A Classification of Chaos for Linear Fractional Transformations of the Backward Shift on l^p

A Senior Comprehensive Project
by
Lucas B. Hawk
Allegheny College
Meadville, PA

April 28, 2017

Submitted to the Department of Mathematics in partial fulfillment of the requirements for the degree of Bachelor of Science.

Project Advisor: Dr. Brent Carswell
Second Reader: Dr. Rachel Weir

I hereby recognize and pledge to fulfill my responsibilities, as defined in the Honor Code, and to maintain the integrity of both myself and the College community as a whole.

Pledge:

Lucas B. Hawk

Abstract

We overview and expand on the results of a paper by Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris titled *Chaos for Linear Fractional Transformations of Shifts* [7]. Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris provide a characterization of chaos for $\varphi(B)$ on Banach sequence spaces, where φ is a linear fractional transformation and B is the backward shift operator. Before exploring this characterization, we first review some basic concepts from topology, complex analysis, and linear algebra. Furthermore, we explore more advanced topics such as conformal mappings in $\mathbb{C}$, l^p and Banach spaces, and chaos.

Contents

1 Introduction

The purpose of this project is to explain the results in *Chaos for Linear Fractional Transformations of Shifts* by Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris. We provide an overview of several topics which must be covered before turning to the main results. The goal of this is to allow a student with only a background in calculus and perhaps some real analysis to fully understand the advanced topics covered by Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris.

Before beginning this project, my knowledge of the topics required to understand this paper was limited only to that of some terms from linear algebra (e.g. eigenvalues), and basic notions of metric and vector spaces. This project is a roadmap of what I learned in order to understand the paper, ending with a dissection of Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris's results which make them much more clear to a reader who has a similar background to mine. For such a reader, very few words in the title of my thesis make sense. The purpose of all of the work before the main results section is essentially to allow the reader to understand these words, and ultimately understand the main results.

The first section of this paper covers some of the more general, basic topics which must be understood before moving on to more advanced topics. Specifically, we overview metric spaces, the complex plane, and a few notions from linear algebra such as eigenvalues. References [1] and [3] were used for the subsection on metric spaces, [2] was used for the complex plane, and [13] was used for the subsection titled Eigenvalues.

The next section of the paper deals with analytic functions of the complex variable and the notion of a conformal mapping. In this section, the reader will become comfortable with complex-valued functions, and will become familiar with linear fractional transformations and the backward shift operator. For this section, [2] is used as the main source.

After that we introduce Banach spaces, a special type of vector space, and a family of Banach spaces called l^p spaces. l^p is the space we work in for our main results, and the necessary background of l^p is provided in this section. [9] is used extensively for this section.

The final section before the main results is on chaos. This section is less about the analysis of chaos, and is focused more-so on simply understanding the notion of chaos, and what chaotic behavior looks like. This is because our main results do not require a strong ability to, say, prove an operator is chaotic — Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris provide a new way to compute chaos for linear fractional transformations of shifts. Hopefully, readers will be able to fully understand this new classification by the time they reach the main results section.

2 Preliminaries

The purpose of this section of the paper is to review and expand on some of the more basic mathematical topics which are important in Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris's paper. We review topics in metric spaces, complex analysis, and a few things from linear algebra. These topics are the basic building blocks we need to understand some of the more advanced concepts in the paper.

2.1 Metric Spaces

Here we provide an overview of some of the elementary topological concepts about metric spaces. We specifically talk about metric and vector spaces, norms, and the completion of metric spaces. Note that many definitions, theorems, etc. come from Gerald Edgar's *Measure, Topology, and Fractal Geometry*[1] and Kaplansky's *Set Theory and Metric Spaces*[3].

Definition 2.1. A **metric space** is a set S together with a function $d : S \times S \to [0, \infty)$ satisfying the following:

(1) $d(x, y) = 0 \Leftrightarrow x = y$

(2) $d(x, y) \geq 0$ for all $x, y \in X$

(3) $d(x, y) = d(y, x)$

(4) $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle inequality)

The nonnegative real number $d(x, y)$ is called the *distance* between x and y, while the function d itself is known as the *metric* of the set S. A metric space is written as (S, d), but oftentimes the metric is implied and the space is simply referred to as S. Note that the last property, the triangle inequality, is a very important property as it is used very often in proofs.

Example 2.2. The set of real numbers $\mathbb{R}$, with $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$d(x, y) = |x - y|$$

is a metric space. This is the usual metric used with $\mathbb{R}$.

The complex plane $\mathbb{C}$ has a similar usual metric:

Example 2.3. The complex numbers $\mathbb{C}$ with $d : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ defined by

$$d(z, w) = |z - w|,$$

where $|z|$ is the modulus of z, is a metric space.

Generally, algebraic operations are not defined on a metric space, just a distance function. Meanwhile, a vector space provides the operations of vector addition and scalar multiplication, but without a notion of distance. We can combine a vector space with a *norm*, though, to create a normed vector space — note that all normed vector spaces are also metric spaces.

Definition 2.4. A **norm** on a vector space V is a function $||x|| : V \to$

$[0, \infty) \subset \mathbb{R}$ which satisfies the following:

(1) $||x|| \geq 0$ and $||x|| = 0$ if and only if $x = 0$;

(2) $||kx|| = |k| \, ||x||$ (scaling property);

(3) $||x + y|| \leq ||x|| + ||y||$.

A vector space V together with a norm $|| \cdot ||$ is called a **normed vector space**, and is denoted $(V, || \cdot ||)$.

The properties of a norm on a vector space are rather intuitive. Property (1) says that a vector has nonnegative length, and the length of x is 0 if and only if x is the 0-vector; property (2) states that multiplying a vector by a scalar k multiplies its length by k; finally property (3) is the triangle inequality, which is analogous to property (4) of definition 3.1.

While a norm is defined rather similarly to a metric, the two are not the same. However, we often define a metric on a normed metric space using the norm.

Proposition 2.5. *If $(X, ||\cdot||)$ is a normed vector space X, then $d : X \times X \to \mathbb{R}$, defined by $d(x, y) = ||x - y||$, is a metric on X.*

The necessary properties for d to be a metric follow immediately from properties (1) and (3) of a norm. If X is a normed vector space, we always use the metric associated with its norm, unless specifically stated otherwise.

A metric defined by a norm has all the properties of a metric discussed earlier, as well as two more — for all $x, y, z \in X$ and $k \in \mathbb{R}$

$$d(x+z, y+z) = d(x, y), \qquad d(kx, ky) = |k|d(x, y).$$

These properties are called *translation invariance* and *homogeneity*, respectively. These properties are not included in a simple metric space because they do not even make sense in that framework — recall that in a space which is only a metric space, we can not add points together or multiply them by scalars.

While there are a variety of norms which can be used on a vector space, the *Euclidean norm* is the most common for $\mathbb{R}^n$ and is perhaps the most intuitive.

Example 2.6. On $\mathbb{R}^n$, the length of a vector $x = (x_1, x_2, ..., x_n)$ is given by

$$||x|| := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

This gives the distance from the origin to the point x and is known as the *Euclidean norm.* This should be familiar as it is the "straight-line" distance between points in space and is a result of the Pythagorean theorem.

Example 2.7. On $\mathbb{C}^n$, the most common norm is given by

$$||z|| := \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} = \sqrt{z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n}$$

Note that while a metric is often derived from a norm, the existence of a metric does not imply a norm — a metric does not even necessarily need to make geometric sense. Take for example what is known as the *discrete metric*:

Example 2.8. On any set X, the discrete metric is defined as

$$d(x, y) = \begin{cases} 0 & x = y, \\ 1 & x \neq y. \end{cases}$$

The discrete metric does not satisfy the homogeneity property we briefly discussed earlier, so we know this metric was not induced by a norm.

We now review a couple of concepts about sequences in a metric space. Many of the definitions and proofs in the rest of this section are based on excerpts from a real analysis textbook[8], which we have adapted for the more general case of metric spaces.

Definition 2.9. A **sequence** is a function whose domain is a set of the form $\{n \in \mathbb{Z} : n \geq m\}$ for some $m \in \mathbb{Z}$, and whose codomain is some topological space. A sequence is denoted $(s_n)_{n=m}^{\infty}$ or just (s_n).

Example 2.10. The following are a couple of examples of sequences:

- Let $s_n = \frac{1}{n^2}, n \geq 1$. Then $(s_n) = (1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \dots)$.
- Let $s_n = \frac{1}{2}(1 + (-1)^n), n \geq 1$. Then, $(s_n) = (0, 1, 0, 1, \dots)$.

Definition 2.11. A sequence (s_n) is *nondecreasing* if $s_n \leq s_{n+1}$ for all n. A sequence (s_n) is *nonincreasing* if $s_n \geq s_{n+1}$ for all n. We call a sequence **monotonic** if it is either nonincreasing or nondecreasing.

Definition 2.12. In a normed linear space, a sequence (s_n) is **bounded** if there exists a number M for which $||s_n|| < M$ for all n.

Example 2.13. We provide a few examples of bounded and/or monotonic sequences:

- Let $s_n = \frac{1}{n^2}, n \geq 1$. Then (s_n) is both bounded and monotonic.
- Let $s_n = \frac{1}{2}(1+(-1)^n), n \geq 1$. Then (s_n) is bounded but not monotonic.
- Let $s_n = n$. Then (s_n) is monotonic but unbounded.

Theorem 2.14. *Every sequence has a monotonic subsequence.*

The formal proof of this result is far more involved than what is necessary to understand why this is true. As such, we only outline the general idea of how to construct such a subsequence in a very informal proof.

Informal proof. Let $(s_n)_{n=1}^{\infty}$ be any sequence. Define a term s_n to be **dominant** if $s_n > s_m$ for all $m > n$. Now, if (s_n) has infinitely many dominant terms, we can construct a subsequence solely of dominant terms which is nonincreasing. Similarly, if (s_n) has a finite number of dominant terms, we can construct a subsequence solely out of terms beyone the last dominant term which is nondecreasing. Thus, we can construct a monotonic subsequence for every sequence. $\square$

Definition 2.15. A sequence (s_n) in metric space (X, d) **converges** to a point s provided that for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n > N$, then $d(x_n, s) < \epsilon$. This point s is called the *limit* of the sequence. We denote the limit as $\lim(s_n) = s$.

Essentially, a sequence converges to a point s if, given an arbitrary distance, we can find a point in the sequence after which all the terms are within that distance from the limit. We do not provide a proof, but note that all bounded monotonic sequences converge. Specifically, if a sequence is nondecreasing and bounded, it will converge to the smallest number that bounds it. Similarly, if a sequence is nonincreasing and bounded, it will converge to the largest number which bounds it.

Definition 2.16. Let (X, d) be a metric space. A sequence $(x_n)_{n=1}^{\infty}$ in X is a **Cauchy sequence** if, for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$n, m \geq N \Rightarrow d(x_n, x_m) < \epsilon.$$

The definition of a Cauchy sequence is very close to that of a convergent sequence — a Cauchy sequence, though, says that beyond some point in the sequence, all the terms are close to one another. The two are closely related, and in fact convergence of a sequence implies the sequence is Cauchy.

Theorem 2.17. *Convergent sequences are Cauchy.*

The idea here is that, if a sequence converges, all the terms are eventually close to the same limit. Since the terms are all close to the same limit, they are also very close to one another. In the following proof, note the use of the triangle inequality property of metric spaces.

Proof. We show that all convergent sequences are Cauchy.
Let $s = \lim(s_n)$. Let $\epsilon > 0$. Since $\lim(s_n) = s$, there exists $N \in \mathbb{N}$ for which

$n > N \Rightarrow d(s_n, s) < \frac{\epsilon}{2}$. So suppose $m, n > N$. Then

$$\begin{aligned} d(s_n, s_m) &\leq d(s_n, s) + d(s, s_m) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon. \end{aligned}$$

Thus by definition (s_n) is Cauchy. □

We can also easily show that all Cauchy sequences are bounded. Also, since all convergent sequences are Cauchy, we can say that all convergent sequences are bounded as well. This fact can be combined with the *Bolzano-Weierstrass Theorem* which states every bounded sequence has a convergent subsequence — we will introduce Bolzano-Weierstrass more formally when we need it later.

Theorem 2.18. *Every Cauchy sequence is bounded.*

Proof. Suppose (s_n) is a Cauchy sequence. Then there exists an $N \in \mathbb{N}$ for which $||s_n - s_m|| < 1$. In particular, if $n > N$, then $||s_n - s_{N+1}|| < 1$ (since $N + 1$ is obviously greater than N). Therefore, if $n > N$, then

$$\begin{aligned} ||s_n|| &= ||s_n - s_{N+1} + s_{N+1}|| \\ &\leq ||s_n - s_{N+1}|| + ||s_{N+1}|| \\ &< 1 + ||s_{N+1}||. \end{aligned}$$

Now, let $M = \max\{||s_1||, ||s_2||, \cdots, ||s_N||, 1 + ||s_{N+1}||\}$. Note that we know M exists since the set is finite. Observe that if $1 \leq n \leq N$ then $||s_n|| \leq M$.

If $n > N$, we have $||s_n|| < 1 + ||s_{N+1}|| \leq M$. So by definition, (s_n) is bounded, specifically bounded by M. $\square$

Theorem 2.19. *If a subsequence of a Cauchy sequence converges to s, then the whole sequence converges to s.*

Proof. Let (X, d) be a metric space. Suppose (s_n) is a Cauchy sequence with subsequence (s_{n_k}) which converges to s. We show (s_n) also converges to s. Let $\epsilon > 0$ and pick $N > 0$ such that for all $n, m > N$, we have

$$d(s_n, s_m) < \frac{\epsilon}{2}.$$

Since (s_{n_k}) converges to s, we can pick $K > 0$ such that for all $n_k > K$, we have

$$d(s_{n_k}, s) < \frac{\epsilon}{2}.$$

Now let $M = \max\{N, K\}$. Fixing $n_k > M$, we have

$$\begin{aligned} d(s_n, s) &\leq d(s_n, s_{n_k}) + d(s_{n_k}, s) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon. \end{aligned}$$

Thus (s_n) converges to s as desired. $\square$

These concepts will be used later on in the section on Banach spaces in order to prove *completeness* of a metric space.

2.2 The Complex Plane

Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris's work is done entirely in the complex plane. In this section we review the complex numbers and some topological concepts of the complex plane. Most definitions and theorems are borrowed from *Fundamentals of Complex Analysis* by Saff and Snider [2].

If we picture the real numbers $\mathbb{R}$ as a one-dimensional number line, the set of complex numbers $\mathbb{C}$ can be thought of as a plane with $\mathbb{R}$ as its x-axis and the set of imaginary numbers as its y-axis. So, an element of $\mathbb{C}$ consists of both a real part and an imaginary part.

Definition 2.20. A **complex number** is an expression of the form $a + bi$ where $a, b \in \mathbb{R}$, and i is the imaginary number defined by $i^2 = -1$. Two complex numbers $a + bi$ and $c + di$ are said to be equal if and only if $a = c$ and $b = d$.

So, for a complex number $z = a+bi$, a is said to be the real part (denoted Re z) and b is the imaginary part (denoted Im z). With this notation we can write $z = \text{Re } z + i \text{ Im } z$. Note that if $b = 0$, then z is a real number, while if $a = 0$, then z is a pure imaginary number.

Recall that the complex plane $\mathbb{C}$ can be combined with a distance function d:

$$d(z, w) = |z - w|$$

where $z, w \in \mathbb{C}$ and $|z|$ is the modulus of z defined as below. Furthermore, this metric is induced by the usual norm of $\mathbb{C}$: $||z|| = \sqrt{x^2 + y^2}$ where

$z = x + iy$ for $x, y \in \mathbb{R}$. From this point on, this metric will be implied when we refer to $\mathbb{C}$.

Definition 2.21. The **modulus** or (**absolute value**) of a number $z = a+bi$ is denoted $|z|$ and is given by

$$|z| := \sqrt{a^2 + b^2}.$$

Note that $|z|$ is always a nonnegative real number, and the only complex number whose modulus is zero is the number 0 itself.

The reflection of a point $z = a+bi$ across the real axis is the point $a-bi$. The relationship between a number and its reflection plays a large role in complex analysis.

Definition 2.22. The **complex conjugate** of a number $z = a + bi$ is denoted $\bar{z}$ and is given by

$$\bar{z} := a - bi.$$

The conjugate is important due to the numerous properties regarding the possible interactions between a complex number and its conjugate.

Example 2.23. First, it is clear that $z = \bar{z}$ if and only if z is a real number. Furthermore, the conjugate of the sum/difference of two complex numbers is equal to the sum/difference of their conjugates. That is,

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 - z_2} = \overline{z_1} - \overline{z_2}.$$

Beyond these two properties are a number of other ones as well:

- $\overline{(z_1 z_2)} = \overline{z_1}\,\overline{z_2}$

  Indeed, if $z_1 = a + bi$ and $z_2 = c + di$, then

$$\begin{aligned}\overline{z_1 z_2} &= \overline{ac - bd + (ad + bc)i} \\ &= ac - bd - (ad + bc)i \\ &= ac - bd - adi - bci \\ &= (a - bi)(c - di) \\ &= \overline{z_1}\,\overline{z_2}.\end{aligned}$$

- In addition, the following can be seen:

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}},\ (z_2 \neq 0);$$

$$\begin{aligned}\text{Re } z &= \frac{z + \bar{z}}{2}; \\ \text{Im } z &= \frac{z - \bar{z}}{2i}.\end{aligned}$$

  The last two properties demonstrate that the sum of a number and its conjugate is real, while the difference is imaginary.

- The conjugate of a conjugate is, of course, the original number:

$$\overline{(\bar{z})} = z.$$

  The final two properties are specifically used in the main results of [7]:

$$|z| = |\bar{z}|, \quad z\bar{z} = |z|^2.$$

The first of these is easy to see geometrically (Figure 1).

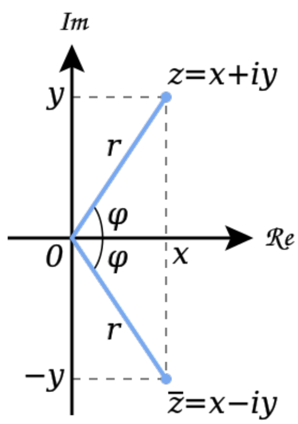

Figure 1: Complex number and its modulus[4]

The second can be proved easily:

$$z\bar{z} = (a+bi)(a-bi) = a^2 + b^2 = |z|^2.$$

For functions of a real variable, we typically work with functions defined on an *interval*, but this concept does not work for $\mathbb{C}$. We must instead define some topological concepts for $\mathbb{C}$.

Definition 2.24. The **open disk** or **circular neighborhood** of a point z_0 is the set of all points z which satisfy the inequality

$$|z - z_0| < p,$$

where p is a positive real number. This set consists of every point that lies inside the circle of radius p around the center z_0.

Example 2.25. The solution sets of the inequalities

$$|z-2|<5, \quad |z+i|<\frac{1}{2}, \quad |z|<8$$

are open disks centered at 2, $-i$, and 0 respectively.

A frequently used neighborhood is the *open unit disk*:

Definition 2.26. The **open unit disk**, denoted $\mathbb{D}$, is defined as follows:

$$\mathbb{D} := \{z : |z| < 1\}.$$

We will now define several closely related topological terms regarding sets. We start with two terms in a single definition — a two-for-one special, if you will.

Definition 2.27. For any set S, a point z_0 is called an **interior point** of S if there is some open disk centered at z_0 which is completely contained in S. If every point in S is an interior point of S, we describe S as an **open set**.

Definition 2.28. An open set S is said to be **connected** if every pair of points can be joined by a piecewise linear curve which does not leave the set. Alternatively, S is connected if for all $p, q \in S$ there exists a finite collection of line segments contained in S which join p and q

Essentially, a set is connected if it is a "single piece", geometrically speaking. A *convex set* is slightly more than this:

Definition 2.29. A set $S \in \mathbb{C}$ is said to be **convex** if

$$tp + (1-t)q \in S$$

for all $p, q \in S$ and for all $0 \leq t \leq 1$. More simply, a set S is convex if, for every pair of points in S, the line connecting the two points is contained within S.

Note that all convex open sets are connected, but a connected set is not necessarily convex. Figure 2 shows this with geometric representations of sets — (a) is both convex and connected, (b) and (c) are connected but not convex, and (d) is neither connected nor convex.

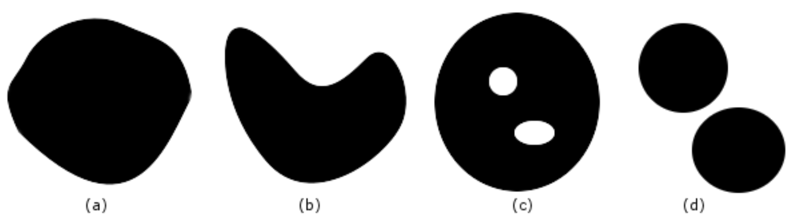

Figure 2: Convex vs. Connected Sets

Definition 2.30. A point z_0 is said to be a **boundary point** of a set S if every neighborhood of z_0 contains at least one point of S and at least one point not in S. Simply put, z_0 is a boundary point if it exists on the edge of the set. The set of all boundary points is called the **boundary** of S, and we will denote this $\partial(S)$.

Definition 2.31. A set S is said to be **closed** if it contains all of its boundary points. Equivalently, S is closed if its complement $\mathbb{C} \backslash S$ is open.

Example 2.32. Let $\overline{\mathbb{D}} = \{z : |z| \leq 1\}$. This is a closed set, as it contains its boundary $\partial(\mathbb{D}) = \{z : |z| = 1\}$. $\overline{\mathbb{D}}$ is known as the *closure of* $\mathbb{D}$ or the *closed unit disk*.

The final topological concepts we need for $\mathbb{C}$ are those of *boundedness* and *compactness*.

Definition 2.33. A set of points S is said to be **bounded** if there exists some $r \in \mathbb{R}^+$ such that $|z| < r$ for every z in S. Equivalently, S is bounded if it is contained in a neighborhood of the origin. S is **unbounded** if it is not bounded.

Definition 2.34. A set that is both closed and bounded is said to be **compact**.

Now that we have an understanding of the topology of $\mathbb{C}$, we can define a few terms regarding functions of the complex variable. Here, we will only define and demonstrate a few key concepts of complex functions which are analogous to their counterparts in real calculus. Later on, in the section on Analytic Functions and Conformal Mappings, we explore complex functions more deeply — specifically, we will define what it means for a function to be analytic, and we will describe these functions from both an algebraic and geometric standpoint.

Definition 2.35. A **function** $f : A \to B$ is a mapping that assigns each element of a set A to one (and only one) element in a set B. If f maps an element $a \in A$ to an element $b \in B$, we write $f(a) = b$. We call b the **image** of a under f. The set A is the domain of f, while the set B is the codomain of f.

Just as we can separate a complex number z into real and imaginary parts $z = x + iy$, we can write a function $f(z) = w$ as $w = u(x, y) + iv(x, y)$, with u, v denoting the real and imaginary parts of w, respectively.

Example 2.36. We can write the function $w = f(z) = z^2 + 2z$ in the form $w = u(x, y) + iv(x, y)$.
Set $z = x + iy$. Then

$$\begin{aligned} f(z) &= z^2 + 2z \\ &= (x + iy)^2 + 2(x + iy) \\ &= x^2 + y^2 + i2xy + 2x + i2y \\ &= (x^2 - y^2 + 2x) + i(2xy + 2y). \end{aligned}$$

Hence, $w = (x^2 - y^2 + 2x) + i(2xy + 2y)$ is the desired form.

This representation allows us to better represent the range of a function.

Example 2.37. Suppose $f(z) = w = x^2 + 2i$ whose domain is the closed unit disk $\overline{\mathbb{D}} = \{z : |z| \leq 1\}$.
We have that $u(x, y) = x^2$ and $v(x, y) = 2$. So, as z varies over $\overline{\mathbb{D}}$, u varies between 0 and 1, and v is constant at 2. Thus the range is the line segment from $w = 2i$ to $w = 1 + 2i$.

Definition 2.38. Let f be a function defined in the neighborhood of z_0, but not necessarily defined at the point z_0 itself. We say that the *limit* of f as z approaches z_0 is the number w_0, denoted

$$\lim_{z \to z_0} f(z) = w_0,$$

if for any $\epsilon > 0$, there exists a positive number δ such that

$$|f(z) - w_0| < \epsilon \quad \text{whenever} \quad 0 < |z - z_0| < \delta$$

Example 2.39. We show that

$$lim_{z \to i} z^2 = -1.$$

We must show that for all $\epsilon > 0$, there is a positive δ such that

$$|z^2 - (-1)| < \epsilon \quad \text{whenever} \quad 0 < |z - i| < \delta.$$

So, we express $|z^2 - (-1)|$ in terms of $|z - i|$. So,

$$z^2 - (-1) = (z - i)(z + i) = (z - i)(z - i + 2i).$$

Using the triangle inequality, we have $|z - i|\,|z - i + 2i| \leq |z - i|(|z - i| + 2)$. Now, choose $\delta > 0$ such that $\delta < \frac{\epsilon}{3}$ and $\delta < 1$. Then, if $0 < |z - i| < \delta$, we have

$$|z - i|(|z - i| + 2) < \frac{\epsilon}{3}(1 + 2) = \epsilon.$$

Definition 2.40. Let f be a function defined in a neighborhood of z_0. We say f is **continuous** at z_0 if

$$\lim_{z \to z_0} f(z) = f(z_0).$$

We say the function is continuous on a set S if f is continuous at every point in S.

We have demonstrated all the preliminary concepts regarding the complex plane. We will revisit the complex plane in Section 3 when we look at specific properties of complex functions.

2.3 Eigenvalues

Our main results rely on a theorem (which we will introduce later) called the *Eigenvalue Criterion for Chaos*. In order to understand what exactly this theorem means, we must have an understanding of what an eigenvalue is, as well as a couple of other concepts (such as the span of a set) from linear algebra. This section provides that necessary background, with [13] being the main source for these definitions.

Definition 2.41. A mapping T from a vector space V_1 to a vector space V_2, i.e. $T : V_1 \to V_2$, is a **linear transformation** if

$$T(c\vec{u} + c\vec{v}) = cT(\vec{u}) + cT(\vec{v}),$$

for all $\vec{u}, \vec{v} \in V_1$, and all $c \in \mathbb{R}$. A **linear operator** is a linear transformation which maps a vector space to itself.

Example 2.42. Let V^n be the space of all $n \times n$ matrices, and let A be a fixed $n \times n$ matrix. Then $T : V^n \to V^n$ by $T(v) = Av$ is a linear operator.

Definition 2.43. Let $T : V \to V$ be a linear operator. If there is a scalar λ and a vector $v \in V$, with $v \neq 0$, such that $T(v) = \lambda v$, then λ is called an **eigenvalue** of T. We call v an **eigenvector** corresponding to λ.

Example 2.44. Let $T : \mathbb{C}^2 \backslash \{0\} \to \mathbb{C}^2 \backslash \{0\}$ be a linear operator by

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

We find the eigenvalues of T. So,

$$\begin{aligned}
&T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&\text{iff } \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}. \\
&\text{iff } -x_2 = \lambda x_1 \quad \text{and} \quad x_1 = \lambda x_2. \\
&\text{iff } x_1, \text{ we have } -x_2 = \lambda^2 x_2, \\
&\text{iff } -1 = \lambda^2, \\
&\text{iff } \lambda = \pm i.
\end{aligned}$$

So, T has eigenvalues of $\lambda = \pm i$. We can see that the vector $\begin{pmatrix} x \\ -ix \end{pmatrix}$ is an eigenvector corresponding to i, and $\begin{pmatrix} x \\ ix \end{pmatrix}$ is an eigenvector corresponding to $-i$. Indeed,

$$T\begin{pmatrix} x \\ -ix \end{pmatrix} = \begin{pmatrix} ix \\ x \end{pmatrix} = \begin{pmatrix} ix \\ -i^2x \end{pmatrix} = i\begin{pmatrix} x \\ -ix \end{pmatrix};$$

$$T\begin{pmatrix} x \\ ix \end{pmatrix} = \begin{pmatrix} -ix \\ x \end{pmatrix} = \begin{pmatrix} -ix \\ -i^2x \end{pmatrix} = -i\begin{pmatrix} x \\ ix \end{pmatrix}.$$

Definition 2.45. Let V be a vector space and let $v_1, v_2, \cdots, v_n \in V$. Then,

1. v is a **linear combination** of $v_1, v_2, \cdots, v_n$ if there exist scalars

$c_1, c_2, \cdots, c_n$ such that $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$.

2. The **span** of $\{v_z, v_2, \cdots, v_n\}$ is the set

$$\text{Sp}\{v_1, v_2, \cdots, v_n\} = \{c_1v_1 + c_2v_2 + \cdots + c_nv_n : c_i \text{ are scalars}\}$$

In other words, the span is the set of all linear combinations of the set.

3. $\{v_1, v_2, \cdots, v_n\}$ is **linearly independent** if for all scalar $c_1, c_2, \cdots, c_n$, the following holds

$$\text{if } c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0, \text{ then } c_1 = c_2 = \cdots = c_n = 0.$$

4. $\{v_1, v_2, \cdots, v_n\}$ is **linearly dependent** if there exist scalars $c_1, c_2, \cdots, c_n$, not all 0, such that

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0.$$

Example 2.46. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -3 \end{pmatrix}; B_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; B_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}; B_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A is a linear combination of B_1, B_2, B_3, since $A = 4B_1 - B_2 - 2B_3$.

Example 2.47. In the vector space of real functions $\mathbb{F} = \{f | f : \mathbb{R} \to \mathbb{R}\}$, let $f(x) = \cos 2x$, $g(x) = \cos^2 x$, $h(x) = \sin^2 x$. Then, $f \in \text{Sp}\{g, h\}$ since $\cos 2x = \cos^2 x - \sin^2 x$. That is, f can be represented by a linear combination of g and h, namely $f = g - h$.

Example 2.48. Let

$$A_1 = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix};\ A_2 = \begin{pmatrix} 4 & -2 \\ 0 & 6 \end{pmatrix};\ A_3 = \begin{pmatrix} 6 & 4 \\ 4 & 8 \end{pmatrix}.$$

Let $a_1, a_2, a_3 \in \mathbb{R}$. Observe that if

$$a_1 = -2a_3,$$

and

$$a_2 = -a_3,$$

where

$$a_3 \in \mathbb{R},$$

then $a_1A_1 + a_2A_2 + a_3A_3 = 0$. For example, if $a_3 = 1$, then $a_1 = -2, a_2 = -1$, and $-2A_1 - A_2 + A_3 = 0$. Thus $\{A_1, A_2, A_3\}$ is linearly dependent.

3 Analytic Functions and Conformal Mappings

Previously, we very briefly visited complex functions. In this section, we define what it means for a function to be *analytic*. This section focuses on what this property means first from an algebraic perspective, and we will then look at a couple of geometric properties of analytic functions. Before we give a formal definition of what an analytic function is, we first provide an overview of what the goal of this classification is.

Recall that we can split any function $f(z)$ where $z = x + iy$ into its real and imaginary parts by $w = u(x, y) + iv(x, y)$, where u, v are real functions. So, suppose we have a pair of functions

$$u_1(x, y) = x^2 - y^2, \qquad v_1(x, y) = 2xy.$$

It is easy to see that this pair of equations "respects" the structure of $z = x + iy$. That is, it is a representation of a function of $z = x + iy$ since $u_1(x, y) + v_1(x, y) = x^2 - y^2 + i2xy = (x + iy)^2 = z^2$. These types of functions are "admissible" to the analytic classification. Any algebraic combination (addition, multiplication, powers, etc.) of admissible functions is also admissible.

However, consider the functions

$$u_2(x, y) = x^2 - y^2, \qquad v_2(x, y) = 3xy.$$

When combined as $u_2 + iv_2 = x^2 - y^2 + i3xy$, this pair of functions does not appear to be able to be represented in terms of $z = x + iy$. These types functions are "inadmissible" to the analytic classification. Note that the function must be able to be represented only in terms of z — a function w which can only be represented using $\bar{z}$ or $|z|$ are inadmissible. The functions $\bar{z}$ and $|z|$ are in fact inadmissible themselves.

Example 3.1. We provide a few examples of admissible and inadmissible functions.

- $w = x + iy$ is admissible since, clearly, $w = z$;
- $w = x^3 - 3xy^2 + i(3x^2y - y^3)$ is admissible since $w = z^3$;
- $w = x - iy = \bar{z}$ is inadmissible, which we will prove later;
- $|z|$ is inadmissible as well since $\bar{z} = |z|^2/2$, so admitting $|z|$ would mean $\bar{z}$ is admissible as well.
- For a final example, suppose we have a function

$$f(z) = \frac{z^2\bar{z}^2 + z^2 + \bar{z}^2 - 2\bar{z}z^2 - 2\bar{z} + 1}{10\bar{z} + z\bar{z}^2 - 2z\bar{z} - 5\bar{z}^2 + z - 5}.$$

  $f(z)$ is admissible, because we can in fact cancel the common factor of $(\bar{z} - 1)^2$ (if $\bar{z} \neq 1$) in the numerator and denominator to leave us with a function represented only by z.

The last function in the example above is representative of a problem in our current classification of analytic functions — simply checking if a function can be represented solely in terms of z is not reliable, since more complicated functions might be incredibly difficult to reduce. Thankfully, we are able to check if a function is analytic using derivatives.

Definition 3.2. Let f be a complex-valued function defined in a neighborhood of z_0. Then the **derivative** of f at z_0 is defined by

$$\frac{df}{dz}(z_0) := f'(z_0) := \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z},$$

provided that this limit exists. If this limit exists, f is said to be **differentiable** at z_0.

This definition is analogous to that of real-valued functions. Since Δz is a complex number, though, it is important to note that it can approach zero in several different ways. For example, Δz can move to zero horizontally, vertically, or perhaps even spiral towards 0. However, for a function to be considered differentiable at z_0, the above quotient must tend to the same, unique limit regardless of how Δz approaches 0.

Definition 3.3. A complex-valued function $f(z)$ is said to be **analytic** on an open set G if it has a derivative at every point in G.

Example 3.4. We show that $f(z) = \bar{z}$ is nowhere differentiable, and hence not analytic. The limit quotient for $f(z)$ is

$$\frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{\overline{(z_0 + \Delta z)} - \overline{z_0}}{\Delta z} = \frac{\overline{\Delta z}}{\Delta z}.$$

So, suppose we let $\Delta z \to 0$ along the real values. That is, $\Delta z = \Delta x$. Then,

$$\frac{\overline{\Delta z}}{\Delta z} = \frac{\overline{\Delta x}}{\Delta x} = 1.$$

But, if we let $\Delta z \to 0$ along the imaginary values, that is, $\Delta z = i\Delta y$, then

$$\frac{\overline{\Delta z}}{\Delta z} = \frac{\overline{i\Delta y}}{i\Delta y} = \frac{-i\Delta y}{i\Delta y} = -1.$$

Since the limit quotient gives us different values depending on how z_0 approaches 0, we cannot assign a value to the derivative of $\bar{z}$ at any point. Thus $\bar{z}$ is not differentiable at any point.

Definition 3.5. A point where f is not analytic but which is the limit of points where f is analytic is called a **singularity** of f.

Using this definition, we can say that a rational function of z is analytic at all points for which its denominator is nonzero — i.e. the points at which the the function does not have a singularity.

Definition 3.6. A function f has a **pole of order** n at z_0 if n is the smallest positive integer for which $(z-z_0)^n f(z)$ is analytic at z_0. An analytic function f is said to have a pole at infinity if $\lim_{z\to\infty} f(z) = \infty$. A pole of order 1 is called a **simple pole**. A pole is a type of singularity.

Definition 3.7. A point z_0 is called a **zero** of f if z_0 is point at which f is analytic and $f(z_0) = 0$.

Example 3.8. Suppose $f(z) = \frac{z+1}{z(z^2-4)^3}$. Then, f has a zero at $z = -1$, and singularities at $z = \pm 2$ and $z = 0$. The singularities at $z = \pm 2$ are poles of order 3, and the singularity at $z = 0$ is a simple pole.

We now move on to discuss a couple of geometric properties of analytic functions. In general, we can split the geometry of analytic mappings into two groups called *local* properties and *global* properties. As the names of the groups imply, a local property need only hold in sufficiently small neighborhoods, while a global property must hold over the whole domain.

Example 3.9. We demonstrate the function $f(z) = z^2$ is only locally one-to-one. For any open disk containing the origin, there will be distinct points z_1 and z_2 such that $z_2 = -z_1$. Hence, the function will not be one-to-one. However, at any point other than the origin, we are able to find an open disk which does not include the origin. $f(z) = z^2$ will be one-to-one in any such disk, and hence $f(z)$ is locally one-to-one at every point except the origin.

A second important local property is *conformality.*

Definition 3.10. A mapping $f : U \to V$ is called **conformal** at a point $u_0 \in U$ if it preserves angles between curves through u_0 with respect to their orientation.

This definition is easier to understand if we describe a situation where f is conformal. So, suppose f is analytic and one-to-one in a neighborhood of the point z_0. Let γ_1 and γ_2 be two smooth curves intersecting at z_0. Then the mappings $f(\gamma_1) = \gamma_1'$ and $f(\gamma_2) = \gamma_2'$ are also two smooth curves which intersect at $w_0 = f(z_0)$. Then, at the point z_0, we construct vectors v_1, v_2 which are tangent to γ_1 and γ_2, respectively. Define the angle θ as the angle between v_1 and v_2. Define the angle θ' similarly for γ_1' and γ_2'.

Now, the mapping f is **conformal** at z_0 if $\theta = \theta'$ for every pair of smooth curves intersecting at z_0. Figure 3 shows what this preservation of angles looks like. Essentially, conformality defines a consistency in how the mapping behaves between different curves.

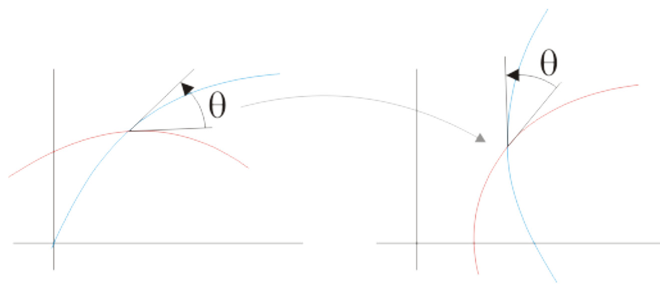

Figure 3: Preservation of angles between two curves [12]

Theorem 3.11. *An analytic function f is conformal at every point z_0 for which $f'(z_0) \neq 0$.*

The condition that $f'(z_0) \neq 0$ is very important for this theorem. For example, the function $f(z) = z^2$ has $f'(0) = 0$. This function not only does not preserve the angles of curves passing through the origin, but it in fact doubles them. However, we commonly call mappings conformal maps even if $f'(z) = 0$ for some points — we usually overlook these violations. We will now look at several simple examples of conformal maps.

Definition 3.12. The following mappings are some of the most basic conformal maps.

- The **translation** mapping is defined by

$$w = f(z) = z + c$$

  for some $c \in \mathbb{C}$. This mapping shifts every point by a vector corresponding to c.

- The **rotation** mapping is defined by

$$w = f(z) = e^{i\phi} z,$$

where $\phi \in \mathbb{R}$. This mapping rotates every point about the origin by angle ϕ.

- The **magnification** mapping is defined by

$$w = f(z) = \rho z,$$

where $\rho \in \mathbb{R}^+$. This mapping enlarges or contracts the distance of every point from the origin by the factor ρ.

- The **inversion** mapping is defined by

$$w = f(z) = \frac{1}{z}.$$

Each of these basic mappings map the complex plane one-to-one onto itself. We can combine these basic mappings to create linear transformations for specific problems. Note here a linear transformation refers to any mapping of the form

$$w = f(z) = az + b,$$

and is the composition of a rotation, magnification, and translation.

Example 3.13. Suppose we have a circle $C_1 = \{z : |z - 1| = 1\}$. We construct a transformation which maps this circle onto $C_2 = |z - 3i/2| = 2$.

First, we translate by -1 to get a circle centered at the origin:

$$w_1 = z - 1.$$

Then, we magnify by a factor of 2:

$$w_2 = 2w_1 = 2z - 2.$$

Finally, we translate by $3i/2$:

$$w = w_2 + 3i/2 = 2z - 2 + 3i/2.$$

Now, we introduce a slightly more complex conformal mapping. This specific mapping is used in our main results.

Definition 3.14. A **Linear Fractional Transformation (LFT)** is a function of the form

$$w = f(z) = \frac{az + b}{cz + d},$$

where $ad \neq bc$.

Note that

$$f'(z) = \frac{ad - bc}{(cz + d)^2} \neq 0,$$

since $ad \neq bc$, so LFTs are conformal at every point at which they are analytic. Since LFTs are analytic at every point except for a simple pole at $z = -d/c$, they are conformal at every point except $z = -d/c$. Also observe that an LFT can be expressed as a combination of elementary mappings.

Example 3.15. We construct the formula for an LFT through a combination of elementary mappings. Begin with a linear transformation

$$w_1 = cz + d,$$

we then invert to get

$$w_2 = \frac{1}{w_1} = \frac{1}{cz + d},$$

and perform another linear transformation

$$w = \left(b - \frac{ac}{d}\right) w_2 + \frac{a}{c} = \left(b - \frac{ac}{d}\right) \frac{1}{cz + d} + \frac{a}{c}. \tag{1}$$

From here, we can perform some algebra to get to our familiar formula

$$\begin{aligned} w &= \left(b - \frac{ac}{d}\right) \frac{1}{cz + d} + \frac{a}{c} \\ &= \frac{a}{c} + \frac{b - \frac{ad}{c}}{cz + d} \\ &= \frac{\frac{a}{c}(cz + d) - \frac{ad}{c} + b}{cz + d} \\ &= \frac{az + b}{cz + d}. \end{aligned}$$

A result of this is a number of properties of LFTs.

Theorem 3.16. *Let f be any LFT. Then*

1. *f can be expressed as the composition of a finite sequence of translations, magnifications, rotations, and inversions;*

2. *f maps the complex plane one-to-one onto itself;*

3. *f maps the class of circles and lines to itself;*

4. *f is conformal at every point except its pole.*

We have already demonstrated parts (1) and (4) of this theorem, and we can easily deduce (2) since each elementary mapping has this property. We must clarify what (3) means, though. If any circle or line passes through the pole ($z = -d/c$) of the LFT, it will be mapped to an unbounded figure — a line. However, any line or circle which does not pass through the pole of the LFT will be mapped to a circle.

4 Banach and l^p Spaces

The main results from [7] which we are exploring in this paper deal with a special type of space called a *Banach space* — recall we very briefly mentioned Banach spaces at the end of section 2.1. More specifically, a special subset of Banach spaces, called l^p spaces, are used extensively in [7].

Definition 4.1. A metric space (X, d) is called **complete** if every Cauchy sequence in X converges in X.

Definition 4.2. A normed vector space $(V, ||\cdot||)$ is called a **Banach space** if V is complete with respect to the metric induced from the norm.

Our first example of a Banach space requires a result known as the Bolzano-Weierstrass theorem to prove completeness.

Theorem 4.3 (Bolzano-Weierstrass)**.** *Every bounded sequence in /mathbbR has a convergent subsequence.*

Proof. Suppose (s_n) is a bounded sequence. By Theorem 2.14, (s_n) has a monotonic subsequence, (s_{n_k}). Since (s_n) is a bounded sequence, (s_{n_k}) is bounded as well. Since all bounded monotonic subsequences converge, we have that (s_{n_k}) converges. $\square$

We now have all the necessary tools in order to easily show certain spaces are complete.

Example 4.4. Consider the normed vector spaces $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.

- $\mathbb{R}$ with the standard Euclidean norm is a Banach space [9].

We already know that $\mathbb{R}$ is a normed vector space. So to prove it is a Banach space, we must show it is complete. That is, we must show that every Cauchy sequence in $\mathbb{R}$ converges to a point in $\mathbb{R}$. This follows immediately from three previous theorems. First, note every Cauchy sequence is bounded (Theorem 2.17). By Bolzano-Weierstrass, every bounded sequence in $\mathbb{R}$ has a convergent subsequence (with limit in $\mathbb{R}$). Finally, if a subsequence of a Cauchy sequence converges, then the whole sequence converges to the same point (Theorem 2.19). Thus $\mathbb{R}$ is complete, and is hence a Banach space.

- $\mathbb{C}$ with its standard norm is a Banach space. Indeed, if $(z_n) = (x_n) + i(y_n)$ is Cauchy in $\mathbb{C}$, then (x_n) and (y_n) are Cauchy in $\mathbb{R}$. Since $\mathbb{R}$ is complete, we have (x_n) converges to x and (y_n) converges to y, $x, y \in \mathbb{R}$. Thus we have that (z_n) converges to $z = x + iy$. Hence $\mathbb{C}$ is complete and is thus a Banach space.

- The metric space $(\mathbb{Q}, d)$ with its usual metric is not complete. This is because there are Cauchy sequences in $\mathbb{Q}$ which converge to irrational limits. For example,
$$(s_n) = \left(1 + \frac{1}{n}\right)^n$$
is contained in $\mathbb{Q}$, but converges to e which is not in $\mathbb{Q}$.

The following lemma is often useful in proving that a metric space is complete. It follows immediately from Theorem 2.19.

Lemma 4.5. *Let (X, d) be a metric space. If every Cauchy sequence has a convergent subsequence in X, then (X, d) is complete.*

Now we will provide a criterion for completion of a normed space in terms of series.

Definition 4.6. There are a number of terms we must define for series.

- If $(E, ||\cdot||)$ is a normed space then a **series** in E is just the summation a sequence (x_n) of terms $x_n \in E$ denoted

$$\sum_{k=m}^{n} x_k;$$

- The sum

$$s_n = \sum_{k=m}^{n} x_k$$

  is called the **nth partial sum** of the series;

- The series **converges** if the sequence of partial sums has a limit. That is, if there exists $s \in E$ such that

$$\lim_{n\to\infty} \left|\left| \left(\sum_{k=m}^{n} x_k \right) - s \right|\right| = 0.$$

  If the sequence of partial sums has no limit in E, we say the series **diverges**;

- We say a series $\sum_{n=m}^{\infty} x_n$ is **absolutely convergent** if $\sum_{n=m}^{\infty} ||x_n|| < \infty$. This is because $\sum_{n=m}^{\infty} ||x_n||$ is the summation of a series of real positive terms, so the sequence of partial sums either converges in $\mathbb{R}$ or increases to ∞.

Definition 4.7. A series satisfies the **Cauchy criterion** if its sequence of partial sums forms a Cauchy sequence.

Theorem 4.8 (Series criterion for completion)**.** *Let* $(E, ||\cdot||)$ *be a normed vector space.* E *is complete (and thus is a Banach space) if and only if each absolutely convergent series* $\sum_{n=1}^{\infty} x_n$ *of terms* $x_n \in E$ *is convergent in* E*.*

Proof. We begin with the forward direction. Suppose E is complete and $\sum_{n=1}^{\infty} ||x_j|| < \infty$. Then the individual partial sums of this series

$$S_n = \sum_{j=1}^{n} ||x_j||$$

must satisfy Cauchy criterion. That is, for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n, m \geq N$, then $|S_n - S_m| < \epsilon$. So suppose $n > m \geq N$. Then,

$$|S_n - S_m| = \left| \sum_{j=1}^{n} ||x_j|| - \sum_{j=1}^{m} ||x_j|| \right| = \sum_{j=m+1}^{n} ||x_j|| < \epsilon.$$

Now, consider the partial sums $s_n = \sum_{j=1}^{n} x_j$. If $n > m \geq N$, then

$$||s_n - s_m|| = \left|\left| \sum_{j=1}^{n} x_j - \sum_{j=1}^{m} x_j \right|\right| = \left|\left| \sum_{j=m+1}^{n} x_j \right|\right| \leq \sum_{j=m+1}^{n} ||x_j|| < \epsilon.$$

Hence the sequence (s_n) is Cauchy in E. Since E is complete, $\lim_{n\to\infty} s_n$ exists in E, and so $\sum_{n=1}^{\infty} x_n$ converges. This completes the forward direction.

For the other direction, assume all absolutely convergent series in E are convergent. Let (u_n) be a Cauchy sequence in E. Then we can find $n_1 > 0$ such that if $n, m \geq n_1$ then

$$||u_n - u_m|| < \frac{1}{2}.$$

We can also find $n_2 > 1$ such that if $n, m > n_2$ then

$$||u_n - u_m|| < \frac{1}{4} = \frac{1}{2^2}.$$

Without loss of generality we can assume that $n_2 > n_1$. By continuing to pick n_j in this way, we can find $n_1 < n_2 < n_3 < \cdots$ such that if $n, m \geq n_j$ then

$$||u_n - u_m|| < \frac{1}{2^j}.$$

Now, let $x_j = u_{n_{j+1}} - u_{n_j}$ for $j \in \mathbb{N}$. Then,

$$\begin{aligned}\sum_{j=1}^{\infty} ||x_j|| &= \sum_{j=1}^{\infty} ||u_{n_{j+1}} - u_{n_j}|| \\ &\leq \sum_{j=1}^{\infty} \frac{1}{2^j} \\ &= 1 < \infty.\end{aligned}$$

Thus $\sum_{j=1}^{\infty} x_j$ is absolutely convergent and, by our assumption, must also be convergent. Thus its sequence of partial sums

$$s_J = \sum_{j=1}^{J} (u_{n_{j+1}} - u_{n_j}) = u_{n_{J+1}} - u_{n_1}$$

has a limit in E. Thus,

$$\lim_{J \to \infty} u_{N_{J+1}} = u_{n_1} + \lim_{J \to \infty} (u_{N_{J+1}} - u_{n_1})$$

exists in E. We have shown that (u_n) has a convergent subsequence in E.

By Lemma 4.4, we finally have that E is complete. $\square$

We now look at a special family of Banach spaces known as l^p-spaces. l^p-spaces are the main type of space we will be working in in the main results section of the paper.

Definition 4.9. For $1 \leq p < \infty$, l^p denotes the set of all sequences $x = \{a_n\}_{n=1}^{\infty}$ which satisfy

$$\sum_{n=1}^{\infty} |a_n|^p < \infty.$$

If $p = \infty$, we define l^∞ as the set of all sequences $x = \{a_n\}_{n=1}^{\infty}$ which satisfy

$$\sup_{n \geq 1} |a_n| < \infty$$

In other words, for $1 \leq p < \infty$, l^p is the set of all sequences whose series converge when each of its individual elements is raised to the p^{th} power. l^∞ is the set of all bounded sequences. We commonly combine l^p with the following norms, called the *p-norms*:

For $x = (a_n) \in l^p$, define

$$||x||_p = \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}, \quad \text{for } p < \infty;$$

$$\text{and } ||x||_\infty = \sup_{n \geq 1} |a_n|.$$

It is very easy to see that the l^p spaces are nested; as p increases, l^p become more permissive. That is,

$$l^1 \subset l^2 \subset l^3 \subset \cdots \subset l^\infty.$$

Example 4.10. We demonstrate the nesting of l^p. Consider the harmonic sequence $(s_n) = \frac{1}{n}$. It is well known that

$$\sum_{n=1}^{\infty} \left| \frac{1}{n} \right| = \infty.$$

Thus $(s_n) \notin l^1$.

However,

$$\sum_{n=1}^{\infty} \left| \frac{1}{n} \right|^2 = \sum_{n=1}^{\infty} \left| \frac{1}{n^2} \right| < \infty.$$

Thus $(s_n) \in l^2$. Similarly, $(s_n) \in l^p$ for $p > 2$.

We can also easily see that the p-norms are nested as well. We show this by drawing the unit ball

$$\{x \in \mathbb{R}^2 : ||x||_p \leq 1\}$$

for each p-norm in $\mathbb{R}^2$ (Figure 4).

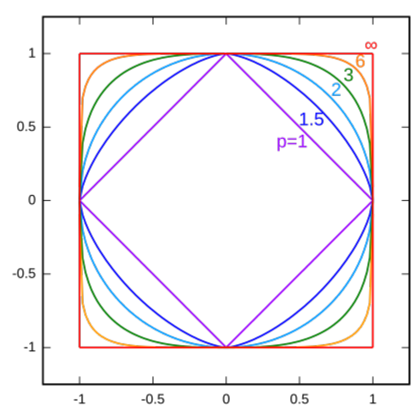

Figure 4: The unit ball of various p-norms [5]

We now show that l^p is a Banach space under its appropriate p-norm. We do this in several steps, first showing l^p is a vector space. Then, we prove the properties of a norm hold for the p-norms. Finally, we show that l^p is complete. Note we often omit the cases for $p = 1$ and $p = \infty$, as those cases will sometimes require separate arguments which are simple and direct.

Proposition 4.11. *$l^p(1 \leq p \leq \infty)$ is a vector space.*

Proof. If $x \in l^p$ and $c \in \mathbb{C}$, then clearly $cx \in l^p$. Then, if $x, y \in l^p$, since $|x_n + y_n| \leq (2|x_n|^p) + (2|y_n|^p)$, we also have that $x + y \in l^p$. So addition and scalar multiplication can be defined on l^p. Now, algebraic operations on l^p are defined componentwise; for example, $x + y$ is the sequence whose nth element is $x_n + y_n$. Since all operations are defined this way, the algebraic rules for l^p are simply inherited from $\mathbb{C}$. $\square$

Proposition 4.12. *$||\cdot||_p$ is a norm on l^p.*

The first two properties of a norm are obvious. To prove the triangle inequality, though, we must first prove an inequality known as Hölder's inequality. For Hölder's inequality, we need the following lemma, which we will take as given.

Lemma 4.13. *Suppose $1 < p < \infty$ and q is defined by $\frac{1}{p} + \frac{1}{q} = 1$. Then,*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \qquad \text{for } a, b \geq 0.$$

Example 4.14 (of previous Lemma)**.** Suppose $p = 3$. Then $q = \frac{3}{2}$ since $\frac{1}{p} + \frac{1}{q} = 1$. Consider several values for a and b.

- $a = 1, b = 1$

  $(1)(1) = 1$. $\frac{1^3}{3} + \frac{1^{3/2}}{3/2} = \frac{1}{3} + \frac{2}{3} = 1$

- $a = 4, b = 9$

  $(4)(9) = 36$. $\frac{4^3}{3} + \frac{9^{3/2}}{3/2} = \frac{64}{3} + \frac{54}{3} = \frac{118}{3} > 36$.

- $a = 20, b = 64$

  $(20)(64) = 1280$. $\frac{20^3}{3} + \frac{64^{3/2}}{3/2} = \frac{8000}{3} + \frac{1024}{3} = \frac{9024}{3} = 3008$.

Lemma 4.15 (Hölder's Inequality)**.** *Let $x = (a_n) \in l^p, y = (b_n) \in l^q$, where $1 \leq p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then,*

$$||xy||_1 = \sum_{n=1}^{\infty} |a_n b_n| \leq ||(a_n)||_p ||(b_n)||_q.$$

This inequality also shows that $xy \in l^1$ under the same conditions. Note that if $p = 1$, we interpret this to mean that $q = \infty$.

Proof. If $p = 1$ and $q = \infty$, this is obvious. So, suppose $p > 1$. Let A and B be as follows:

$$\begin{aligned} A &= ||x||_p = \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p} \\ B &= ||y||_q = \left(\sum_{n=1}^{\infty} |b_n|^q \right)^{1/q} \end{aligned}$$

If either $A = 0$ or $B = 0$, the inequality is obviously true; both sides must equal 0. So, assume $A \neq 0, B \neq 0$ and use Lemma 4.13 with $a = |a_n|/A$ and $b = |b_n|/B$ to see

$$\frac{|a_n b_n|}{AB} \leq \frac{1}{p} \frac{|a_n|^p}{A^p} + \frac{1}{q} \frac{|b_n|^q}{B^q}$$

So

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{|a_n b_n|}{AB} &\le \frac{1}{p} \sum_{n=1}^{\infty} \frac{|a_n|^p}{A^p} + \frac{1}{q} \sum_{n=1}^{\infty} \frac{|b_n|^q}{B^q} \\ &= \frac{1}{p} + \frac{1}{q} \\ &= 1.\end{aligned}$$

Hence

$$\sum_{n=1}^{\infty} |a_n b_n| \le AB = ||x||_p ||y||_q.$$

□

Lemma 4.16 (Minkowski's inequality or triangle inequality for l^p). *If $x, y \in l^p (1 \le p \le \infty)$, then $x + y \in l^p$ and*

$$||x + y||_p \le ||x||_p + ||y||_p$$

Proof. This inequality is trivial for $p = 1$ and $p = \infty$, so suppose $1 < p < \infty$. We already know $(x_n + y_n)_n \in l^p$ (Proposition 4.11). For the inequality, observe that

$$\begin{aligned}\sum_n |x_n + y_n|^p &= \sum_n |x_n + y_n| \, |x_n + y_n|^{p-1} \\ &\le \sum_n |x_n| \, |x_n + y_n|^{p-1} + \sum_n |y_n| \, |x_n + y_n|^{p-1}.\end{aligned}$$

Let $a_n = |x_n|$ and $b_n = |x_n + y_n|^{p-1}$. So $(a_n) \in l^p$. Also, $(b_n) \in l^q$ since

$\frac{1}{p} + \frac{1}{q} = 1$ and

$$\begin{aligned}\sum_n (b_n)^q &= \sum_n |x_n + y_n|^{(p-1)q} \\ &= \sum_n |x_n + y_n|^p \\ &< \infty\end{aligned}$$

since we know $(x_n + y_n)_n \in l^p$. Now, from Hölder's inequality, we have

$$\begin{aligned}\sum_n |x_n|\,|x_n + y_n|^{p-1} &\leq \left(\sum_n |x_n|^p\right)^{1/p} \left(\sum_n |x_n + y_n|^{(p-1)q}\right)^{1/q} \\ &= ||x||_p\,||x + y||_p^{p/q}\end{aligned}$$

Similarly, we have

$$\sum_n |y_n|\,|x_n + y_n|^{p-1} \leq ||y||_p||x + y||_p^{p/q}.$$

If we add the two inequalities, we are left with

$$||x + y||_p^p \leq ||x||_p||x + y||_p^{p/q} + ||y||_p||x + y||_p^{p/q}.$$

Now, if $||x+y||_p = 0$, then Minkowski's inequality is true since a property of norms is that $||\cdot|| \geq 0$. So suppose $||x + y||_p \neq 0$. Then we can divide both sides of the inequality by $\left(||x + y||_p\right)^{p/q}$ to get

$$||x + y||_p^{p-p/q} \leq ||x||_p + ||y||_p.$$

Since, $p - \frac{p}{q} = 1$, we have achieved our goal. $\square$

At this point, we have that l^p is a normed vector space for $1 \leq p \leq \infty$. To show that l^p is a Banach space, all that remains is to show that l^p is complete. The proof of this requires some awkward notation, since the elements of l^p are sequences themselves — so, a sequence of elements in l^p is a sequence of sequences. So in a sequence of sequences in l^p, we will use a superscript to denote the nth sequence — i.e. $x^{(n)}$ is the nth sequence.

Theorem 4.17. *l^p is complete under the norm*

$$||x||_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}$$

for $x \in l^p, 1 < p < \infty$.

Proof. Let $(x^{(n)})$ be a Cauchy sequence in l^p. Let $\epsilon > 0$. Since $(x^{(n)})$ is Cauchy, there exists $n_0 \in \mathbb{N}$ such that if $n, m > n_0$, then

$$\sum_{j=1}^{\infty} |x_j^{(n)} - x_j^{(m)}|^p < \epsilon^p.$$

So, for each fixed $j \in \mathbb{N}$,

$$|x_j^{(n)} - x_j^{(m)}| \leq \left(\sum_{j=1}^{\infty} |x_j^{(n)} - x_j^{(m)}|^p\right)^{1/p} < \epsilon.$$

That is, $(x_j^{(n)})_n$ is a Cauchy sequence in $\mathbb{C}$. Since $\mathbb{C}$ is complete, these sequences have limits in $\mathbb{C}$. Define

$$x_j = \lim_{n\to\infty} x_j^{(n)}.$$

Let $x = (x_j)$. So, we have shown $x^{(n)} \to x = (x_j)$. Now we must verify that $x \in l^p$. Observe that if we arbitrarily fix $N \in \mathbb{N}$,

$$\begin{aligned} \sum_{j=1}^{N} |x_j|^p &= \lim_{n\to\infty} \sum_{j=1}^{N} \left|x_j^{(n)}\right|^p \\ &\leq \limsup_{n\to\infty} ||x^{(n)}||^p. \end{aligned}$$

Since $x^{(n)} \in l^p$, this is sufficient to show that $x = (x_j) \in l^p$. □

To end this section, we will introduce an operator known as the *backward shift*. This concept isn't directly related to Banach or l^p spaces, but it is often used on sequences in l^p, as we do in our main results.

Definition 4.18. The **backward shift operator** is a mapping on an infinite sequence by

$$B : (a_1, a_2, a_3, \cdots) \to (a_2, a_3, a_4, \cdots).$$

Essentially, the backward shift removes the first entry from the sequence.

Example 4.19. Suppose we have a sequence $(a_n) = (-1, 0, 1, 2, \cdots)$. Then, $B(a_n) = (0, 1, 2, 3, \cdots)$.

We can also multiply the backward shift: $2B(a_n) = (0, 2, 4, \cdots)$.

Proposition 4.20. *If $x \in l^p$, then $B^n x \in l^p$ and $||B^n x||_p \leq ||x||_p$.*

Proof. Suppose $x = (x_1, x_2, \cdots) \in l^p$. Then $B(x) = (x_2, x_3, \cdots)$. Now,

$$||x||_p^p = \sum_{n=1}^{\infty} |x_n|^p.$$

Observe that

$$||Bx||_p^p = \sum_{n=2}^{\infty} |x_n|^p \leq \sum_{n=1}^{\infty} |x_n|^p = ||x||_p^p.$$

In general,

$$||B^n x||_p^p = \sum_{k=n}^{\infty} |x_k|^p \leq \sum_{n=1}^{\infty} |x_n|^p = ||x||_p^p.$$

Thus $B^n x \in l^p$ and $||B^n x||_p \leq ||x||_p$. $\square$

Lemma 4.21. *The operator $B : l^p \to l^p$ is continuous.*

Proof. Suppose $x^{(n)}$ converges to $x \in l^p$. Then, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n > N$, then $||x^{(n)} - x||_p < \epsilon$. Now, since B is linear,

$$||B(x^{(n)} - Bx)||_p = ||B(x^{(n)} - x)||_p.$$

Then, by Proposition 4.20, we have

$$\begin{aligned} ||B(x^{(n)} - x)||_p &\leq ||x^{(n)} - x)||_p \\ &< \epsilon. \end{aligned}$$

Thus B is continuous on l^p. $\square$

5 Chaos

In this section we introduce the notion of chaos. This section is slightly different from the previous sections in that we are not proving or even providing any results relating to chaos. Rather, we are simply providing the definitions needed to understand Devaney's definition of chaos, as well as this definition itself. This is because our main results require an understanding of the definition of chaos, but understanding chaos on an analytical level is not so important. For this section, Devaney's *A First Course in Chaotic Dynamical Systems*[10] was widely used.

While there is no universally accepted definition of chaos, one of the most popular descriptions of chaos is given by Devaney in [10]. This definition gives three features for a continuous map $f : X \to X$ to be considered chaotic on X. Before we look at this definition, though, we need a few preliminary notions — specifically *denseness* of a set, the *orbit* of a point, and a couple of other terms related to an orbit.

Definition 5.1. Suppose X is a set, and Y is a subset of X. We say that Y is **dense** in X if, for any point $x \in X$, there is a sequence of point $\{y_n\} \in Y$ that converges to x. Less formally, Y is dense in X if for any point $x \in X$, we can find points in Y that are arbitrarily close to x.

To prove that a subset $Y \subset X$ is dense in X, we must exhibit a sequence of points in Y that converges to an arbitrary point in X. Specifically, we must show there is a sequence of points in Y that converges to an arbitrary point in $X \setminus Y$, since it is trivial to show there is a sequence that converges to a point in Y.

Example 5.2. We show the set of rationals $\mathbb{Q}$ is dense in $\mathbb{R}$. Let $x \in \mathbb{R} \setminus \mathbb{Q}$. That is, let x be irrational. Then, x has an infinite decimal expansion of the form

$$x = a_n \cdots a_0.b_1b_2b_3 \cdots$$

where a_j, b_j are integers between 0 and 9 (note the decimal point between a_0 and b_1). Then, for $j \in \mathbb{N}$, set

$$x_j = a_n \cdots a_0.b_1 \cdots b_j.$$

Then, each x_j has a finite decimal expansion and thus is a rational number. Clearly, the sequence $(x_j)_{j=1}^{\infty}$ has limit x. Thus, $\mathbb{Q}$ is dense in $\mathbb{R}$.

The next definition is pivotal in defining orbits and periodic points.

Definition 5.3. The **iteration** of a function F is the process of repeatedly composing F with itself. The nth iteration of F refers to the process of composing F with itself n times. We denote the nth iteration of F as F^n.

Example 5.4. Suppose $F : X \to X$ is defined by $F(x) = x^2 + 1$ for $x \in X$. Then

$$\begin{aligned}
F^2(x) &= F(F(x)) = (x^2+1)^2 + 1 \\
F^3(x) &= F(F^2(X)) = ((x^2+1)^2+1)^2 + 1 \\
&\vdots \\
F^n(x) &= F(F^{n-1}(x))
\end{aligned}$$

Definition 5.5. Given a function $F : X \to X$ and $x_0 \in X$, we define the

orbit of x_0 to be the sequence of points (x_n), where $x_n = F^n(x_0)$ for each $n \geq 0$. The point x_0 is called the **seed** of the orbit.

Example 5.6. In $\mathbb{R}$, if $F(x) = \sqrt{x}$ and $x_0 = 256$, the first several points of the orbit of x_0 are

$$\begin{aligned}
x_0 &= 256 \\
x_1 &= \sqrt{256} = 16 \\
x_2 &= \sqrt{16} = 4 \\
x_3 &= \sqrt{4} = 2 \cdots
\end{aligned}$$

There are several classifications of orbits, which we define now.

Definition 5.7. A **fixed point** is a point x_0 which satisfies $F(x_0) = x_0$.

A consequence of a fixed point is that its orbit is just the constant sequence $(x_0, x_0, x_0, \cdots)$. We can find the fixed points of a function by solving $F(x) = x$ for x.

Example 5.8. In $\mathbb{R}$, suppose $F(x) = x^2 + x - 4$. Then $F(x)$ has fixed points at the solutions of

$$x^2 + x - 4 = x,$$

which are $x = \pm 2$.

Definition 5.9. A point $x_0 \in X$ is called a **periodic point** of a function $F : X \to X$ if $F^n(x_0) = x_0$ for some $n > 0$. The least such n for which this is true is called the **period** of the orbit of x_0.

Example 5.10. In $\mathbb{R}$, suppose $F(x) = x^2 - 1$. The point $x_0 = 0$ has period 2, since $F(0) = -1$ and $F(-1) = 0$. Thus, the orbit of 0 is just $(0, -1, 0, -1, \cdots)$.

Definition 5.11. A point x_0 is called **eventually fixed** or **eventually periodic** if x_0 itself is neither fixed nor periodic, but some point on the orbit of x_0 is fixed or periodic, respectively.

Example 5.12. We provide an example of an eventually fixed point, and an eventually periodic point.

- Suppose $F(x) = x^2$. Then $x_0 = -1$ is eventually fixed, because $F(-1) = 1$ and $F^2(-1) = F(1) = 1$, which is then fixed. Its orbit is $(-1, 1, 1, 1, \cdots)$.
- Suppose $F(x) = x^2 - 1$. Then $x_0 = 1$ is eventually periodic since $F(1) = 0$ and 0 is periodic on F with period 2. Its orbit is $(1, 0, -1, 0, -1, 0, \cdots)$.

The examples we have provided so far may make the functions we have been using seem very simple when iterated. However, the existence of very simple orbits does not mean the function will always behave so simply when iterated.

Example 5.13. Suppose $F(x) = x^2 - 2$. Then, the orbit of 0 is eventually fixed: $(0, -2, 2, 2, 2, \cdots)$. However, if we look at points very close to 0, the functions performs rather erratically when iterated. For example, Figure 5 compares several iterations of 0 and 0.1 for this function. Observe that the elements of the orbit of 0.1 seem to randomly wander between -2 and

n	$x=0$	$x=0.1$
1	0	0.1
2	-2	-1.99
3	2	1.960
4	2	1.842
5	2	1.393
6	2	-0.597
7	2	-1.996
8	2	1.986
9	2	1.943

Figure 5: Several orbits of $F(x) = x^2 - 2$

2. Other values close to 0 have erratic behavior similar to 0.1. This is an example of what chaotic behavior looks like.

Definition 5.14. An operator $T : X \to X$ on a Banach space X is called **hypercyclic** if there exists a vector $x \in X$ such that the sequence $\{T^n(x) : n = 0, 1, 2, \cdots\}$ is dense in X. We refer to such an x as a **hypercyclic vector**.

As mentioned earlier, Devaney gives three conditions which must be met for a system to be considered chaotic. The first condition is that for a mapping $F : X \to X$, the periodic points of F are dense in X. The other two conditions are that F is topologically transitive, and F has sensitive dependence on initial conditions. These last two conditions still require definition.

Definition 5.15. $F : X \to X$ is **topologically transitive** if and only if for all nonempty open subsets U, V of X, there exists some $k \in \mathbb{N}$ such that $F^k(U) \cap V$ is nonempty. Alternatively, $F : X \to X$ is transitive if for any $x, y \in X$ and any $\epsilon > 0$, there exists a point $z \in X$ with $d(z, x) < \epsilon$ for which $d(F^n(z), y) < \epsilon$ for some $n \in \mathbb{N}$. That is, there exists a point within ϵ of x whose orbit comes with ϵ of y.

Definition 5.16. $F : X \to X$ is said to have **sensitive dependence on initial conditions** (SDIC) if there is some $\delta > 0$ (called the sensitivity constant) such that for every $x \in X$ and any $\epsilon > 0$, there exists a point $y \in X$ and $k \in \mathbb{N}$ such that $d(x, y) < \epsilon$ and $d(F^k(x), F^k(y)) > \delta$.

Essentially, SDIC says that, no matter what x we begin with and no matter how small a region we choose around x, we can always find some y in this region whose orbit eventually separates from that of x by at least δ. In other words, for every x, there are points arbitrarily close to x whose orbits are eventually "far" from the orbit of x.

We can now provide a formal definition of chaos in one piece:

Definition 5.17. A continuous mapping F on a metric space X is **chaotic** if

1. The periodic points for F are dense in X;
2. F is topologically transitive;
3. F has sensitive dependence on initial conditions.

Interestingly, it was shown in [11] that the first two conditions imply the third; it is not necessary to prove SDIC to show a mapping is chaotic. However, SDIC is really the central idea of chaos. It says that if a mapping is chaotic, minute errors in computation of the start state (e.g., due to rounding) may lead to a large divergence from the intended orbit upon iteration.

A chaotic mapping often produces a *fractal* — a set which exhibits self-similarity. For example, the famous Mandelbrot set (Figure 6) is generated by iterating the complex quadratic function $f_c(z) = z^2 + c$ at $z = 0$. The Mandelbrot set is the set of all c for which $f_c^n(0)$ does not go to ∞ as n goes to ∞. Take note of the similar structure that appears at different magnifications, and also observe the behavior at the edge of the set.

Figure 6: Several magnifications of the Mandelbrot set [14]

Having an understanding of what chaos entails is important, but we do not provide any examples of proving a mapping is chaotic. Unfortunately, proving chaos relies on a lot more theory in linear dynamics which is outside the scope of our main results. Rather than use this definition to prove chaos, our main results rely on theorem known as the *Eigenvalue Criterion for Chaos* to prove a mapping is Devaney chaotic. We will go over this criterion in the main results section.

To summarize, a chaotic mapping has three features, the third of which is seemingly contradictory of the other two: unpredictability, indecomposability, and regularity. A chaotic system is unpredictable due to the SDIC condition. Transitivity dictates that its behavior cannot be separated into two distinct subsystems. Finally, chaotic mappings nevertheless have an element of regularity in that their set of periodic points is dense.

6 Main results

The main goal of this section is to state the chaotic behavior of $\varphi(B)$ on l^p, where φ is an LFT and $B : l^p \to l^p$ is the backward shift operator. In order to avoid the more trivial cases where $\varphi(z)$ reduces to a constant or a first degree polynomial, we will restrict the LFT by $ad - bc \neq 0$ and $c \neq 0$. In other words, φ is defined by

$$\varphi(z) = \frac{az+b}{cz+d},\ a, b, c, d \in ,\ ad - bc \neq 0,\ c \neq 0. \tag{2}$$

Before we define chaos for $\varphi(B)$, we must first define exactly what $\varphi(B)$ is. We do this using the Taylor series expansion of φ, which we derive using equation (1) from Example 3.15. From (1), we have that

$$\varphi(z) = \left(b - \frac{ac}{d}\right)\frac{1}{cz+d} + \frac{a}{c}.$$

Then, by dividing the numerator of the $\frac{1}{cz+d}$ term by d, we have

$$\begin{aligned}\left(b - \frac{ac}{d}\right)\frac{1}{cz+d} + \frac{a}{c} &= \left(b - \frac{ac}{d}\right)\frac{1/d}{\frac{c}{d}z+1} + \frac{a}{c} \\ &= \left(b - \frac{ac}{d}\right)\frac{1/d}{1-(-\frac{c}{d})z} + \frac{a}{c}.\end{aligned}$$

If we then add and subtract $\frac{b}{d}$ from this formula, we are left with a nicely structured Taylor series representation of $\varphi(B)$ after a little algebra.

$$\begin{aligned}
\varphi(z) &= \left(\frac{b}{d} - \frac{ac}{d^2}\right)\frac{1}{1-(-\frac{c}{d})z} + \frac{a}{c} \\
&= \left(\frac{b}{d} - \frac{ac}{d^2}\right)\frac{1/d}{1-(-\frac{c}{d})z} + \frac{a}{c} + \frac{b}{d} - \frac{b}{d} \\
&= \varphi(z) = \frac{b}{d} + \sum_{n=1}^{\infty}(-1)^{n+1}\frac{ad-bc}{cd}\left(\frac{c}{d}\right)^n z^n
\end{aligned}$$

This is our Taylor series expansion of φ.

Then, for $|d| > |c|$, and for each $x \in l^p$, we have

$$\sum_{n=1}^{\infty}\left(\frac{c}{d}\right)^n ||B^n x||_p \leq \sum_{n=1}^{\infty}\left(\frac{c}{d}\right)^n ||x||_p,$$

since, by Proposition 4.20,

$$\sum_{k=1}^{\infty}|B^n x_k| = \sum_{k=n}^{\infty}|x_k| \leq \sum_{k=1}^{\infty}|x_k|.$$

As a result of this, the series

$$\frac{b}{d} + \sum_{n=1}^{\infty}(-1)^{n+1}\frac{ad-bc}{cd}\left(\frac{c}{d}\right)^n B^n$$

converges on l^p. We denote the limit of partial sums of this series as $\varphi(B)$. That is,

$$\varphi(B) = \frac{b}{d} + \lim_{k\to\infty}\sum_{n=1}^{k}(-1)^{n+1}\frac{ad-bc}{cd}\left(\frac{c}{d}\right)^n B^n.$$

Now, to show chaos of $\varphi(B)$ on l^p, we will use the Eigenvalue Criterion for Chaos, which we mentioned earlier in the paper.

Theorem 6.1 (Eigenvalue Criterion)**.** *Let $T : X \to X$ be an operator on a separable complex Banach space X. Consider the subspaces*

$$X_0 := \mathrm{Span}\{\mathrm{x} \in \mathrm{X} : \mathrm{T(x)} = \lambda\mathrm{x} \text{ for some } \lambda \in \mathbb{C} \text{ with } |\lambda| < 1\},$$

$$Y_0 := \mathrm{Span}\{\mathrm{x} \in \mathrm{X} : \mathrm{T(x)} = \lambda\mathrm{x} \text{ for some } \lambda \in \mathbb{C} \text{ with } |\lambda| > 1\},$$

$$Z_0 := \mathrm{Span}\{\mathrm{x} \in \mathrm{X} : \mathrm{T(x)} = \mathrm{e}^{\alpha\pi\mathrm{i}}\mathrm{x} \text{ for some } \alpha \in \mathbb{Q}\}.$$

If X_0, Y_0, and Z_0 are all dense in X, then T is chaotic.

In other words, for an operator $T : X \to X$ to be chaotic, consider the following sets:

1. The of the set of eigenvectors of T with eigenvalues in the open unit disk;
2. The of the set of eigenvectors of T with eigenvalues outside the open unit disk;
3. The set of vectors in X for which Tx is a rotation by $\alpha\pi$, where α is a rational number.

If the spans of these three sets are dense in X, then T is chaotic. Let's deduce what exactly this means for our framework. First, we find the set of eigenvalues for B.

Proposition 6.2. *The set of eigenvalues of B on l^p is the unit disk $\mathbb{D}$. We will denote this as $\sigma_p(B) = \mathbb{D}$.*

Proof. We need to find the set of $\lambda \in \mathbb{C}$ such that $Bx = \lambda x$ for some $x \in l^p$. So, observe that

$$
\begin{aligned}
&Bx = \lambda x \\
&\text{iff } B(x_1, x_2, x_3 \cdots) = \lambda(x_1, x_2, x_3, \cdots) \\
&\text{iff } (x_2, x_3, x_4, \cdots) = \lambda(x_1, x_2, x_3, \cdots) \\
&\text{iff } x_{j+i} = \lambda x_j \quad \text{ for all } j \in \mathbb{N} \\
&\text{iff } (x_1, x_2, x_3, \cdots) = x_1(1, \lambda, \lambda^2, \cdots).
\end{aligned}
$$

Since $x = (x_1, x_2, x_3, \cdots) \in l^p$, we must have that $x_1(1, \lambda, \lambda^2, \cdots) \in l^p$. This is only true if $(1, \lambda, \lambda^2, \cdots) \in l^p$. Hence $|\lambda| < 1$. So we have that the set of eigenvalues of B in l^p is the set of all complex numbers whose modulus is less than 1. In other words,

$$\sigma_p(B) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}.$$

□

Proposition 6.3. *If λ is an eigenvalue for B, then $\varphi(\lambda)$ is an eigenvalue for $\varphi(B)$.*

Proof. Suppose λ is an eigenvalue for B with corresponding eigenvector $x \in l^p$; that is, $Bx = \lambda x$. Then,

$$
\begin{aligned}
\varphi(B)(x) &= \frac{b}{d} + \lim_{k\to\infty}\sum_{n=1}^{k}(-1)^{n+1}\frac{ad-bc}{cd}\left(\frac{c}{d}\right)^{n} B^n x \\
&= \frac{b}{d} + \lim_{k\to\infty}\sum_{n=1}^{k}(-1)^{n+1}\frac{ad-bc}{cd}\left(\frac{c}{d}\right)^{n} \lambda^n x \\
&= \varphi(\lambda)x.
\end{aligned}
$$

Thus, if λ is an eigenvalue of B then $\varphi(\lambda)$ is an eigenvalue of $\varphi(\lambda)$. $\square$

Since $\sigma_p(B) = \mathbb{D}$, we have that $\sigma_p(\varphi(B)) = \varphi(\mathbb{D})$. So $T = \varphi(B)$ is chaotic if our three sets from Theorem 6.1 are dense in l^p, with regards to $\varphi(B)$. Let's think about what the first two sets, X_0 and Y_0, are.

Essentially, an eigenvector x of a linear operator T is a vector for which Tx is a simple magnification — that is, T enlarges or contracts the length of x with regards to the origin. The eigenvalue corresponding to the eigenvector x is the magnitude of that magnification. So, the set X_0 is the set of eigenvectors which are contracted by $T = \varphi(B)$, and Y_0 is the set of eigenvectors which are enlarged by $T = \varphi(B)$.

So, we must have eigenvalues λ for which $|\lambda| < 1$, as well as eigenvalues λ for which $|\lambda| > 1$; otherwise, one of X_0 or Y_0 will be the empty set, which cannot be dense. Since we already know our set of eigenvalues for $\varphi(B)$ is $\varphi(\mathbb{D})$, this is equivalent to saying that $\varphi(B)$ is chaotic if and only if $\varphi(\mathbb{D})$ intersects the unit circle.

So, to prove that $\varphi(B)$ is chaotic in l^p, we need to show that $\varphi(\mathbb{D})$ intersects the unit circle. To do this, we first provide a geometrical description of $\varphi(\mathbb{C})$ when the pole of φ lies outside the unit disk.

Lemma 6.4. *Let φ be an LFT as defined in (2) with the added restriction $|d| > |c|$. Then $\varphi(\mathbb{D})$ is the disk $P + r\mathbb{D}$ with center P and radius r given by*

$$P = \frac{b\bar{d} - a\bar{c}}{|d|^2 - |c|^2}, \quad r = \frac{bc - ad}{|d|^2 - |c|^2}.$$

Proof. First, recall that LFTs map the class of circles and lines to circles and lines. Since $\overline{\mathbb{D}}$ is clearly bounded and convex, and $|d| > |c|$ (i.e. the pole of φ is outside of $\overline{\mathbb{D}}$), we know that $\varphi(\overline{\mathbb{D}})$ is a circle whose boundary is $\varphi(\partial\mathbb{D})$. Now, take three distinct points in the unit circle ($\partial\mathbb{D}$). We choose $z_1 = 1$, $z_2 = -1$, $z_3 = i$ for the sake of simplicity. Since φ is linear-fractional, it is also one-to-one. Thus $A = \varphi(z_1)$, $B = \varphi(z_2)$, $C = \varphi(z_3)$ are three distinct points. Since z_1, z_2, z_3 are on $\partial\mathbb{D}$, their respective images A, B, and C are three distinct points on the circle $\varphi(\partial\mathbb{D})$. That is, the circle circumscribed over A, B, and C coincides with $\varphi(\partial\mathbb{D})$.

To verify that $\varphi(\partial\mathbb{D})$ indeed has center P and radius r, we need to show that

$$|A - P| = |B - P| = |C - P| = r.$$

This only requires some simple algebra. We need three properties of the complex numbers, which we described earlier, to achieve this:

$$|z|^2 = z\bar{z} \tag{3}$$

$$|c + d| = |\bar{c} + \bar{d}| \tag{4}$$

$$|ci+d| = |\bar{c}+\bar{d}i| \tag{5}$$

So,

$$\begin{aligned}
|A-P| &= \left|\frac{a+b}{c+d} - \frac{b\bar{d}-a\bar{c}}{|d|^2-|c|^2}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(a+b)(|d|^2-|c|^2)-(b\bar{d}-a\bar{c})(c+d)}{c+d}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{ad\bar{d}-bc\bar{c}-bc\bar{d}+ad\bar{c}}{c+d}\right|, \text{ by using (3)} \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(\bar{c}+\bar{d})(bc-ad)}{c+d}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left(\frac{|\bar{c}+\bar{d}||bc-ad|}{|c+d|}\right) \\
&= \frac{1}{|d|^2-|c|^2}\left(\frac{|c+d||bc-ad|}{|c+d|}\right), \text{ by using (4)} \\
&= \frac{bc-ad}{|d|^2-|c|^2} \\
&= r.
\end{aligned}$$

Showing $|B-P| = r$ and $|C-P| = r$ is very similar:

$$\begin{aligned}
|B-P| &= \left|\frac{-a+b}{-c+d} - \frac{b\bar{d}-a\bar{c}}{|d|^2-|c|^2}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(-a+b)(|d|^2-|c|^2)-(b\bar{d}-a\bar{c})(-c+d)}{-c+d}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{-ad\bar{d}-bc\bar{c}+bc\bar{d}+ad\bar{c}}{-c+d}\right|, \text{ by using (3)} \\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(-\bar{c}+\bar{d})(bc-ad)}{-c+d}\right| \\
&= \frac{1}{|d|^2-|c|^2}\left(\frac{|-\bar{c}+\bar{d}||bc-ad|}{|-c+d|}\right).
\end{aligned}$$

Using (4), we then see that

$$\begin{aligned}
\frac{1}{|d|^2-|c|^2}\left(\frac{|-\bar{c}+\bar{d}||bc-ad|}{|-c+d|}\right) &= \frac{1}{|d|^2-|c|^2}\left(\frac{|-c+d||bc-ad|}{|-c+d|}\right)\\
&= \frac{bc-ad}{|d|^2-|c|^2}\\
&= r.
\end{aligned}$$

Finally, we have

$$\begin{aligned}
|C-P| &= \left|\frac{ai+b}{ci+d}-\frac{b\bar{d}-a\bar{c}}{|d|^2-|c|^2}\right|\\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(ai+b)(|d|^2-|c|^2)-(b\bar{d}-a\bar{c})(ci+d)}{ci+d}\right|\\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{ad\bar{d}i-bc\bar{c}-bc\bar{d}i+ad\bar{c}}{ci+d}\right|, \text{ by using (3)}\\
&= \frac{1}{|d|^2-|c|^2}\left|\frac{(\bar{c}+\bar{d}i)(ad-bc)}{ci+d}\right|\\
&= \frac{1}{|d|^2-|c|^2}\left(\frac{|\bar{c}+\bar{d}i||ad-bc|}{|ci+d|}\right).
\end{aligned}$$

By using (5), we then have

$$\begin{aligned}
\frac{1}{|d|^2-|c|^2}\left(\frac{|\bar{c}+\bar{d}i||ad-bc|}{|ci+d|}\right) &= \frac{1}{|d|^2-|c|^2}\left(\frac{|ci+d||ad-bc|}{|ci+d|}\right)\\
&= \frac{bc-ad}{|d|^2-|c|^2}\\
&= r.
\end{aligned}$$

$\square$

Now that we know what $\varphi(\mathbb{D})$ is, we can state and easily prove our main theorem.

Theorem 6.5. *Let φ be an LFT as defined in (2) and let $|d| > |c|$. The operator $\varphi(B)$ is chaotic on l^p if and only if*

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

Proof. We showed in Lemma 6.4 that $\varphi(\mathbb{D}) = P + r\mathbb{D}$ with center P and radius r given where

$$P = \frac{b\bar{d} - a\bar{c}}{|d|^2 - |c|^2},\ r = \frac{|bc - ad|}{|d|^2 - |c|^2}.$$

We used the Eigenvalue criterion to show $\varphi(B)$ is chaotic on l^p if and only if $\varphi(\mathbb{D})$ intersects the unit circle.

So suppose $\varphi(B)$ is chaotic. Then $\varphi(\mathbb{D}) = P + r\mathbb{D}$ intersects the unit circle. We have two possibilities for the disk $P + r\mathbb{D}$: the center of the disc P is contained within the unit circle, or P is outside the closed unit disc. If P is inside the unit disk, then $|P| + |r| > 1$; if P is outside the closed unit disk, then we must have $|P| - |r| < 1$.

These conditions lead to

$$-|r| < 1 - |P| < |r|.$$

After substituting in the values of P and r, we have

$$-\frac{|bc - ad|}{|d|^2 - |c|^2} < 1 - \left|\frac{|b\bar{d} - a\bar{c}|}{|d|^2 - |c|^2}\right| < \frac{|bc - ad|}{|d|^2 - |c|^2}.$$

Multiplying by $|d|^2 - |c|^2$ gives

$$-|bc - ad| < |d|^2 - |c|^2 - |b\bar{d} - a\bar{c}| < |bc - ad|.$$

So finally, we have

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

So, we have shown that if $\varphi(B)$ is chaotic, then $\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|$. The other direction is done similarly, using the same algebra, as shown below.

Suppose that $\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|$. Then,

$$-|bc - ad| < |d|^2 - |c|^2 - |b\bar{d} - a\bar{c} < |bc - ad|.$$

So by dividing everything by $|d|^2 - |c|^2$,

$$-\frac{|bc - ad|}{|d|^2 - |c|^2} < 1 - \left|\frac{|b\bar{d} - a\bar{c}|}{|d|^2 - |c|^2}\right| < \frac{|bc - ad|}{|d|^2 - |c|^2}.$$

So we have that

$$-|r| < 1 - |P| < |r|.$$

So we have shown that if $\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|$, then $\varphi(\mathbb{D}) = P + r\mathbb{D}$ intersects the unit circle. Hence $\varphi(B)$ is chaotic. $\square$

This theorem makes it very easy to check if an LFT of the backward shift is chaotic in l^p. The next two examples use this theorem to check if $\varphi(B)$ is chaotic for a given LFT, and then we further verify the result using previous results. These examples only use real numbers for the sake of simplicity.

Example 6.6. Let φ be an LFT as in 2, with $a = 1$, $b = 2$, $c = 2$, $d = 3$. That is,

$$\varphi(z) = \frac{z+2}{2z+3}.$$

By Theorem 6.5, $\varphi(B)$ is chaotic in l^p if and only if

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

Observe that

$$\begin{aligned} \big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| &= \big||3|^2 - |2|^2 - |2\bar{3} - 1\bar{2}|\big| \\ &= \big|9 - 4 - |6 - 2|\big| \\ &= 1. \end{aligned}$$

Meanwhile,

$$\begin{aligned} |bc - ad| &= |(2)(2) - (1)(3)| \\ &= 1. \end{aligned}$$

Since our inequality does not hold (i.e. $1 \not< 1$), $\varphi(B)$ is not chaotic.

We can check this further by checking if $\varphi(\mathbb{D})$ intersects the unit circle. By Lemma 6.4, we know $\varphi(\mathbb{D})$ is the disk $P + r\mathbb{D}$ given by

$$P = \frac{b\bar{d} - a\bar{c}}{|d|^2 - |c|^2}, \quad r = \frac{bc - ad}{|d|^2 - |c|^2}.$$

So, plugging in our values for φ, we get $P = \frac{4}{5}$, $r = \frac{1}{5}$. So, $\varphi(\mathbb{D})$ touches the unit circle, but does not intersect it — it is contained completely inside the unit circle. Since $\varphi(\mathbb{D})$ is the set of eigenvalues for $\varphi(B)$, the set

$$Y_0 := \mathrm{Span}\{\mathrm{x} \in \mathrm{X} : \mathrm{T}(\mathrm{x}) = \lambda \mathrm{x} \text{ for some } \lambda \in \mathbb{C} \text{ with } |\lambda| > 1\}$$

from our original eigenvalue criterion is empty. Thus it cannot be dense, further showing the φ is not chaotic with these values.

Example 6.7. Let φ be an LFT as in 2, with $a = 2$, $b = -2$, $c = 1$, $d = 2$. That is,

$$\varphi(z) = \frac{2z - 2}{z + 2}.$$

By Theorem 6.5, $\varphi(B)$ is chaotic in l^p if and only if

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

Observe that

$$\begin{aligned}\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| &= \big||2|^2 - |1|^2 - |-2\bar{2} - 2\bar{1}|\big| \\ &= \big|4 - 1 - |-4 - 2|\big| \\ &= 3.\end{aligned}$$

Then,

$$\begin{aligned}|bc - ad| &= |(-2)(1) - (2)(2)| \\ &= 6.\end{aligned}$$

Since $3 < 6$, our inequality holds and thus $\varphi(B)$ is chaotic. We can again check if $\varphi(\mathbb{D})$ intersects the unit circle. Using Lemma 6.3, we can find that $\varphi(\mathbb{D})$ is the disk with center $P = -2$ and radius $r = 2$. Hence $\varphi(\mathbb{D})$ intersects the unit circle.

The following corollary shows that $\varphi(B^n)$ for any positive integer n is chaotic on l^p exactly when $\varphi(B)$ is chaotic on l^p.

Corollary 6.8. *Let φ be an LFT as in (2), $|d| > |c|$, and $g(z) = z^n$, with n being any positive integer. Consider the composition $\varphi \circ g(z) = (az^n + b)/(cz^n + d)$. Then the operator $(\varphi \circ g)(B)$ is chaotic on l^p if and only if*

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

Proof. For this, we just need to observe that $\varphi(\mathbb{D})$ intersects the unit circle if and only if $\varphi \circ g\mathbb{D}$ does as well. Well $(\partial\mathbb{D})^n = \partial\mathbb{D}$ for any positive integer n. So, $(\varphi \circ g)(\mathbb{D}) = \varphi(\mathbb{D})$. Hence we can follow an analogous argument to Theorem 6.5 to prove $(\varphi \circ g)(B)$ is chaotic if and only if

$$\big||d|^2 - |c|^2 - |b\bar{d} - a\bar{c}|\big| < |bc - ad|.$$

□

We have demonstrated a way of computing Devaney chaos for LFTs of the backward shift n l^p. Although only Devaney chaos was considered for this classification, Jimènez-Munguìa, Galán, Martínez-Giménez, and Peris state that, interestingly, other forms of chaos can also happen under the same conditions presented here.

References

[1] Edgar, Gerald, *Measure, Topology, and Fractal Geometry*, Second Ed., Springer Science+Business Media, 2008.

[2] Saff, E.B. and Snider, A.D., *Fundamentals of Complex Analysis with Applications to Engineering and Science*, Third Ed., Pearson Education, 2003.

[3] Kaplansky, Irving, *Set Theory and Metric Spaces*, reprint, American Mathematical Society, 2001.

[4] Alexandrov, Oleg, *Illustration of Complex Conjugate*, Wikipedia, 2007.

[5] User Quartl, *Illustration of the unit circles of p-norms*, Wikipedia, 2011.

[6] Johnson, Lee W., *Introduction to Linear Algebra*, Fifth Ed., Pearson Education, 2001.

[7] Jimënez-Munguïa, Ronald R., Galän, Vïctor J., Martïnez-Gimënez, and Fëlix, Peris, Alfredo, *Chaos for Linear Fractional Transformations of Shifts*, Journal of Topology and its Applications, Elsevier B.V., 2016

[8] Abbott, Stephen, *Understanding Analysis*, Second Ed., Springer, 2015

[9] Chen, Kuo-Chang, *Introduction to Banach Spaces*, `http://www.math.nthu.edu.tw/~kchen/teaching/5131week1.pdf`

[10] Devaney, Robert L., *A first course in chaotic dynamical systems*, Perseus Books, 1992

[11] Banks, J., Brooks, J., Cairns, G., Davix, G., and Stacey, P., *On Devaney's Definition of Chaos*, Mathematical Association of America, 1992

[12] `https://ece.uwaterloo.ca/~math212/Lectures/30/`, University of Waterloo, 2016

[13] Johnson, Riess, and Arnold, *Introduction to Linear Algebra*, Fifth ed., Pearson, 2002

[14] User Maksim, *Images of the Mandelbrot set at multiple zooms*, Wikipedia, 2005