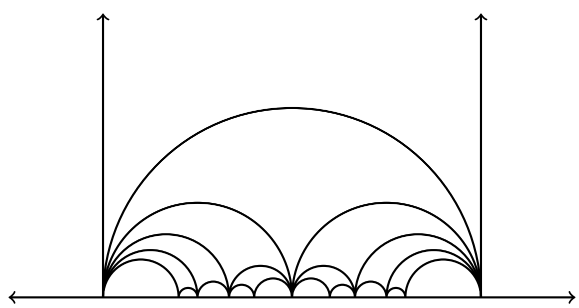
Wesleyan University

Continued Fractions

A Geometric Perspective

by

Max Ryan



Advisor: Felipe A. Ramírez

Assistant Professor of Mathematics

A Thesis in Mathematics
submitted in partial fulfillment of the
requirements for the degree of Master of Arts
at Wesleyan University

Middletown, CT May, 2016

Acknowledgements

First of all I would like to thank my advisor Felipe A. Ramírez. There were many times that I struggled, and he was always patient and willing to provide assistance however he could to get me back on the right track. I could not have done this without Professor Ramírez, and I have nothing but gratitude for his efforts.

I would also like to thank Adam Fieldsteel and David Constantine for being on my committee and helping eliminate any errors from this paper. I also want to thank the Wesleyan Mathematics Department as a whole, faculty and other graduate students, for providing such a helpful environment.

Finally, special thanks to my family for all the support they have given me. Your encouragement was instrumental in getting me to where I am today, and I will not forget it.

Abstract

In this paper we explore the relationship between continued fractions and Diophantine approximation using an alternative geometric view developed by Caroline Series in her 1985 paper *The Modular Surface and Continued Fractions*. Continued fractions provide many useful tools for answering problems in Diophantine approximation, and we begin by giving an overview of this relationship from the classical perspective. We then explore ways to bring dynamical systems into the discussion, specifically the action of $\mathrm{PSL}(2, \mathbb{R})$ on the space $\mathcal{L}$ of unimodular lattices in $\mathbb{R}^2$. This lays a foundation for Series' work which also deals with certain dynamical flows.

We let $\mathbb{H}$ be the hyperbolic plane, and instead of tesselating $\mathbb{H}$ in the usual way by copies of the fundamental region $\{|\Re(z)| \leq 1/2, |z| \geq 1\}$ of $\mathrm{PSL}(2, \mathbb{Z})$, we construct an alternate tesselation using Farey fractions. Then Series showed that the way in which a geodesic γ cuts across the triangles of this tesselation is intimately linked to the continued fraction expansions of the endpoints of γ on the real line. We utilize this connection to provide nice visual observations of known properties of continued fractions and their relations to problems in Diophantine approximation that we saw earlier in the paper.

Contents

Chapter 1

Introduction

This thesis is about interactions between number theory and dynamics. Although on the face of it these may appear to be two disparate fields, there are methods by which one can translate a problem from one field into the language of another, as is often the case in mathematics. In doing so we give ourselves access to new machinery and a new perspective with which to tackle a problem, which can prove invaluable to coming up with a solution.

Number theory is a broad field; in this paper we are specifically interested in Diophantine approximation, which is concerned with the density of rationals in the reals. Its main goal is to classify real numbers based on various approximation properties, as some numbers lend themselves more easily to certain types of approximation, as well as understand what exactly makes these different classes of numbers behave in these different ways. Similarly, dynamics is also a rather broad subject of which we will only be dealing with a small portion. We concern ourselves here with the dynamical behavior of the geodesic flow on the modular surface, a quotient space of the hyperbolic plane.

Over the past 30 or so years, there have been a number of important advances in number theoretic problems using dynamical techniques. Sprindžuk's Extremality Conjecture was posed in 1980 and was solved by Kleinbock–Margulis in 1998. The Littlewood Conjecture was (and still is to some extent) an open problem since the 1930s, until some great progress was made in only the last decade by Einsiedler–Katok–Lindenstrauss through use of dynamical methods.

There are many different interactions between number theory and dynamical systems, of which the above examples are only a taste. In this thesis we focus on a beautiful connection between the continued fraction expansion of a number, and the behavior of a particular flow on the modular surface. Specifically, we will survey a coding by Caroline Series [Ser85] that establishes a dictionary between continued fractions and cutting sequences for the Farey tesselation. We will also interpret some well-known properties of continued fractions using the geometric point-of-view established in [Ser85].

Chapter 2

Diophantine approximation and continued fractions

This chapter deals with fundamental facts about Diophantine approximation and continued fractions that can be found in any standard text on the subject, such as [Cas57] and [RS92].

That the rationals $\mathbb{Q}$ are dense in the reals $\mathbb{R}$ is a well known fact. For any $x \in \mathbb{R}$ and any $\varepsilon > 0$, we can find a rational number $a/n \in \mathbb{Q}$ such that $|x - a/n| < \varepsilon$. Diophantine approximation seeks to further quantify the density of $\mathbb{Q}$ in $\mathbb{R}$. For instance, what if we require a/n to be within a bound that also depends on n?

The first such bound we might consider is simply $1/n$. If we consider multiples of $1/n$, we recognize there is some $a \in \mathbb{Z}$ such that $a/n \leq x < (a+1)/n$, and so $|x - a/n| < 1/n$. Thus we have just shown that for any real number x and natural number n, we can find an approximation a/n such that $|x - a/n| < 1/n$. Clearly, however, arriving at this bound is a trivial exercise, and Diophantine approximation would not be a very interesting field if things stopped here, so we will improve on this bound shortly. In fact by noting which of a/n and $(a+1)/n$ is closer to x, we have already found a rational within a bound (inclusive this time) of $1/2n$.

Another way we can find such a/n that fit the trivial bound is via the decimal expansion of x. For instance, if we take $x = \pi = 3.1415926...$ and $n = 10^6$, then truncating at the sixth decimal place gives $3.141592 = 3141592/1000000$ which is within $1/1000000$ of π. Notice however that since $3141592/1000000$ reduces to $392699/125000$, we have actually found a/n such that $|x - a/n| < 1/8n$ for $n = 125000$. So it appears that we can do better than just $1/n$ in some cases. Our first big result on how much better we can do is provided by Dirichlet.

Theorem 1 (Dirichlet's Approximation Theorem)**.** *For any $x \in \mathbb{R}$, $N \in \mathbb{N}$, there exists $(a, n) \in \mathbb{Z} \times \mathbb{N}$ with $1 \leq n \leq N$ such that* $\left|x - \frac{a}{n}\right| < \frac{1}{nN}$.

Proof. The proof of Dirichlet is rather simple. Consider $\mathbb{R}/\mathbb{Z}$, identified with its fundamental domain $[0, 1)$. Partition the interval $[0, 1)$ into intervals of length $1/N$:

$$[0,1) = \left[0, \frac{1}{N}\right) \cup \left[\frac{1}{N}, \frac{2}{N}\right) \cup \cdots \cup \left[\frac{N-1}{N}, 1\right).$$

Now consider $\{0, \{x\}, \{2x\}, \ldots, \{Nx\}\} \subset [0, 1)$ (here $\{x\}$ denotes the *fractional part* of x, i.e. $x - \lfloor x \rfloor$). Since there are $N+1$ elements in this set and we partitioned $[0, 1)$ into N subintervals, by the pigeonhole principle there exist $0 \leq k < l \leq N$ such that $\{kx\}, \{lx\}$ are in the same subinterval.

In other words, $\{lx\} - \{kx\} = \{(l - k)x\} < 1/N$. Thus there exists some $a \in \mathbb{Z}$ such that $|(l - k)x - a| < 1/N$, or

$$\left|x - \frac{a}{l-k}\right| < \frac{1}{(l-k)N},$$

so $a/(l - k)$ works. □

In particular we can improve our bound of $1/n$ to $1/n^2$ for infinitely many n as we see in the following immediate corollary.

Corollary 1. *For any $x \in \mathbb{R}$, there exist infinitely many (a, n) such that $\left|x - \frac{a}{n}\right| < \frac{1}{n^2}$. Furthermore, for all $x \in \mathbb{R} \setminus \mathbb{Q}$, there exist infinitely many (a, n) such that the above holds and $\gcd(a, n) = 1$.*

Proof. The first statement follows directly from the fact that Dirichlet holds for all $N \in \mathbb{N}$ and $1/nN \leq 1/n^2$ as $1 \leq n \leq N$. To show there are infinitely many that are coprime, suppose for the sake of contradiction that there exists $x \in \mathbb{R} \setminus \mathbb{Q}$ with only finitely many, $a_1/n_1, ..., a_k/n_k$ with $\gcd(a_i, n_i) = 1$. Let N be large enough that $1/N < \min_i |x - a_i/n_i|$. By Dirichlet's Theorem, there must exist some $n \leq N$ such that $|x - a/n| < 1/nN$, and putting a/n into reduced form $\bar{a}/\bar{n}$ obviously preserves this inequality. But then we have $|x - \bar{a}/\bar{n}| < 1/\bar{n}^2$ as noted earlier in the proof, and $\gcd(\bar{a}, \bar{n}) = 1$. However, this $(\bar{a}, \bar{n})$ cannot be a member of our original list, as $|x - \bar{a}/\bar{n}| < 1/\bar{n}N < 1/N$. □

This is already *much* better than the trivial bound because it tells us that we can find very good approximations with relatively simple fractions. For instance, in our π example when we wanted an approximation within 1/1000000, our naïve strategy motivated by the trivial bound required an approximation itself with a denominator of 1000000 (which then reduced to a denominator of 125000, but that is still very large). Now, while Dirichlet doesn't tell us exactly that there will be such an approximation with denominator 1000 since it only claims infinitely many n and $n = 1000$ may not be one of them, we can expect that we will be able to find some approximation to π that does the job while only needing a denominator with 3, maybe 4 digits, which is much simpler. Indeed, the reader may be familiar with the famous approximation 355/113 to π, which we will discuss in more detail further on.

So now that we have one bound, the natural question to ask is whether we can improve it. The answer is yes, but only to a certain degree.

Theorem 2 (Hurwitz's Theorem)**.** *For any $x \in \mathbb{R} \setminus \mathbb{Q}$, there are infinitely many coprime (a, n) such that $\left|x - \frac{a}{n}\right| < \frac{1}{\sqrt{5}n^2}$. Furthermore, $1/\sqrt{5}$ is the best we can do. That is, there exist $x \in \mathbb{R} \setminus \mathbb{Q}$ such that for any $c < 1/\sqrt{5}$, there are only finitely many coprime (a, n) such that $\left|x - \frac{a}{n}\right| < \frac{c}{n^2}$.*

On the other hand, it is also known that there are numbers for which we *can* improve Dirichlet arbitrarily well. This leads us to define new ways we can classify real numbers according to their approximability.

Definition 1. We call $x \in \mathbb{R}$ a *badly approximable number* if there exists $c > 0$ such that $\left|x - \frac{a}{n}\right| < \frac{c}{n^2}$ has only finitely many solutions (and we know such numbers exist thanks to Hurwitz's Theorem). If x is not badly approximable, x is said to be *well approximable.* We denote the set of badly approximable numbers as BAD.

One known fact about BAD is that it is an uncountable set, yet despite this it has Lebesgue measure 0, which already suggests some interesting behavior.

For the next definition we introduce some notation. We define $\|x\| := \min_{a\in\mathbb{Z}} |x - a|$, i.e. $\|x\|$ is the distance between x and the nearest integer. With this notation, in solving an equation like $|x - \frac{a}{n}| < \frac{1}{n^2}$ we can multiply through by n to get $|nx - a| < \frac{1}{n}$, and since $a \in \mathbb{Z}$, minimizing $|nx - a|$ is the same as minimizing $\|nx\|$.

Definition 2. We call $x \in \mathbb{R}$ a *very well approximable number* if there exists $\tau > 1$ such that $\|nx\| < n^{-\tau}$ for infinitely many $n \in \mathbb{N}$.

Note that very well approximability of x implies well approximability of x, and bad approximability of x implies x is neither well nor very well approximable, which is what we would intuitively expect. So now that we have theorems such as Dirichlet and Hurwitz guaranteeing infinitely many "good" approximations to x, how can we find them explicitly? The answer is continued fractions.

Definition 3. A (finite, simple) *continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}},$$

where $a_0 \in \mathbb{Z}$ and $a_n \in \mathbb{N}$ for $n \geq 1$, denoted $[a_0; a_1, a_2, ..., a_n]$. An infinite continued fraction is one which continues forever:

$$[a_0; a_1, a_2, ...] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}$$

which we interpret as the sequence of finite continued fractions $[a_0; a_1, ..., a_n]_{n=0}^{\infty}$. The a_i are called the *partial quotients* of the continued fraction. It is easy to see that since the partial quotients are integers, $[a_0; a_1, ..., a_n] \in \mathbb{Q}$, so we can represent it as a fraction p_n/q_n. These p_n/q_n are called the *convergents*.

As one might guess from the name, the p_n/q_n do in fact converge as $n \to \infty$. This is a nontrivial fact, however, and it is required in order to justify viewing an infinite continued fraction as actually representing a real number. For convenience we will omit the proof for now, as we will see it as a consequence of a later lemma. It is also a fact that every $x \in \mathbb{R}$ has a continued fraction expansion, and in this case the proof is constructive, as we actually have an algorithm to find it which we will see in a moment.

First, however, let's take a look at an example of a continued fraction and some of its convergents, for instance $\pi = [3; 7, 15, 1, 292, 1, 1, 1, 2, \ldots]$. Computing the first few convergents, we get

$$
\begin{aligned}
[3] &= \frac{3}{1} = \frac{p_0}{q_0} \\
[3; 7] &= 3 + \frac{1}{7} = \frac{22}{7} = \frac{p_1}{q_1} \\
[3; 7, 15] &= 3 + \frac{1}{7 + \frac{1}{15}} = \frac{333}{106} = \frac{p_2}{q_2} \\
[3; 7, 15, 1] &= 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1}}} = \frac{355}{113} = \frac{p_3}{q_3} \\
[3; 7, 15, 1, 292] &= 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac{1}{292}}}} = \frac{103993}{33102} = \frac{p_4}{q_4}
\end{aligned}
$$

Already we can make some interesting observations. For one, we can note that all of the convergents we've calculated here satisfy Dirichlet's theorem. For example, $|\pi - 333/106| = .0000832... < 1/106^2 = .0000889....$ This will indeed hold in general, which we will show later. Another thing to notice, however, is some of the convergents go above and beyond just satisfying Dirichlet. Here in just the first four convergents we see two famous approximations of π, 22/7 and the previously mentioned 355/113. The value 22/7 first appeared in the 3rd century BCE when Archimedes proved the inequalities $223/71 < \pi < 22/7$. 22/7 was independently arrived at by the Chinese mathematician Zu Chongzhi in the 5th century AD, who also derived 355/113. The reason these approximations are famous is precisely because of how abnormally close they come to π for how simple of fractions they are. $22/7$ is accurate to two decimal places, well beating $1/7^2$. Furthermore, $355/113$ is accurate to an astonishing *six* decimal places, which is a much smaller margin than $1/113^2 = 1/12769$.

What about these specific convergents makes them such good approximations? We won't answer that right away, but notice that both these convergents appear right before there is a relatively large jump in the size of the partial quotients. For instance $22/7 = [3; 7]$, while the next partial quotient is 15, more than double either of the previous two. Our even better approximation $355/113 = [3; 7, 15, 1]$ appears just before a positively massive jump in the next partial quotient to 292. This corresponds directly to a particular type of behavior we will see when looking from our geometric point of view.

Furthermore, any $x \in \mathbb{R}$ has a continued fraction expansion, and in fact we have an algorithm to produce it.

Definition 4 (The Continued Fraction Algorithm)**.** Step 1: Set $x_0 = x$ and $a_0 = \lfloor x_0 \rfloor$.
Step 2: For $k = 0, 1, 2, ...$, if $x_k \neq a_k$, then set $x_{k+1} = (x_k - a_k)^{-1}$ and $a_{k+1} = \lfloor x_{k+1} \rfloor$. If $x_k = a_k$, stop.

Although initially this alogrithm may look a little odd, note that these are the exact steps we would do if we were simply presented with the equation

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}$$

and began "back solving" for each a_i in turn. We have $a_0 = \lfloor x \rfloor$, and

$$(x - a_0)^{-1} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}$$

so that $a_1 = \lfloor (x - a_0)^{-1} \rfloor$, just as the algorithm says, and we would do the exact same steps again to solve for further a_i.

Example 1 (Continued Fraction Algorithm for $\sqrt{2}$.)*.* According to the algorithm, our first step is to set $x_0 = \sqrt{2}$ and $a_0 = \lfloor \sqrt{2} \rfloor = 1$. Since $a_0 \neq x_0$, we set $x_1 = (x_0 - a_0)^{-1} = (\sqrt{2} - 1)^{-1} = 1 + \sqrt{2}$ and $a_1 = \lfloor 1 + \sqrt{2} \rfloor = 2$. Continuing on, $x_2 = (1 + \sqrt{2} - 2)^{-1} = (\sqrt{2} - 1)^{-1} = 1 + \sqrt{2}$. But notice now that $x_2 = x_1$, so $a_2 = a_1$ and so our calculation of x_3 will be the same, and x_4, x_5, and so on. Thus $\sqrt{2} = [1; 2, 2, 2, ...] := [1; \overline{2}]$. This is a simple example of a *periodic continued fraction.*

Definition 5. A *periodic continued fraction* is a continued fraction with a repeating tail segment, written $[a_0; a_1, ..., \overline{a_k, ..., a_{k+t}}]$ for some $t \geq 0$.

This behavior is not a coincidence, either. It is a combined result of work by Euler and Lagrange that the numbers with periodic continued fraction expansions are precisely the quadratic irrationals. Euler showed how to take a periodic continued fraction and construct a corresponding quadratic equation that it must satisfy, while later Lagrange was able to prove the converse that any quadratic irrational must have a periodic continued fraction expansion.

Theorem 3 (Euler–Lagrange)**.** *Let $x \in \mathbb{R}$. Then x has a periodic continued fraction expansion if and only if x is a quadratic irrational.*

Something else to notice about $\sqrt{2}$ is that there is never a big jump in the size of the partial quotients like we saw in the continued fraction expansion of π. As these jumps seemingly corresponded to particularly good approximations of π, from this fact one may already suspect that something of the opposite behavior is occuring for $\sqrt{2}$. Indeed, this is the case:

Theorem 4. *The badly approximable numbers are exactly those whose partial quotients are bounded.*

So $\sqrt{2}$ (and, indeed, *any* periodic continued fraction) is a badly approximable number, and this fact is directly tied to the lack of larger and larger jumps in its partial quotients.

Before we prove this, we will need a few more tools for dealing with convergents. We know one way to calculate convergents simply by their definition (truncate the continued fraction of our number and then calculate the resulting finite continued fraction). However this task quickly becomes onerous after just the first few convergents; we don't want to have to calculate a continued fraction with 100 terms just to get p_{100}/q_{100}, after all. Is there a quicker way?

Lemma 5 (Recursion formula for convergents)**.** *Suppose* $\frac{p_n}{q_n} = [a_0; a_1, ..., a_n]$ *are the convergents for a continued fraction. Then for any* $n \geq 0$*,*

$$p_n = a_n p_{n-1} + p_{n-2}, \text{ and}$$

$$q_n = a_n q_{n-1} + q_{n-2}$$

where we define $p_{-1} := 1$*,* $p_{-2} := 0$*,* $q_{-1} := 0$*, and* $q_{-2} := 1$*.*

Let's consider the golden ratio φ as an example. The continued fraction expansion of φ is easily determined: from the relation $\varphi = 1 + 1/\varphi$, we see that if

$\varphi = [a_0; a_1, a_2, ...]$, then

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}} = 1 + \cfrac{1}{a_0 + \cfrac{1}{a_1 + \ddots}}.$$

Thus we see that $a_i = a_{i+1}$ for all i, and $a_0 = 1$, so that $\varphi = [1; 1, 1, 1, ...]$. Notice then that since all the a_i are equal to 1, the recursion formulae become simply $p_n = p_{n-1} + p_{n-2}$ and $q_n = q_{n-1} + q_{n-2}$, the same recurrence relation as for the Fibonacci numbers. So we have $p_0 = 1 + 0 = 1$, $p_1 = 1 + 1 = 2$, $p_2 = 2 + 1 = 3, ...$ and $q_0 = 0 + 1 = 1$, $q_1 = 1 + 0 = 1$, $q_2 = 1 + 1, ...$ so that $p_n = F_{n+1}$ and $q_n = F_n$, where F_n is the nth Fibonacci number. This is unsurprising, as it is a well-known fact that $\lim_{n\to\infty} F_{n+1}/F_n = \varphi$.

Back to our characterization of BAD, there was seemingly a correlation between large increases in partial quotients of x and good approximations to x, and the lack of arbitrarily large such jumps is what forces a number with bounded partial quotients to be badly approximable. In this sense, φ is somehow the "most" badly approximable number, as it exhibits no increase in partial quotients at all. This is reflected in the fact that φ is one of the numbers mentioned in Hurwitz's Theorem that prevent us from improving on the bound of $1/\sqrt{5}n^2$.

This draws some attention to the question of these problem numbers from Hurwitz's Theorem for which $c = 1/\sqrt{5}$ cannot be improved. First of all, what are they? That $\varphi = [1, 1, 1, \dots]$ is one gives us a hint, and indeed it is the case that the numbers for which $1/\sqrt{5}$ is the best we can do are exactly those whose continued fraction expansions are periodic with a tail of all 1's, i.e. whose tails match that of the golden ratio after some point. What is even more interesting is that if x is not one of these numbers, then for *any* c in between $1/\sqrt{5}$ and

$1/\sqrt{8}$ we have infinitely many solutions for $|x - a/n| < c/n^2$, but when we get to $1/\sqrt{8}$ we run into some more problem points. The points for which $1/\sqrt{8}$ cannot be improved are exactly those whose tails match that of one of the roots of the polynomial $x^2 - 2x - 1 = 0$ (which must also have periodic continued fraction expansions, being quadratic irrationals). Then when we throw these points out, we can again improve our c from $1/\sqrt{8}$ to $5/\sqrt{221}$. This behavior continues, and the resulting sequence $\{1/\sqrt{5}, 1/\sqrt{8}, 5/\sqrt{221}, \ldots\}$ has an accumulation point at 1/3 [Cas57]. It is contained in the *Lagrange spectrum*, which is the set of all possible values of $\liminf_{n\to\infty} n\|nx\|$, $x \in \mathbb{R}$ [RS92]. Since we know that BAD is uncountable, and the matching tail conditions only ever produce countably many badly approximable numbers, we may deduce that the Lagrange spectrum must contain uncountably many points between 0 and 1/3. The Lagrange spectrum is closely related to the Markov spectrum, which has to do with the theory of quadratic forms. This is explored further in Cassels' book.

The following lemma describes two nice properties of the convergents. Notice that with this lemma in hand we finally have proof that the convergents do indeed converge to x, as $\left|x - \frac{p_n}{q_n}\right| \leq 1/q_n q_{n+1}$ which goes to 0 as $n \to \infty$.

Lemma 6 (Properties)**.**

- *The convergents of x alternately over/under estimate x. More specifically, we have that $\dfrac{p_n}{q_n} < x$ if n is even, and $\dfrac{p_n}{q_n} > x$ if n is odd.*
- $\dfrac{1}{q_n(q_n + q_{n+1})} < \left|x - \dfrac{p_n}{q_n}\right| \leq \dfrac{1}{q_n q_{n+1}}$.

This also shows that the convergents satisfy Dirichlet's theorem, since we always have $1/q_n q_{n+1} < 1/q_n^2$. But more than that, this is at the heart of why a jump in the partial quotients corresponds to a more accurate approximation: if there is a large jump from a_n to a_{n+1}, then by the recursion formula in Lemma 5 this creates a large jump from q_n to q_{n+1} so that $1/q_nq_{n+1}$ is much less than $1/q_n^2$. Armed with this vital tool, we can now go ahead with our proof of the theorem.

Proof of Theorem 4. First we show that any continued fraction with unbounded partial quotients cannot be badly approximable. Suppose $x = [a_0; a_1, a_2, ...] \in \mathbb{R} \setminus \mathbb{Q}$ where the a_i are unbounded. Then there exists a subsequence $\{a_{n_j}\}_{j=1}^{\infty}$ that strictly increases to infinity. By the second property in the lemma,

$$\begin{aligned} \left| x - \frac{p_{n_j}}{q_{n_j}} \right| &< \frac{1}{q_{n_j} q_{n_{j+1}}} \\ &= \frac{1}{q_{n_j}(a_{n_{j+1}} q_{n_j} + q_{n_{j-1}})} \text{ by Lemma 5,} \\ &< \frac{1}{q_{n_j}^2 a_{n_{j+1}}}. \end{aligned}$$

Since $a_{n_{j+1}} \to \infty$, $\frac{1}{a_{n_{j+1}}} \to 0$, so there cannot exist $c > 0$ such that $\left|x - \frac{p}{q}\right| \geq \frac{c}{q^2}$ for all (p, q). Thus x is not badly approximable.

To show containment in the other direction, suppose $x \in \mathbb{R} \setminus \mathbb{Q}$ has bounded partial quotients, i.e. there exists $c > 0$ such that $a_n \leq c$ for all $n \in \mathbb{N}$. Let $q \in \mathbb{N}$, $p \in \mathbb{Z}$, and let n be such that $q_{n-1} \leq q < q_n$. Then

$$\left| x - \frac{p}{q} \right| \geq \left| \frac{p_n}{q_n} - \frac{p}{q} \right| - \left| x - \frac{p_n}{q_n} \right|$$

by the triangle inequality, which we further bound

$$\geq \frac{|p_n q - p q_n|}{q_n q} - \frac{1}{q_n q_{n+1}}.$$

We will see later (Lemma 14) that convergents are always reduced fractions. Therefore, since $q < q_n$, the expression $|p_n q - p q_n|$ is a nonzero integer. Hence, we may continue to bound

$$\begin{aligned}
&\geq \frac{1}{q_n q} - \frac{1}{q_n q_{n+1}} = \frac{q_{n+1} - q}{q q_n q_{n+1}} \\
&\geq \frac{q_{n+1} - q_n}{q q_n q_{n+1}}
\end{aligned}$$

And by the recursion formula $q_{n+1} = a_{n+1}q_n + q_{n-1} \geq q_n + q_{n-1}$,

$$\begin{aligned}
&\geq \frac{q_{n-1}}{q q_n q_{n+1}} \\
&\geq \frac{q_{n-1}}{q q_n^2 (a_{n+1} + 1)} \\
&\geq \frac{q_{n-1}}{q q_{n-1}^2 (a_{n+1} + 1)(a_n + 1)^2} \\
&\geq \frac{1}{q^2 (a_{n+1} + 1)(a_n + 1)^2} \\
&\geq \frac{1}{q^2 (c+1)^3},
\end{aligned}$$

which finishes the proof. □

We mentioned above the result of Lagrange that all quadratic irrationals have periodic continued fraction expansions and are therefore badly approximable. Surprisingly, however, when we move to higher degrees we know almost nothing. It is an open problem whether there even exist any algebraic irrationals of degree 3 or greater that are badly approximable, let alone being able to classify all of them nicely as in the degree 2 case. There is clearly some deep theory involved with the badly approximable numbers behind the scenes. Another seemingly odd fact about the set BAD is that one can show it has Lebesgue measure 0, while simultaneously maintaining full Hausdorff dimension.

There are also various other approximability conditions we can classify real numbers by and study the properties of these sets; we mentioned the notion of very well approximability above as one such classification. Our goal in this paper is to interpret these ideas in a geometric way using dynamical systems. As it turns out, these approximation properties correspond to the behavior of certain dynamical trajectories with respect to the cusp in the space $\mathcal{L}$ of unimodular lattices and $M = \mathbb{H}/\mathrm{SL}(2, \mathbb{Z})$.

Chapter 3

Hyperbolic geometry and geodesic flows

In this chapter we refer to standard information about the hyperbolic plane and its geometric properties that can be found in [EW11] and [Dal11].

So far we have been discussing purely abstract number theory, but like many areas of math it is often helpful to translate the problem to another perspective. For instance, we mentioned above that one common treatment is to take $|x - a/n|$ and multiply through by n, so that instead we are now studying how small $|nx - a|$ can get. This means for a good approximation we are looking for some multiple nx of x that gets very close to some integer a. That is, we are studying the collection $\{\|nx\|\}$ where $\|x\| := \min_{a\in\mathbb{Z}}|x - a|$, the distance from x to the nearest integer.

Notice however that since all we care about is distance to the integers and not which integer in particular we get close to, we can instead view x as an element of $\mathbb{R}/\mathbb{Z} \cong S^1$. Then the set $\{nx : n \in \mathbb{N}\}$ becomes the set of orbit points of 0 under rotation of $2\pi x$ radians (see Figure 3.1). In this way we have translated our problem into one involving aspects of geometry and dynamics, allowing us to use tools from those fields. This chapter will discuss a different geometric approach to Diophantine approximation, this time with regards to the hyperbolic plane.

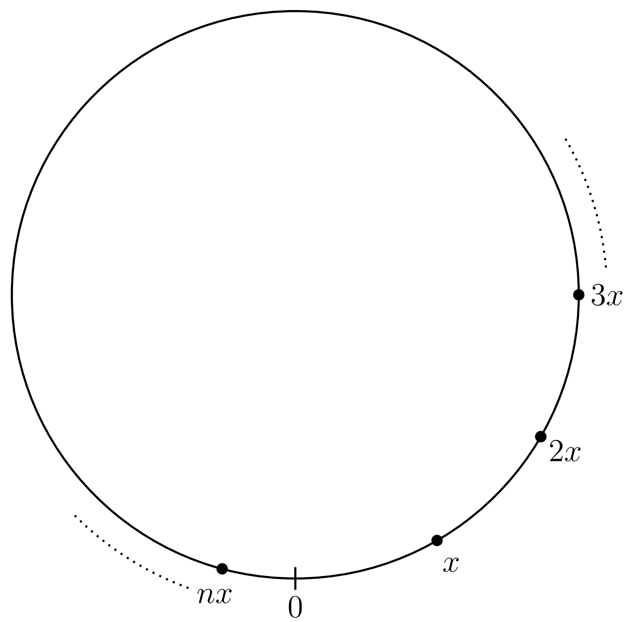


Figure 3.1: Orbits of a circle rotation.

3.1 Hyperbolic plane and its isometries

The hyperbolic plane is

$$\mathbb{H} := \{x + iy \in \mathbb{C} : y > 0\},$$

i.e. the upper half of the complex plane, but endowed with a certain metric d defined as follows. For $z, w \in \mathbb{H}$, let $p(t) = (x(t), y(t))$ be a smooth path in $\mathbb{H}$ with $p(0) = z$ and $p(1) = w$. Let

$$L(p) = \int_0^1 \frac{|p'(t)|}{y(t)} dt.$$

Then $d(z, w) = \inf L(p)$, with the infimum taken over all paths in $\mathbb{H}$ from z to w.

It is a standard fact that the group of isometries of $\mathbb{H}$ with respect to the hyperbolic metric d is isomorphic to $\mathrm{PSL}(2, \mathbb{R})$. An element $g \in \mathrm{SL}(2, \mathbb{R})$ acts on $\mathbb{H}$ by the action

$$g \cdot z = \frac{az + b}{cz + d},$$

where $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Notice however that the action of $-I_2$ is trivial:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} z = \frac{-z + 0}{0 - 1} = z.$$

Thus we can mod out by $\{\pm I_2\}$, leaving us with $\mathrm{PSL}(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})/\{\pm I_2\}$. The following lemma is standard and can be found, for example, in [EW11, Page 280].

Lemma 7 (Properties of the $\mathrm{PSL}(2, \mathbb{R})$ action)**.**

1. *For any $z, w \in \mathbb{H}$ and $g \in \mathrm{PSL}(2, \mathbb{R})$,*

$$d(gz, gw) = d(z, w).$$

2. *The action is transitive. That is, for any $z, w \in \mathbb{H}$, there is some $g \in \mathrm{PSL}(2, \mathbb{R})$ such that $gz = w$.*

Now that we have our metric, it is natural to ask what distance-minimizing paths in $\mathbb{H}$ look like, or even if they exist at all. Our method will be to identify a particular geodesic directly, and then use the transitivity of the $\mathrm{PSL}(2, \mathbb{R})$ action to map it to the other curves, which will also have to be geodesics because $\mathrm{PSL}(2, \mathbb{R})$

acts by isometries.

Lemma 8. *[EW11, Page 282] Let* $z = iy_1$ *and* $w = iy_2$*,* WLOG $0 < y_1 < y_2$*. Then* $p : [0,1] \to \mathbb{H}$ *defined by*

$$p(t) = iy_1 \left(\frac{y_2}{y_1}\right)^t$$

is a distance-minimizing path from z *to* w*.*

Proof. We first calculate the speed of our path $p(t) = (x(t), y(t))$ at time t with respect to our hyperbolic metric, given by $\dfrac{|p'(t)|}{y(t)}$:

$$p'(t) = iy_1 \left(\frac{y_2}{y_1}\right)^t \log\frac{y_2}{y_1} \text{ so that } |p'(t)| = y_1 \left(\frac{y_2}{y_1}\right)^t \log\frac{y_2}{y_1}, \text{ and}$$
$$y(t) = y_1 \left(\frac{y_2}{y_1}\right)^t \text{as } p \text{ is simply travelling along the imaginary axis.}$$

Thus we see that p has constant speed

$$\frac{|p'(t)|}{y(t)} = \frac{y_1 \left(\frac{y_2}{y_1}\right)^t \log\frac{y_2}{y_1}}{y_1 \left(\frac{y_2}{y_1}\right)^t} = \log\frac{y_2}{y_1} = \log y_2 - \log y_1$$

so that $L(p) = \log y_2 - \log y_1$ as well. Therefore $d(z, w) \leq \log y_2 - \log y_1$ since distance is the infemum over the lengths of all paths from z to w. All that remains to show is that any other path has greater length. To that end, let $q(t) = (q_x(t), q_y(t)) : [0,1] \to \mathbb{H}$ be some other path from z to w. Then

$$\begin{aligned} L(q) &= \int_0^1 \frac{|q'(t)|}{q_y(t)} dt \geq \int_0^1 \frac{\left|q_y'(t)\right|}{q_y(t)} dt \\ &\geq \int_0^1 \frac{q_y'(t)}{q_y(t)} dt = \log q_y(1) - \log q_y(0) = \log y_2 - \log y_1. \end{aligned}$$

Note that our first inequality where we dropped the real component of q becomes equality if and only if the real component is identically zero, i.e. if q lies entirely on the imaginary axis. $\square$

Now, the action of $\mathrm{PSL}(2,\mathbb{R})$ is not only transitive in that any one point can be sent to any other point; it is actually transitive on equidistant pairs, so that we can send any two points z_1, z_2 to any other pair of points w_1, w_2 such that $d(z_1, z_2) = d(w_1, w_2)$ (remember, $\mathrm{PSL}(2,\mathbb{R})$ acts by isometries so the distance must be preserved). Thus if we want to find a geodesic arc between z and w (i.e. a path of length $d(z, w)$ between them), we can return to $i\mathbb{R}$ where we understand the geodesic arcs and map the points i and $i\exp(d(z, w))$ to z and w via some g. Then the image under g of the geodesic arc between i and $i\exp(d(z, w))$ will be the geodesic arc between z and w. We restate this as a proposition.

Proposition 1. *For any $z, w \in \mathbb{H}$, there is a unique path $p(t) : [0, d(z, w)] \to \mathbb{H}$ of unit speed with $p(0) = z$ and $p(d(z, w)) = w$. Furthermore, there is a unique isometry $g \in \mathrm{PSL}(2,\mathbb{R})$ such that $p(t) = g(ie^t)$.*

Since any geodesic arc is the image of some arc of $i\mathbb{R}$ under some $g \in \mathrm{PSL}(2,\mathbb{R})$, it follows that the full set of geodesics is $\{g \cdot i\mathbb{R} : g \in \mathrm{PSL}(2,\mathbb{R})\}$. As g acts by fractional linear transformations, we know that it takes lines and circles to lines and circles so that $g \cdot i\mathbb{R}$ will be a line or circle in $\mathbb{C}$ for any g. Furthermore, fractional linear transformations are conformal, meaning they preserve angles; in particular, since $g \cdot \mathbb{R} = \mathbb{R}$ for any $g \in \mathrm{PSL}(2,\mathbb{R})$, the 90 degree angle between $\mathbb{R}$ and $i\mathbb{R}$ is preserved. Thus the geodesics on $\mathbb{H}$ are the vertical lines along with circles that intersect the real axis at 90 degrees, i.e. are centered on the real axis.

Now it becomes important for us to turn our focus to a special subgroup of $\mathrm{PSL}(2,\mathbb{R})$, the modular group $\mathrm{PSL}(2,\mathbb{Z})$. This group has several applications to

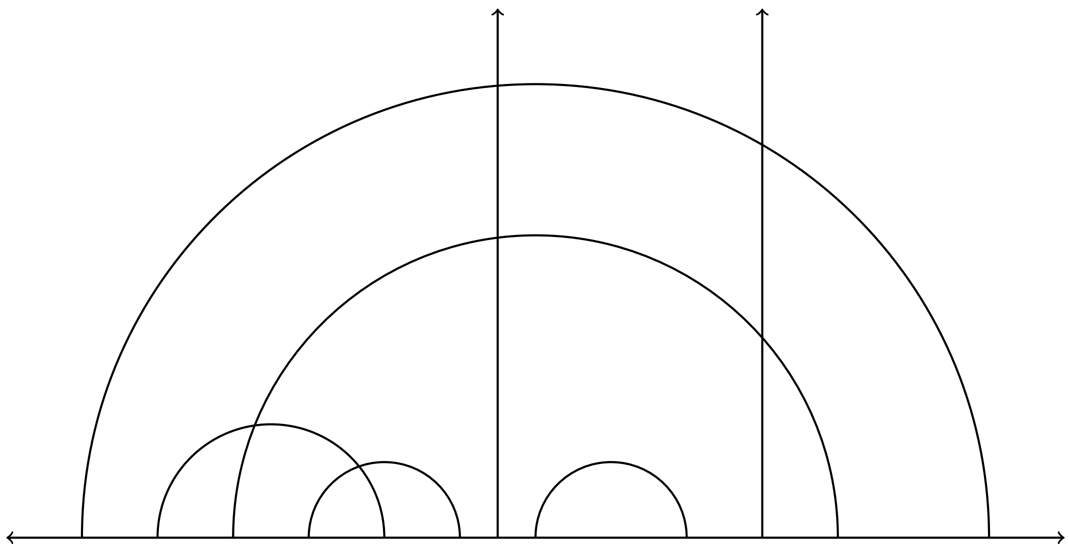


Figure 3.2: Some geodesics on $\mathbb{H}$.

our discussion. For one, as a discrete subgroup of the group of isometries of $\mathbb{H}$, we can find a fundamental domain for $\mathrm{PSL}(2, \mathbb{Z})$ in $\mathbb{H}$. The fundamental domain most commonly used is the region $\{z \in \mathbb{H} : |z| \geq 1, |\Re(z)| \leq 1/2\}$ (see Figure 3.3). We can then consider the modular surface $M := \mathbb{H}/\mathrm{PSL}(2, \mathbb{Z})$, which we will return to in a later section.

Next we will show the connection between Diophantine approximation and certain flows on the quotient space $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$. Note that $\mathrm{SL}(2, \mathbb{Z})$ is a lattice in $\mathrm{SL}(2, \mathbb{Z})$, which implies that $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$ has finite volume. However, this quotient space is noncompact, meaning in the geometric picture of this space we see a cusp going off to infinity. We will examine the crucial relationship of this cusp to Diophantine approximation further in the next section.

3.2 Space of lattices

In this section we delve more deeply into the ties between Diophantine approximation and dynamics on $\mathrm{PSL}(2, \mathbb{R})/\mathrm{PSL}(2, \mathbb{Z})$.

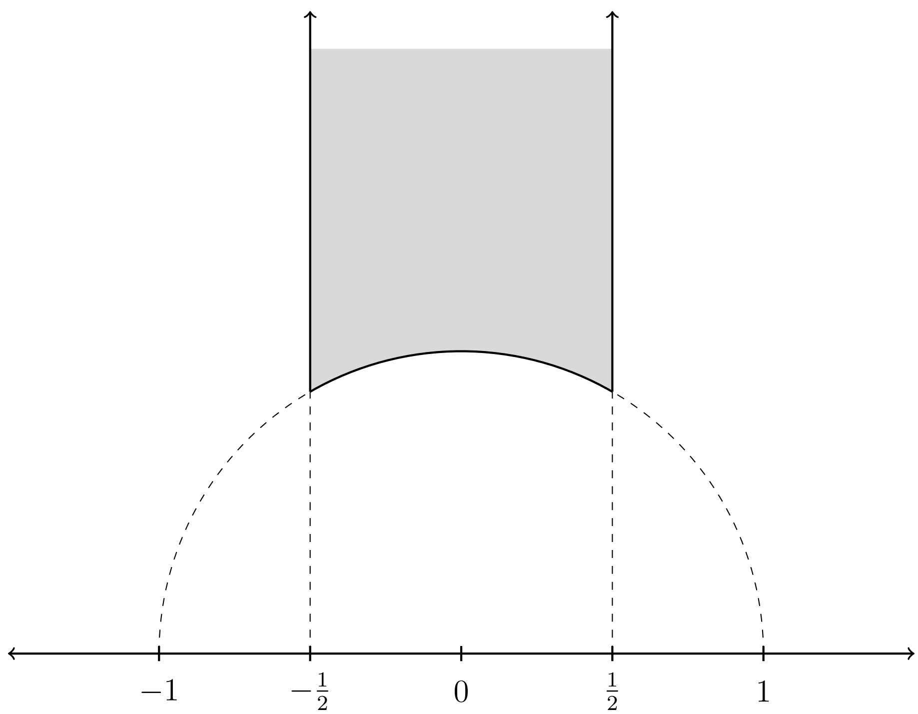


Figure 3.3: The standard fundamental domain $\{|\Re(z)| \leq 1/2, |z| \geq 1\}$ for $\mathrm{PSL}(2,\mathbb{Z})$ in $\mathbb{H}$.

For $x \in \mathbb{R}$ we define the matrix

$$u_x := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Now consider the product of u_x and an element $\binom{a}{n}$ of $\mathbb{Z}^2$:

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ n \end{pmatrix} = \begin{pmatrix} nx + a \\ n \end{pmatrix}.$$

Recall our goal in Diophantine approximation - for really good approximations to x we want to find (a, n) such that $|x - a/n|$ (or equivalently $|nx - a|$) is very small while n is not too large. Then we see that this condition for (a, n) is equivalent to the vector $u_x\binom{a}{n} = \binom{nx+a}{n}$ being very close to the y-axis while not being too far up. This means for a given x, we are examining the unimodular lattice $\Lambda_x := u_x\mathbb{Z}^2$ looking for such points (note that since n is the denominator of our fraction we need $n \neq 0$, so we ignore lattice points on the horizontal axis).

However we don't want to have to take every $x \in \mathbb{R}$ and manually draw the lattice Λ_x so that we can examine its points. Rather we hope for some quicker way that we can identify whether Λ_x has the points we seek or not. Consider the matrix

$$g_t := \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$

and its effect on the point $u_x\binom{a}{n}$:

$$\begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ n \end{pmatrix} = \begin{pmatrix} e^{t/2}(nx+a) \\ e^{-t/2}n \end{pmatrix}.$$

Notice that as t increases from 0, this traces out a hyperbola in $\mathbb{R}^2$; furthermore, the smaller $nx + a$ and/or n is (i.e. the better of an approximation (a, n) is), the closer this hyperbola will come to the origin (see Figure 3.4). This is the heart of the connection between approximation and dynamics - now the problem is about finding lattices with points whose trajectories under g_t take them close to 0. As above, we again don't want to try to deal with each point individually, so instead we examine the action of g_t on the full lattice Λ_x. As $g_t\Lambda_x$ is again a unimodular lattice for any $t \geq 0$, we see that $\{g_t\Lambda_x\}_{t\geq 0}$ is tracing out a trajectory in the space $\mathcal{L}$ of unimodular lattices.

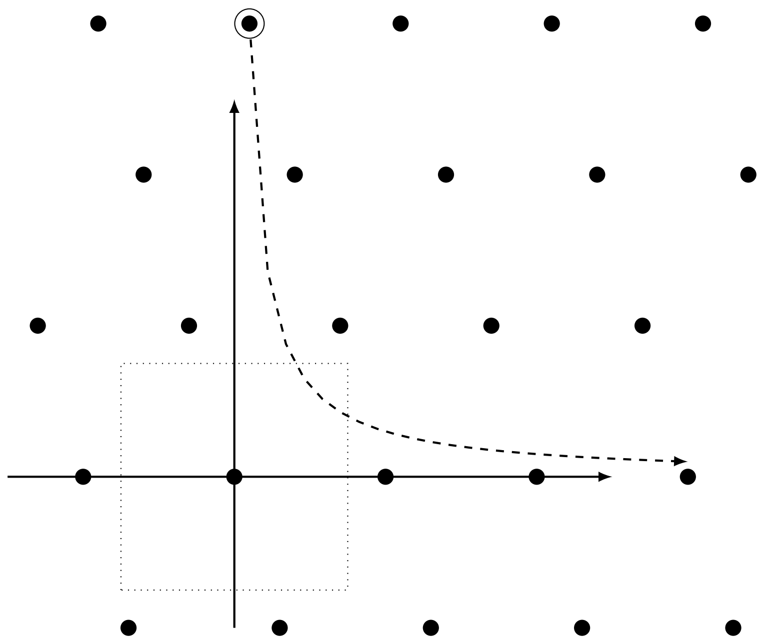


Figure 3.4: The trajectory of a good approximation point passing relatively close to the origin under g_t.

Note that as

$$\mathcal{L} = \left\{ v\mathbb{Z} + w\mathbb{Z} : \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} = \pm 1 \right\} = \left\{ \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \mathbb{Z}^2 \right\},$$

we can act on $\mathcal{L}$ by multiplication by an element of $\mathrm{SL}(2,\mathbb{R})$. Then as $\mathrm{stab}\mathbb{Z} = \mathrm{SL}(2,\mathbb{Z})$, we get an identification of $\mathcal{L}$ with $\mathrm{SL}(2,\mathbb{R})/\mathrm{SL}(2,\mathbb{Z})$.

As mentioned earlier, this space has a cusp (see Figure 3.5). So what does the cusp correspond to in terms of lattices in $\mathcal{L}$? Mahler gives us the answer: lattices living deeper and deeper into the cusp correspond to those lattices having shorter and shorter shortest vectors.

Theorem 9 (Mahler's Compactness Criterion, [Mah46]). *A set $K \subseteq \mathcal{L}$ has compact closure if and only if there is some $s > 0$ with $\Lambda \cap B_s(0) = \{0\}$ for all $\Lambda \in K$.*

Thus limiting ourselves to lattices that are only allowed to live some fixed distance into the cusp in $\mathcal{L}$ means every point in every one of these lattices is bounded away from the origin by some fixed distance s. This together with our perspective on "good" approximations coming close to the origin leads us to the following result:

Theorem 10 (Dani's Correspondence Principle, [Dan85]).

- *$x \in \mathbb{R}$ is badly approximable if and only if $\{g_t\Lambda_x\}_{t\geq 0}$ is bounded, i.e. contained in some compact set $K \subset \mathcal{L}$.*
- *$x \in \mathbb{R}$ is singular if $\{g_t\Lambda_x\}_{t\geq 0}$ is divergent, i.e. for each compact set $K \subset \mathcal{L}$ there is some $T_K \in \mathbb{Z}^+$ such that $g_t\Lambda_x \notin K$ for all $t \geq T_K$.*

A real number x is said to be *singular* if for all $\varepsilon > 0$ there exists N_ε such that for all $N \geq N_\varepsilon$, there exists $(a, n) \in \mathbb{Z} \times \{1, ..., N\}$ with $|x - a/n| < \varepsilon/nN$.

So now with these two theorems we have a way to view approximability just by looking at trajectories of lattices in $\mathcal{L}$. If the trajectory of Λ_x only ever goes a bounded distance into the cusp (see Figure 3.5), then that means the trajectories of every point in Λ_x stay away from the origin and so have a limit on how well the corresponding (a, n) can approximate x, meaning x is badly approximable. If we see the trajectory take a dive deep into the cusp before coming back out again, that corresponds to an abnormally good approximation to x, like we saw in our example with π and $355/113$; recall that this behavior was also the direct result of a large increase in successive partial quotients, suggesting there is some relationship between the partial quotients themselves and the behavior of the trajectory $g_t\Lambda_x$ with respect to the cusp.

There is a direct relationship between this flow on $\mathcal{L}$ and $T^1\mathbb{H}$, the unit tangent bundle of $\mathbb{H}$. (And therefore the unit tangent bundle of the modular surface $T^1M = T^1(\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})))$. We have an identification of $\mathrm{PSL}(2,\mathbb{R})$ with $T^1\mathbb{H}$, given by observing where (i,i) (the unit vector based at i pointing up along the imaginary axis) is sent by each element $g \in \mathrm{PSL}(2,\mathbb{R})$. Therefore to understand the action of g_t on $T^1\mathbb{H}$ it is enough to consider what happens when g_t acts on (i,i). This gives us the geodesic going upwards from i, and recall that it was exactly this geodesic that we mapped around by elements of $\mathrm{PSL}(2,\mathbb{R})$ to find all the other geodesics on $\mathbb{H}$. Therefore g_t acts on $T^1\mathbb{H}$ as the geodesic flow. In the next section we explore the relation of continued fractions and approximation to this geodesic flow, and indeed there is an analogous relation to the depth these geodesic trajectories reach into the cusp on M.

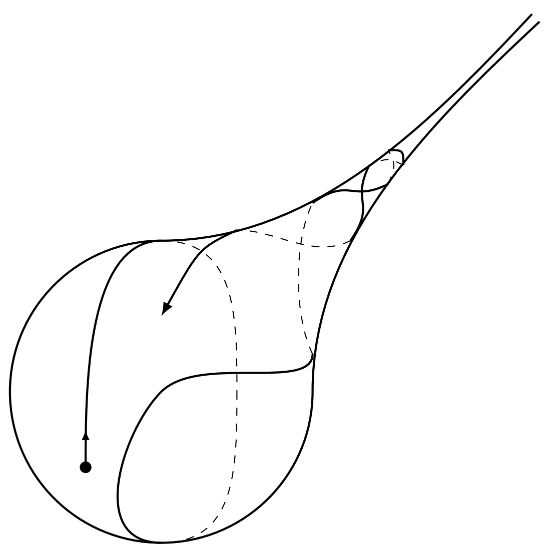


Figure 3.5: Geometric picture of $\mathcal{L}$, with some trajectory $\{g_t\Lambda_x\}$ reaching a bounded distance into the cusp.

3.3 Some modern history

Before moving on, let us briefly discuss some modern results in Diophantine approximation that have similarly utilized a dynamical point of view. Mahler's Compactness Criterion and Dani's Correspondence let us interpret badly approximable (and singular) numbers and points from a dynamical perspective; they correspond to bounded (resp. divergent) orbits in $\mathcal{L}$. What about other types of trajectories? Bounded and divergent orbits are certainly not the only kind of orbits that exist. Indeed, this has been the subject of further study. Refined versions of Dani's Correspondence appear in the literature due to work by Kleinbock and Margulis. They were inspired by a long history of results beginning with a problem of Mahler.

Specifically, Mahler asked the question in 1932, [Mah32] (or rather, a question formulated differently but equivalent to this one): Is almost every point on the Veronese curve $(x, x^2, \ldots, x^d) \subset \mathbb{R}^d$ not very well approximable[1]? This remained an open problem for several decades, until Sprindžuk answered in the affirmative in 1969, [Spr69]. In 1980, he posed the related Sprindžuk's Conjecture, [Spr80] that almost every point on any analytic nondegenerate submanifold of $\mathbb{R}^d$ is not very well approximable. This conjecture was also answered in the affirmative, proved by Kleinbock and Margulis using the tools we have been developing in sections 3.1 and 3.2. In 1998, [KM98] they were able to improve on Dani's Correspondence and were used this to arrive at a proof of Sprindžuk's Conjecture, showing that in some sense being not very well approximable is somehow "in between" badly approximable and singular. In fact, they further developed things in 1999, [KM99]

[1]The definition of very well approximability in $\mathbb{R}^d$ is analogous to the one-dimensional case. We have $\mathbf{x} \in \mathbb{R}^d$ is very well approximable if there exists τ such that $\|n\mathbf{x}\| < n^{-\tau}$ for infinitely many n, where now $\tau > 1/d$ and $\|\mathbf{x}\|$ denotes distance to $\mathbb{Z}^d$ (and this does indeed match up with our original definition when $d = 1$).

using dynamical machinery to offer a new proof of Khintchin's Theorem, which deals with the notion of ψ-*approximability.* For a given non-increasing function $\psi : \mathbb{Z}^+ \to \mathbb{R}^+$, x is said to be ψ-*approximable* if there are infinitely many $a/n \in \mathbb{Q}$ such that $|x - a/n| < \psi(n)/n$. What Khintchin's Theorem states is that if the sum $\sum_n \psi(n)$ diverges, then almost every $x \in \mathbb{R}$ (with respect to Lebesgue measure) is ψ-approximable, and if $\sum_n \psi(n)$ converges, then almost every $x \in \mathbb{R}$ is *not* ψ-approximable.

This is just one series of results, however, and there are quite a few others. Geometric/dynamical ideas are useful in Diophantine approximation in many other ways. For example, one of the biggest open problems is the Littlewood Conjecture from the 1930s that for every $x, y \in \mathbb{R}$, we have $\liminf n\|nx\|\|ny\| = 0$. This is a very hard problem, reflected in the fact that little progress had been made (all attempted through analytic methods) until as recently as 2006, when Einsiedler, Katok, and Lindenstrauss, [EKL06] used a type of "measure rigidity" for the diagonal action on $\mathcal{L}_3$ to prove the Littlewood conjecture up to a set of exceptions of Hausdorff dimension 0.

We can see how there are clearly some some very strong ties between the two fields of dynamics and Diophantine approximation, and there are more than just the ones we have mentioned here. In the next chapter we examine another such relationship, this time between the geodesic flow on $\mathbb{H}$ and continued fractions, which as noted in Section 2 are themselves very strongly tied to Diophantine approximation.

Chapter 4

Continued fractions and the Farey tesselation

In this chapter, we discuss an interpretation of continued fractions using hyperbolic geometry, developed by Caroline Series, [Ser85]. The connection begins with the Farey tesselation.

4.1 Farey sequences and Farey tesselation

Definition 6. The *Farey sequences* are defined in the following way:

$$\begin{aligned}
\mathcal{F}_1 &= \left\{-\infty, \frac{-1}{1}, \frac{0}{1}, \frac{1}{1}, \infty\right\} \\
\mathcal{F}_2 &= \left\{-\infty, \frac{-2}{1}, \frac{-1}{1}, \frac{-1}{2}, \frac{0}{1}, \frac{1}{2}, \frac{1}{1}, \frac{2}{1}, \infty\right\} \\
&\vdots
\end{aligned}$$

That is, the Farey sequence of order n contains all the rationals p/q, listed in increasing order, such that $|p|, |q| \leq n$. (If we need to justify including $\pm\infty$, we can view them as being of the form $\pm 1/0$ so that they appear in every sequence.)

Then the Farey tesselation of $\mathbb{H}$ is a tesselation that is constructed using the Farey sequences. First, draw a vertical line at 0. Now suppose p/q and p'/q' are such that $pq' - p'q = \pm 1$. Then draw a geodesic arc in $\mathbb{H}$ connecting the real points p/q and p'/q'. This is an edge in the Farey tesselation. Notice that our condition on p, p', q, q' is met exactly when p/q and p'/q' are adjacent in some level of the Farey sequence. Notice also that at the end of the sequence of order n we get an edge between $\pm n/1$ and $\pm 1/0$, as $\pm n(0) - \pm 1(1) = \mp 1$; this appears in our tesselation as a vertical line extending up from each integer point on the real axis.

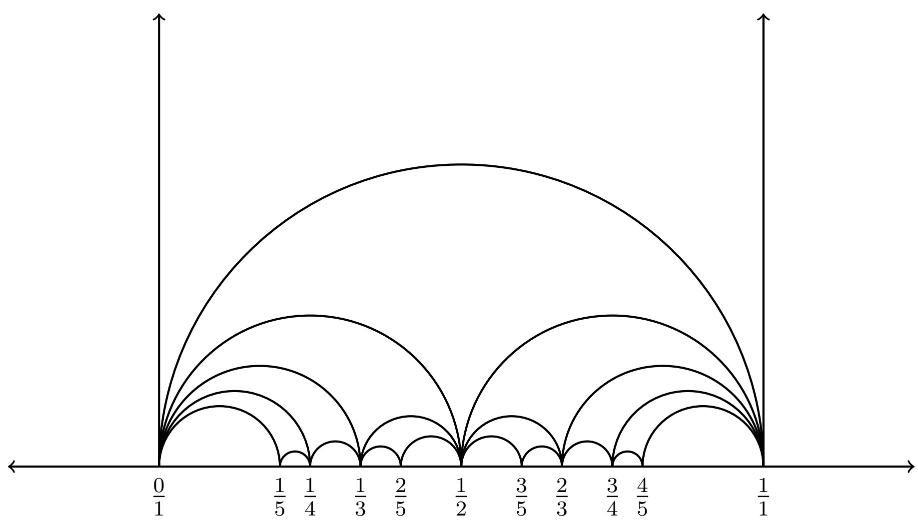


Figure 4.1: The fifth step of the Farey tesselation.

Remark. From another point of view, a tile of the Farey tesselation is a three-fold cover of the fundamental domain of $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$. We begin with our standard fundamental domain, $\{|\Re(z)| \leq 1/2, |z| \geq 1\}$. Take the left half $\{\Re(z) < 0\} \subseteq \{|\Re(z)| \leq 1/2, |z| \geq 1\}$, and shift it to the right by 1. Then mapping this new fundamental domain by the matrix $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}^2$ (this matrix has order 3) gives 3 copies of the fundamental domain filling exactly the triangle with vertices at $0, 1$, and ∞. Note the appearance of this triangle in the first level of our previous construction of the Farey tesselation. Then mapping this triangle by all elements of $\mathrm{SL}(2, \mathbb{Z})$ gives us the Farey tesselation (note that the only times this triangle overlaps with its image are under the action of the identity and the two above matrices, and each of these respect the three copies of the fundamental domain, so there are no problems here). However our previous description gives us a nicer way to actually draw the Farey tesselation, as we are simply connecting in order the elements of each Farey sequence.

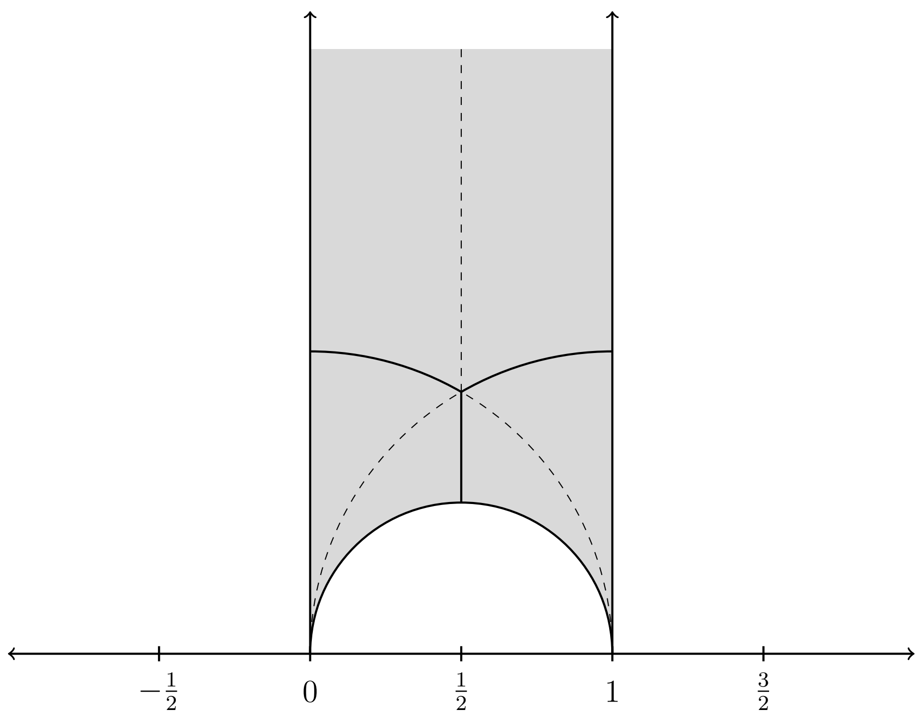
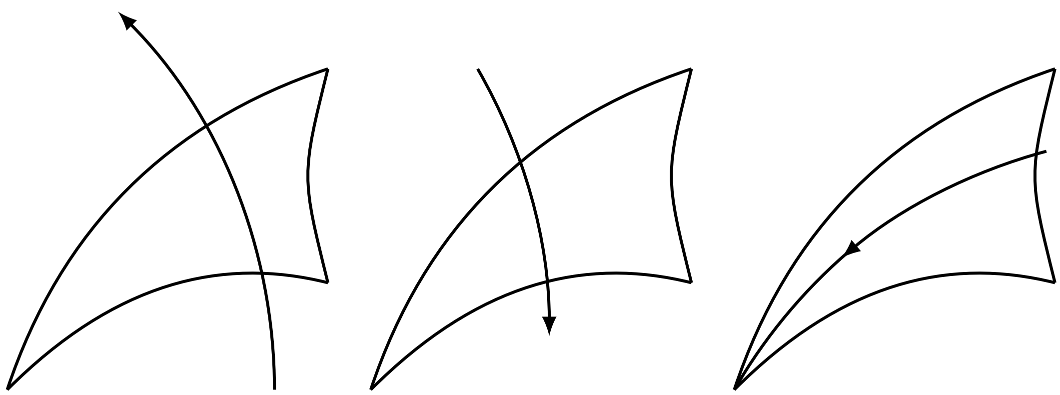


Figure 4.2: Three copies of fundamental region of $\mathrm{SL}(2, \mathbb{Z})$.

4.2 Cutting sequences

Now, let $x \geq 1$, and consider a geodesic in $\mathbb{H}$ that emanates from some point in $(-1, 0)$, and terminates at x. Note that any geodesic of this form must intersect the vertical line at 0 at some point, which we will denote y. Call this geodesic γ, and denote its forward and backward endpoints as γ_∞ and $\gamma_{-\infty}$ respectively. Now consider how this geodesic cuts through the triangles in the Farey tesselation. Generically, when the geodesic passes through a triangle, it divides the triangle into two sections. There are three cases that can occur (see Figure 4.3). If one section contains one vertex of the triangle and one contains two vertices, then we note on which side the single vertex appears; if it appears to the right of γ we call this a "right cut", and likewise if the vertex appears on the left we call it a "left cut." The third possibility is that γ falls directly into one of the vertices of the triangle, in which case both sides of γ have only one vertex, so this is not a true right or left cut. We call this case an "ambiguous cut" as we are free to record it as either right or left, for reasons we will explain shortly.



(a) Left cut. (b) Right cut. (c) Ambiguous cut.

Figure 4.3: Different types of triangle cuts.

So, the geodesic trajectory defines a sequence of "left-right" cuts recording the type of cut γ makes as it passes through each triangle in turn. An example of a cutting sequence is $(\dots L, R, R, R, R, y, L, L, L, R, L, \dots)$, where y denotes the "midway point" of the geodesic, i.e. the intersection of γ with the imaginary axis (so the cut directly following y corresponds to the triangle with vertices $0, 1, \infty$).

We can abbreviate our cutting sequences slightly by noting the length of each sequence of cuts of the same type:

$$(\underbrace{\ldots L,}_{a_{-2}}\underbrace{R,R,R,R,}_{a_{-1}}y,\underbrace{L,L,L,}_{a_0}\underbrace{R,}_{a_1}\underbrace{L,\ldots}_{a_2}).$$

In our example, $a_{-1} = 4$, $a_0 = 3$, and $a_1 = 1$, so our cutting sequence is written $...L^{a_{-2}}R^4yL^3RL^{a_2}...$ Series shows the following remarkable connection between this sequence and the continued fraction expansion of x:

Theorem 11 (Series, [Ser85])**.**
Let $x \geq 1$ and γ as above, with cutting sequence $\ldots R^{a_{-1}}yL^{a_0}\ldots$. Then

$$\gamma_\infty = x = [a_0; a_1, a_2, \ldots] \qquad \frac{-1}{\gamma_{-\infty}} = [a_{-1}; a_{-2}, a_{-3}, \ldots].$$

Some discussion is in order. For one, since we already know that continued fraction expansions are unique, we can infer from the above theorem that the cutting sequence of the forward half does not depend on our choice of $\gamma_{-\infty} \in (-1, 0)$. Notice also that since the Farey tesselation is symmetric about the imaginary axis, we have an analogous result for a geodesic γ with $\gamma_\infty \leq -1$ and $\gamma_{-\infty} \in (0, 1)$. In this case the cutting sequence of γ is of the form $\ldots L^{a_{-1}}yR^{a_0}\ldots$, and we have $\gamma_\infty = -[a_0; a_1, a_2, \ldots]$, $1/\gamma_{-\infty} = [a_{-1}; a_{-2}, a_{-3}, \ldots]$.

Another note is what happens when $x \in \mathbb{Q}$. We know the rationals are precisely those continued fractions of finite length, $x = [a_0; a_1, ..., a_n]$ for some $n \in \mathbb{N}$. This means the cutting sequence of γ must terminate at some finite step, and the only way for this to happen is if γ makes an ambiguous cut somewhere. If it falls directly into the vertex of some triangle, then it will not cut across any further triangles and so the cutting sequence will end there. But this is perfect, as the

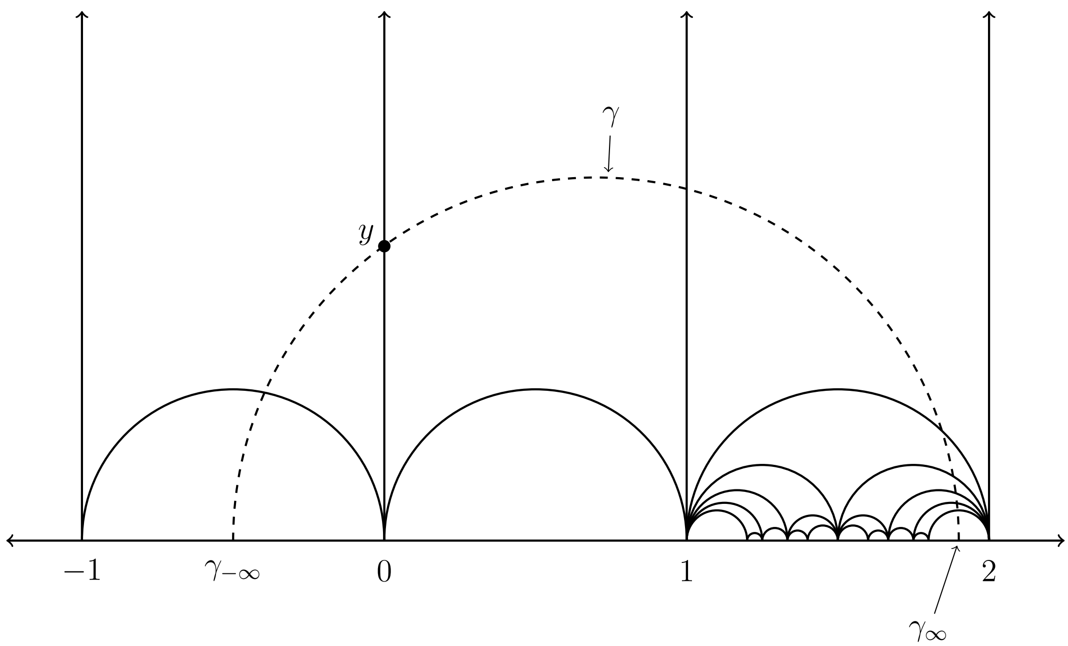


Figure 4.4: The geodesic terminating at the point γ_∞. From the cutting sequence, we may deduce that $\gamma_\infty = [1; 1, a_2, \ldots]$, where $a_2 \geq 4$. Notice how much "zooming in" would be required to determine even the second partial quotient a_2. This is a testament to the precision of approximation that continued fractions provide.

vertices of the triangles of the Farey tesselation are only at rationals on the real axis (and ∞, but γ must end at $x \in \mathbb{R}$ so this is irrelevant). Thus if the cutting sequence of γ terminates, it must be that $x \in \mathbb{Q}$, which aligns with the theorem.

This is also where our freedom of choice when presented with an ambiguous cut comes in. We know that every $x \in \mathbb{R}$ has a unique continued fraction expansion, with the slight caveat that $x \in \mathbb{Q}$ actually has two expansions. If $x = [a_0; \ldots, a_n]$ and $a_n > 1$, notice that x is also equal to $[a_0; \ldots, a_n - 1, 1]$:

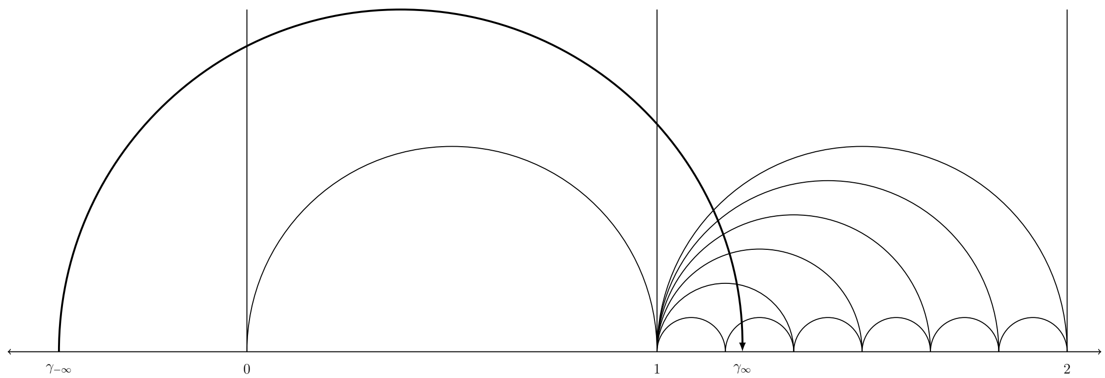
$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n - 1 + \frac{1}{1}}}}}.$$

This is reflected in the final cut of a geodesic ending in a rational being an ambiguous cut. We have $a_n - 1$ cuts of one type followed by an ambiguous cut, which we can either record as a cut of the same type as the preceding cuts, or as a change in cut type, giving us expansions ending in $[\dots, a_n]$ and $[\dots, a_n - 1, 1]$ respectively. For simplicity's sake we will generally stick to the former case of recording an ambiguous cut as matching the type of cut immediately preceding it.

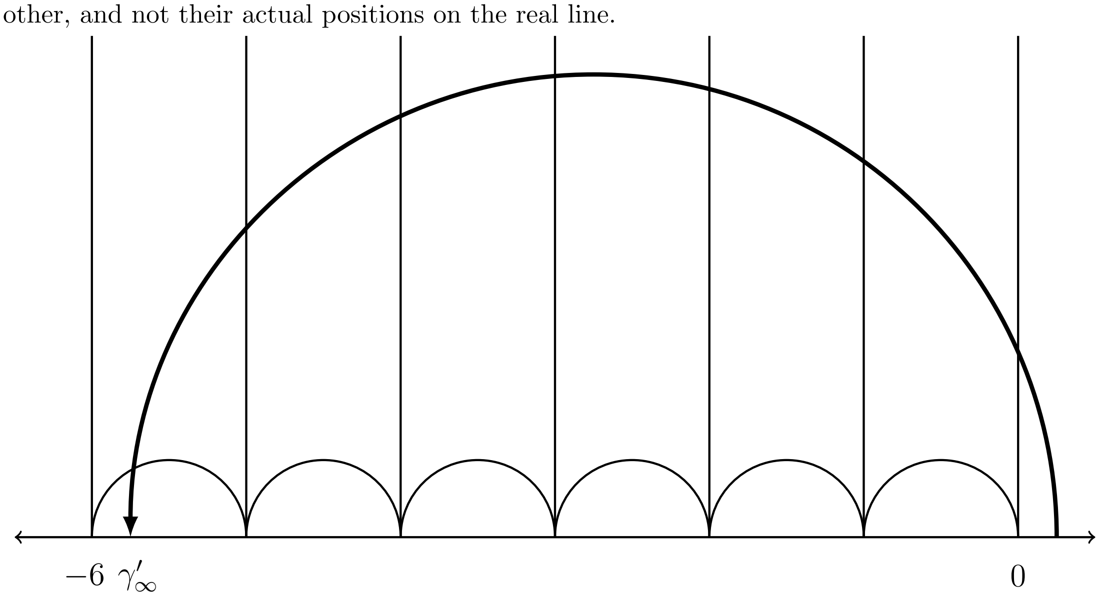
Series has a number of other results in [Ser85] considering dynamics on the modular surface M and the continued fraction map $T(\theta) = 1/\theta - \lfloor 1/\theta \rfloor$, but let us focus on the main theorem above, and try to derive some consequences. Series does some of this herself towards the end of her paper, where she uses cutting sequences to form a proof of the earlier theorem of Euler–Lagrange that a number x has periodic continued fraction expansion if and only if it is a quadratic irrational. We will continue this line of thought and attempt to use this geometric picture to provide alternative proofs of other known properties of continued fractions.

First, let us give an overview of the proof of Theorem 11. Note that $\gamma_\infty \geq 1$ implies that the first cut immediately following y must be of type L, and similarly $-1 < \gamma_{-\infty} < 0$ implies the cut immediately preceding y must be of type R. Let $p_0 := \max\{n \in \mathbb{N} : n < \gamma_\infty\}$, and let $\eta_\gamma = \gamma \cap \{\Re(z) = p_0\}$. By construction we have that the cutting sequence of the segment of γ between y and η_γ is L^{a_0}. Now apply the transformation $\rho_0 : z \mapsto -1/(z - p_0)$ (notice how this mirrors the continued fraction algorithm, where we have $a_0 = \lfloor x_0 \rfloor$ and $x_1 = (x_0 - a_0)^{-1}$). Series' formulation of the theorem involves looking at γ as a lift of a geodesic on the modular surface $M = \mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$; since $\rho_0 \in \mathrm{SL}(2,\mathbb{Z})$ itself, we get that $\rho_0(\gamma)$ is a lift of the same geodesic on M as γ. Thus we see that $\rho_0(\gamma)_\infty \leq -1$, $0 < \rho_0(\gamma)_{-\infty} \leq 1$, and $\rho_0(\eta_\gamma) = y_{\rho_0(\gamma)}$ (i.e. the intersection of the geodesic $\rho_0(\gamma)$

with the imaginary axis). Furthermore the cutting sequence of γ beginning at η_γ is the same as the cutting sequence of $\rho_0(\gamma)$ beginning at $y_{\rho_0(\gamma)}$. Similarly defining $p_1 = \max\{n \in \mathbb{N} : -n > \rho_0(\gamma)_\infty\}$ and $\rho_1 : z \mapsto -1/(z + p_1)$ repeats the argument and brings us back to a geodesic with backwards endpoint in $(-1, 0)$, and forward endpoint ≥ 1 with continued fraction expansion $[a_2; a_3, ...]$ (see Figure 4.5).



(a) A picture of γ with cutting sequence $\ldots yLR^5L^{a_2}\ldots$. This indicates that $\gamma_\infty = [1; 5, a_2, \ldots]$. Note that for convenience we have not drawn the tesselation to scale as we are only concerned with the positions of γ_∞ and the vertices of the triangles in relation to each other, and not their actual positions on the real line.



(b) The image of γ after applying the transformation $z \mapsto -1/(z-1)$, denoted γ'. Notice that the five right cuts from the previous figure now define the first partial quotient of γ'_∞.

Figure 4.5

Now, with Theorem 11 in hand, let us see what properties we can observe about continued fractions simply by taking this geometric perspective.

Considering how important the convergents of a continued fraction are, it is natural that one of our first questions would be how to recognize them in the picture. Recall that the nth convergent of $x = [a_0; a_1, a_2, ...]$ is obtained by truncating the continued fraction: $p_n/q_n = [a_0; a_1, ..., a_n]$. With the view of Series, this means we have a geodesic with finite forward cutting sequence $L^{a_0}R^{a_1}...L^{a_n}$ (if n is even) or $L^{a_0}R^{a_1}...R^{a_n}$ (if n is odd). To get this geodesic, we simply take our original geodesic γ and look for the cut corresponding to the last cut in this truncated sequence (that is, the one right before the $(n+1)$st type change). Then we perturb γ so that this final L or R cut becomes an ambiguous cut (see Figure 4.6).

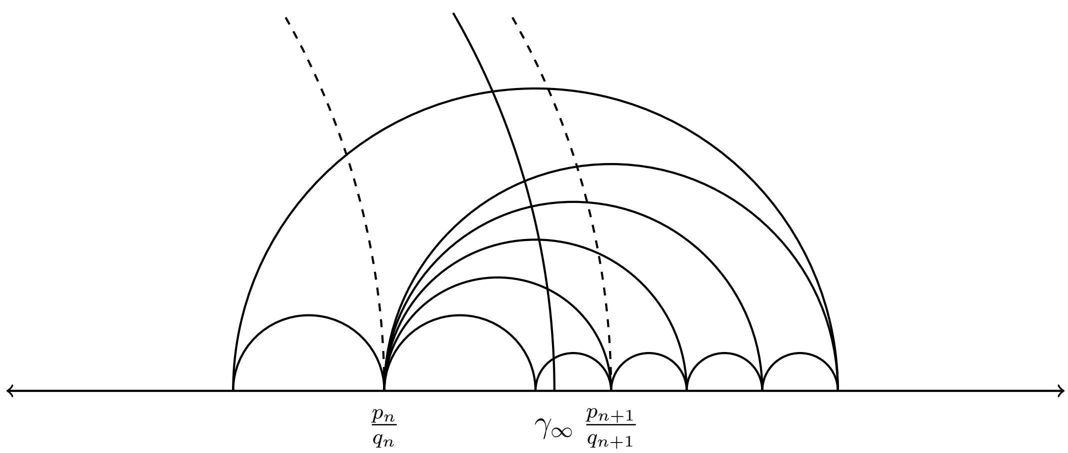


Figure 4.6: Perturbing γ to p_n/q_n and p_{n+1}/q_{n+1}. Again note the tesselation is not to scale.

Notice that if the cut we are changing is a left cut, then our perturbation shifts γ in the negative direction, while changing a right cut to an ambiguous cut means we perturb γ in the positive direction. Since the cutting sequence of the nth convergent ends in a left cut if n is even and a right cut if n is odd, we have already proved the following lemma from earlier:

Lemma 12. *The convergents of x alternately over/under estimate x. More specifically, $\frac{p_n}{q_n} < x$ if n is even, and $\frac{p_n}{q_n} > x$ if n is odd.*

There are further immediate observations we can make about consecutive convergents. Referring back to Figure 4.6, notice that whenever we have a sequence of cuts of the same type, the isolated vertex in each triangle lies on the same point on the real axis. This is due to the fact that the triangles corresponding to two consecutive cuts share a side, and if the other edge of the second triangle emanating from this shared vertex did not reach γ that would mean we had two vertices on that side of γ which would force a type change. Since the cut we perturb to find p_n/q_n occurs right before a type change, the next a_{n+1} cuts are all the same type as each other. Thus the isolated vertices of each of these cuts lies at the same point in $\mathbb{R}$, which we recognize from the picture as exactly the same point as the vertex that we perturbed γ into, i.e. at p_n/q_n. The cut we need to perturb to find p_{n+1}/q_{n+1} is one of these, so that triangle's isolated vertex is at p_n/q_n; therefore the vertex we perturb γ into at p_{n+1}/q_{n+1} is connected by a side of this triangle to the vertex at p_n/q_n. This proves the next lemma.

Lemma 13. *Any two consecutive convergents of $x > 1$ are connected by a triangle side of the Farey tesselation.*

Furthermore, since two rationals p/q and p'/q' are connected by an arc of the Farey tesselation if and only if $pq' - p'q = \pm 1$, we get the following known property of convergents for free.

Lemma 14. *For all $n \geq 0$, we have that $p_n q_{n+1} - p_{n+1} q_n = \pm 1$.*

It is also clear from the picture that the convergents get ever closer to x. After all, as we proceed further along γ the triangles we cut across become arbitrarily small, and therefore the perturbations of γ to the convergents that we make inside those triangles become arbitrarily small. But we can do better. The lemmas we have observed so far are enough to prove the upper bound on $|x - p_n/q_n|$ that we noted back in Section 2:

Lemma 15. *For all $n \geq 0$ we have $|x - p_n/q_n| < \dfrac{1}{q_n q_{n+1}}$.*

Proof. We use the fact that consecutive convergents appear on opposite sides of x to note that

$$\left| x - \frac{p_n}{q_n} \right| < \left| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right| = \left| \frac{p_n q_{n+1} - p_{n+1} q_n}{q_n q_{n+1}} \right|.$$

But by the previous lemma, $|p_n q_{n+1} - p_{n+1} q_n| = 1$, so we have $|x - p_n/q_n| < 1/(q_n q_{n+1})$. □

As we discussed earlier, this upper bound of $1/(q_n q_{n+1})$ relates directly to the fact that if a_{n+1} is large, then p_n/q_n is a very good approximation. But even this observation can be seen directly in the picture. As noted before, the sequence of a_{n+1} cuts of the same type have each of their respective isolated vertices at the same point p_n/q_n on the real axis. Notice however that each of these triangles' respective middle vertices (i.e. the vertex we would perturb γ into to make an ambiguous cut) get closer and closer to p_n/q_n as we progress along γ (see Figure 4.6). Therefore if a_{n+1} is very large, this means as we cut through each of these smaller and smaller triangles our geodesic γ remains between p_n/q_n and the middle vertex of the triangle for a very long time, and so the endpoint x of γ must itself be very close to p_n/q_n.

There is even another way we can arrive at this observation from the picture. Recall that our justification for this connection between large a_{n+1} and quality of p_n/q_n as an approximation to x back in Section 2 was based on the recursion formula for q_{n+1} involving a_{n+1}, so that if a_{n+1} is large it forces q_{n+1} to be large, which forces the bound of $1/(q_n q_{n+1})$ to be small. So if we can somehow read the recursion formula off of the picture, then we will have arrived at the same justification purely from our geometric perspective.

It is a curious known fact that one way to construct the Farey fractions is to start at the first level with $0/1$ and $1/1$, and then fill in the gap between two consecutive fractions p/q, p'/q' by performing a sort of naïve addition p/q "+" $p'/q' = (p+p')/(q+q')$. Observe the first few steps of this construction:

$$\begin{array}{ccccccccc}
\frac{0}{1} & & & & & & & & \frac{1}{1} \\
\frac{0}{1} & & & & \frac{0+1}{1+1}=\frac{1}{2} & & & & \frac{1}{1} \\
\frac{0}{1} & & \frac{0+1}{1+2}=\frac{1}{3} & & \frac{1}{2} & & \frac{1+1}{1+2}=\frac{2}{3} & & \frac{1}{1} \\
\frac{0}{1} & \frac{0+1}{1+3}=\frac{1}{4} & \frac{1}{3} & \frac{1+1}{3+2}=\frac{2}{5} & \frac{1}{2} & \frac{1+2}{2+3}=\frac{3}{5} & \frac{2}{3} & \frac{2+1}{3+1}=\frac{3}{4} & \frac{1}{1}
\end{array}$$

So what is the analogue in the Series picture? It means that when we have an arc in the Farey tesselation between points p/q and p'/q', the two arcs that appear beneath it that together form a triangle are connecting the points p/q to $(p+p')/(q+q')$ and $(p+p')/(q+q')$ to p'/q'. Now, recall that any two consecutive convergents are connected by an arc of the Farey tesselation; in particular, we have an arc between p_n/q_n and p_{n-1}/q_{n-1} which then bounds a triangle with middle vertex at $(p_n + p_{n-1})/(q_n + q_{n-1})$ (see Figure 4.7). Notice that γ cuts through this triangle, and the triangle direcly preceding this one in the cutting sequence is the triangle in which we perturb γ to find p_n/q_n. The triangle in which we next perturb γ (to find p_{n+1}/q_{n+1}) appears a_{n+1} cuts after this one. But where are the middle vertices of these a_{n+1} triangles that γ cuts through? We know the first one is at $(p_n+p_{n-1})/(q_n+q_{n-1})$, so then the arc between this point and p_n/q_n is the top arc of the second triangle. Following our naïve addition, the middle vertex of this second triangle is at $(p_n+(p_n+p_{n-1}))/(q_n+(q_n+q_{n-1})) = (2p_n+p_{n-1})/(2q_n+q_{n-1})$. Continuing this logic, the middle vertex of the a_{n+1}st triangle is at the point $(a_{n+1}p_n + p_{n-1})/(a_{n+1}q_n + q_{n-1})$. But the middle vertex of the a_{n+1}st triangle is exactly where we perturb γ in order to find p_{n+1}/q_{n+1}.

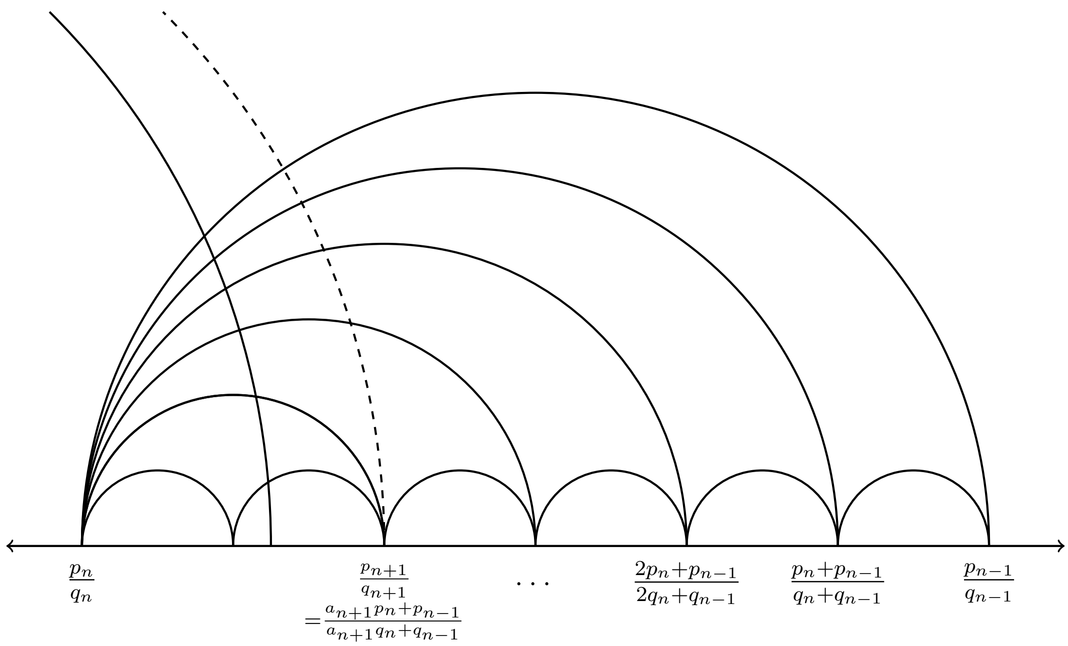


Figure 4.7: Observing the recursion formula in the picture. Note again that the tesselation is not to scale.

Thus we have proved exactly the recursion formula from Section 2:

Lemma 16. *For all* $n \geq 0$ *we have* $p_{n+1} = a_{n+1}p_n+p_{n-1}$ *and* $q_{n+1} = a_{n+1}q_n+q_{n-1}$.

Now that we have seen this formula in the picture, we can say that our geometric perspective has both allowed us to observe this implication of a large a_{n+1}

as well as provided the justification for why it should be so. Furthermore, it can be used in conjunction with other lemmas from earlier in this section to prove even more properties of convergents, such as the lower bound on $|x - p_n/q_n|$ that was also mentioned back in Section 2.

Lemma 17. *For all $n \geq 0$ we have $|1/(q_n(q_n + q_{n+1}))| < |x - p_n/q_n|$.*

Proof. We use the fact that the convergents alternately over/under estimate x to note that p_n/q_n and p_{n+2}/q_{n+2} must both appear on the same side of x. Moreover, since the convergents are also coming ever closer to x, we have that p_{n+2}/q_{n+2} must be in between p_n/q_n and x. In other words, we have the inequality

$$\left|x - \frac{p_n}{q_n}\right| > \left|\frac{p_n}{q_n} - \frac{p_{n+2}}{q_{n+2}}\right|.$$

Then applying the recursive formula for p_{n+2} and q_{n+2}, along with $|p_n q_{n+1} - q_n p_{n+1}| = 1$, we get

$$\begin{aligned}
\left|x - \frac{p_n}{q_n}\right| &> \left|\frac{p_n}{q_n} - \frac{a_{n+2}p_{n+1} + p_n}{a_{n+2}q_{n+1} + q_n}\right| \\
&= \left|\frac{p_n(a_{n+2}q_{n+1} + q_n) - q_n(a_{n+2}p_{n+1} + p_n)}{q_n(a_{n+2}q_{n+1} + q_n)}\right| \\
&= \left|\frac{p_n q_n - q_n p_n + a_{n+2}(p_n q_{n+1} - q_n p_{n+1})}{q_n(a_{n+2}q_{n+1} + q_n)}\right| \\
&= \left|\frac{a_{n+2}}{q_n(a_{n+2}q_{n+1} + q_n)}\right| \\
&\geq \frac{a_{n+2}}{q_n(a_{n+2}q_{n+1} + a_{n+2}q_n)} = \frac{1}{q_n(q_{n+1} + q_n)},
\end{aligned}$$

finishing the proof. □

Note that just as large a_{n+1} implied a large q_{n+1} which made for a small upper bound forcing x to be close to p_n/q_n, we similarly have the lower bound becoming small with large q_{n+1}, so x is also "allowed" to be close to p_n/q_n.

Chapter 5

Conclusion

The results we have discussed so far are just the fundamental properties of continued fractions. One can push this further, of course, and explore via the geometry many other important qualities of continued fractions. For instance, we can use the fact of the badly approximable numbers being those with bounded partial quotients to characterize them with respect to the behavior of the geodesic γ on $M = \mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$.

Looking at the Farey tesselation on $\mathbb{H}$ itself, the easiest part of γ to visualize on M is the initial series of a_0 cuts from the imaginary axis across the triangles with vertices at integer points and ∞. We see that γ reaching higher in the plane while making these cuts corresponds directly to the projection of γ on M reaching deeper into the cusp, for that segment at least. However by the methods we saw in the proof of Series' main theorem, we can find another lift of the same geodesic with the cutting sequence shifted so that the sequence of a_n cuts occurs immediately after crossing the imaginary axis, for whatever n we desire (see Figure 4.5). Therefore if x has bounded partial quotients, γ or any other lift of the corresponding geodesic on M can only ever make a limited number of cuts across these particular triangles, and so the geodesic on M can only ever travel a bounded distance into the cusp. This gives us a characterization of BAD in yet another setting.

In our quest to rely solely on the geometric view wherever possible, however, the next question to ask would be if we can somehow observe the relationship between only being able to make a bounded number of cuts of the same type back to back (i.e. bounded partial quotients) and the original definition of bad approximability vis-à-vis there existing only finitely many rationals beating a certain bound. Suppose x is badly approximable, with constant c. Where does c appear in the picture? Is there a way to recognize some or all of the finitely many approximations that do come within c/n^2? This could be one good subject for further inquiry.

Another phenomenon to study could be that of best approximations. A number $q \in \mathbb{N}$ is a best approximation to x if for all $q' < q$ we have $\|q'x\| > \|qx\|$, and it is well known that these q are exactly the denominators of the convergents of x. The phenomenon in the picture should therefore be each record breaker corresponding to a change in cut type in some way. Specifying exactly how to see this connection would be useful as well.

Bibliography

[Cas57] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45, Cambridge University Press, New York, 1957.

[Dal11] Françoise Dal'Bo, *Geodesic and horocyclic trajectories*, Universitext, Springer-Verlag London, Ltd., London; EDP Sciences, Les Ulis, 2011, Translated from the 2007 French original. MR 2766419

[Dan85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[EKL06] Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), no. 2, 513–560.

[EW11] Manfred Einsiedler and Thomas Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer-Verlag London, Ltd., London, 2011.

[KM98] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360.

[KM99] ______, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494.

[Mah32] Kurt Mahler, *Über das Maß der Menge aller S-Zahlen*, Math. Ann. **106** (1932), no. 1, 131–139. MR 1512754

[Mah46] K. Mahler, *On lattice points in n-dimensional star bodies. I. Existence theorems*, Proc. Roy. Soc. London. Ser. A. **187** (1946), 151–187. MR 0017753

[RS92] Andrew M. Rockett and Peter Szüsz, *Continued fractions*, World Scientific Publishing Co., Inc., River Edge, NJ, 1992.

[Ser85] Caroline Series, *The modular surface and continued fractions*, J. London Math. Soc. (2) **31** (1985), no. 1, 69–80.

[Spr69] V. G. Sprindžuk, *Mahler's problem in metric number theory*, Translated from the Russian by B. Volkmann. Translations of Mathematical Monographs, Vol. 25, American Mathematical Society, Providence, R.I., 1969.

[Spr80] ______, *Achievements and problems of the theory of Diophantine approximations*, Uspekhi Mat. Nauk **35** (1980), no. 4(214), 3–68, 248. MR 586190