

The transformation $z \mapsto z+1$ has one fixed point: ∞. It isn't exactly attracting or repelling.

4. Day Four

Definition 25. An element of $\mathrm{PGL}_2(\mathbb{C})$ is called:

- Elliptic, if it is conjugate to $z \mapsto e^{i\theta}z$ with $e^{i\theta} \neq 1$;
- Loxodromic, if it is conjugate to $z \mapsto re^{i\theta}z$ with $|r| \neq 1$;
- Parabolic, if it is conjugate to $z \mapsto z + 1$.

So, to draw fractals, we would like to construct elements a and b of $\mathrm{PGL}_2(\mathbb{C})$ such that $aba^{-1}b^{-1}$ is parabolic. But how do we determine which elements of $\mathrm{PGL}_2(\mathbb{C})$ are parabolic? To answer, that question, let's look at Jordan normal form again, and its relation to *eigenvectors* and *eigenvalues*.

Definition 26. Let M be a square matrix, let v be a vector, and let $\lambda \in \mathbb{C}$. If $Mv = \lambda v$, the we say that v is an *eigenvector* of M and λ is the corresponding *eigenvalue*.

There is a close connection between the eigenvalues and eigenvectors of a matrix, and the fixed points of the corresponding Möbius transformation.

Proposition 27. *Let* $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})$*, and let* $w \in \mathbb{C}$*. Then* w *is a fixed point of the Möbius transformation* $z \mapsto \frac{az+b}{cz+d}$ *if and only if* $\begin{pmatrix} w \\ 1 \end{pmatrix}$ *is an eigenvector of* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$*. Likewise,* ∞ *is a fixed point of* $z \mapsto \frac{az+b}{cz+d}$ *if and only if* $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *is an eigenvector of* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$*.*

Anyway, conjugating a matrix preserves eigenvalues. It moves eigenvectors around, but does not change the dimension of the space of eigenvectors. So, for $x \neq y$, a matrix has Jordan normal form $\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$ if and only if its eigenvalues are x and y. A matrix has Jordan normal form $\begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$ if and only if x is its only eigenvalue and it is not equal to $\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$.

The following proposition gives us a way of computing eigenvalues.

Proposition 28. *Let A be a square matrix. The eigenvalues of A are precisely the zeros of the characteristic polynomial* $\det(Ix - A)$*. Here I is the identity matrix.*

If A is an $n \times n$ matrix, then its characteristic polynomial has degree n. So it has n eigenvalues, up to multiplicity.

In the 2×2 case, we have

$$\det\left(\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}x - \begin{pmatrix}a & b\\ c & d\end{pmatrix}\right) = \det\begin{pmatrix}x-a & -b\\ -c & x-d\end{pmatrix} = (x-a)(x-d)-bc = x^2-(a+d)x+(ad-bc)\,.$$

Observe that $ad-bc$ is the determinant of A. The quantity $a+d$ is called the trace of A, denoted $\operatorname{tr} A$. So the characteristic polynomial can be written as

$$x^2 - (\operatorname{tr} A)x + \det A\,.$$

In order for an element to be parabolic, it must have only one eigenvalue, so the discriminant $(\operatorname{tr} A)^2 - 4\det A$ must be zero.

We would like to find matrices A, B such that $ABA^{-1}B^{-1}$ represents a parabolic transformation, i.e. $\operatorname{tr}(ABA^{-1}B^{-1})^2 - 4\det(ABA^{-1}B^{-1}) = 0$. Note that $\det(ABA^{-1}B^{-1}) = \det A \det B(\det A)^{-1}(\det B)^{-1} = 1$, so this is equivalent to $\operatorname{tr}(ABA^{-1}B^{-1})^2 - 4 = 0$, which in turn is equivalent to

$$\operatorname{tr}(ABA^{-1}B^{-1}) = \pm 2\,.$$

5. Day Five

Now we are finally ready to construct some fractals! We want to choose transformations a, b such that $aba^{-1}b^{-1}$ is parabolic. For the moment, let us assume that a is loxodromic. Since conjugating a and b just moves the fractal around $\hat{\mathbb{C}}$, there is no harm in assuming that a fixes 0 and ∞. Therefore, we may assume a is represented by the matrix $\begin{pmatrix}x & 0\\ 0 & 1\end{pmatrix}$. Let $\begin{pmatrix}r & s\\ t & u\end{pmatrix}$ be a matrix representing b. Then we compute

$$\operatorname{tr} aba^{-1}b^{-1} = \frac{2st - ru(x+x^{-1})}{ru - st}\,.$$

(It's feasible to do this by hand but I'm happy to use a computer algebra system.) Then

$$\operatorname{tr} aba^{-1}b^{-1} - 2 = \frac{ru(2-x-x^{-1})}{ru-st}\,.$$

$$\operatorname{tr} aba^{-1}b^{-1} + 2 = \frac{4st - ru(2+x+x^{-1})}{ru-st}\,.$$

So we set

$$ru(2-x-x^{-1}) = 0 \text{ or}$$

$$4st - ru(2+x+x^{-1}) = 0\,.$$

The first equation has a problem: any solution will have the property that a and b have a common fixed point. (They both fix ∞ if $t = 0$, they both fix 0 if $s = 0$, and a is the identity if $2 - x - x^{-1} = 0$.) But the second equation is perfectly fine. Choose a solution, plot some orbits, and you get fancy fractals!

Well, not always. If we chose a and b arbitrarily, there is no guarantee that there is any choice of Schottky discs D_a, $D_{a^{-1}}$, D_b, $D_{b^{-1}}$ compatible with a and b. So there is no guarantee that the group generated by a and b is free or discrete. Sometimes it is, sometimes it isn't. But this procedure works a lot of the time.

The above construction is a bit ad hoc. For example, it's not entirely obvious when two different sets of parameters correspond to genuinely different fractals, and when they are related by conjugation. So how many parameters do we need to specify a fractal?

Four complex numbers are needed to specify a matrix, so $\mathrm{GL}_2(\mathbb{C})$ has (complex) dimension four. Since a Möbius transformation is determined by a matrix up to scaling, $\mathrm{PGL}_2(\mathbb{C})$ is three-dimensional. So the space of all possible pairs $a, b \in \mathrm{PGL}_2(\mathbb{C})$ is six-dimensional. But conjugating a and b by an element of $\mathrm{PGL}_2(\mathbb{C})$ gives an equivalent pair, so up to equivalence, there are only three dimensions' worth of pairs. The equation $\operatorname{tr} aba^{-1}b^{-1} = -2$ imposes one constraint. So, the space of possible fractals up to equivalence has complex dimension 2.

This means, for example, that if we use freedom to scale matrices to fix $\det A = \det B = 1$, then specifying a fractal is (more or less) the same as specifying values of $\operatorname{tr} A$ and $\operatorname{tr} B$. This is what the program `https://guests.mpim-bonn.mpg.de/gulotta/kleinian/` does: given values of $\operatorname{tr} A$ and $\operatorname{tr} B$, it finds matrices with $\det A = \det B = 1$, $\operatorname{tr} ABA^{-1}B^{-1} = -2$, and attempts to plot the limit set of the Kleinian group generated by the transformations A and B.

Below is a formula for finding A and B in terms of $\operatorname{tr} A$ and $\operatorname{tr} B$, taken from the book *Indra's Pearls.*

$$t_a := \operatorname{tr} A, \quad t_b := \operatorname{tr} B$$

$$t_{ab} := \frac{t_a t_b - \sqrt{t_a^2 t_b^2 - 4t_a^2 - 4t_b^2}}{2}$$

$$z_0 := \frac{(t_{ab} - 2)t_b}{t_b t_{ab} - 2t_a + 2it_{ab}}$$

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

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

Since we are free to conjugate A and B, many other formulas are possible. This particular formula has the property that it generates fractals that have 180 degree rotational symmetry about the origin. (This actually means that the symmetry group of the fractals is larger than the Schottky group. This is because we are uniformizing a torus with a puncture (along with its mirror image), and every torus with a puncture has an angle-preserving symmetry that is essentially a 180-degree rotation around the puncture.)

So, given Möbius transformations a and b, how do we plot the limit set? A relatively convenient way is the following. First, we find one point in the limit set. For example, we can choose a fixed point z of a. This amounts to solving an equation of the form $\frac{rz+s}{tz+u} = z$. After clearing denominators, this becomes a quadratic equation. Then we plot $z, az, a^{-1}z, bz, b^{-1}z, a^2z, \ldots$ until we get tired of plotting. Alternatively, we can just repeatedly apply a random element of $\{a, a^{-1}, b, b^{-1}\}$.

The downside of the above method is that it doesn't cover the fractal very evenly, so we will get more detail in some areas than others.

What my program actually does is to create some "fake Schottky discs" D_a, D_b, $D_{a^{-1}}$, $D_{b^{-1}}$. In general, these will intersect each other. It then defines smaller discs using the usual rules $D_{a^2} = aD_a$, $D_{ab} = aD_b$, etc. It repeatedly replaces the largest disc with three smaller ones. After a fixed number of iterations, it plots the centers of the remaining discs. I don't have a proof that the limit set is contained in the "fake Schottky discs" or that the discs shrink to zero size, but this seems to work well in practice.

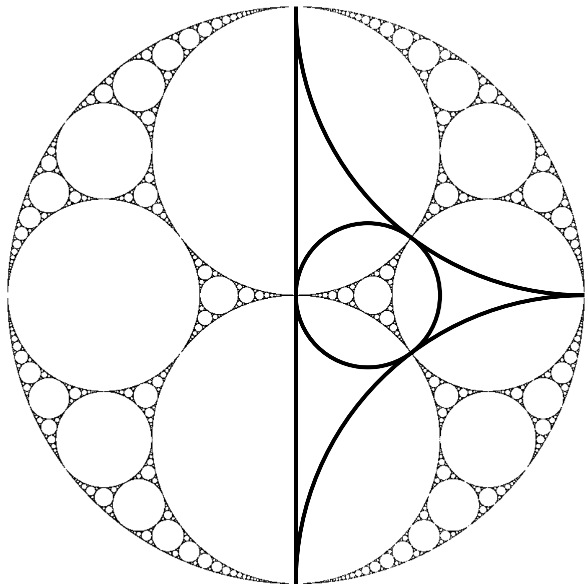

FIGURE 25. A set of Schottky "discs" for the Apollonian gasket. In this case, one of the "discs" is actually a half plane.

So now we have a procedure for producing fractals! Some typical examples are found in the top row of Figure 1. They were generated with the parameters $\operatorname{tr} a = \operatorname{tr} b = 1.91 + 0.05i$ and $\operatorname{tr} a = 1.87 + 0.1i$, $\operatorname{tr} b = 1.87 - 0.1i$, respectively.

The fractal on the bottom left of Figure 1 is a special case known as an *Apollonian gasket*. An Apollonian gasket is a figure constructed by drawing a set of four pairwise tangent circles, and then repeatedly drawing circles that are tangent to three existing circles. All sets of four pairwise tangent circles are related by Möbius transformations. So, up to Möbius transformations, there is a unique Apollonian gasket.

It admits a set of Schottky circles. These are shown in Figure 25. Because there are extra tangency points between the circles, the Apollonian gasket group has more parabolic elements than usual. In fact, it was generated using the parameters $\operatorname{tr} a = \operatorname{tr} b = 2$.

We don't have much time to discuss the Apollonian gasket, but I want to mention one more interesting fact about it. Define the *curvature* of a circle to be the reciprocal of its radius. In the figure that we have drawn, if the curvature of the large circle is taken to be 1, then the curvature of every circle is a integer! Some of these curvatures are shown in Figure 26. In fact, if the four initial circles used to construct the gasket have integer curvature, then all of the circles will have integer curvature.

The last few chapters of *Indra's Pearls* contain some other interesting special cases that we unfortunately don't have time to cover in detail. I'll briefly describe one of these. Figure 27 is another fractal composed of tangent circles. It was generated by choosing a and b so that $aba^{-1}b^{-1}$, b, and $a^3b^{-1}a^2b^{-1}$ are parabolic. This fractal is part of an infinite family. For any rational number $\frac{p}{q}$, it is possible to generate a fractal such that there is a chain of p tangent circles between the

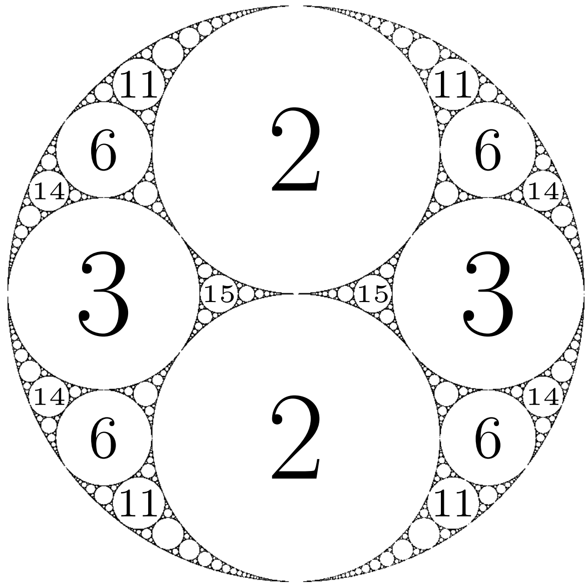

FIGURE 26. The curvatures of circles in an Apollonian gasket. The curvatures are all integers. Moreover, if k_1, k_2, k_3, k_4 are curvatures of a pair of mutually tangent circles, then $(k_1+k_2+k_3+k_4)^2 = 2(k_1^2 + k_2^2 + k_3^2 + k_4^2)$. Given the curvatures of the initial four circles, the remaining curvatures can be computed using a technique known as "Vieta jumping".

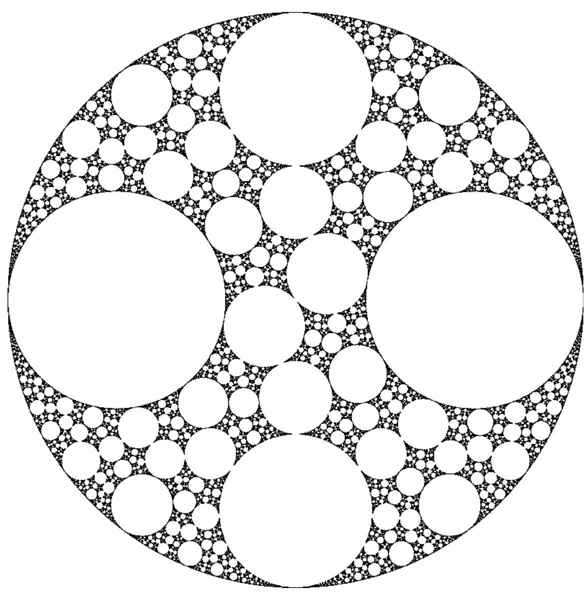

FIGURE 27. Another fractal composed of tangent circles. It was generated with parameters $\operatorname{tr} a = 1.64213876 - 0.76658841i$, $\operatorname{tr} b = 2$.

topmost circle and the rightmost circle, and a chain of $q-1$ tangent circles between the topmost circle and the bottommost circle. (So the fractal shown has $p = 2$, $q = 5$.) Correspondingly, there is a word with q a's and p b^{-1}'s that is parabolic. The fractal on the bottom right of Figure 1 is a limiting case of these fractals, as $\frac{p}{q}$ approaches the golden ratio!

One last thing I'd like to mention is that, although we have only considered the action of $\mathrm{PGL}_2(\mathbb{C})$ on $\hat{\mathbb{C}}$, it is also the symmetry group of three-dimensional hyperbolic space. There is a lot of interesting research on three-manifolds formed by taking a quotient of hyperbolic space by Kleinian groups!