CHAPTER 1

Introduction

1.1. From Geometry to Dynamics

Figure 1.1 shows the constructions for three classical theorems in geometry.

![Figure 1.1. Three classic theorems](image)

(1) Start with an arbitrary quadrilateral. Then the midpoints of the edges of the quadrilateral are the vertices of a parallelogram. I don’t know the origins of this result.

(2) Start with an arbitrary triangle and construct equilateral triangles on each of the sides. Then the centers of the three equilateral triangles themselves make an equilateral triangle. This is known as Napoleon’s Theorem, and it is attributed (perhaps not in all seriousness) to the famous emperor.

(3) Start with an arbitrary pentagon. Connect the vertices in a star pattern and consider the smaller pentagon in the middle. The inner pentagon and the outer pentagon are equivalent by a projective transformation. In some sense this result was known to Darboux, and it is studied explicitly in [Mot] and [S1].

All these results have probably been rediscovered many times. I will give proofs of the first two of these results in §2, and the third one in §5.1. (The first one is very easy.)

Figure 1.2 shows the same constructions for polygons having more vertices, and one can ask whether the theorems above generalize. Strictly speaking, the results above do not generalize as configuration theorems. For instance, the quadrilateral joining the centers of the equilateral triangles is not typically a square.
However, there are dynamical generalizations of the results. For each of the constructions, let \( P \) be the initial polygon and let \( P' \) denote the polygon defined by the construction. Let \( P^{(2)} = P'' \) and \( P^{(3)} = P''' \), etc.

In the first case, it is useful to work in \( \mathcal{A}_N \), the space of \( N \)-gons modulo affine transformations. An affine transformation is the composition of a linear isomorphism and a translation. Two \( N \)-gons are equivalent if there is an affine transformation carrying one to the other. The midpoint map commutes with affine transformations, so it makes sense to consider \( \{P^{(n)}\} \) as a sequence in \( \mathcal{A}_n \). The general result is that \( \{P_n\} \) converges to the equivalence class of the regular \( N \)-gon for almost all choices of \( P \). This is a classical result which is closely related to heat flow and the discrete Fourier transform. I will give a proof in \( \S 2 \).

In the second case, it is useful to work in the space \( \mathcal{S}_N \) of \( N \)-gons modulo similarity. Napoleon’s construction commutes with similarities and so it makes sense to consider \( \{P^{(n)}\} \) as a sequence in \( \mathcal{S}_N \). What typically happens to this sequence depends on \( N \). For example, when \( N = 12 \) the sequence \( \{P^{(n)}\} \) converges to the class of the 12-gon which wraps twice around the regular hexagon for almost all choices of \( P \). \( \S 2.4 \) I’ll give an analysis of most cases. See also [Bo] and [Gr].

The third construction is the pentagram map, which I introduced in [S1]. In recent years, there have been many papers on the pentagram map. See the references listed in \( \S 5.2 \). The natural setting for the pentagram map is the space \( \mathcal{P}_N \) of projective equivalence classes of \( N \)-gons. Here, the \( N \)-gons are considered subsets of the projective plane, and two \( N \)-gons are equivalent if there is a projective transformation carrying one to the other.

In brief, the projective plane \( \mathbb{RP}^2 \) is the space of lines through the origin in \( \mathbb{R}^3 \), and a projective transformation is a map on \( \mathbb{RP}^2 \) induced by an invertible linear transformation on \( \mathbb{R}^3 \). In \( \S 3 \) I will give a primer on projective geometry and explain this in more detail. The basic result here, due to V. Ovsienko, myself, and S. Tabachnikov, [OST1], [OST2], and independently due to F. Soloviev [Sol], is that the pentagram map is (in the appropriate sense) a discrete completely integrable system on \( \mathcal{P}_N \). I will discuss this and other results about the pentagram map in \( \S 5 \).

All the results above make statements about polygon iterations. One starts with an \( N \)-gon \( P \) and produces a new \( N \)-gon \( P' \) by some geometric construction. One then asks about the behavior of the sequence \( \{P^{(n)}\} \), perhaps modulo some group
1.2. The Projective Heat Map

The projective heat map is based on the following construction. Given 4 points \( a_{-3}, a_{-1}, a_1, a_3 \in \mathbb{RP}^2 \) in general position – i.e. no three lie on a line – there is a canonical choice of a point \( b_0 \) on the line \( \overline{a_{-1}a_1} \). The construction is shown in Figure 1.3. One might call \( b_0 \) the projective midpoint of \( a_{-3}, a_{-1}, a_1, a_3 \). Note that \( b_0 \) is typically not the actual midpoint of the segment \( \overline{a_{-1}a_1} \). One situation where this does happen is when there is some isometry that swaps \( a_{-3} \) with \( a_3 \) and \( a_{-1} \) with \( a_1 \), but this is not the typical situation.

![Figure 1.3. The construction of \( b_0 \) from \( a_{-3}, a_{-1}, a_1, a_3 \)](image)

Starting with an \( N \)-gon \( P \), with vertices \( \ldots a_{-3}, a_{-1}, a_1, a_3, \ldots \) we form the \( N \)-gon \( P' \) with vertices \( \ldots b_{-2}, b_0, b_2, \ldots \), where \( b_k \) is the projective midpoint of \( a_{k-3}, a_{k-1}, a_{k+1}, a_{k+3} \). The indices are taken cyclically in this construction. We define \( H(P) = P' \). The map \( H \) commutes with projective transformations and thus \( H \) induces a map on the quotient space \( \mathcal{P}_N \) of projective equivalence classes of \( N \)-gons. The subspace \( \mathcal{C}_N \) of equivalence classes of convex \( N \)-gons is an invariant subspace. It is worth remarking that \( H \), like the pentagram map, is not entirely defined on \( \mathcal{P}_N \). The points of the polygon need to be in sufficiently general position for this to make sense.

Numerical evidence supports the following conjecture.

**Conjecture 1.1.** For any \( N \geq 5 \), and for almost all \( P \in \mathcal{P}_N \), the sequence \( \{P^{(n)}\} \) converges to the projectively regular class.

I will discuss Conjecture 1.1 in §7.5. What makes this conjecture difficult is that, in contrast to the midpoint map and Napoleon’s construction, the map \( H \) is nonlinear. Even in the case \( N = 5 \), where the basic space \( \mathcal{P}_5 \) is a 2 dimensional space, there is a lot of complexity. In the case \( N = 5 \), the projective heat map gives rise to a two-variable rational map

\[
H(x, y) = (x', y'), \\
x' = \frac{(xy^2 + 2xy - 3) (x^2y^2 - 6xy - x + 6)}{(xy^2 + 4xy + x - y - 5) (x^2y^2 - 6xy - y + 6)}
\]
\begin{equation}
\begin{aligned}
y' & = \frac{(x^2y + 2xy - 3)(x^2y^2 - 6xy - y + 6)}{(x^2y + 4xy - x + y - 5)(x^2y^2 - 6xy - x + 6)} \\
\end{aligned}
\end{equation}

Our notation is such that $H$ stands both for the map on $\mathcal{P}_5$ and the above rational map on $\mathbb{R}^2$.

\section*{1.3. A Picture of the Julia Set}

The main goal of this monograph is to prove Conjecture 1.1 for the case $N = 5$. At this point we set $\mathcal{P} = \mathcal{P}_5$, etc., because this is the only case we consider. To be precise, $\mathcal{P}$ consists of projective equivalence classes of pentagons whose points are in general position.

We let $\mathcal{J}$ be the subset of $\mathcal{P}$ consisting of those points with well-defined orbits that do not converge to the regular class. The set $\mathcal{J}$ is akin to the Julia set from complex dynamics, a topic we discuss in §6.1. Our proof of Conjecture 1.1 amounts to showing that $\mathcal{J}$ has measure 0. However, given the beauty of $\mathcal{J}$, I couldn’t resist analyzing it and getting finer information about it. Almost all my motivation for this monograph came from wanting to rigorously justify the computer pictures of $\mathcal{J}$ which I produced.

Figure 1.4 shows the most interesting portion of $T \circ B(\mathcal{J})$, where $T$ is a certain linear transformation and $B$ is a birational map discussed below. The points are colored according to how many iterates of $\mathcal{H}$ it takes to map them into $\mathcal{C}$, the space of convex classes, and then these colored points are mapped into the picture plane via $T \circ B$. Once a point gets into $\mathcal{C}$ it converges under iteration to the regular class. (See Theorem 1.3 below.) The darker the color, the longer it takes. So, $T \circ B(\mathcal{J})$ would be the black points – or at least the black points with well defined orbits.

\section*{1.4. The Core Results}

Our first result is topological in nature.

\textbf{Theorem 1.2.} The map $H$ is generically 6-to-1 when it acts on $\mathcal{C}^2$.

Our next result deals with the action of $H$ on $\mathcal{C}$.

\textbf{Theorem 1.3.} There is a smooth and rational function $f : \mathcal{C} \to \mathbb{R}$ such that $f \circ H(P) \geq f(P)$ for all $P \in \mathcal{C}$, with equality if and only if $P$ is the regular class. Moreover, the level sets of $f$ are compact.

Given that $f$ has compact level sets, Theorem 1.3 has the following corollary. See the proof of Corollary 9.7 for details.

\textbf{Corollary 1.4.} $\{H^n(P)\}$ converges to the regular class for all $P \in \mathcal{C}$.

Now we turn to Conjecture 1.1. The way we understand $\mathcal{J}$ is that we first understand a certain Cantor set in $\mathcal{J}$ and then we understand the rest. Accordingly, here are our two main structural results.

\textbf{Theorem 1.5.} $\mathcal{J}$ contains a measure 0 forward-invariant Cantor set $\mathcal{J}C$. The restriction of $H$ to $\mathcal{J}C$ is conjugate to the 1-sided shift on 6-symbols.

\textbf{Theorem 1.6.} $\mathcal{J}$ contains a measure 0 forward-invariant Cantor band $\mathcal{J}A$ such that

$$\mathcal{J} = \mathcal{J}C \cup \bigcup_{k=0}^{\infty} H^{-k}(\mathcal{J}A).$$
1.4. THE CORE RESULTS

Figure 1.4. A subset of \( \mathcal{J} \).

\( \mathcal{J} \) is an open subset of \( \mathcal{J} \) in the subspace topology. The action of \( H \) in a neighborhood of \( \mathcal{J} \) is the 10-fold covering of a quasi-horseshoe.

Remarks:
(i) A Cantor band is a space homeomorphic to the product of a Cantor set and an open interval. If you look at Figure 1.4 you can see a lot of Cantor bands.
(ii) We discuss the one sided shift in §6.2.
(iii) In §6.5 we define what we mean by a quasi-horseshoe. Such maps are pretty close to the Smale horseshoe, which we describe in §6.4. When we prove theorem 1.6 we will explain what we mean by a 10-fold covering.

Corollary 1.7. Conjecture 1.1 holds for \( N = 5 \).

Proof: \( \mathcal{P} \) inherits a smooth structure from its inclusion into \( \mathbb{R}^2 \). Recall that a smooth map is regular at \( p \) if it is a local diffeomorphism at \( p \). Almost every point in \( \mathcal{P} \) has a well defined \( H \)-orbit in which every power of \( H \) is regular at \( p \). (The
set of points which do not have this property is contained in a countable union of algebraic curves.) Call such points \textit{totally regular}.

Let \(X = J - JC\) and let \(Y\) denote the set of totally regular points in \(X\). Since almost every point is totally regular, and since \(JC\) has measure 0, it suffices to prove that \(Y\) has measure 0.

Call an open disk \(\Delta\) \textit{clean} if there is some \(k\) such that the restriction \(H^k|\Delta\) is a diffeomorphism and if \(H^k(\Delta \cap X) \subset JA\). In this case \(\Delta \cap X\) is the image of a subset of \(JA\) under a diffeomorphism, and hence has measure 0. If follows from Theorem 1.6 that \(Y\) is covered by a countable union of clean disks. Hence \(Y\) is the countable union of sets of measure 0. Hence \(Y\) has measure 0. \(\square\)

1.5. Deeper Structure

For most of the results above, and all the results in this section, we will instead use the map \(H = BHB^{-1}\), where

\[
(1.2) \quad B(x, y) = (b(x), b(y)), \quad b(t) = \phi^3 \left( \frac{\phi + t}{-1 + \phi t} \right)
\]

Here \(\phi = (1 + \sqrt{5})/2\) is the golden ratio.

After we change coordinates, we will replace the space \(R^2\) by the more global space \(M\) obtained by blowing up \((R \cup \infty)^2\) at 3 specially chosen points. The space \(M\) is a nice compact moduli space of projective classes of labeled pentagons – not necessarily in general position. It turns out that there is an order 10 group \(\Gamma\) of birational diffeomorphisms of \(M\) corresponding to dihedral relabelings of the pentagons. Beautifully, there is a fundamental domain for \(\Gamma\) acting on \(M\) that is obtained by blowing up one corner of a particular Euclidean triangle. This fundamental domain turns out to be extremely useful to us.

Another advantage of using \(H\) and \(M\) is that the fixed points of \(H\) in \(M\) have a particularly nice form. One fixed point is \((\infty, \infty)\), corresponding to the regular class. Another fixed point is \((0, 0)\), corresponding to the star regular class. Finally, there are 5 additional fixed points corresponding to the \(\Gamma\) orbit of \((1, 1)\). These 5 points represent various labelings of the star convex pentagon shown in Figure 7.4. I discovered the map \(B\) by trying to move the fixed points of \(H\) to the nicest possible locations, and then the triangular fundamental domain turned out to be a happy surprise.

Figure 1.4 really shows the Julia set for \(H\), up to a linear transformation which is chosen so that the differentials of elements in \(\Gamma\) act isometrically at \((0, 0)\).

We let \(\lozenge J\) denote \(^3\) the closure of the set of points in \(M\) which have well defined \(H\)-orbits but which do not converge to \((\infty, \infty)\). We think of \(\lozenge J\) as a kind of completion of \(J\).

We say that a \textit{cone point} is a point in \(\lozenge J\) having arbitrarily small neighborhoods which intersect \(\lozenge J\) in the cone on a Cantor set. Intuitively, the cone points are

\(^1\)A pentagon is \textit{star regular} if the relabeling relabeling \((1, 2, 3, 4, 5) \rightarrow (1, 3, 5, 2, 4)\) makes it regular, and \textit{star convex} if this relabeling makes it convex.

\(^2\)This linear transformation makes the geometric picture as nice as possible, but we don’t use it in our analysis because it is defined over a fairly high degree number field.

\(^3\)We put the symbol \(\lozenge\) in front of subsets of the Julia set in \(M\) to avoid notational clashes what the many other objects that get letter names.
where the Cantor bands pinch down to single points. See Figure 1.4. Here is a general structural result.

**Theorem 1.8.** $\cup J$ is the union of a Cantor set $\cup JC$, a countable collection of Cantor bands, and a countable collection of cone points.

Here $\cup JC = B(JC)$, where $JC$ is as in Theorem 1.5.

Now we describe structures in $\cup J$ which elaborate the ones from Theorems 1.5 and 1.6. Consider the following infinite graph. One starts with the finite graph shown on the left side of Figure 1.5 below and then puts the same graph inside each of the 6 shaded pentagonal “holes”. This produces a more complicated graph with 36 pentagonal holes, shown on the right hand side of Figure 1.5. One then repeats indefinitely. Call the limiting “graph” $G_\infty$. This space is a variant of the Sierpinski triangle.

**Figure 1.5.** The seed for $G_\infty$ and the second step in the construction.

**Theorem 1.9.** $\cup J$ contains a subset $\cup G$ which is homeomorphic to $G_\infty$.

**Theorem 1.10.** $\cup J$ contains a forward invariant subset $\cup S$ which, when blown up at all its cone points, is homeomorphic to the connected 5-fold cover of the 2-adic solenoid.

**Remarks:**
(i) $\cup JC$ is the subset of $\cup G$ comprised of the nested intersections of pentagonal holes in the iterated construction.
(ii) In local coordinates, $\cup G$ is a compact subset of $R^2$. The filled-in version $\text{Fill} \cup G$, i.e. the complement of the unbounded component of $R^2 - \cup G$, is a “solid pentagon” whose 5 “vertices” are the orbit $\Gamma(1,1)$. The center of $\cup G$ is $(0,0)$.
(iii) The connected 5-fold cover of the 2-adic solenoid is the quotient

$$(R \times \mathbb{Z}_2)/\sim, \quad (x,y) \sim (x + 5n, y + 5n), \quad n \in \mathbb{Z}.$$
Here \( \mathbb{Z}_2 \) is the topological group of 2-adic integers. We will discuss this space in more detail in §6.6.

(iv) \( \Diamond S \) is the closure of the union of the maximal \( C^1 \) arcs of \( \Diamond J \) which intersect the Cantor band \( \Diamond JA = B(JA) \). Here \( JA \) is the Cantor band from Theorem 1.6.

(v) Our last picture in the monograph, Figure 20.7, is the culmination of all our analysis. It shows a detailed schematic picture of \( \Diamond G \) and \( \Diamond S \) sitting inside \( M \). We get the more precise result that \( \Diamond S = (\Diamond J - \text{Fill}\Diamond G) \cup \Gamma(1,1) \).

All this structure contributes to our final result:

**Theorem 1.11.** The Julia set \( \Diamond J \) is path connected.

Note that \( \Diamond J \) is obviously not locally connected, on account of all the Cantor bands. So, the connectivity comes about in a complicated way.

### 1.6. A Few Corollaries

The Cantor set \( JC \) from Theorem 1.5 is contained entirely within the set of star convex classes, and the one-sided shift has periodic points of all orders. Hence \( H \) has periodic points of all orders, and we can even find such periodic points where every point in the orbit represents a star convex pentagon.

For the interested dynamics expert, I will sketch a proof in §15.6 that \( H \) is not post-critically finite. That is, the forward images of the set where \( dH \) is singular cannot be contained in a finite union of algebraic curves. What happens is that the singular set gets mapped transversely across the quasi-horseshoe and then it gets wrapped around like crazy, making it impossible for the forward image to be contained in a finite union of algebraic curves. If \( H \) were post-critically finite, there would be additional tools available to investigate \( H \), as in [N], so the result here rules out one possible shortcut to the analysis of \( H \).

Also for the dynamics expert, I will sketch a proof in §15.7 that \( H \) is not rationally conjugate to a one-dimensional rational map. That is, there is no pair \((f, h)\) where \( f : \mathbb{R}^2 \to \mathbb{R} \) and \( h : \mathbb{R} \to \mathbb{R} \) are rational and \( fH = hf \). This situation is impossible because some fibers of \( f \) would either cross the horseshoe transversely or run along the leaves of the solenoid. Either case leads to a contradiction. The lack of a rational semi-conjugacy rules out another possible shortcut to the analysis of \( H \).

### 1.7. Sketch of the Proofs

Theorem 1.2 has an elementary algebraic geometry proof. We count roots of an associated pair of polynomials using Bezout’s theorem.

Theorem 1.3 is a direct calculation once we identify the increasing quantity. The increasing quantity turns out to be \( E_5O_5 \), the simplest of the pentagram map invariants. See §5.3.

For the remaining results, we partition \( \mathcal{P} \) into a finite union of polygonal pieces on which the action on the pieces is simple enough to analyze. Our partition will have roughly the following structure.

- Some pieces will map into \( C \) after finitely many steps.
- Some pieces will map over themselves in an expanding way. This will give rise to the Cantor set \( JC \) from Theorem 1.5.
• Some pieces will map over themselves in a hyperbolic way. Roughly, they will be stretched in one direction and contracted in another. This will give rise to the quasi-horseshoe \( JA \) from Theorem 1.6.

The proofs of the results mentioned in §1.5 build on the properties of \( JC \) and \( JA \) and also make use of our partition.

The novel part of our approach is how we rigorously prove that pieces in the partition move as we think that they do. Essentially, we boil down every step to verifying that the image of some (solid) polygonal subset of \( P \) under \( f \) is contained in, or disjoint from, another (solid) polygonal subset. We reduce both questions to statements that certain finite collections of polynomials are positive (or non-negative) on certain finite collections of polygons. We then use a divide-and-conquer algorithm to establish this positivity. We call our algorithm the method of positive dominance. We explain it in §11.

Everything involved in our construction is defined over the ring \( \mathbb{Z}[1/2, \sqrt{5}, \sqrt{13}] \), and so all our calculations are exact integer calculations. There are no round-off errors. Our results sometimes require the analysis of polynomials having total degree about 45 and coefficients whose integer components are about 50 digits long. To handle the enormous polynomials we get, we implement our calculations in Java, using the BigInteger class. The BigInteger class allows one to do arithmetic involving integers which are thousands of digits long. Our calculations don’t come anywhere near the size limit imposed by the hardware of modern computers.

1.8. Some Comparisons

In spite of the complexity of the equation for the projective heat map, the results we get for this real variable map on \( \mathbb{R}^2 \) are almost comparable in detail to the kinds one sees in one dimensional complex dynamics. See [Mil] for an introduction to that vast subject. For instance, our Theorem 1.9 is similar in spirit to the combinatorial models of Julia sets called Hubbard trees. In §6.1 we will discuss Julia sets of one dimensional rational maps and compare them to our \( J \).

The projective heat map is defined over almost any field, and in particular makes sense on the complexified version of \( P \). Thus, one could consider the projective heat map in the context of 2-variable complex dynamics. There is a large literature on rational maps on \( \mathbb{C}^2 \) or on other complex surfaces. For instance, the paper [BLS] is one of a long series of papers written by the authors on the case of polynomial maps. The projective heat map is not a polynomial map, so works like [BLS] would probably not help with the proofs with the results above, but they might be a beacon for future research.

Our quasi-horseshoe result is akin to results about the Hénon map:

\[
F_{a,b}(x, y) = (1 - ax^2 + y, by).
\]

Here \( a \) and \( b \) are parameters which influence the nature of the map. The classic case is \( a = 3/10 \) and \( b = 14/10 \). In this case, there is an attracting Cantor band. As long as \( a \neq 0 \), the Hénon map is a polynomial diffeomorphism of \( \mathbb{R}^2 \). The papers [A1] and [A2] deal with computational techniques for finding subsets of parameters \((a, b)\) for which \( F_{a,b} \) acts as the Smale horseshoe on the set of bounded orbits. These techniques are somewhat like ours except that they use interval arithmetic and they deal with an entire family of maps. The paper also [BS] has a discussion of this problem.
In terms of the method of proof, one could also compare our results to those in [Tu] concerning the existence of the Lorenz attractor. This paper also uses a kind of partition approach to deal with a single dynamical system.

While not giving a comprehensive list, let me mention some other papers which get detailed dynamical pictures for real rational maps of the plane. The results in these papers, while certainly inspired by computer experiments, involve traditional proofs. One thing about all these other papers is that the formulas for the maps are considerably simpler than the formula for the projective heat map. Either the high degree nature of the projective heat map makes it too difficult to study in a traditional way – witness the difference in the amount known about the dynamics of quadratic polynomials and the dynamics of higher degree polynomials – or else a smarter author could do a better job.

The paper [BD] studies the map $\tau \circ \sigma$, where

\begin{equation}
\sigma(x, y) = \left(1 - x + \frac{x}{y}, 1 - y + \frac{y}{x}\right), \quad \tau(x, y) = (x, bx + a + 1 - y).
\end{equation}

This is a 2-parameter family of bi-rational maps depending on parameters $a$ and $b$.

The paper [BLR] deals with the single rational map coming from (a case of) the Migdal Kadanoff RG equations

\begin{equation}
(x, y) \rightarrow \left(\frac{x^2 + y^2}{x^{-2} + y^2}, \frac{x^2 + x^{-2} + 2}{x^2 + x^{-2} + y^2 + y^{-2}}\right).
\end{equation}

These equations have to do with the Ising model.

The paper [N] studies the complex dynamics of the map

\begin{equation}
(z, w) \rightarrow \left((1 - 2z/w)^2, (1 - 2/w)^2\right)
\end{equation}

and constructs a combinatorial model for the Julia set. Our Theorem 1.9 is similar in spirit to this, though we get less information. In this map, the forward orbit of the critical set – i.e. where the map is not a local diffeomorphism – is just a finite union of lines. The map is an example of a post-critically finite map, unlike the projective heat map.

The paper [HP] studies the dynamics of Newton’s method when it is used to solve two simultaneous quadratic equations. (It is somewhat difficult to extract as concrete formula as the ones given above.) This leads to a rational map on $\mathbb{C}^2$ which the authors analyze in detail.

Finally, the paper [BDM] also detailed information about the structure of some concrete rational self-maps of $\mathbb{CP}^2$.

1.9. Outline of the Monograph

This monograph comes in 3 parts.

Part 1: Context: Here I place the projective heat map in a broader context. I include some background on projective geometry and dynamical system, and analyze some dynamical systems related to the projective heat map. Some readers might like §5, which has an account of some of the main features of the pentagram map, including integrability. The sections directly relevant to the projective heat map are §3, §4.1-4.3, §5.3, §5.6, §6.2 and §6.5.
Part 2: The Core Results: In this part of the monograph, I prove Theorems 1.2, 1.3, 1.5, and 1.6. I also explain the method of positive dominance, which is used throughout this part and Part 3. This is the core material. At the end of Part 2, the proof of the pentagon case of Conjecture 1.1 is done.

Part 3: Deeper Structure: This part of the monograph studies the structure of $J$ in more detail. In particular, I prove Theorems 1.11, 1.8 and 1.9 and 1.10. (This is not the order in which the results are proved.) The arguments in this part are considerably more intricate than the ones in Part 2. To help guide the reader through the thicket of details, I have included an introductory chapter, §15, which gives detailed sketches of all the proofs.

At the end of the monograph, I have included a reference chapter in an attempt to make this monograph easier to read. §21.1 lists many of the basic definitions and objects in the monograph and points to where they are discussed. The rest of the chapter lists out important formulas, including coordinates for the vertices of the polygons in our partition. We mention §21.3 especially. This section has the formulas for all our basic maps.

1.10. Companion Program

I discovered practically everything in this monograph by writing a java computer program which implements the dynamics of the projective heat map. The computer-assisted part of my proof resides in the program. The reader can launch the proofs using the program, and survey them down to a fine level of detail.

I strongly encourage the reader of this monograph to download and use the program. The heavily documented program illustrates practically everything about our results, and also has a tutorial section which teaches the user how to run the computational tests used in our proofs. I have tried to write the monograph so that it stands on its own, but I think that the reader will have a much more satisfying experience reading the monograph while operating the program and seeing vivid illustrations of what is going on. I would say that this program relates to the monograph much like a movie relates to its screenplay.

One can download the program from

http://www.math.brown.edu/~res/Java/HEAT2.tar

The program has a README file which explains how to compile and run the program.
CHAPTER 6

Some Related Dynamical Systems

In this chapter we discuss some dynamical systems related to the projective heat map.

6.1. Julia Sets of Rational Maps

Here I will give some very basic material on the notion of a Julia set for a single (complex) variable rational map. I include this section just so that the reader can see the classical definition of a Julia set and compare it to the definition of our set $J$. This section barely scratches the surface of this vast topic. For much more information, see J. Milnor’s book [Mil].

First we'll consider the case of polynomials. Let $n \geq 2$ and let

$$P(z) = a_n z^n + \cdots + a_1 z + a_0$$

be a polynomial of degree $n$. The notation $P^m$ denotes the $m$-fold composition of $P$ with itself. Let $U_P$ denote the set of points $z$ such that

$$\lim_{m \to \infty} P^m(z) = \infty.$$

**Lemma 6.1.** $U_P$ is an open set which contains a neighborhood of $\infty$.

**Proof:** There is some constant $C$ so that $|z| > C$ implies that $|P(z)| > 2|z|$. Hence, $U_P$ contains a neighborhood of $\infty$. Also, if $z \in U_P$, then there is some $n$ such that $|P^n(z)| > C$. But then $|P^m(w)| > C$ for all $w$ sufficiently close to $z$. Hence $U_P$ contains an open neighborhood about $z$. $\Box$

Certainly $P(U_P) \subset U_P$ and $P^{-1}(U_P) \subset U_P$, and from these equations it is not hard to see that $P(U_P) = U_P$ and $P^{-1}(U_P) = U_P$. That is, $U_P$ is fully invariant. The *Julia set* $J_P$ has a simple definition: It is the boundary of $U_P$. Since $U_P$ is open, $J_P$ does not intersect $U_P$. Hence, every point in $J_P$ has a bounded orbit. Moreover, $J_P$ is bounded. Hence, $J_P$ is compact. Since $U_P$ is fully invariant, so is $J_P$. The Julia set is always a compact set without isolated points.

**Remark:** This definition of $J_P$ lines up well with our set $J = \mathcal{J}_5$, which we are calling the Julia set of the projective heat map $H$ acting on $\mathcal{P}_5$. We will see that the set $\mathcal{U}$ of projective classes $x$ such that $\{H^n(x)\}$ converges to the regular class is open. This follows from the easy fact that the regular class is an attracting fixed point of $H$. Thus, our set $\mathcal{J}$ is contained in the boundary of $\mathcal{U}$. We will see that $\mathcal{J}$ has measure zero, and in particular no interior. Hence $\mathcal{J}$ is exactly the boundary of $\mathcal{U}$. So, the definition is quite analogous to definition of the Julia set of a polynomial.
Julia sets in the quadratic case have been very well studied. In the quadratic case, one typically makes a change of variables so that $P(z) = z^2 + a$, and then considers the entire family as $a \in \mathbb{C}$ varies. The famous Mandlebrot set consists of those values $a \in \mathbb{C}$ for which $J_F$ is connected.

Now we’ll consider the case of rational maps. A single variable rational map is an expression of the form $f = P/Q$, where $P$ and $Q$ are polynomials. The simplest examples are the case when $P$ and $Q$ are linear maps. In this case $f$ is a linear fractional transformation, or Mobius transformation. This case is rather boring from the point of view of complex dynamics.

Let $S^2 = \mathbb{C} \cup \infty$ denote the Riemann sphere. Let $U \subset S^2$ be some open set. A family $F$ of holomorphic maps from $U$ into $S^2$ is a normal family if every sequence in $F$ has a subsequence which converges, uniformly on compact subsets to another holomorphic map on $U$. Here are some examples, all taking $U$ to be the open unit disk.

- If $f(z) = z/2$ then the family $\{f^n\}$ forms a normal family. Any limiting map is either a contraction by a factor of $2^{-n}$ or the 0-map.
- If $f(z) = z + 1$, then the family $\{f^n\}$ forms a normal family. Any limiting map is either a translation or the map which sends all points to $\infty$.
- If $f(z) = uz$ for some unit complex number $z$ then the family $\{f^n\}$ is a normal family. Any limiting map is a rotation.
- If $f(z) = 2z$ then $\{f^n\}$ is not a normal family. The sequence itself has no convergent subsequence.

Given a rational map $f : S^2 \to S^2$, the Fatou set is defined to be the set of points $p \in S^2$ such that there exists an open neighborhood $U$ of $p$ such that the family of iterates of $f|_U$ forms a normal family. The Julia set $J_f$ is the complement of the Fatou set.

Here is another characterization of the Julia set $J_f$. Let $z$ be a periodic point for $P$. Say that $P^m(z) = z$. Setting $f = P^m$, we call $z$ repelling if $|f'(z)| > 1$. From the discussion of normal families above, it is not hard to see that $J_f$ must contain all the repelling periodic points. It turns out that the repelling periodic points are dense in $J_f$. So, one can characterize $J_f$ as the closure of the set of repelling periodic points.

**Remark:** One can use normal families to define the Julia set of a rational map on $\mathbb{CP}^n$. For the case $n = 2$ see [BDM] for instance. We can define the projective heat map over $\mathbb{C}$, and perhaps the general notion of a Julia set for a rational map coincides with the closure of the set of projective classes of polygons which do not converge to the regular class upon iteration.

### 6.2. The One-Sided Shift

**6.2.1. Basic Definition.** The construction of the one-sided shift is based on some finite set $F$. For concreteness, we take $F = \{1, 2, 3, 4, 5, 6\}$. This is the example that comes up in connection with Theorem 1.5. The shift space is the space of all infinite sequences $\{a_i\}_{i=0}^{\infty}$ with $a_i \in F$. Call this space $\Sigma_F$. This space is naturally a metric space. Define the distance from $\{a_i\}$ to $\{b_i\}$ to be $6^{-K}$ where
$K$ is the smallest integer such that $a_K \neq b_K$. The space $Σ_F$ has diameter 1, and is homeomorphic to a Cantor set. The one-sided shift is the map $φ : Σ_F → Σ_F$ defined by: $φ(a_i) = b_i$ where $b_i = a_{i+1}$ for all $i$. In other words $φ$ just chops the zeroth term off the sequence. By definition, $φ$ expands distances by a factor of 6 and is 6-to-1.

A periodic point in $Σ_F$ is a point $p$ such that $φ^n(p) = p$ for some $n$. The smallest $n$ for which this holds is called the period of the point. This definition makes sense in any dynamical system.

**Lemma 6.2.** $Σ_F$ has periodic points of all orders, and the periodic points are dense in $Σ_F$.

**Proof:** Any sequence in $Σ_F$ which repeats after $n$ steps is a periodic point of period $n$. Hence $Σ_F$ has periodic points of all orders. Any point in $Σ_F$ can be approximated arbitrarily well by a periodic point. We just take the first $n$ terms of the point and then make it repeat endlessly. Hence the periodic points are dense. □

$Σ_F$ naturally has a measure: The measure of the set of all sequences starting $i_0, \ldots, i_k$ has measure $6^{-k}$. These sets are called cylinder sets. Equipped with this measure, $φ$ is a measure-preserving map in the sense that $φ^{-1}(S)$ and $S$ have the same measure for any measurable subset of $Σ_F$. It suffices to check this on the cylinder sets. The inverse image of a cylinder set of size $6^{-k}$ is a disjoint union of $6$ cylinder sets of size $6^{-k-1}$.

A subset $S ⊂ Σ_F$ has measure 0 if, for every $ε > 0$, we have

$$S ⊂ \bigcup C_i, \quad \sum \mu(C_i) < ε.$$  

Here $\{C_i\}$ is a countable collection of cylinder sets.

**Lemma 6.3.** Almost every point of $Σ_F$ has a dense orbit.

**Proof:** A point in $Σ_F$ has a dense orbit provided that it contains every finite sequence. We will show that the set of points which do not have a particular finite sequence has measure zero. Since the countable union of sets of measure zero has measure zero, the set of points which avoid some subsequence has measure zero. Hence, almost all points contain all finite sequences.

Let $S$ be a finite sequence. Suppose that $S$ has length $n$. Let $Σ_F(S)$ denote the set of points which do not have $S$. Suppose that $S$ has length $n$. Let $Σ'_F(S)$ denote the set of points with the following property: There is no $k$ for which $a_{kn}, \ldots, a_{(k+1)n-1}$ is $S$. In other words, elements of $Σ'_F(S)$ might contain $S$ but just not in positions starting at $0, n, 2n, 3n, \ldots$.

Obviously $Σ_F(S) ⊂ Σ'_F(S)$. So, it suffices to prove that $Σ'_F(S)$ has measure 0. Note that $Σ'_F(S)$ is contained in $6^n - 1$ cylinder sets of size $6^{-n}$. Each of these cylinder sets intersects $Σ'_F(S)$ in exactly $6^n - 1$ cylinder sets of size $6^{-2n}$, and so on. Hence, for any $k$, we see that $Σ'_F(S)$ is contained in a union of $(6^n - 1)^k$ cylinder sets of size $6^{-kn}$. The total measure of these cylinder sets is less than any desired $ε$ provided that $k$ is taken large enough. □
6.2.2. The Shift in Action. Now we show the 1-sided shift arises in a 2-dimensional context. Suppose $K$ is a closed topological disk and $K_1, \ldots, K_6 \subset D$ are 6 pairwise disjoint topological disks. Suppose we have a rational map $h$ on $\mathbb{R}^2$ such that

- For $k = 1, \ldots, 6$, we have $h(K_k) = D$ and the restriction of $h$ to some open neighborhood of $K_k$ is a diffeomorphism.
- For every point of $D - K_k$ where $h$ is defined, we have $h(p) \notin D$.

We define $\varnothing K \subset K_1 \cup \cdots \cup K_6$ to be those $p$ such that $h^n(p) \in D$ for all $n$. To analyze this set, we let $K(n)$ denote those points $p$ such that $h^n(p) \in D$. Evidently $\varnothing K = \bigcap K(n)$.

Each disk $K_j$ contains 6 smaller topological disks $K_{ij} = K_j \cap h^{-1}(K_j)$. Each of the 36 disks $K_{ij}$ contains 6 smaller disks $K_{ijk} = K_i \cap h^{-1}(K_{jk})$. And so on. From this description we see that $K(n)$ consists of $6^n$ pairwise disjoint disks, each of which contains 6 disks of $K(n+1)$.

To get more information about $\varnothing K$ we need some control over the sizes of these disks. The assumption we make is that there exists some uniform constant $C > 0$ such that the restriction of $h$ to each disk of $K(n_0)$ expands distances by at least $\eta$. From this assumption we see that any nested family of disks in the intersection $\bigcap K(n)$ shrinks to a point.

Thus, we recognize $\varnothing K$ as a Cantor set. To be more precise, we have a well-defined map $\Psi : \Sigma_F \to \varnothing K$. Each point in the shift space $\Sigma_F$ corresponds to a nested sequence of disks in this construction and $\Psi$ maps this point to the intersection of the corresponding disks. $\Psi$ is surjective and, thanks to the expanding property, continuous. Any continuous surjective map from a compact space to a subset of the plane is a homeomorphism. Hence $\Psi$ is a homeomorphism. Moreover, by construction $\Psi$ is a conjugacy: $\Psi \Phi \Psi^{-1} = h$. Thus, we recognize the action of $h$ on the set $\varnothing K$ as (conjugate to) the one-sided shift on 6 symbols.

6.2.3. A Contrived Generalization. When it comes time to consider an application of the analysis above in §12.5 we will have a slightly more contrived situation. We explain it here. We assume that we have a piecewise smooth and everywhere continuous function $\rho$ on the tangent bundle of $D$ that is comparable to the Euclidean metric in the sense that there is some uniform constant $C > 0$ such that $C^{-1}||V|| \leq \rho(V) \leq C||V||$ for all vectors $V$. We call $\rho$ a tangent bundle function.

We say that $h$ is $\eta$-expanding on a set $S$, with respect to $\rho$, if

$$\rho_{h(p)}(dh(V)) \geq \eta \rho_p(V),$$

for all $p \in S$ and all tangent vectors $V$ based at $p$. We get the same shrinking disks conclusion if we assume that there is some $n_0$ and some $\eta > 1$ so that $h$ is $\eta$-expanding on each disk of $K(n_0)$ with respect to $\rho$.

Here’s the proof. Choose some arbitrary disk $\kappa_n$ of $K(n_0 + n)$ and consider $\kappa_0 = h^n(\kappa_n)$. The disks $h(\kappa_n), \ldots, h^n(\kappa_n) = \kappa_0$ all lie in $K(n_0)$. The function $\rho$ allows us to define the lengths of paths in $D$, and by the expansion property, the length of any path in $\kappa_n$ is at most $\eta^{-n}$ as long as its image in $\kappa_0$. This combines with the comparison to the Euclidean metric to show that the Euclidean diameter of $\kappa_n$ is at most $C\eta^{-n}$ for some constant $C$ that does not depend on any choices.
6.3. The Two-Sided Shift

The two-sided shift is based on some finite set $F$. We define the space $\Theta_F$ to be the set of bi-infinite sequences $\{a_i\}_{i=-\infty}^{\infty}$. The metric on $\Theta_F$ is defined in a similar way to the one on $\Sigma_F$. Here $d(\{a_i\}, \{b_i\}) = |F|^{-K}$, where $K$ is the smallest integer such that $a_k = b_k$ for all $|k| < K$. The space $\Theta_F$ is compact.

The two-sided shift is the map $f(\{a_k\}) = \{b_k\}$ where $b_k = a_{k+1}$. The map $f$ is a homeomorphism from $\Theta_F$ to itself. As with the one-sided shift, $\Theta_F$ has a natural measure with respect to which $f$ is measure preserving. $f$ has periodic points of all orders, and the periodic points are dense, and almost every point has a dense orbit. The proofs are essentially the same as for the one-sided shift. When we look at the Smale Horseshoe, we will see how the two-sided shift arises naturally in the context of a planar diffeomorphism.

6.4. The Smale Horseshoe

Let $X = [0, 1] \times [-1/2, 3/2]$. The Smale horseshoe is a smooth map $f : X \to X$ which has the topological features shown in Figure 6.1.

![Figure 6.1. The Smale horseshoe](image)

The map also has the following additional properties.

- $f : X \to f(X)$ is a diffeomorphism.
- $f$ is a contraction on $X_1$ and on $X_7$. In particular, $f$ has an attracting fixed point $p_\infty \in X_7$.
- $f$ is an affine map on $X_3$ and on $X_5$. Here the derivative $df$ is a diagonal matrix of the form
  \[
  \begin{bmatrix}
  \lambda^{-1} & 0 \\
  0 & \lambda
  \end{bmatrix}
  \]
  $\lambda > 1$.

Note that $f^3(p) \in X_7$ when $p \in X - X_3 - X_5$, and then for such points $f^n(p) \to p_\infty$. So, the only points not attracted to $p_\infty$ under iteration are the points $p$ such that $f^n(p) \in X_3 \cup X_5$ for all $n$. Let $A$ denote the set of these points.
Let $X' = X_3 \cup X_5$. The set of points $p \in X'$ such that $f(p) \in X'$ is the union of 4 rectangles. These rectangles have the form $X_{ij} = X_i \cap f^{-1}(X_j)$ for $i, j \in \{3, 5\}$. These 4 rectangles are shown in Figure 6.2.

The pattern continues, the set $p \in X$ such that $f^k(p) \in X'$ for $k = 0, \ldots, n - 1$ consists of $2^n$ rectangles, indexed by all sequences of length $n$ in $F = \{3, 5\}$. The intersection $A$ of all these sets is $\Sigma_F \times [0,1]$, the product of a Cantor set and an interval. The action of $f$ on $X_a$ maps horizontal line segments to horizontal line segments, and the action on these line segments is the 1 sided shift on $\Sigma_F$.

The inverse map $f^{-1}$ is defined on $Y_3 = f(X_3)$ and $Y_5 = f(X_5)$. (These are the shaded rectangles on the right side of Figure 6.1.) The same analysis shows that the set $B$ of points $p \in Y_3 \cup Y_5$ such that $f^{-n}(p) \in Y_3 \cup Y_5$ for all $n$ is the product $[0,1] \times \Sigma_F$, and the action of $f^{-1}$ on the vertical line segments of $B$ is the 1-sided shift. $B$ is just $A$ turned sideways. The set $B$ is the attracting set for the restriction of $f$ to $A$. For every $p \in A$, the forward orbit $\{f^n(p)\}$ accumulates on $A \cap B$. Likewise, $A \cap B$ is the attracting set for the restriction of $f^{-1}$ to $B$. Figure 6.3 hints at the nature of the set $A \cap B$.

Each point $p \in A \cap B$ has two infinite sequences attached to it. The sequence $s_A$ describes the position of $p$ in $A$ and the sequence $s_B$ describes the sequence in $B$. form the bi-infinite sequence $s_B.s_A$, where $s_B$ is written right to left and $s_A$ is written left to right. The decimal point indicates that term 0 of this bi-infinite sequence is the first term in $s_A$. The map $f$ moves the decimal point one unit to the right and the map $f^{-1}$ moves the decimal point one unit to the left. Thus, the restriction of $f$ to $A \cap B$ is conjugate to the 2-sided shift.

### 6.5. Quasi Horseshoe Maps

Here I describe the kind of map that appears in Theorem 1.6, something I call a quasi-horseshoe. The notion of a quasi-horseshoe is a relaxation of the notion of
6.5. QUASI HORSESHOE MAPS

Figure 6.3. A hint of the set $A \cap B$

a Smale horseshoe. The main point of making the relaxation is that it is easier to establish the existence of a quasi-horseshoe than it is to establish the existence of a Smale horseshoe. I do not know if my definition of a quasi-horseshoe arises in the literature.

6.5.1. Adapted Quadrilaterals. Let $\vee \subset \mathbb{R}^2$ denote the standard light cone. Here $\vee$ is the set of vectors in $\mathbb{R}^2$ whose slope exceeds 1 in absolute value. We say that a curve is timelike if all of its chords lie in $\vee$. (A chord of the curve is a vector pointing from one point on the curve to another.) We say that a curve is spacelike if reflection in the diagonal maps it to a timelike curve. These are the usual definitions.

We say that an adapted quadrilateral (or quad for short) is a piecewise smooth embedded loop with 4 distinguished vertices we call corners, such that one pair of opposite sides is timelike and the other pair is spacelike. Here a side is an arc of the quad connecting two consecutive corners. We say that two quads are interlaced if the following holds.

- Each spacelike edge of one quad crosses each timelike edge of the other.
- The spacelike edges of one quad are disjoint from the spacelike edges of the other.
- The timelike edges of one quad are disjoint from the timelike edges of the other.

Figure 6.4 shows a quad, two interlaced quads, and one quad interlacing four others.

6.5.2. Interlacing and Quasi-Hyperbolicity. Let $P_0$ be a quad and let $P_1, \ldots, P_k$ be a finite list of quads. We say that a continuous map $F : P_0 \to \mathbb{R}^2$ interlaces $P_1, \ldots, P_k$ if

1. $F(P_0) \subset P'$, where $P'$ is a quad which interlaces $P_1, \ldots, P_k$. 
(2) $F$ maps one of the spacelike edges of $P_0$ above $P_1, \ldots, P_k$ and one below. In each case, we mean that some horizontal line separates the relevant sets.

Figure 6.5 shows what we mean.

Now we add some more structure. Given $\lambda > 1$, we call $F$ \emph{\lambda-quasi-hyperbolic} on $P_0$ if
- $F$ is a local diffeomorphism in a neighborhood of $P_0$.
- $dF(\mathcal{V}) \subset \mathcal{V}$ strictly at all points in a neighborhood of $P_0$.
- $\|dF(V)\| \geq \lambda \|V\|$ for all $V \in \mathcal{V}$.

We call $\lambda$ the \emph{stretch factor}. When the choice of $\lambda$ is not important, we simply say that $F$ is \emph{quasi-hyperbolic}.

\textbf{Remark:} Being quasi-hyperbolic is weaker than what people usually mean by \emph{hyperbolic}. In the hyperbolic case, one would want $dF$ to have one eigenvalue less than 1 and one eigenvalue greater than 1.

\textbf{6.5.3. A Preliminary Definition.} I hope the reader will forgive the upcoming terminology. I first want to describe a map which has many of the features of a
quasi-horseshoe but lacks one complicating definition. Most of the analysis we care about does not depend on the extra definition.

Let \( \Omega = P_1 \cup \cdots \cup P_k \). Here \( P_1, \ldots, P_k \) are pairwise disjoint quads. We say that a quasi quasi-horseshoe is a map \( F : \Omega \to \mathbb{R}^2 \) with the following properties.

- \( F(P_j) \) interlaces \( P_1, \ldots, P_k \) for each \( j \).
- The restriction of \( F \) to each \( P_j \) is \( \lambda \) quasi-hyperbolic for some \( \lambda > k \) that does not depend on the index \( j \).

Let \( A \subset \Omega \) be the set of points \( x \) such that \( F^n(x) \in \Omega^o \) for all \( n = 0, 1, 2, \ldots \). Note that \( A \subset \Omega^o \). Here \( \Omega^o \) is the interior of \( \Omega \).

**Lemma 6.4.** \( A \) has (2-dimensional Lebesgue) measure 0.

**Proof:** Let \( L_1 \) denote length. Consider the family \( \mathcal{T} \) of all smooth timelike curves which intersect \( A \). Let \( M \) be the supremum of \( L_1(\gamma \cap A) \), where \( \gamma \in \mathcal{T} \). Certainly \( M < 2 \text{ diam}(\Omega) \). If \( A \) has positive measure then, by Fubini’s Theorem, we can find a vertical line which intersects \( A \) in a set of positive length. Hence \( M > 0 \).

Recall that \( k/\lambda < 1 \). Here \( \lambda \) is the stretch factor. Choose some \( \gamma_1 \in \mathcal{T} \) such that \( L_1(\gamma_1 \cap A) > (k/\lambda)M \). But then there is some index \( j \) such that

\[
L_1(\gamma_1 \cap A \cap P_j) > M/\lambda.
\]

Given the the interlacing property of \( F \), we see \( \gamma_2 = F(\gamma_1 \cap P_j) \in \mathcal{T} \). Given the \( \lambda \)-stretching property, and the forward-invariance of \( A \), we see that \( L_1(\gamma_2 \cap A) > M \). This is a contradiction. \( \square \)

Recall that a Cantor band is a space homeomorphic to the product of a Cantor set and an open interval. Note that a Cantor band is naturally a union of maximal embedded open arcs. We call these arcs the *strands* of the Cantor band.

**Lemma 6.5.** \( A \) is a Cantor band. All the strands of \( A \) are spacelike.

**Proof:** Let \( A' \) be the set of points \( x \) such that \( F^n(x) \in \Omega \) for all \( n = 1, 2, 3, \ldots \). The difference between \( A' \) and \( A \) is that \( A = A' \cap \Omega^o \). We will show that \( A' \) is homeomorphic to \([0, 1] \times Y\), where \( Y \) is a Cantor set. We get \( A \) by removing the endpoints of all the arcs of \( A' \).
Let $\Omega(0) = \Omega$ and let $\Omega(n+1) = F^{-1}(\Omega(n))$. Note that
\[
A' = \bigcap_{n=0}^{\infty} \Omega(n).
\]
The interlacing property combines with induction to show that $\Omega(n)$ consists of $k^n$ quads, and each quad of $X_{n-1}$ contains $k$ quads of $\Omega(n)$. These quads are pairwise disjoint and have their timelike sides in the timelike sides of $X_0$. Figure 6.7 shows $X_0$ and $X_1$ when $k = 2$.

\textbf{Figure 6.7.} $X_0$ and $X_1$ when $k = 2$

Let $Y_n$ denote a collection of $k^n$ pairwise disjoint segments so that each segment of $Y_n$ contains $k$ segments of $Y_{n-1}$. Choose these segments so that $Y = \bigcap Y_n$ is a Cantor set – i.e. every infinite nested intersection is a point. There is a homeomorphism $h_n$ between $\Omega(n)$ and $[0,1] \times Y_n$ which has constant speed along the spacelike edges of the quads in $\Omega(n)$.

The same argument as in the measure 0 case shows that each vertical line intersects each quad of $\Omega(n)$ in a segment of length at most $O((k/\lambda)^n)$. Hence, any infinite nested intersection of these quads is a single spacelike arc. From this property, our sequence of homeomorphisms induces a homeomorphism between $A'$ and $[0,1] \times Y$.

Finally, there is a dense set of strands in $A$ corresponding to the tops and bottoms of the nested quads. These strands are all spacelike. Hence, their limit strands are spacelike in the weak sense that their chords are never timelike. If such a limit strand $\gamma$ has a chord of slope $\pm 1$ then the image $F(\gamma)$ has a timelike chord by the quasi-hyperbolicity. Hence $\gamma$ only has spacelike chords. Hence all strands in $A$ are spacelike. □

We can get one more result without too much trouble. The basic idea behind the next result is that the quasi-hyperbolicity exaggerates a hypothetical point of non-differentiability (or non-continuous differentiability) to the point where it contradicts the spacelike nature of the arcs.

\textbf{Lemma 6.6.} The strands of $A$ are continuously differentiable.

\textbf{Proof:} As in the previous result, we work with the larger set $A'$. We will show that the strands of $A'$ are differentiable. At the endpoints of the strands of $A'$ we mean to use the one-sided definition of the derivative. We work with this larger
space because it is compact. Exactly the same argument shows that the strands are in fact continuously differentiable.

Let \( \alpha \) be the line through the origin of slope \(-1\) and let \( \delta \) be the line through the origin of slope \(1\). These are the two boundaries of the lightcone \( \lor \). Suppose some arc \( \mu \) of \( A' \) is not differentiable at some point \( x \in \mu \). Then there are sequences of points \( \{y_n\} \) and \( \{x_n\} \), all in \( \mu \), so that \( y_n \rightarrow x \) and \( z_n \rightarrow x \) and the angles between the lines \( \beta_n = (xy_n) \) and \( \gamma_n = (xz_n) \) do not converge to \(0\). We label so that the slopes of \( \alpha, \beta_n, \gamma_n, \delta \) come in order. Consider the cross ratios of the slopes:

\[
(6.4) \quad [x] = \limsup_{n \to \infty} [\alpha, \beta_n, \gamma_n, \delta].
\]

By compactness, we can choose \( x \) so that \([x] \) is as large as possible.

Suppose we apply \( F \) to this whole picture. Let \( x' = F(x) \), etc. Let \( \beta'_n \) be the line \((x'y'_n)\) and let \( \gamma'_n \) be the line \((x'z'_n)\). Recall that \( dF(\lor) \) is contained strictly inside \( \lor \). If \([x] = \infty\) then either \( \beta_n \to \alpha \) or \( \gamma_n \to \delta \). In either case, one of \( \beta'_n \) or \( \gamma'_n \) lies strictly in \( \lor \) for \( n \) large. This contradicts the spacelike nature of the curves.

Now we know that \([x] \) is finite and moreover \( \beta_n \) does not converge to \( \alpha \) and \( \gamma_n \) does not converge to \( \delta \). Hence, for \( n \) sufficiently large we have

\[
(6.5) \quad [\alpha, \beta_n, \gamma_n, \delta] = [\alpha', \beta'_n, \gamma'_n, \delta'] \leq (1 - \epsilon) [\alpha, \beta'_n, \gamma'_n, \delta],
\]

for some \( \epsilon > 0 \) independent of \( n \). This gives \([x] \leq (1 - \epsilon) [x']\), a contradiction. \( \square \)

### 6.5.4. Main Definition.

Here we motivate our main definition by pointing out one shortcoming of the definition above. Let \( F \) be a quasi quasi-horseshoe as above. Let \( B \) denote the set of accumulation points of \( F \)-orbits in \( A \). Let \( \alpha \) be a strand of \( A \). We might have a situation where \( F \) maps each \( P_j \) onto a thin neighborhood of a single timelike curve and contracts the strands of \( A \) in this neighborhood. In this case, \( B \) would intersect each strand of \( A \) in a single point. This seems rather unlike what happens with the horseshoe.

We say that \( F \) is a **quasi-horseshoe** if we can partition each \( P_j \) into a left and a right half

\[
(6.6) \quad P_j = P^1_j \cup P^2_j,
\]

so that the following things are true.

1. \( F(P_j) \) interlaces \( P^1_1, \ldots, P^1_k \) or \( P^2_1, \ldots, P^2_k \) for each \( j \).

2. Each of the options in Item 1 occurs for some index.

What we are saying is that sometimes \( F \) maps the quads over the left half and sometimes \( F \) maps the quads over the right half.

**Remark:** Unlike the Smale horseshoe, a quasi-horseshoe need not be an injective map.

Given a subset \( S \subset A \), let \( S^* \) denote the set of accumulation points of \( S \). Inductively define \( S^{(n)} = (S^{(n-1)})^* \). Say that \( S \) has infinite Cantor-Bendixson rank if \( S^{(n)} \) is nonempty for all \( n \).

**Lemma 6.7.** If \( F \) is a quasi-horseshoe, then \( B \) intersects each strand of \( A \) in a set of infinite Cantor-Bendixson rank.
Proof: Since $F$ is a local diffeomorphism and $F$ maps each strand of $A$ into a strand of $A$, the restriction of $F$ to each strand is injective.

The action of $F$ on the strands of $A$ is conjugate to the one-sided shift on $k$ symbols. In particular, this strand-action has dense orbits. Hence $B$ intersects a dense set of strands of $A$. Recall that $A'$ is the compact set obtained by adjoining the endpoints of each strand of $A$. Since $B$ is closed and $A'$ is compact, we see that $B$ intersects every strand of $A'$. Since $F(A') \subset A$ we see, finally, that $B$ intersects every strand of $A$.

But now we can apply $F$ and use the definition above to say that $B$ intersects each strand of $A$ in two points, one on the left and one on the right. Now we can iterate. Since the restriction of $F$ to each strand is injective, we can say that $B$ intersects $A$ in at least $2^n$ points for every $n$. Hence $B^*$ intersects every strand of $A$. But now we can repeat the same argument as above to show that $B^{**}$ intersects every strand of $A$. And so on. □

Remark: One can cook up examples where $B$ intersects the strands of $A$ in sets which are not Cantor sets.

6.6. The 2-adic Solenoid

A 2-adic integer is an infinite sequence of the form $a_1, a_2, a_3, \ldots$ where $a_j \in \mathbb{Z}/2^j$ and

$$a_{j+1} \equiv a_j \mod 2^j \quad \forall j.$$  

The set $\mathbb{Z}_2$ of such sequences forms a ring. One does coordinatewise addition and multiplication in the corresponding finite rings and the compatibility given by equation (6.7) makes this well defined. We only care about the addition and not the multiplication in what we say below, though people interested in $p$-adic dynamics generally care about both.

The set $\mathbb{Z}_2$ also has a natural metric, the 2-adic metric. The distance between two points $\{a_k\}$ and $\{b_k\}$ is $2^{-N}$ where $N$ is the smallest integer where the sequences disagree. This makes $\mathbb{Z}_2$ into a metric ring. There is a natural inclusion of $\mathbb{N}$, the natural numbers, into $\mathbb{Z}_2$. The point $n$ maps to the sequence $\{a_j\}$ where $a_j$ is the reduction of $n \mod 2^j$. We simply think of $\mathbb{N}$ as a subset of $\mathbb{Z}_2$. The set $\mathbb{N}$ consists of those sequences which are eventually constant. Clearly $\mathbb{N}$ is dense in $\mathbb{Z}_2$. Hence $\mathbb{Z}_2$ is the completion of $\mathbb{N}$ with respect to 2-adic metric. With the 2-adic metric, the space $\mathbb{Z}_2$ is homeomorphic to a Cantor set.

The 2-adic odometer is the map $x \rightarrow x+1$ acting on $\mathbb{Z}_2$. The orbit of 0 under this map is $\mathbb{N}$. Hence 0 has a dense orbit. The orbit of any other point under the 2-adic odometer is just a translate of $\mathbb{N}$. Hence, every point has a dense orbit.

The 2-adic solenoid is the mapping cylinder for the 2-adic odometer. That is, we take the space $[0,1] \times \mathbb{Z}_2$ and we identify the points $(0,y)$ with $(1,y+1)$. More canonically we take $\mathbb{R} \times \mathbb{Z}_2$ and we take quotient out by the $\mathbb{Z}$ action generated by the map $(x,y) \rightarrow (x+1,y+1)$.

The 2-adic solenoid is partitioned into infinite curves. Each infinite curve is dense in the solenoid. The $M$-fold cyclic cover of the 2-adic solenoid is the quotient
$R \times \mathbb{Z}_2$ by the subgroup $M \mathbb{Z}$. Here $M \mathbb{Z}$ is generated by the map

$$(x, y) \rightarrow (x + M, y + M).$$

Theorem 1.10 mentions the 5-fold cover.

Finally, we mention that we can think of the 2-adic solenoid as a fibration over the circle. The fibration map is just the projection onto the first coordinate. The fibers (a.k.a. cross-sections) are copies of $\mathbb{Z}_2$ and hence homeomorphic to Cantor sets. Thus, one can picture the 2-adic solenoid as a kind of twisted version of the product of a circle and a Cantor set.

### 6.7. The BJK Continuum

Here is another point of view on the 2-adic integers. One can alternatively represent a 2-adic integer as a formal infinite series

$$(6.8) \quad b_0 + 2b_1 + 4b_2 + \ldots$$

Here the $b_i$ are either 0 or 1. We get a 2-adic sequence by taking the partial sums $a_1 = b_0$ and $a_2 = b_0 + 2b_1$, and so on. This gives us an identification of $\mathbb{Z}_2$ with the set of infinite binary sequences. With this identification, addition is done the way one learns it in elementary school, except that one works in base 2 and carries to the right. We call this identification the series identification.

**Remark:** Here is how to see the homeomorphism between $\mathbb{Z}_2$ and the middle third Cantor set. We take the series in equation (6.8) and map it to the base 3 expansion

$$(2b_0), (2b_1), (2b_2), \ldots$$

This expansion contains only 0s and 2s and hence defines a point in the middle third Cantor set.

We identify $\mathbb{Z}_2$ with the set of infinite binary strings, using the series identification. We consider the following two involutions on $\mathbb{Z}_2$. The involution $I_0$ reverses all the digits after the first 1 is encountered. For instance,

$$I_0(0010101110 \ldots) = 0011010001 \ldots$$

The involution $I_1$ reverses all the digits.

We can use these involutions to construct the *Brouwer-Janiszewski-Knaster continuum*. This space is obtained as follows. We start with the $[0, 1] \times \mathbb{Z}_2$ and we identify the points $(j, y)$ to $(j, I_j(y))$ for $j = 0, 1$. Figure 6.8 shows the first few steps in a concrete construction of the BJK continuum.

**Figure 6.8. Constructing the BJK Continuum**

Since $\mathbb{Z}_2$ is a ring, we can interpret the involutions $I_0$ and $I_1$ algebraically. We do this by working formally with the series representation of points in $\mathbb{Z}_2$. 
Lemma 6.8. $I_1(x) = -1 - x$.

Proof: Using the rule for addition mentioned above, we have

$$(1, 0, 0, 0, \ldots) + (1, 1, 1, 1, \ldots) = 0.$$ 

Here is what is going on: When we perform the addition, we get $1 + 1 = 0$ in the leftmost position and then we carry the 1 to the right. This gives 0 in the second position and again we carry the 1 to the right. And so on. The remainder of 1 sweeps rightward, leaving all 0s in its wake. What this means is that $(1, 1, 1, 1, \ldots)$ is the series representation for $-1$. Now, given $x = (b_0, b_1, b_2, \ldots) \in \mathbb{Z}_2$, we have

$$I(x) = (1 - b_0, 1 - b_1, 1 - b_2, \ldots).$$

Adding coordinatewise, we get $x + I(x) = (1, 1, 1, \ldots) = -1$. □

Lemma 6.9. $I_0(x) = -x$.

Proof: Taking the series representation for $x$ as in equation (6.8), we have

$$x + I_0(x) = 1(0) + 2(0) + \cdots + 2^k(0) + 2^{k+1}(1 + 1) + 2^{k+2}(1) + 2^{k+3}(1) + \cdots = 0.$$ 

Here $k + 1$ is the index of the first 1 in the series representation of $x$. What is going on here when we add is that we first produce $k$ consecutive 0s and then we have $1 + 1$, which produces another 0 with a remainder of 1. Following this, the remainder sweeps rightward, leaving an infinite string of 0s in its wake. □

We use our algebraic knowledge of the maps $I_0$ and $I_1$ to define the BJK continuum in a different way. We let $I_0$ and $I_1$ act on $\mathbb{R} \times \mathbb{Z}_2$ by the diagonal action $I_0(x, y) = (-x, -y)$ and $I_1(x, y) = (-1 - x, -1 - y)$. Then $\langle I_1, I_2 \rangle$ is the infinite dihedral group and this group acts on $\mathbb{R} \times \mathbb{Z}^2$. The quotient space is the BJK continuum.

Observe that

$$(I_0I_1)^M(x, y) = (x + M, y + M).$$ 

In this way we recognize the 2-adic solenoid as the double cover of the BJK continuum. Moreover, the $M$-fold cover of the 2-adic solenoid is an order $2M$ dihedral cover of the BJK continuum. This is how we will recognize the space from Theorem 1.10 in the case $M = 5$. 

60 6. SOME RELATED DYNAMICAL SYSTEMS