The Maximal Lyapunov Exponent of a Time Series

Mark Goldsmith

A Thesis

in

The Department

of

Computer Science

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montreal, Quebec, Canada

December 2009

Abstract

The Maximal Lyapunov Exponent of a Time Series
Mark Goldsmith

Techniques from dynamical systems have been applied to the problem of predicting epileptic seizures since the early 90's. In particular, the computation of Lyapunov exponents from a series of electrical brain activity has been claimed to have great success. We survey the relevant topics from pure dynamical systems theory and explain how Wolf et al. adapted these ideas to the practical situation of trying to extract information from a time series. In doing so, we consider instances of time series where we may visually extract properties of the maximal Lyapunov exponent in an attempt to cultivate some intuition for more complicated and realistic situations.

Contents

Introduction

This thesis is first and foremost a primer on dynamical systems and Lyapunov exponents. The literature is scattered with inconsistencies and vague explanations, which will be pointed out and hopefully clarified throughout this thesis. The goal we set out to achieve is not simply to rigorously introduce dynamical systems, but to explain how Lyapunov exponents can be used to extract information from a given time series. The motivation for this stems from trying to understand how Lyapunov exponents are currently being used in an attempt to predict epileptic seizures (see [19], [18] and [24]). On the whole, we expand on the notes found in [6], which in turn are a clarification of the celebrated paper by Wolf et. al ([40]).

Chapter 1 begins with a treatment of dynamical systems as pure mathematical objects. We introduce typical notions such as fixed points, periodic trajectories and stability, and then proceed to define the Lyapunov exponent (of a point under a map). We will then explore the counterparts of these ideas in the multi-dimensional case. Throughout this treatment we will illustrate the topics at hand with demonstrative examples, some of which have become classical staples in the literature of dynamical systems. Chapter 1 ends with a rigorous treatment of the algorithm provided in [40] for estimating the maximal Lyapunov exponent of a trajectory.

In Chapter 2 we explain how [40] deals with the problem of estimating Lyapunov exponents in a minimal setting, where only a time series is provided. Finally, we present various results about strictly monotonic time series that are also strictly convex or strictly concave. These results will demonstrate how to use the definitions and the theory that has been built up along the way. Furthermore, through the exploration of such simple time series, we hope to build some intuition as to when the Lyapunov exponent can be easily deduced and therefore understood, so that clarity in more complex situations may ultimately be achieved.

Chapter 1

Discrete Dynamical Systems

1.1 Introduction.

A *discrete-time dynamical system* on a set X is just a function $\Phi : X \to X$. This function, often called a *map,* may describe deterministic evolution of some physical system: if the system is in state x at time t, then it will be in state $\Phi(x)$ at time $t+1$. Study of discrete-time dynamical systems is concerned with iterates of the map: the sequence

$$x, \Phi(x), \Phi^2(x), \ldots$$

is called the *trajectory of* x and the set of its points is called the *orbit of* x. These two terms are often used interchangeably, although we will remain consistent in their usage. The set X in which the states of the system exist is referred to as the *phase space* or *state space*. We will restrict our attention to maps $\Phi : X \to X$ such that X is a subset of $\mathbf{R}^d$.

We shall begin by studying one-dimensional maps.

1.2 One-dimensional discrete dynamical systems.

1.2.1 Examples

The following examples of one-dimensional maps will be used throughout this thesis to illustrate new ideas and techniques as they are introduced.

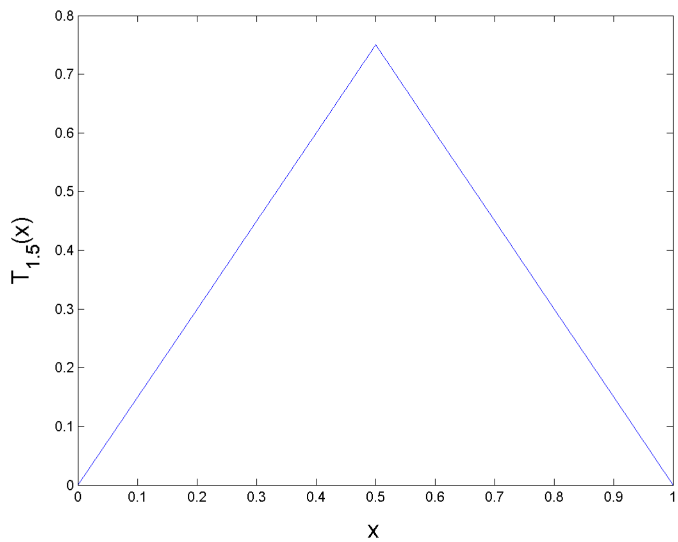
Example 1.2.1:
The tent map with parameter r (ranging from 0 to 2), is a one-dimensional map $T_r : [0,1] \to [0,1]$

defined as follows:

$$T_r(x) = \begin{cases} rx & 0 \leq x \leq \frac{1}{2}, \\ r(1-x) & \frac{1}{2} \leq x \leq 1. \end{cases}$$

The reason for its name is obvious once we plot it.

Figure 1.1: The tent map with $r = 1.5$.

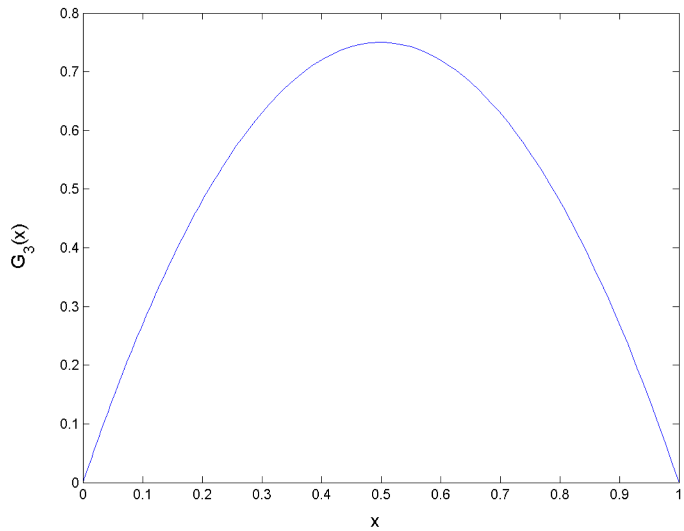


The parameter is sometimes locked to $r = 2$ (for example in [8] and [10]) in order to show its similarities with our next example. ■

Example 1.2.2:
The logistic map with parameter μ (ranging from 0 to 4) is a one-dimensional map $G_\mu : [0,1] \to [0,1]$ defined as

$$G_\mu(x) = \mu x(1-x). \tag{1.1}$$

Figure 1.2: Logistic map with $\mu = 3$.



The logistic map is likely the most celebrated of all one-dimensional maps. It was originally introduced by Verhulst in 1845 ([39]) in continuous form, and later unearthed as a discrete map by the biologist Robert May in [23], where he introduces many of its classical features. ■

1.2.2 Cobwebbing.

It is often cumbersome and uninformative to work out the trajectory of a point by hand. Fortunately, for one-dimensional maps there is a technique known as cobwebbing (or less commonly known as a Verhulst diagram, see [28]) that allows us to visually work out the trajectory of a point. It works as follows:

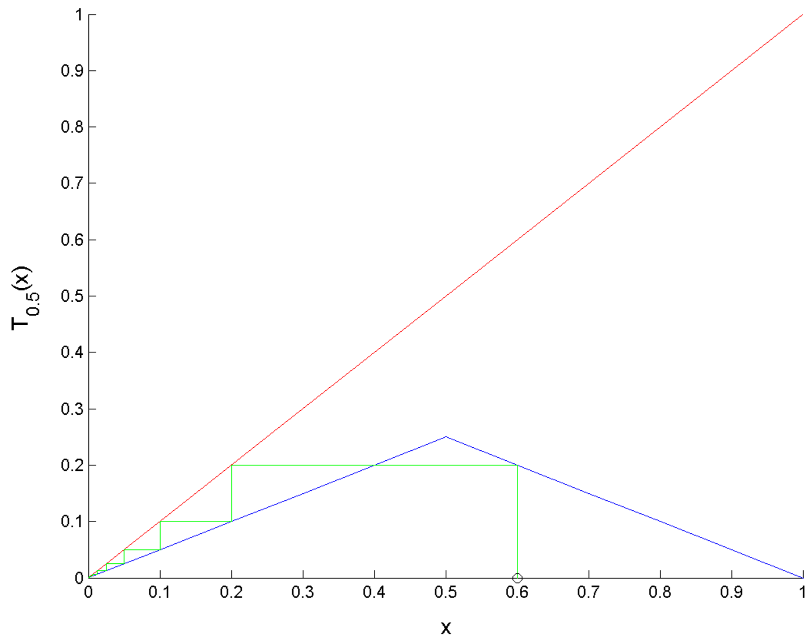
1: plot the map $\Phi(x)$ vs. x
2: plot the diagonal line $y = x$, we will call this line L
3: begin at initial condition x_0, draw a vertical line (either up or down) until the plot for $\Phi(x)$ is met
4: **while** you feel like iterating **do**
5: draw a horizontal line left or right until L is met at some point (p, p)
6: (p corresponds to the next point on the trajectory, output it if you wish)
7: draw vertically up or down until $\Phi(x)$ is met

Example 1.2.3:
Recall the tent map from Example 1.2.1 with $r = 0.5$, $x = 0.6$. We can use cobwebbing to visually verify the trajectory

$$0.6, 0.2, 0.1, 0.05, 0.025, \ldots.$$

as follows:

Figure 1.3: Tent map cobwebbing with $r = 0.5$, $x_0 = 0.6$.



In the above figure, we start the system off at $x = 0.6$. Following the cobwebbing algorithm, we move vertically upwards until we hit the map, which is the point that corresponds to $T_{0.5}(x)$. Shifting vertically to the line $y = x$ essentially sets the new x to the old x, and we repeat. ■

1.2.3 Stable and attracting fixed points.

A fixed point of a map $\Phi : X \to X$ is any point x^* such that $x^* \in X$ and $\Phi(x^*) = x^*$. Fixed points x^* of a map Φ are special in that once a trajectory lands on x^*, it will stay there forever. Note that not all maps have fixed points. For example, the map $\Phi : \mathbf{R} \to \mathbf{R}$ defined by $\Phi(x) = x + 1$ has no fixed points.

Example 1.2.4:
Recall the tent map, which was defined as

$$T_r(x) = \begin{cases} rx & 0 \le x \le \frac{1}{2}, \\ r(1-x) & \frac{1}{2} \le x \le 1. \end{cases}$$

When $r = 0$, the only fixed point of T_0 is 0 and indeed the map is quite boring. When $r = 1$, every point in $[0, 1/2]$ is a fixed point of T_1, and there are no others.

When $0 < r < 1$, consider the equation $x = r(1 - x)$. Solving for x yields $x = r/(1 + r)$, and then $r < 1$ yields $x < 1/2$. Thus, fixed points of the system must lie in $[0, 1/2)$. In this case, we note that every fixed point must be a solution to $x = rx$, and thus $x = 0$ is the only fixed point of the system when $0 < r < 1$.

Finally, in the case where $1 < r \leq 2$ the map has two fixed points: 0 and $r/(r+1)$ corresponding to the solutions of $x = rx$ and $x = r(1-x)$, respectively. ■

Example 1.2.5:
The fixed points of the logistic map, $G_\mu(x) = \mu x(1-x)$, are trivially 0 when $\mu = 0$. In order to find the fixed points when $0 < \mu \leq 4$, we solve for x in $x = \mu x(1-x)$. Thus, the fixed points are the roots of $\mu x^2 + (1-\mu)x$. The roots are given by

$$\frac{\mu - 1 \pm (1-\mu)}{2\mu},$$

which are 0 and $(\mu-1)/\mu$. Note that $(\mu-1)/\mu \notin [0,1]$ when $0 < \mu < 1$. Therefore, 0 is a fixed point of G_μ when $0 < \mu \leq 4$, and $(\mu-1)/\mu$ is a fixed point of G_μ when $1 \leq \mu \leq 4$. ■

A trajectory that lands on a fixed point becomes rather boring. However, numerous questions may be asked about the points close to a fixed point x^*. Do these nearby points get sucked in to x^*? Are they repelled by it? Is it possible that neither of these two cases occur? These questions are dealt with by what is loosely known as *stability analysis*. The analysis depends heavily on the definitions being used, and unfortunately there does not appear to be much of a convention in place. We will begin by defining what we mean for a fixed point x^* to be *stable* or *unstable*.

Definition 1.2.1:
A fixed point x^* of a map F is *stable* if for all positive ε, there exists a positive δ such that for all positive integers t, we have that for all points x,

$$\text{if } x \in X \text{ and } |x - x^*| < \delta \text{ then } |F^t(x) - x^*| < \varepsilon.$$

Definition 1.2.2:
A fixed point x^* of a map F is *unstable* if it is not stable.

More explicitly, this means that there exists a positive ε such that for all positive δ there exist a positive integer t and a point x in X for which

$$|x - x^*| < \delta \text{ and } |F^t(x) - x^*| \geq \varepsilon.$$

Example 1.2.6:
Let us return to the fixed points of the tent map. We have seen in example 1.2.4 that when $r = 0$, the only fixed point is 0. We will now show that 0 is stable. Let a positive ε be given, and for simplicity let us set $\delta = 1$. If x is any point in $[0, 1]$ such that $|x - 0| < \delta$, then for every positive integer t we have $T_0^t(x) = 0$, so that $|T_0^t(x) - 0| < \varepsilon$. Thus, 0 is a stable fixed point.

Let us move to case where $r = 1$. The fixed points of T_1 are all the points in $[0, 1/2]$, and they are all stable as well. First we will show that if x^* is any point in $[0, 1/2)$, then x^* is stable. To this end, let any positive ε be given. Set

$$\delta = \min\left\{\varepsilon, \frac{\frac{1}{2} - x^*}{2}, \frac{x^*}{2}\right\}.$$

Now every point x in $(x^* - \delta, x^* + \delta)$ remains fixed under T_1. Thus, for every positive integer t, we have

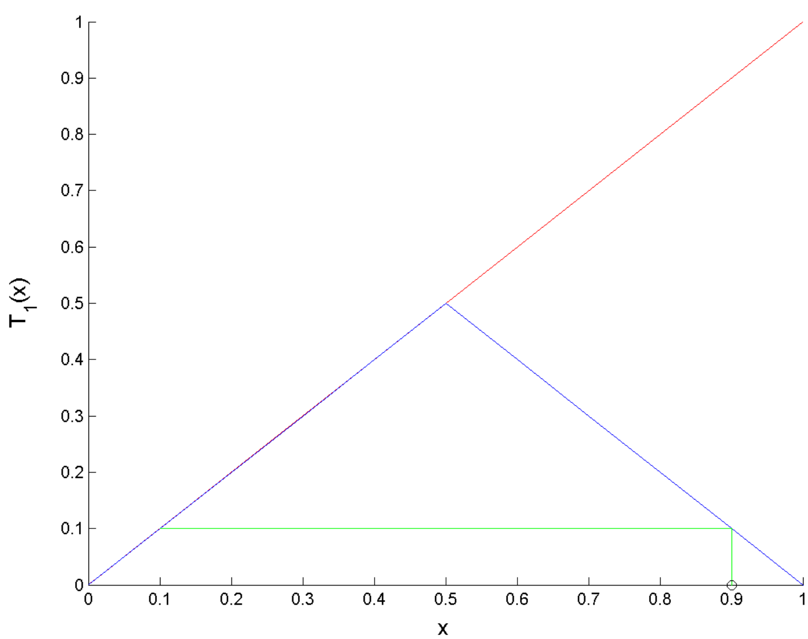
$$|T_1^t(x) - x^*| = |x - x^*| < \delta \leq \varepsilon$$

and we have stability. To see that the point $x^* = 1/2$ is stable, we let ε be given and set $\delta = \varepsilon$. Then every point in $(1/2 - \delta, 1/2]$ is fixed, and every point x such that $x \in (1/2, 1/2 + \delta)$ is sent to $(1/2 - \delta, 1/2)$ and remains fixed there. Thus, for all positive integers t we have

$$|T_1^t(x) - 1/2| < \delta = \varepsilon$$

and x^* is stable.

Figure 1.4: Tent map cobwebbing with $r = 1$, $x_0 = 0.9$.



Next, suppose r is such that $0 < r < 1$. We have seen that 0 is the only fixed point in this case.

Once again, it is stable. Let a positive ε be given and let us set $\delta = \min\{\varepsilon, 1/2\}$. Let x be any point such that $|x - 0| < \delta$. In this case, note that $T^t(x) < 1/2$ for all positive integers t, and therefore

$$|T_r^t(x) - 0| = T_r^t(x) = r^t x < x < \delta < \varepsilon,$$

for all positive integers t, so 0 is stable.

Finally, when $1 < r \leq 2$ there are two fixed points: 0 and $r/(r+1)$. Let us show that each of these points is unstable. We will deal with 0 first. Consider $\varepsilon = 1/2$ and let a positive δ be given. Consider any point x such that $0 < x < \min\{1/2, \delta\}$. Since $r > 1$, there must be some t for which

$$T_r^t(x) = r^t(x) > 1/2 = \varepsilon.$$

Lastly, let us show that $x^* = r/(r+1)$ is unstable. Let us choose $\varepsilon = x^* - 1/2$ and let us be given a positive δ. Take any x in $(x^*, x^* + \min\{\delta, 1 - x^*\})$ and let us denote the distance between x and x^* by d. Note that

$$\begin{aligned}
T_r(x) &= T_r(x^* + d) \\
&= r(1 - x^* - d) \\
&= r\left(1 - \frac{r}{r+1} - d\right) \\
&= \frac{r(r+1) - r^2 - rd(r+1)}{r+1} \\
&= \frac{r - rd - r^2 d}{r+1},
\end{aligned}$$

and therefore

$$|T_r(x) - x^*| = |\frac{r - rd - r^2 d}{r+1} - \frac{r}{r+1}| = dr. \tag{1.2}$$

Suppose that there exists no integer t for which $|T_r^t(x) - x^*| > \varepsilon$. This implies that the trajectory

$$x, T_r(x), T_r^2(x), \ldots$$

never enters $[0, 1/2)$. Since equation (1.2) holds for any x such that $x \in [1/2, 1]$, we may iterate it to yield $|T_r^t(x) - x^*| = dr^t$. However, note that $r > 1$ implies $dr^t \to \infty$ as $t \to \infty$, which is a contradiction. We conclude that there must be some positive integer t for which $|T_r^t(x) - x^*| > \varepsilon$, and therefore x^* is unstable. ■

The definition of stability requires that points within a certain neighborhood of a fixed point x^* cannot drift too far away from x^*. Our next notion is that of an *attracting* fixed point.

Definition 1.2.3:
A fixed point x^* of a map F is *attracting* if there exists a positive δ such that for all points x,

$$\text{if } x \in X \text{ and } |x - x^*| < \delta \text{ then } \lim_{t\to\infty} F^t(x) = x^*.$$

Definition 1.2.4:
A fixed point x^* of a map F is *non-attracting* if it is not attracting.

Let us be explicit and note that this means that for all positive δ there exists a positive ε such that for all positive integers t_0 there exist an integer t and a point x in X such that $t > t_0$ and

$$|x - x^*| < \delta \text{ and } |F^t(x) - x^*| \geq \varepsilon.$$

Example 1.2.7:
Let us return to the tent map. When $r = 0$, the only fixed point is 0 and we will show that it is attracting. Setting $\delta = 1$ and taking any point x such that $|x - 0| < \delta$, we have $T_0^t(x) = 0$ for all positive integers t, and therefore $\lim_{t\to\infty} T_0^t(x) = 0$.

Next, we will consider r such that $0 < r < 1$ and show that 0, which is the only fixed point, is attracting. Choose $\delta = 1/2$. Given any x such that $|x - 0| < \delta$, we have $T_r^t(x) = r^t x$ and since $0 < r < 1$ we have $\lim_{t\to\infty} T_r^t(x) = 0$. So 0 is attracting.

When $r = 1$, we have seen that the fixed points are all the points in $[0, 1/2]$. Let us pick any fixed point x^* in $[0, 1/2]$ and show that it is non-attracting. Let a positive δ be given. If $x^* \neq 1/2$, set

$$\varepsilon = \min\left\{\frac{1/2 - x^*}{3}, \frac{\delta}{3}\right\}$$

and choose $x = x^* + 2\varepsilon$. If $x^* = 1/2$, set

$$\varepsilon = \min\left\{\frac{x^*}{3}, \frac{\delta}{3}\right\}$$

and choose $x = x^* - 2\varepsilon$. Let a positive integer t_0 be given and set $t = t_0 + 1$. Then $|T_1^t(x) - x^*| = |x - x^*| > \varepsilon$. Therefore every fixed x^* in $[0, 1/2]$ is non-attracting. Note, however, that they are all stable as we have seen in example 1.2.6.

When $1 < r \leq 2$ we have two fixed points to deal with: 0 and $r/(1 + r)$. We will show that both of these points are non-attracting. Let us deal with 0 first. Let a positive δ be given and choose $\varepsilon = 1/2$. Let t_0 be given and let us choose

$$x = \min\{\delta, \frac{1}{2r^{t_0}}\}.$$

Note that $T_r^t(x) \leq 1/2$ for all positive integers t such that $t \leq t_0$. Since $T_r^{-1}(0) = 1$, we can note for later that $T_r^{t_0}(x) \neq 0$. Now suppose that $|T_r^t(x) - 0| < \varepsilon = 1/2$ for all positive integers t such that $t > t_0$. Then by how the tent map is defined, we know that

$$T_r^t(x) = r^{t-t_0} T_r^{t_0}(x)$$

for all integers t such that $t > t_0$. Since $r > 1$ and $T_r^{t_0}(x) \neq 0$, we have $T_r^t(x) \to \infty$ as $t \to \infty$ and this is a contradiction. Therefore, there must be some t for which $t > t_0$ and $|T_r^t(x) - 0| \geq \varepsilon$.

Finally, let us show that the fixed point $r/(1+r)$ is non-attracting. We will repeat the proof from example 1.2.6 with some slight modifications. Let us be given a positive δ. Let us choose $\varepsilon = x^* - 1/2$. Let a positive integer t_0 be given. Let us choose a point x such that

$$x = x^* + \min\{\delta, \frac{x^* - 1/2}{r^{t_0}}\}.$$

Let us denote the distance between x and x^* by d. Note that $d = |x - x^*| \leq (x^* - 1/2)/r^{t_0}$. We have

$$\begin{aligned} T_r(x) &= T_r(x^* + d) \\ &= r(1 - x^* - d) \\ &= r\left(1 - \frac{r}{r+1} - d\right) \\ &= \frac{r(r+1) - r^2 - rd(r+1)}{r+1} \\ &= \frac{r - rd - r^2 d}{r+1}, \end{aligned}$$

and therefore

$$|T_r(x) - x^*| = |\frac{r - rd - r^2 d}{r+1} - \frac{r}{r+1}| = dr. \tag{1.3}$$

Note that equation (1.3) holds for every $x \in [1/2, 1]$.

We claim that $T_r^t(x) \geq 1/2$ for all positive integers t such that $t \leq t_0$. Suppose this is not the case. Let t' be the smallest positive integer t such that $T_r^t(x) < 1/2$. Then $T_r^t(x) \geq 1/2$ for all positive integers t such that $t < t'$ and we may iterate equation (1.3) to get

$$|T_r^{t'}(x) - x^*| = dr^{t'} \leq \frac{(x^* - 1/2)}{r^{t_0}} r^{t'} \leq x^* - 1/2,$$

which is a contradiction.

Now suppose that there does not exist a positive integer t for which $|T_r^t(x) - x^*| > \varepsilon$. This implies that the trajectory

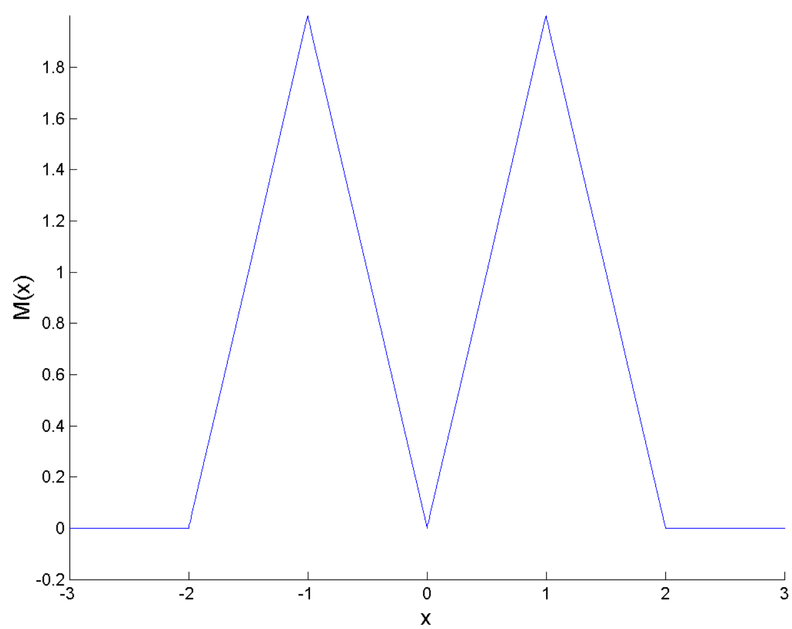
$$x, T_r(x), T_r^2(x), \ldots$$

never enters $[0, 1/2)$. Again, since equation (1.3) holds for every $x \in [1/2, 1]$, we may iterate it to yield $|T_r^t(x) - x^*| = dr^t$. Since $r > 1$ implies $dr^t \to \infty$ as $t \to \infty$, we conclude that there must be some integer t (which is greater than t_0) for which $|T_r^t(x) - x^*| > \varepsilon$, and therefore x^* is unstable.■

Example 1.2.8:
The case where $r = 1$ in the preceding example showed us that it is possible to have a fixed point that is stable and non-attracting. Is it possible to have a fixed point that is attracting and unstable? Let us try to construct one. For simplicity we will try to construct a map with a fixed point at 0 that is unstable and attracting. Bearing in mind what the definitions mean, we need a map that that forces every point near 0 away, and eventually maps it back to x. Consider the map $M : \mathbf{R} \to \mathbf{R}$ defined by

$$M(x) = \begin{cases} 2x + 4 & -2 \leq x \leq -1, \\ -2x & -1 \leq x \leq 0, \\ 2x & 0 \leq x \leq 1, \\ 4 - 2x & 1 \leq x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 1.5: M(x) vs. x.



Indeed 0 is an unstable fixed point of M as we will now show. Let us choose $\varepsilon = 1$ and let a positive δ be given. Let us take t to be any integer such that $t > 1 - \lg \delta$. Let us also choose $x = \min\{\delta/2, 1\}$. Note that $x \in [0, 1]$ so that we have $M^t(x) = 2^t x > 2^{1-\lg \delta}(\delta/2) = 1$. Therefore 0 is unstable.

Thus we have a map that sends points near 0 away from it. Must they come back to 0? Unfortunately, if we take any point in $[-2, 2]$, its trajectory will forever remain in that interval. Let us consider only the positive domain of the map and note that

$$[0,1] \xrightarrow{2x} [0,2]$$

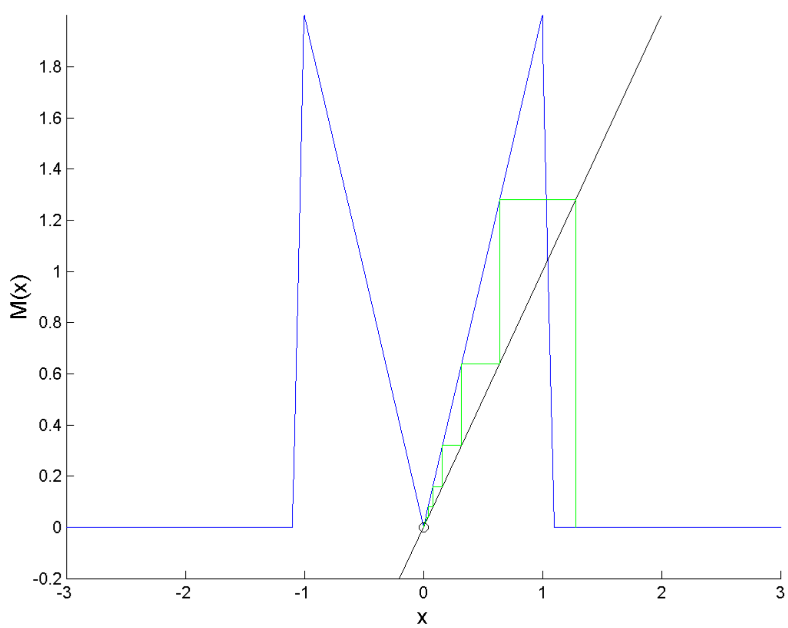
and

$$[1,2] \xrightarrow{4-2x} [0,2].$$

We may try to remedy this situation by increasing the slope of the lines in $[-2,-1]$ and $[1,2]$:

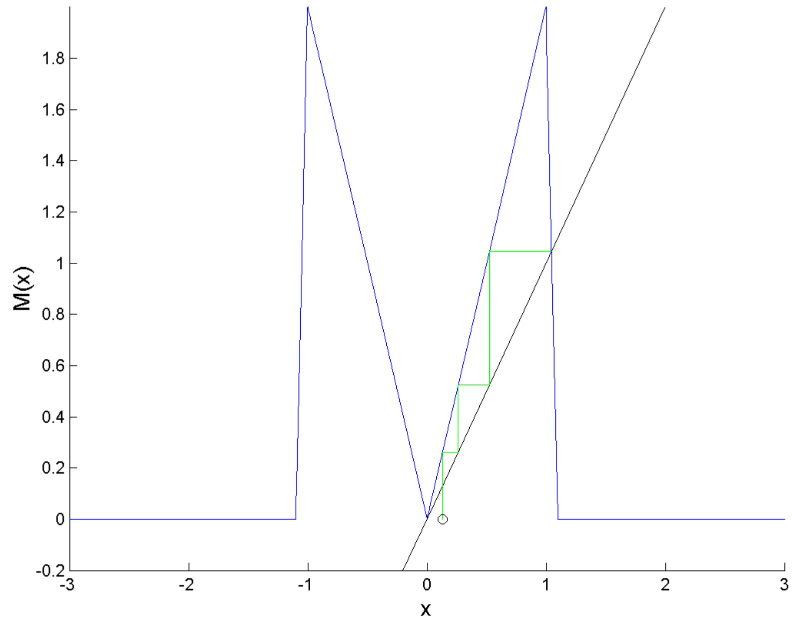
$$M(x) = \begin{cases} 20x+22 & -\frac{22}{20} \leq x \leq -1, \\ -2x & -1 \leq x \leq 0, \\ 2x & 0 \leq x \leq 1, \\ 22-20x & 1 \leq x \leq \frac{22}{20}, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 1.6: M(x) cobweb.



Points near 0 may now leave $[-2,2]$ and return. However, this is still not good enough, since we need this to hold for all points. Consider, say, the point $11/84$.

Figure 1.7: Unfortunate fixed point.



Its trajectory is

$$\frac{11}{84}, \frac{11}{42}, \frac{11}{21}, \frac{22}{21}, \frac{22}{21}, \frac{22}{21}, \ldots.$$

There is an unfortunate fixed point at $22/21$ on the line $22 - 20x$. This squashes the possibility of 0 being attracting, since we need *all* points near it to get sucked back in.

We may try to remedy this situation by creating a discontinuity at $22/21$ as follows:

Let the map $M : \mathbf{R} \to \mathbf{R}$ be defined by

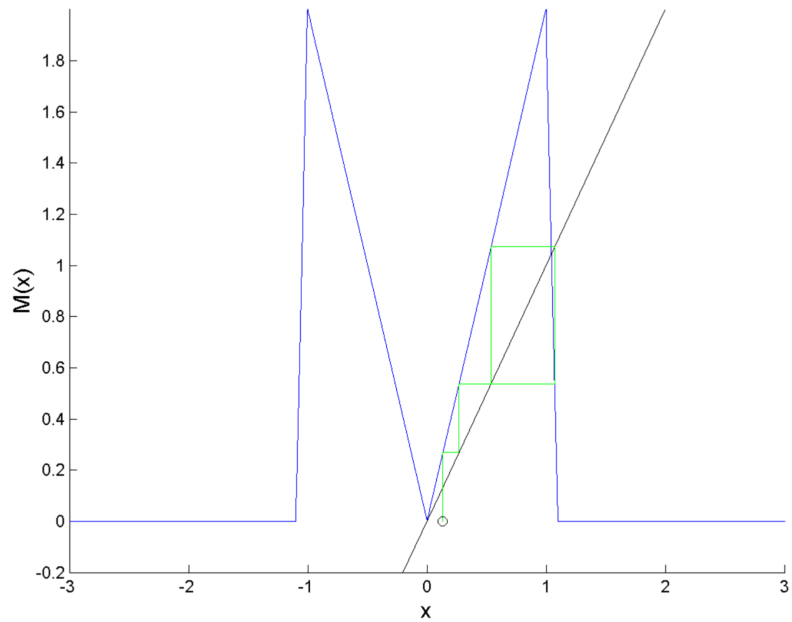
$$M(x) = \begin{cases} 20x + 22 & -\frac{22}{20} \leq x \leq -1, \\ -2x & -1 \leq x \leq 0, \\ 2x & 0 \leq x \leq 1, \\ 22 - 20x & x \in [1, 22/20] \setminus \{\frac{22}{21}\}, \\ 0 & \text{otherwise.} \end{cases}$$

Our proof that 0 is an unstable fixed point still holds, and the problematic trajectory of the point $11/84$ is now

$$\frac{11}{84}, \frac{11}{42}, \frac{11}{21}, 0, 0, 0, \ldots.$$

Unfortunately we are not off the hook yet. Consider, say, the point 11/82.

Figure 1.8: Unfortunate periodic points.



Its trajectory is

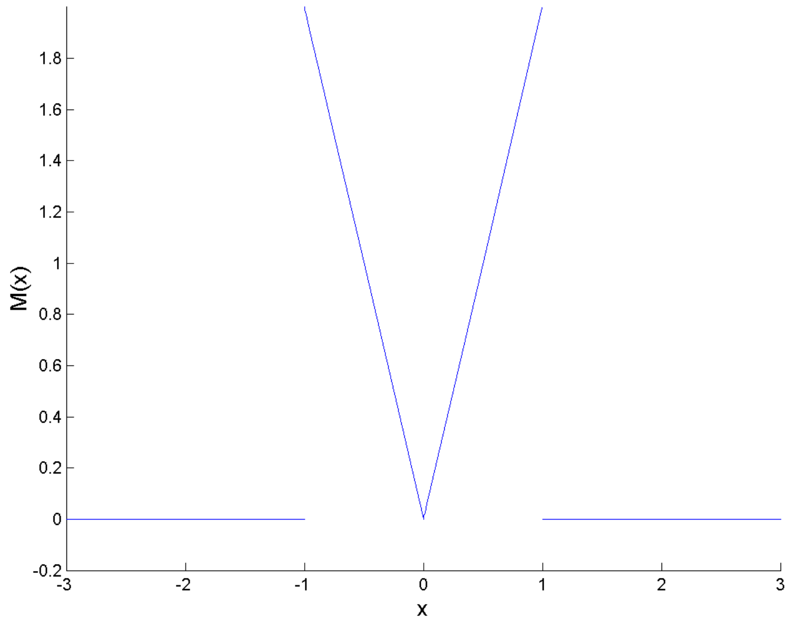
$$\frac{11}{82}, \frac{11}{41}, \frac{22}{41}, \frac{44}{41}, \frac{22}{41}, \frac{44}{41}, \frac{22}{41}, \frac{44}{41}, \ldots$$

and thus we have a point that leaves the neighborhood of 0 and never gets mapped back to 0.

At this point, rather than trying to patch each hole as it becomes apparent, let us simplify our map to

$$M(x) = \begin{cases} 2x & 0 \le x \le 1, \\ -2x & -1 \le x \le 0, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 1.9: Problem points removed.



Our fixed point 0 is still unstable and we are left to show that it is attracting. Let us set $\delta = 1$ and let x be any point in $(0, 1)$.

If $1/2 \leq x < 1$ then $M^t(x) = 0$ for all integers t such that $t > 1$, which allows us to conclude that $\lim_{t\to\infty} M^t(x) = 0$.

If $0 < x < 1/2$, then we claim that $1/2 \leq M^t(x) < 1$ for some positive integer t. Supposing this is not the case, we note that $M^t(x) = 2^t x$ and this yields a contradiction since $2^t x \to \infty$ as $t \to \infty$. Thus there exists some positive integer t for which $1/2 \leq M^t(x) < 1$ and $\lim_{t\to\infty} M^t(x) = 0$ by the preceding paragraph. ■

Although our attempts at constructing a continuous map with an unstable attracting fixed point were fruitless, we have seen that such points may exist for discontinuous maps. In fact, [5] contains a proof of the following Theorem, which we will not prove here.

Theorem 1.2.1:
Let x^* be an attracting fixed point of a continuous map $f : I \to \mathbf{R}$, where I is an interval. Then x^* is stable.

Proof:
See [5] or [10]. □

We are now ready to state a theorem that allows us to easily deduce the stability of a fixed point, in most cases.

Theorem 1.2.2:
Let F be a one-dimensional differentiable map ($F : X \to X$ and X is an interval of $\mathbf{R}$) and let x^* be a fixed point of F such that F' is continuous at x^*. We have the following:

1. *If $|F'(x^*)| < 1$, then x^* is stable and attracting.*
2. *If $|F'(x^*)| > 1$ then x^* is unstable.*

Proof (of 1.):
Let a positive ε be given and let us set

$$\varepsilon' = \min\left\{\varepsilon, \frac{1 - |F'(x^*)|}{2}\right\},$$

so that $|F'(x^*)| + \varepsilon' < 1$.

Since F' is continuous at x^*, we know that there is a positive δ such that

$$F'(x^*) - \varepsilon' < F'(y) < F'(x^*) + \varepsilon' \tag{1.4}$$

for all y such that

$$x^* - \delta < y < x^* + \delta. \tag{1.5}$$

Now consider any trajectory starting at x, where x is such that $0 < |x - x^*| < \min\{\delta, \varepsilon'\}$.

By the Mean Value Theorem, there exists a z such that z is between x and x^* and such that

$$\begin{aligned} |F(x) - x^*| &= |F(x) - F(x^*)| \\ &= |F'(z)||x - x^*|. \end{aligned} \tag{1.6}$$

Since z is between x and x^*, and x is less than δ far away from x^*, we know that z is less than δ far away from x^*. Thus, (1.5) holds with z in place of y and we can apply (1.4) to get that

$$-1 < F'(x^*) - \varepsilon < F'(z) < F'(x^*) + \varepsilon < 1,$$

so that $|F'(z)| < M < 1$, where $M = F'(x^*) + \varepsilon$. Thus, we have

$$|F(x) - x^*| < M|x - x^*|.$$

In particular, $|F(x) - x^*| < |x - x^*|$ and we may iterate equation (1.6) (applying the Mean Value Theorem, to $F(x)$ and x^* instead of x and x^*) to get

$$|F^2(x) - x^*| < M^2|x - x^*|$$

by iteration. Proceeding in this fashion allows us to conclude that for every positive integer t, we have
$$|F^t(x) - x^*| < M^t|x - x^*|$$
by iteration.

Since $M < 1$ we can conclude that $|F^t(x) - x^*| < |x - x^*| < \varepsilon$ for all positive integers t, so that x^* is stable. Furthermore, since $M < 1$ we may conclude that $|F^t(x) - x^*| \to 0$ as $t \to \infty$. Thus, x^* is attracting. □

Proof (of 2.):
Suppose that $|F'(x^*)| > 1$. Let us set $\varepsilon_c = |F'(x^*) - 1|/2$. Since F' is continuous at x^* there is a positive δ_c such that $F'(x^*) - \varepsilon_c < F'(y) < F'(x^*) + \varepsilon_c$ for all y such that $x^* - \delta_c < y < x^* + \delta_c$. With this in mind, we will set $\varepsilon = \delta_c$ and show that for all positive δ, there is a point x and positive integer t such that $0 < |x - x^*| < \delta$ and $|F^t(x) - x^*| \geq \varepsilon$.

To this end, let a positive δ be given. Without loss of generality, we may assume that $\delta < \delta_c$. Let x be any point such that $0 < |x - x^*| < \delta$. By the Mean Value Theorem, there is a point z between x and x^* such that
$$F(x) - F(x^*) = F'(z)(x - x^*).$$
Setting $M = |F'(x^*) - \varepsilon_c|$, we have
$$|F'(z)| > |F'(x^*) - \varepsilon_c| = M > 1$$
by the above continuity arguments. We now have
$$|F(x) - F(x^*)| > |M||(x - x^*)|. \tag{1.7}$$
If $|F(x) - F(x^*)| \geq \varepsilon$ then we are done. If not, then we may iterate (1.7) with $F(x)$ in place of x (we are allowed to do this since in this case we must have $|F(x) - F(x^*)| < \varepsilon = \delta_c$). Proceeding in this manner we get
$$|F^t(x) - F^t(x^*)| \geq |M^t||(x - x^*)|,$$
and therefore there must be some t such that $|F^t(x) - x^*| \geq \varepsilon$, since $M > 1$. □

Note that Theorem 1.2.2 does not tell us how to deal with a fixed point x^* of a map F for which $|F'(x^*)| = 1$. Such a fixed point is called non-hyperbolic. The stability of non-hyperbolic fixed points can be studied by looking at higher derivatives (and Schwarzian[1] derivatives in particular). We will not explore such techniques here. See [10] for a detailed account.

Example 1.2.9:
Let us apply Theorem 1.2.2 to the fixed points of the tent map, which we have already studied. Note that $T_r'(x^*) = |r|$ for every fixed point x^* (and indeed every point in $[0, 1]$), so that fixed points are stable and attracting when $r < 1$ and unstable when $r > 1$. This confirms our previous remarks. ■

[1]The Schwarzian derivative of a function f is $\frac{f'''(x)}{f'(x)} - \frac{3}{2}\left[\frac{f''(x)}{f'(x)}\right]^2$

Example 1.2.10:
Recall that the fixed points of the logistic map, $G_\mu(x) = \mu x(1-x)$, are 0 when $0 \leq \mu \leq 4$ and $(\mu-1)/\mu$ when $1 \leq \mu \leq 4$. The derivative of the map at a point x is $\mu - 2\mu x$, so that 0 is stable and attracting for $0 \leq \mu < 1$ and unstable for $1 < \mu \leq 4$. Furthermore, $(\mu-1)/\mu$ is stable and attracting for $1 < \mu < 3$ and unstable when $3 < \mu \leq 4$. We will study the case $3 < \mu \leq 4$ in more detail in the next section. ■

Now that we have an understanding of what the definitions in this section mean, here are some comments on some of the variations that may be found in the literature. Our definitions are borrowed from Elaydi (reference [10]), although he adds the term *asymptotically stable* to refer to fixed points that are both stable and attracting. Our definitions are also similar to those found in [9].

In reference [1], Alligood et al. define an *attracting* fixed point (or *sink*) in the same way we did, and uses *source* or *repelling fixed point* to refer to our unstable fixed points.

The text by Devaney ([8]) states that a fixed point x^* of a map f for which $|f'(x^*)| < 1$ is an *attracting* point or a *sink*. Thus, his attracting fixed point (or sink) is a stable and attracting fixed point for us. Furthermore, he defines a *repellor* or *source* to be a fixed point for which $|f'(x^*)| > 1$. This corresponds to our unstable fixed points.

Similarly, Lynch ([22]) calls a fixed point *stable* if $|f'(x^*)| < 1$, and an *unstable* fixed point is one for which $|f'(x^*)| > 1$. This corresponds to our stable and attracting fixed points and unstable fixed points, respectively.

Schuster ([34]) calls our attracting fixed point a *locally stable* fixed point, and uses *unstable* to refer to points that we call non-attracting.

Scheinerman ([33]) calls our stable and attracting fixed points *stable* fixed points, and uses *marginally stable* to denote fixed points that, in our language, are both stable and non-attracting. His *unstable* fixed points are those that are neither *stable* nor *marginally stable*, which make it equivalent to our notion of an unstable fixed point.

As we can see, the literature is riddled with inconsistencies. Stability and attraction are two separate notions, and we choose to keep them as such.

Similar proofs of Theorem 1.2.2 can be found in most textbooks on dynamical systems, for example [1], [10] and [33].

1.2.4 Periodic trajectories.

Let us now consider trajectories with more than one repeating point.

Definition 1.2.5:
A trajectory

$$x, \Phi(x), \Phi^2(x), \ldots$$

under a map $\Phi : X \to X$ is *eventually periodic* if there exist a nonnegative integer k and a positive integer p such that for all integers t such that $t \geq k$

$$\Phi^{p+t}(x) = \Phi^t(x).$$

The smallest positive integer p for which this holds is the *period of the trajectory*, and the set of points

$$\{x_k, \Phi(x_k), \Phi^2(x_k), \ldots, \Phi^{p-1}(x_k)\},$$

where $x_k = \Phi^k(x)$, is called the *periodic orbit* of the trajectory.

Note that our fixed points from the previous section are simply periodic orbits with period 1. Let us also note here the allowance for k to be nonzero in the definition. It would be tempting to simply require that $\Phi^p(x_i) = x_i$ for some positive integer p and all nonnegative integers i. However, this would exclude trajectories such as

$$0, 1, 2, 3, 2, 3, 2, 3, \ldots$$

from being eventually periodic, which is not what we want. Thus, letting k be positive gives us the *eventually* part of the definition. Such trajectories are possible since the maps under our consideration need not be one-to-one.

We should note that

$$2, 3, 2, 3, 2, 3, \ldots$$

is eventually periodic as well.

What would it mean for us to ask if a periodic orbit is stable or attracting? Let

$$\{x_0, \Phi(x_0), \Phi^2(x_0), \ldots, \Phi^{p-1}(x_0)\}$$

be the periodic orbit in question. We want to know what happens to a nearby point after p iterates of the map Φ. Thus, in order to determine if the periodic orbit is stable, attracting, unstable or non-attracting, we can view Φ^p as a new map Ψ and ask ourselves about x_0 as a fixed point under Ψ. Let us write $x_i = \Phi^i(x_0)$ for every nonnegative integer i. If Φ is a differentiable map, then by the chain rule we have

$$\begin{aligned}\Psi'(x_0) &= \Phi'(\Phi^{p-1}(x_0)) \times {\Phi^{p-1}}'(x_0) \\ &= \Phi'(x_{p-1}) \times \Phi'(\Phi^{p-2}(x_0)) \times {\Phi^{p-2}}'(x_0) \\ &= \vdots \\ &= \Phi'(x_{p-1}) \times \Phi'(x_{p-2}) \times \cdots \times \Phi'(x_0),\end{aligned}$$

so that x_0 is stable and attracting (under Ψ) if $\prod_{i=0}^{p-1} |\Phi'(x_i)| < 1$, by Theorem 1.2.2. Thus,

$$P = \{x_0, \Phi(x_0), \Phi^2(x_0), \ldots, \Phi^{p-1}(x_0)\}$$

is a stable and attracting periodic orbit of Φ if $\prod_{i=0}^{p-1} |\Phi'(x_i)| < 1$. If $\prod_{i=0}^{p-1} |\Phi'(x_i)| > 1$, then P is an *unstable periodic orbit*.

Definition 1.2.6:
A trajectory

$$x, \Phi(x), \Phi^2(x), \ldots$$

under a map $\Phi : X \to X$ is called *asymptotically periodic* if there exists a positive integer s and an eventually periodic trajectory

$$y_0, y_1, y_2, \ldots$$

of points in X with periodic orbit

$$\{y_t, y_{t+1}, \ldots, y_{t+p-1}\}$$

such that

$$\lim_{k\to\infty} |\Phi^k(x_s) - y_{t+k}| = 0.$$

In words: an asymptotically periodic trajectory gets arbitrarily close to an eventually periodic trajectory. Note that an eventually periodic trajectory is also asymptotically periodic.

Example 1.2.11:
The tent map with $r = 2$ has many eventually periodic orbits. For example,

$$\frac{1}{10}, \frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{2}{5}, \frac{4}{5}, \ldots$$

is eventually periodic. The period is unstable since

$$\left| T_2'(\frac{2}{5}) \times T_2'(\frac{4}{5}) \right| = 4 > 1.$$

We may also note that when $r = 1/2$, the trajectory

$$0.25, 0.1250, 0.0625, 0.0313, \ldots$$

is asymptotically periodic (with period 1) since it approaches 0 but never actually lands on it. ■

Example 1.2.12:
We have seen that the logistic map, $G_\mu(x) = \mu x(1-x)$, has two fixed points: 0 and $(\mu-1)/\mu$. Furthermore, we know from example 1.2.10 that 0 is asymptotically stable for $0 \le \mu < 1$, and that $(\mu-1)/\mu$ is asymptotically stable for $1 < \mu < 3$. Once μ becomes slightly larger than 3, the fixed points are both unstable since

$$G'_\mu(0) = \mu > 3 > 1$$

and

$$G'_\mu((\mu-1)/\mu) = \mu - 2\mu(\mu-1)/\mu = 2-\mu < -1.$$

Where do trajectories go? When $\mu = 3$, a periodic trajectory of period 2 that is stable and attracting is born. In order to study the period 2 trajectories, we need to consider the second iterate of the map, which is

$$G^2_\mu(x) = -\mu^2 x(x-1)(1-\mu x+\mu x^2).$$

In order to find the periodic trajectory (of period 2) of G_μ, we must find the fixed points of G^2_μ, and to do that we need to find the roots of $-\mu^2 x(x-1)(1-\mu x+\mu x^2) - x$. This can be rewritten as

$$-x(\mu x + 1 - \mu)(\mu^2 x^2 - (\mu^2+\mu)x + \mu + 1),$$

which immediately yields two of the four roots, namely 0 and $(\mu-1)/\mu$, which are simply the fixed points of G_μ. We are left to consider the roots of $\mu^2x^2 - (\mu^2+\mu)x + \mu + 1$, which are

$$x_2^+ = \frac{\frac{\mu}{2} + \frac{1}{2} + \sqrt{\frac{\mu^2-2\mu-3}{2}}}{\mu},$$
$$x_2^- = \frac{\frac{\mu}{2} + \frac{1}{2} - \sqrt{\frac{\mu^2-2\mu-3}{2}}}{\mu}.$$

Note that these numbers are complex if $\mu < 3$, which confirms that the period 2 trajectory only comes into play once $\mu = 3$. To check stability, we first note that the derivative of G^2_μ is $-4\mu x^3 + 6\mu x^2 - (2\mu^3 + 2\mu^2) + \mu^2$, which yields

$$G^{2\,\prime}_\mu(x_2^+) = -\mu^2 + 2\mu + 4,$$
$$G^{2\,\prime}_\mu(x_2^-) = -\mu^2 + 2\mu + 4.$$

At $\mu = 3$, we have $G^{2\,\prime}_\mu(x_2^+) = G^{2\,\prime}_\mu(x_2^-) = 1$. A small increase in μ causes this number to drop below 1, making it a stable and attracting orbit, until μ is large enough so that $G^{2\,\prime}_\mu(x_2^+) = G^{2\,\prime}_\mu(x_2^+) = -1$. Without going into the details of the analysis (which are similar to the preceding case, see [23] or [37] for details), the value of μ at which the period 2 trajectory becomes unstable is $\mu = 1 + \sqrt{6}$.

Figure 1.10: G_μ^2 with $\mu = 3$.

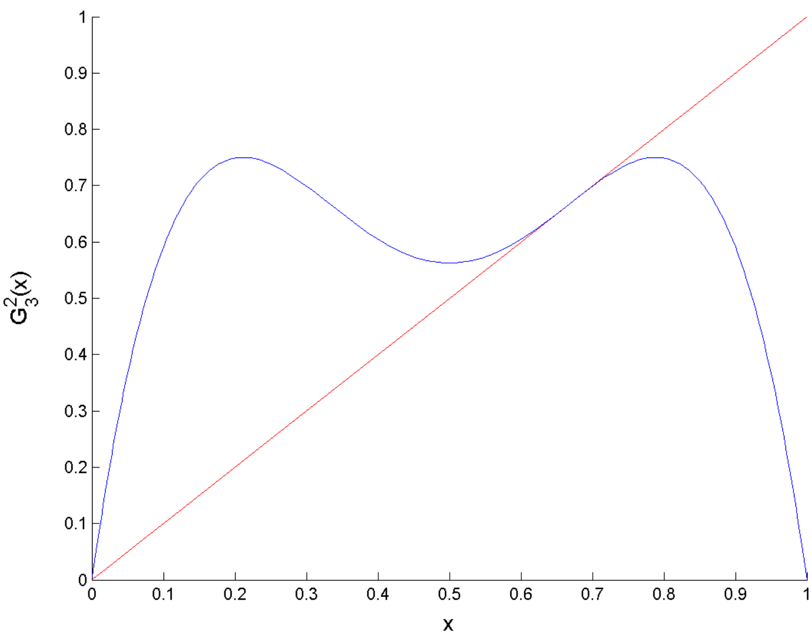
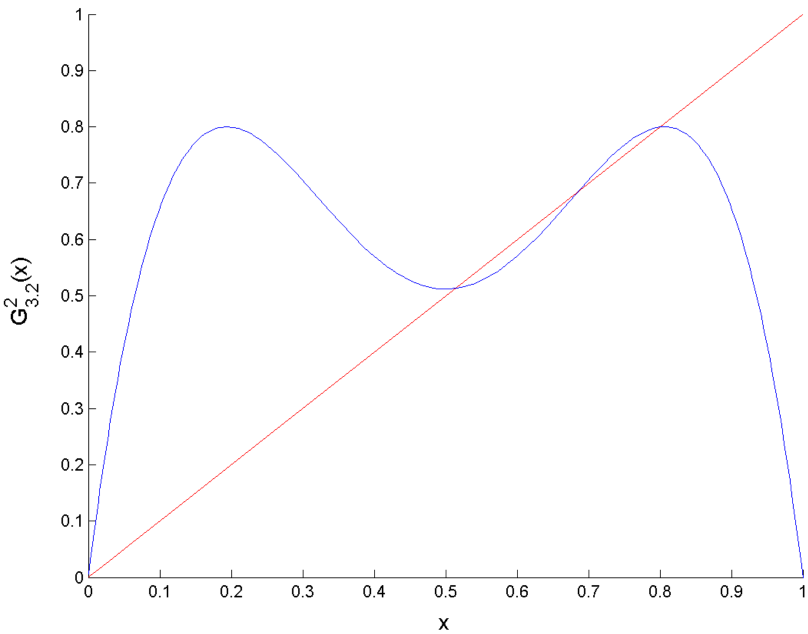


Figure 1.11: G_μ^2 with $\mu = 3.2$.



What happens for $\mu > 1 + \sqrt{6}$? The story seems to repeat itself. Looking back to when $\mu = 3$, the stable and attracting fixed point became unstable and gave rise to a period 2 trajectory that has a stable and attracting orbit. Similarly, at $\mu = 1 + \sqrt{6}$, the period 2 orbit becomes unstable, with each of the two points giving rise to two new points, creating a period 4 trajectory, which has a stable and attracting orbit.

This process continues. As μ increases, the system gives rise to a stable and attracting periodic orbit of period 2^k at some critical value μ_k. This goes until until a critical value μ_∞ is reached where G_{μ_∞} has no orbits that are stable and attracting. ■

The previous example is merely a small sample of the elegant properties of the logistic map. As we have mentioned in the example, May goes into further details in [23]. The definitions in this section are modified from the notions found in [1] and [3], but are common enough to be found in most texts.

1.2.5 The Lyapunov exponent of a one-dimensional map.

Let us return to the general setting where $\Phi : X \to X$ is any map (and not necessarily differentiable), and $X \subseteq \mathbf{R}$. For every point x in the interior of X, the *local Lyapunov exponent* of Φ at x is defined as the limit (if it exists)

$$\lim_{\delta\to 0} \ln \frac{|\Phi(x+\delta)-\Phi(x)|}{\delta};$$

the *global Lyapunov exponent* of Φ at x is defined as the limit (if it exists)

$$\lim_{t\to\infty} \frac{1}{t} \lim_{\delta\to 0} \ln \frac{|\Phi^t(x+\delta)-\Phi^t(x)|}{\delta}.$$

If the global Lyapunov exponent of Φ at x equals λ then, for all sufficiently large values of t for all sufficiently small (with respect to the value of t) values of δ, we have

$$\frac{|\Phi^t(x+\delta)-\Phi^t(x)|}{\delta} \approx e^{\lambda t}.$$

For this reason, emphasis is usually placed on whether the Lyapunov exponent is positive or negative.

1.2.6 The Lyapunov exponent of a one-dimensional differentiable map.

Let $F : X \to X$ be a differentiable map and let x be any point in X, where $X \subseteq \mathbf{R}$. Since

$$F(x+\delta) - F(x) \; = \; F'(x)\delta + o(|\delta|) \quad \text{as } \delta \to 0,$$

the local Lyapunov exponent of a differentiable map Φ at a point x equals

$$\ln|\Phi'(x)|$$

and the global Lyapunov exponent of Φ at x, which we will denote by $\lambda(x)$, equals

$$\lim_{n\to\infty} \frac{1}{n} \ln|\Phi^{n\prime}(x)|, \tag{1.8}$$

if the limit exists. Let us note that

$$\begin{aligned} \frac{1}{n}\ln|\Phi^{n\prime}(x)| &= \frac{1}{n}\ln|\Phi'(x)\Phi'(\Phi^2(x))\Phi'(\Phi^3(x))\cdots\Phi'(\Phi^{n-1}(x))| \\ &= \frac{1}{n}\sum_{t=0}^{n-1}\ln|\Phi'(\Phi^t(x))|. \end{aligned}$$

Therefore,

$$\lambda(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1}\ln|\Phi'(\Phi^t(x))|, \tag{1.9}$$

if the limit exists. This means that the global Lyapunov exponent of Φ at the point x is the average of the local Lyapunov exponents of Φ at points $\Phi^0(x)$, $\Phi^1(x)$, $\Phi^2(x)$, $\ldots$ of the trajectory of x.

Note that the Lyapunov exponent of the trajectory $x_0, x_1, x_2, \ldots$ does not exist if there is some x_n such that $\Phi'(x_n) = 0$. We may also note that the limit in (1.9) is guaranteed to exist under certain mild restrictions by the Birkhoff Ergodic Theorem (see [4] and [7]), which we will not cover here. For future reference, the Birkhoff Ergodic Theorem is the one-dimensional version of a more general theorem known as Oseledets' Multiplicative Ergodic Theorem, which we will mention when we arrive at Section 1.3.6.

Example 1.2.13:
For the tent map, $|T_r'(x)| = |r|$ everywhere except when $x = 1/2$, where the derivative does not exist. Let us consider a point x such that the trajectory $x, T_r(x), T_r^2(x), \ldots$ never hits $1/2$ (letting x be an irrational number on the unit interval when r is rational is one such way to ensure this). In this case, equation (1.9) tells us that the Lyapunov exponent of the trajectory is simply $\ln r$, so that the Lyapunov exponent is positive when $\ln(r) > 0$, or simply when $r > 1$.

Figure 1.12: Tent map cobwebbing with $r = 1.9$, $x_0 = 0.3$.

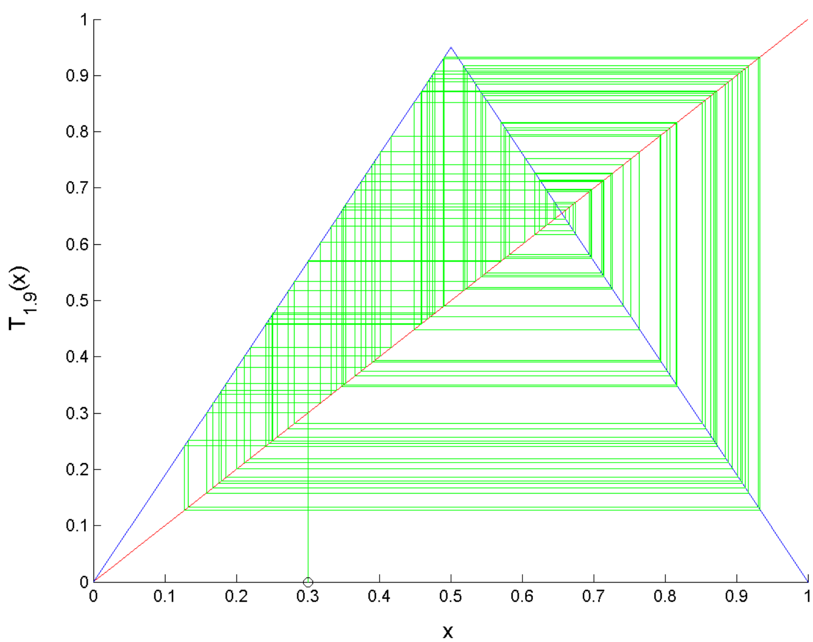
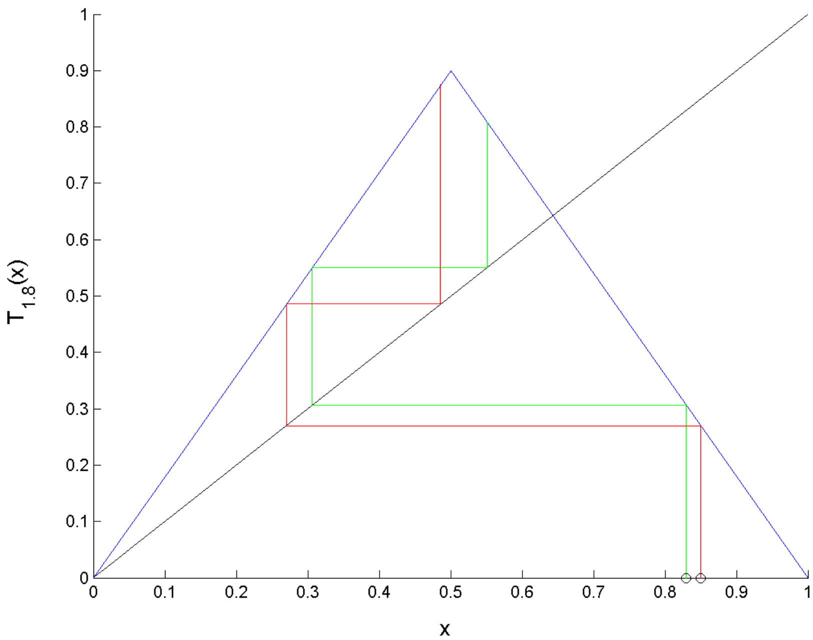


Figure 1.13: Separation of initially nearby points when $r = 1.8$, $x_0 = 0.83, 0.85$.



These two figures show the effect of a positive Lyapunov exponent. In particular, the second figure displays what is commonly referred to as sensitive dependence on initial conditions or the butterfly effect: "Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?". This quote is often misattributed to Edward Lorenz who initially had a similar quote in [21]: "one flap of a sea gull's wings would be enough to alter the course of the weather forever." In fact, a similar quote can be found much earlier involving grasshoppers. See [14] for a more detailed account of the history of the butterfly effect. ■

Note that in general the calculation of the Lyapunov exponent of the trajectory of a map is not so simple. The tent map provided for a particularly simple calculation since the derivative of the map is the same at every point in the phase space (with the exception of 1/2).

Example 1.2.14:
It would be wrong to think that fixed points of a map must have a Lyapunov exponent equal to 0. This is not the case. Consider the tent map $T_2(x)$ and the fixed point 2/3. The (fixed) trajectory $2/3, 2/3, 2/3, \ldots$ has a Lyapunov exponent of $\ln 2$. Similarly, a periodic trajectory need not have a Lyapunov exponent equal to 0. The periodic trajectory $2/5, 4/5, 2/5, 4/5, \ldots$ of the tent map $T_2(x)$ has a Lyapunov exponent of $\ln 2$. We must not be fooled by the fact that nearby points on that same periodic trajectory will not be separating or contracting. Nearby points not on the trajectory must be taken into consideration. ■

1.2.7 Conjugacy.

Definition 1.2.7:
A function $f : A \to B$ is a *homeomorphism* if it is continuous, bijective, and its inverse is continuous.

We may now introduce the notion of *conjugacy*, which will allow us to relate two dynamical systems to each other.

Definition 1.2.8:
Let $X \subseteq \mathbf{R}$, $Y \subseteq \mathbf{R}$ and let $\alpha : A \to A$ and $\beta : B \to B$ be two maps. We say that α and β are *conjugate* if there exists a homeomorphism $f : B \to A$ such that $\alpha(f(x)) = f(\beta(x))$. We call f a *conjugation* between α and β and write $\alpha \overset{f}{\sim} \beta$.

Theorem 1.2.3:
The conjugacy relation is an equivalence relation between maps. Namely, if $\alpha : A \to A$ and $\beta : B \to B$ and $\gamma : C \to C$ are maps then

1. $\alpha \sim \alpha$,
2. $\alpha \sim \beta \quad \Rightarrow \quad \beta \sim \alpha$,
3. $\alpha \sim \beta$ *and* $\beta \sim \gamma \quad \Rightarrow \quad \alpha \sim \gamma$.

Proof:
There are three things to prove:

1. We may take the homeomorphism $f : A \to A$ defined by $f(x) = x$. Then $f(\alpha(x)) = \alpha(f(x))$.
2. Let $f : B \to A$ be a homeomorphism such that $\alpha(f(x)) = f(\beta(x))$ for all x such that $x \in B$. Let $g : A \to B$ be defined by $g(x) = f^{-1}(x)$, which is a homeomorphism since f is a homeomorphism. Then $g(\alpha(f(x))) = g(f(\beta(x))) = \beta(x)$ for all $x \in B$. In particular, for every y such that $y \in A$ there is an x such that $x \in B$ and $g(y) = x$, and therefore $\beta(g(y)) = g(\alpha(y))$.
3. Let $f : B \to A$ and $g : C \to B$ be homeomorphisms such that $\alpha(f(x)) = f(\beta(x))$ and $\beta(g(y)) = g(\gamma(y))$, for all x such that $x \in B$ and all y such that $y \in C$. For every y such that $y \in C$, write $y = g(x)$. We have

$$\alpha(f(g(x))) = f(\beta(g(x))) = f(g(\gamma(x)))$$

so that $g \circ f : C \to A$ (we write $g \circ f$ to denote the function $f(g(x))$) is a homeomorphism such that $\alpha \overset{g \circ f}{\sim} \gamma$. $\square$

Example 1.2.15:
Let $\alpha : A \to A$ be the identity map, defined by $\alpha(x) = x$ for all x in A. Then by the previous Theorem, we know that α is conjugate to itself. Suppose $\alpha \overset{f}{\sim} \beta$ for some map $\beta : B \to B$ and some homeomorphism $f : B \to A$. Then for all x in B we must have

$$f(\beta(x)) = \alpha(f(x)) = f(x).$$

Since f is one-to-one, this means that $\beta(x) = x$ for all x in B. Thus, the only map that is conjugate to an identity map is an identity map. ■

Example 1.2.16:
Let $\alpha : \mathbf{R}^+ \to \mathbf{R}^+$ and $\beta : \mathbf{R}^+ \to \mathbf{R}^+$ be defined by $\alpha(x) = ax$ and $\beta(x) = bx$, where $a > 0$ and $b > 0$. Since we have just seen that identity maps can only be conjugate to other identity maps (or themselves), let us also assume that $a \neq 1$ and $b \neq 1$. We will show that $\alpha \overset{f}{\sim} \beta$, where

$$f(x) = x^{\log_b a}.$$

We have

$$f(\beta(x)) = f(bx) = (bx)^{\log_b a} = ax^{\log_b a},$$

and

$$\alpha(f(x)) = ax^{\log_b a}.$$ ■

We will now work towards relating the Lyapunov exponents of two conjugate maps.

Lemma 1.2.1:
If $\alpha : A \to A$ and $\beta : B \to B$ are one-dimensional maps and $f : B \to A$ is a homeomorphism such that $\alpha \overset{f}{\sim} \beta$, then for all positive integers k we have $\alpha^k \overset{f}{\sim} \beta^k$.

Proof:
Let $f : B \to A$ be a homeomorphism such that $\alpha(f(x)) = f(\beta(x))$ for all x such that $x \in B$. We will prove the claim by induction on k. First we note that $\alpha(\alpha(f(x))) = \alpha(f(\beta(x)))$, so that $\alpha^2(f(x)) = \alpha(f(\beta(x))) = f(\beta^2(x))$. Now we suppose the claim is true for all integers n such that $2 \leq n \leq k-1$. Then $\alpha^n(f(x)) = f(\beta^n(x))$ so that $\alpha^{n+1}(f(x)) = \alpha(f(\beta^n(x))) = f(\beta^{n+1}(x))$. □

Corollary 1.2.1:
Let $\alpha : A \to A$ and $\beta : B \to B$ be one-dimensional maps and let $f : B \to A$ be a homeomorphism such that $\alpha \overset{f}{\sim} \beta$. If

$$x_0, x_1, x_2, \ldots$$

is a trajectory of β, then

$$f(x_0), f(x_1), f(x_2), \ldots$$

is a trajectory of α.

Proof:
For every positive integer i we have $x_i = \beta^i(x_0)$ so that $f(x_i) = f(\beta^i(x_0)) = \alpha^i(f(x_0))$, by the previous Lemma. □

Theorem 1.2.4:
Let $\alpha : X \to X$ and $\beta : Y \to Y$ be one-dimensional differentiable maps, where $X \subseteq \mathbf{R}$ and $Y \subseteq \mathbf{R}$, and let $f : Y \to X$ be a homeomorphism such that $\alpha \overset{f}{\sim} \beta$. If $x_0, x_1, x_2, \ldots$ is a trajectory of β such that

- *$f'(x_i) \neq 0$ for all nonnegative integers i and*
- *$\lim_{n\to\infty} \frac{1}{n} \ln|f'(x_n)| = 0$,*

then the Lyapunov exponent of the trajectory of x_0 under β is the same as that of the trajectory of $f(x_0)$ under α.

Proof:
By the chain rule, $f'(\beta(x)) \times \beta'(x) = \alpha'(f(x)) \times f'(x)$. Therefore, for a trajectory $x_0, x_1, \ldots$ of β, we have

$$\begin{aligned}\beta'(x_0)\beta'(x_1)\cdots\beta'(x_n) &= \frac{\alpha'(f(x_0)) \times f'(x_0)}{f'(x_1)} \frac{\alpha'(f(x_1)) \times f'(x_1)}{f'(x_2)} \cdots \frac{\alpha'(f(x_n)) \times f'(x_n)}{f'(x_{n+1})} \\ &= \alpha'(f(x_0))\alpha'(f(x_1))\cdots\alpha'(f(x_n)) \times \frac{f'(x_0)}{f'(x_{n+1})},\end{aligned}$$

as long as $f'(x_i)$ is never 0, which is guaranteed in our hypothesis.

Next, we note that

$$\sum_{t=0}^{n} \ln|\beta'(x_i)| = \ln|f'(x_0)| - \ln|f'(x_{n+1})| + \sum_{t=0}^{n} \ln|\alpha'(f(x_i))|,$$

and thus

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n} \ln|\beta'(x_i)| = \lim_{n\to\infty} \frac{1}{n} \left(\ln|f'(x_0)| - \ln|f'(x_{n+1})| + \sum_{t=0}^{n} \ln|\alpha'(f(x_i))| \right).$$

Let us note that $\lim_{n\to\infty} \frac{1}{n} \ln|f'(x_0)| = 0$. Furthermore, the trajectories

$$f(x), f(\beta(x)), f(\beta^2(x)), \ldots$$

and

$$f(x), \alpha(f(x)), \alpha^2(f(x)), \ldots$$

are the same by Lemma 1.2.1. Thus, we have shown that the Lyapunov exponent of β starting at x_0 is the same as that of α starting at $f(x_0)$, as long as $\lim_{n\to\infty} \frac{1}{n} \ln|f'(x_{n+1})| = 0$. □

With this Theorem under our belt, we can use it to calculate the Lyapunov exponent of the logistic map when $\mu = 4$.

Example 1.2.17:
Recall the logistic map, $G_\mu(x) = \mu x(1-x)$, and let us focus on the case when $\mu = 4$. Fixing a starting point x, we are interested in calculating the value of

$$\begin{aligned}\lambda(x) &= \lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} \ln|G_4'(G_4^t(x))| \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} \ln|4 - 8G_4^t(x)|.\end{aligned}$$

Unfortunately, this reduces to being able to say something about the trajectory

$$x, G(x), G^2(x), \ldots.$$

Instead, we will calculate the Lyapunov exponent of a trajectory starting at x by showing that G_4 is conjugate to the tent map T_2.

Let

$$f(x) = \frac{1-\cos \pi x}{2},$$

which is a homeomorphism on $[0,1]$. We will first show that $G_4(f(x)) = f(T_2(x))$ for all $x \in [0,1]$. We begin by noting that

$$\begin{aligned}G_4(f(x)) &= 4f(x)(1-(f(x))) \\ &= 4\left(\frac{1-\cos\pi x}{2})(\frac{1+\cos\pi x}{2}\right) \\ &= 1-\cos^2 \pi x \\ &= \sin^2 \pi x.\end{aligned}$$

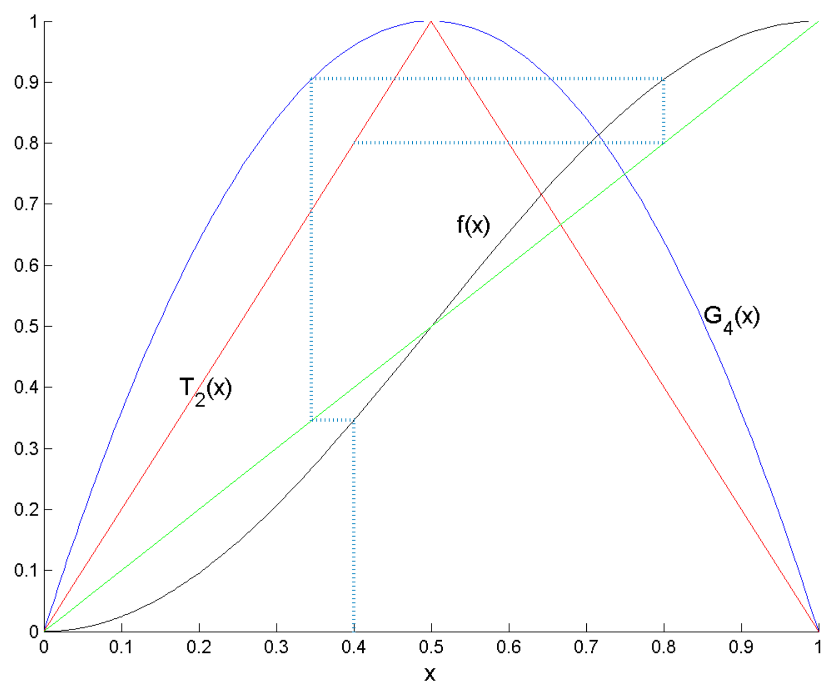
Next, we note that $f(T_2(x)) = (1-\cos(\pi T_2(x)))/2$, so that if $x \in [0, 1/2)$, then

$$\begin{aligned}f(T_2(x)) &= \frac{1-\cos(2\pi x)}{2} \\ &= \frac{1-1+2\sin^2(\pi x)}{2} \\ &= \sin^2(\pi x).\end{aligned}$$

If $x \in [1/2, 1]$, then

$$\begin{aligned} f(T_2(x)) &= \frac{1 - \cos(\pi(2 - 2x))}{2} \\ &= \frac{1 - \cos(2\pi - 2\pi x))}{2} \\ &= \frac{1 - (\cos(2\pi)\cos(2\pi x) + \sin(2\pi)\sin(2\pi x))}{2} \\ &= \frac{1 - \cos(2\pi x)}{2} \\ &= \frac{1 - 1 + 2\sin^2(\pi x)}{2} \\ &= \sin^2(\pi x). \end{aligned}$$

Figure 1.14: $G_4(f(x) = f(T_2(x))$.



Thus, if $x_0, x_1, x_2, \ldots$ is a trajectory of T_2 that doesn't hit $1/2$, and

$$\lim_{n\to\infty} \frac{1}{n} \ln |f'(x_n)| = 0, \tag{1.10}$$

then the Lyapunov exponent of the trajectory $f(x_0), f(x_1), \ldots$ of G_4 is $\ln 2$.

In particular, if $x_0, x_1, \ldots$ is a trajectory of T_2 that never hits 0, then the trajectory never hits 1/2, since $T_2^2(1/2) = 0$. Furthermore, (1.10) is satisfied since $f'(x) = 0$ if and only if $x = 0$. Since $f(0) = 0$, we may conclude by saying that every trajectory of G_4 that does not hit 0 has a Lyapunov exponent of $\ln 2$.

It is worth noting that the trajectory starting at $x = 0$ can be calculated directly, since 0 is a fixed point. We have

$$\begin{aligned}
\lambda(0) &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \ln |G_4'(G_4^t(0))| \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \ln |4 - 8G_4^t(0)| \\
&= \ln 4,
\end{aligned}$$

which shows that not all trajectories have the same Lyapunov exponent under a given map. ■

The definition of conjugacy we provided is the same as that used in [8], [10], [1], and many others texts on dynamical systems, although this notion is referred to as topological conjugacy in some texts. Our sole purpose of introducing Theorem 1.2.4, which was adopted from [1], was for the calculation of the Lyapunov exponent of G_4, which can be found in that reference as well. Conjugacy between G_4 and T_2 seems to be the simplest way to find the exact value of G_4. Alternative methods use ergodic theory to achieve the same results (see [29]), and many references calculate the Lyapunov exponent of the logistic map numerically (see [36] for example). Finally, we note that all of the results from this section are general enough to hold for higher dimensions, with the exception of Theorem 1.2.4.

1.2.8 Computing the global Lyapunov exponent of a map.

How can we compute the global Lyapunov exponent of a prescribed differentiable map $\Phi : X \to X$ (such that $X \subseteq \mathbf{R}$) at a prescribed point x^0 in the interior of X? Answers to this question depend on the way Φ is prescribed; let us assume that it is prescribed by an oracle that, given any x in X, returns $\Phi(x)$.

In this situation, we can compute iteratively

$$x^t = \Phi(x^{t-1}),$$

and take δ_t, a number small enough to ensure that

$$\Phi(x^t + \delta_t) - \Phi(x^t) \approx \Phi'(x)\delta_t,$$

until the sequence of averages

$$\frac{1}{n} \sum_{t=0}^{n-1} \ln \frac{|\Phi(x^t + \delta_t) - \Phi(x^t)|}{\delta_t}$$

shows signs of convergence, at which time we return an estimate of its limit as an estimate of the Lyapunov exponent.

1.3 Multi-dimensional discrete dynamical systems.

We may now proceed to multi-dimensional discrete dynamical systems, which, as we will see, have many similarities to the one-dimensional systems we have just studied. Let us begin by presenting a few examples.

1.3.1 Examples.

Example 1.3.1:
The *delayed logistic map* is a two-dimensional map $D_\mu : \mathbf{R}^2 \to \mathbf{R}^2$ defined as

$$D_\mu(x_1, x_2) = (\mu x_1(1 - x_2), x_1),$$

where μ is any nonzero value in $\mathbf{R}$. ■

Example 1.3.2:
The map $\Phi : \mathbf{R}^+ \times [0, 2\pi) \to \mathbf{R}^+ \times [0, 2\pi)$ defined by $\Phi(r, \theta) = \left(\sqrt{r}, \sqrt{2\pi\theta}\right)$ will be particularly useful when studying stability. In this case we view (r, θ) as a point in the plane in polar coordinates. ■

Example 1.3.3:
Arnold's cat map[2] is a two-dimensional map $\Phi : [0, 1)^2 \to [0, 1)^2$ defined as

$$\Phi(x_1, x_2) = (x_1 + x_2 \bmod 1, x_1 + 2x_2 \bmod 1). \qquad ■$$

Example 1.3.4:
The *Kaplan-Yorke map*[3] is another two-dimensional map $\Phi : [0, 1)^2 \to [0, 1)^2$. It is defined as

$$\Phi(x_1, x_2) = (ax_1 \bmod 1, bx_2 + \cos(4\pi x_1) \bmod 1). \qquad ■$$

1.3.2 Stability of fixed points.

Fixed points of multi-dimensional systems share some similarities with the fixed points of one-dimensional systems. They are defined in the same way, namely if $\Phi : X \to X$ and x^* is some point in X such that $\Phi(x^*) = x^*$, then x^* is a fixed point.

[2] Named after Vladimir Igorevich Arnold who displayed the effects of the map on the image of a cat, see [2].
[3] Introduced by Kaplan and Yorke in 1979, see [20].

Example 1.3.5:
Let us find the fixed points of the delayed logistic map, $D_\mu(x_1, x_2) = (\mu x_1(1 - x_2), x_1)$. In order to do so we must solve for $x = \mu x(1-x)$, which is simply our one dimensional logistic map from before. Thus, solutions for x are 0 and $(\mu - 1)/\mu$, so that the fixed points of the delayed logistic map are

$$(0,0) \quad \text{and} \quad (\frac{\mu-1}{\mu}, \frac{\mu-1}{\mu}).$$

■

Example 1.3.6:
The only fixed points of $\Phi(r,\theta) = \left(\sqrt{r}, \sqrt{2\pi\theta}\right)$ are easily seen to be $(0,0)$ and $(1,0)$. ■

The definitions of section 1.2.3 are easily extended to multi-dimensional maps. Let $\Phi : X \to X$ be a multi-dimensional map, where $X \subseteq \mathbf{R}^d$. We have:

Definition 1.3.1:
A fixed point x^* of a map Φ is *stable* if for all positive ε, there exists a positive δ such that for all positive integers t and for all points x in X we have

$$\text{if } ||x - x^*|| < \delta, \text{ then } \left|\left|\Phi^t(x) - x^*\right|\right| < \varepsilon.$$

Definition 1.3.2:
A fixed point x^* of a map Φ is *unstable* if it is not stable.

Definition 1.3.3:
A fixed point x^* of a map Φ is *attracting* if there exists a positive δ such that for all points x such that $x \in X$ we have

$$\text{if } ||x - x^*|| < \delta, \text{ then } \lim_{t\to\infty} \Phi^t(x) = x^*.$$

Example 1.3.7:
Let us apply these definitions to the polar coordinate map from above, $\Phi : \mathbf{R}^+ \times [0, 2\pi) \to \mathbf{R}^+ \times [0, 2\pi)$ defined by $\Phi(r,\theta) = \left(\sqrt{r}, \sqrt{2\pi\theta}\right)$.

Before proceeding, let us note that for every positive integer t, we have

$$\Phi^t(r,\theta) = \left(r^{2^{-t}}, \frac{2\pi\theta^{2^{-t}}}{(2\pi)^{2^{-t}}} \right). \tag{1.11}$$

Let us begin with the fixed point $(0, 0)$. We will first show that it is attracting. Let $\delta = 1$ and let x be any point such that $||x - (0, 0)|| < \delta$. Since we have $x = (r, \theta)$ for some r such that $r < 1$, equation (1.11) allows us to conclude that $\lim_{t\to\infty} \Phi^t(x) = (0, 0)$.

Next, let us show that $(0, 0)$ is stable. Let a positive ε be given, and set $\delta = \min\{\varepsilon, 1\}$. For every point x such that $||x - (0, 0)|| < \delta$, we necessarily have $x = (r, \theta)$, where $r < \delta$ and θ is some angle. Since for every positive integer t we have $||\Phi^t(x) - (0, 0)|| = r^{2^{-n}} < r < \delta < \varepsilon$, we are done.

Let us now consider the fixed point $x^* = (1, 0)$.

To see that $(1, 0)$ is attracting, take $\delta = 1/2$. Now for every x such that $||x - x^*|| < \delta$ (note that $(0, 0)$ is excluded), by equation (1.11) we have

$$\lim_{t\to\infty} \Phi^t(x) = \lim_{t\to\infty} \left(r^{2^{-t}}, \frac{2\pi\theta^{2^{-t}}}{(2\pi)^{2^{-t}}} \right) = (1, 2\pi) = (1, 0).$$

Finally, we will now show that $(1, 0)$ is unstable. Let us begin by choosing $\varepsilon = 1/2$. Now suppose we are given a positive δ. Without loss of generality, we can assume $\delta < 1/2$. Now we consider the point

$$x = \left(1, \pi 2^{1-2^n}\right).$$

Since this point approaches $(1, 0)$ as n gets larger, we may assume there is some positive integer $n(\delta)$ for which

$$\left|\left|\left(1, \pi 2^{1-2^{n(\delta)}}\right) - (1, 0)\right|\right| < \delta.$$

Let us call this point x_δ. Now we may note that

$$\Phi^n(\delta)(x_\delta) = (1, \pi)$$

by equation (1.11). Thus

$$||\Phi^n(\delta)(x_\delta) - x^*|| = 1 > \varepsilon. \qquad \blacksquare$$

We have just seen a rather simple system in which there is a fixed point that is both attracting and unstable. This provides an interesting contrast against example 1.2.8 and Theorem 1.2.1, where we saw that an attracting fixed point of a continuous one-dimensional map must also be stable. On the one hand, this indicates that we may not simply take what we knew in the one-dimensional case and assume it to be true in higher dimensions. On the other hand, we will now prove a theorem that can be seen as the multi-dimensional case of Theorem 1.2.2. Before doing so, let us introduce some notation.

We will use $J(F, x)$ to denote the Jacobian matrix of F evaluated at x: the entry in the ith row and the jth column of $J(F, x)$ is the value of $\partial y_i/\partial x_j$ at x, where $(y_1, y_2, \ldots, y_d) = F(x_1, x_2, \ldots, x_d)$. The following Lemma also makes use of matrix norms. For their precise definition see section 2.2 of the Appendix.

Lemma 1.3.1:
Let $X \subseteq \mathbf{R}^d$ and let $F : X \to X$ be a differentiable map with a fixed point x^*. If $\lambda_1, \lambda_2, \dots, \lambda_d$ are the eigenvalues of $J(F, x^*)$, then

1. *If $|\lambda_i| < 1$ for all integers i such that $1 \le i \le d$, then $\lim_{t\to\infty} ||J(F^t, x^*)|| = 0$.*
2. *If $|\lambda_i| > 1$ for any integer i such that $1 \le i \le d$, then $\lim_{t\to\infty} ||J(F^t, x^*)|| = \infty$.*

Proof:
By the chain rule,

$$\begin{aligned} J(F^n, x^*) &= J(F, F^{n-1}(x^*))J(F^{n-1}, x^*) \\ &= J(F, F^{n-1}(x^*))J(F, F^{n-2}(x^*))J(F^{n-2}, x^*) \\ &= \vdots \\ &= J(F, F^{n-1}(x^*))J(F, F^{n-2}(x^*)) \cdots J(F, F(x^*))J(F, x^*). \end{aligned}$$

Since x^* is a fixed point of F, we have $J(F^n, x^*) = J(F, x^*)^n$.

Let $\lambda_1, \lambda_2, \dots, \lambda_m$ be the distinct (and possibly complex) eigenvalues of $J(\Phi, x^*)$, where $m \le d$. Since $J(F, x^*)$ is a square matrix with real entries, we can put it into Jordan normal form. In particular, $J(F, x^*) = SJS^{-1}$ where J is a block diagonal matrix consisting of Jordan blocks $J_1, J_2, \dots, J_m$, and S is some $d \times d$ invertible matrix. Note that since $J(F, x^*) = SJS^{-1}$, we have $J(F^t, x^*) = J(F, x^*)^t = SJ^tS^{-1}$. By Lemma .0.7 in the Appendix, we know that

$$J^t = \begin{bmatrix} J_1^t & & \\ & \ddots & \\ & & J_m^t \end{bmatrix}, \tag{1.12}$$

and

$$J_i^t = \begin{bmatrix} \lambda_i^t & a_{1,2} & a_{1,3} & a_{1,4} & \cdots & a_{1,p} \\ 0 & \lambda_i^t & a_{2,3} & a_{2,4} & \cdots & a_{2,p} \\ 0 & 0 & \lambda_i^t & a_{3,4} & \cdots & a_{3,p} \\ 0 & 0 & 0 & \ddots & & \\ 0 & 0 & 0 & 0 & \lambda_i^t & a_{p-1,p} \\ 0 & 0 & 0 & 0 & 0 & \lambda_i^t \end{bmatrix} \tag{1.13}$$

for $i = 1, \dots, m$, and where p is the multiplicity of λ_i. The precise values of the entries of J_i^t are

$$a_{r,s} = \binom{t}{s-r} \lambda_i^{t-(s-r)},$$

for $1 \le r < s \le p$, by Lemma .0.6 in the Appendix. The point of all this is that since $|\lambda_i| < 1$, we have $\lim_{t\to\infty} \lambda_i^t = 0$ and $\lim_{t\to\infty} a_{r,s} = 0$, and therefore $\lim_{t\to\infty} J^t = 0_M$, where 0_M is the zero matrix. The first statement of the Lemma follows. The proof of the second statement is similar, except we now have that $|\lambda_i|^t \to \infty$ as $t \to \infty$ for any λ_i that has magnitude strictly greater than 1. $\square$

Theorem 1.3.1:
Let $X \subseteq \mathbf{R}^d$ and let $F : X \to X$ be a differentiable map with a fixed point x^*. If $\lambda_1, \lambda_2, \ldots, \lambda_d$ are the eigenvalues of $J(F, x^*)$, then

1. *If $|\lambda_i| < 1$ for all integers i such that $1 \le i \le d$, then x^* is stable and attracting.*
2. *If $|\lambda_i| > 1$ for any integer i such that $1 \le i \le d$, then x^* is unstable.*

Proof (of 1.):
Let a positive ε be given, and let us assume that $\varepsilon < 1$ with no loss of generality. By the previous Lemma, there exists a positive integer k such that for all integers t such that $t \ge k$, we have $||J(F^t, x^*)|| < \varepsilon/2$.

We begin by proceeding backwards from k, using the continuity of F to argue as follows.

There exists a positive δ_{k-1} such that

$$\left|\left|F^{k-1}(x) - F^{k-1}(x^*)\right|\right| < \delta_{k-1} \quad \Rightarrow \quad \left|\left|F^k(x) - F^k(x^*)\right|\right| < \varepsilon.$$

There exists a positive δ_{k-2} such that

$$\left|\left|F^{k-2}(x) - F^{k-2}(x^*)\right|\right| < \delta_{k-2} \quad \Rightarrow \quad \left|\left|F^{k-1}(x) - F^{k-1}(x^*)\right|\right| < \min\{\varepsilon, \delta_{k-1}\}$$

$$\vdots$$

There exists a positive δ_1 such that

$$||F(x) - F(x^*)|| < \delta_1 \quad \Rightarrow \quad \left|\left|F^2(x) - F^2(x^*)\right|\right| < \min\{\varepsilon, \delta_2\}.$$

There exists a positive δ_0 such that

$$||x - x^*|| < \delta_0 \quad \Rightarrow \quad \left|\left|F^1(x) - F^1(x^*)\right|\right| < \min\{\varepsilon, \delta_1\}.$$

Thus, setting $\delta = \min\{1, \delta_0, \delta_1, \ldots, \delta_{k-1}\}$, we get that

$$||x - x^*|| < \delta \quad \Rightarrow \quad \left|\left|F^t((x) - x^*\right|\right| < \varepsilon$$

for all integers t such that $1 \le t \le k$.

We now proceed forward from k as follows. Since F^k is differentiable at x^*, there exists a positive δ_k such that if $||x - x^*|| < \delta_k$ then

$$\left|\left|F^k(x) - F^k(x^*) - J(F^k, a)(x - a)\right|\right| < \frac{\varepsilon}{2} \left|\left|x - x^*\right|\right|. \tag{1.14}$$

Therefore, we have

$$\begin{aligned}\left|\left|F^k(x) - F^k(x^*)\right|\right| &= \left|\left|F^k(x) - F^k(x^*) - J(F^k, a)(x-a) + J(F^k, a)(x-a)\right|\right| \\ &\le \left|\left|F^k(x) - F^k(x^*) - J(F^k, a)(x-a)\right|\right| + \left|\left|J(F^k, a)(x-a)\right|\right| \\ &\le \frac{\varepsilon}{2}\left|\left|x - x^*\right|\right| + \frac{\varepsilon}{2}\left|\left|(x-x^*)\right|\right| \\ &\le \varepsilon\left|\left|x - x^*\right|\right|. \end{aligned} \tag{1.15}$$

Note that $F^k(x)$ is closer to x^* than x was. Thus, (1.14) holds with $F^k(x)$ in place of x and may iterate (1.15) to get that

$$\left|\left|F^{2k}(x) - F^{2k}(x^*)\right|\right| < \varepsilon\left|\left|F^k x - x^*\right|\right| < \varepsilon^2\left|\left|x - x^*\right|\right|.$$

Proceeding in this fashion yields

$$\left|\left|F^{mk}(x) - F^{mk}(x^*)\right|\right| < \varepsilon^m\left|\left|x - x^*\right|\right|$$

for every positive integer m. Since $\varepsilon < 1$ we get

$$\lim_{t\to\infty}\left|\left|F^{tk}(x)\right|\right| = x^*.$$

To complete the proof, we note that we can apply this argument to every integer r such that $1 \le r \le k-1$ to get a δ_{k+r} for which

$$\left|\left|F^{mk+r}(x) - F^{mk+r}(x^*)\right|\right| < \varepsilon^m\left|\left|F^r(x) - x^*\right|\right|$$

for every point x such that $||F^r(x) - x^*|| < \min\{\delta, \delta_{k+r}\}$.

We conclude that for every point x such that $||x - x^*|| < \min\{\delta, \delta_k, \delta_{k+1}, \ldots, \delta_{2k-1}\}$ we have $\lim_{t\to\infty} F^{k\times t+r}(x) = x^*$ for every integer r such that $0 < r < k$, and thus $\lim_{t\to\infty} F^t(x) = x^*$. This shows that x^* is attracting, and in fact we have also shown that $||F^t(x) - x^*|| < \varepsilon$ for all positive integers t, so that x^* is stable as well. □

Proof (of 2.):
By the previous Lemma, we know that $||J(F^t, x^*)||$ is unbounded. Thus, we may find a positive integer t' such that $||J(F^t, x^*)|| > M > 1$ for all integers t such that $t \ge t'$. Let us note that by the definition of our matrix norm, there exists a vector v such that $||v|| \le 1$ and $\left|\left|J(F^{t'}, x^*)\right|\right| = \left|\left|J(F^{t'}, x^*)v\right|\right|$.

Since $F^{t'}$ is differentiable at x^*, we know that there exists a positive δ_c such that if x is any point such that

$$||x - x^*|| < \delta_c, \tag{1.16}$$

then

$$\frac{\left|\left|F^{t'}(x) - F^{t'}(x^*) - J(F^{t'}, x^*)(x - x^*)\right|\right|}{||x - x^*||} < \frac{M}{4\,||v||}. \tag{1.17}$$

Let us now set $\varepsilon = \delta_c$ and let a positive δ be given. Without loss of generality, we will assume $\delta < \delta_c$.

Consider the point

$$x = x^* + \frac{v\delta}{2\,||v||}.$$

Note that $||x - x^*|| < \delta$, so we now have

$$\begin{aligned}
\left|\left|F^{t'}(x) - x^*\right|\right| &\geq \left|\left|J(F^{t'}, x^*)(x - x^*)\right|\right| - \left|\left|J(F^{t'}, x^*)(x - x^*) - F^{t'}(x) + F^{t'}(x^*)\right|\right| \\
&= \left|\left|J(F^{t'}, x^*)\frac{v\delta}{2\,||v||}\right|\right| - \left|\left|F^{t'}(x) - F^{t'}(x^*) - J(F^{t'}, x^*)(x - x^*)\right|\right| \\
&> \frac{\delta}{2\,||v||}\left|\left|J(F^{t'}, x^*)v\right|\right| - \frac{M\,||x - x^*||}{4\,||v||} \\
&= \frac{\delta}{2\,||v||}\left|\left|J(F^{t'}, x^*)\right|\right| - \frac{M\,||x - x^*||}{4\,||v||} \\
&> \frac{||x - x^*||}{2\,||v||}\left|\left|J(F^{t'}, x^*)\right|\right| - \frac{M\,||x - x^*||}{4\,||v||} \\
&= M\,||x - x^*||\,.
\end{aligned}$$

If $\left|\left|F^{t'}(x) - x^*\right|\right| \geq \varepsilon$ then we are done. Otherwise, we have $\left|\left|F^{t'}(x) - x^*\right|\right| < \varepsilon = \delta_c$ and thus condition 1.16 holds with $F^{t'}(x)$ in place of x, which allows us to iterate to produce

$$\left|\left|F^{2t'}(x) - x^*\right|\right| > M\left|\left|F^{t'}(x) - x^*\right|\right| > M^2\,||x - x^*||\,.$$

Repeating these arguments yields that $\left|\left|F^{kt'}(x) - x^*\right|\right| > M^k\,||x - x^*||$ for all positive integers k. Since $M > 1$, we must have $\left|\left|F^{kt'}(x) - x^*\right|\right| > \varepsilon$ for some positive integer k, and we are done. □

Example 1.3.8:
Recall from example 1.3.5 that the fixed points of the delayed logistic map, $D_\mu(x_1, x_2) = (\mu x_1(1 - x_2), x_1)$, are

$$(0, 0)$$

and

$$(\frac{\mu - 1}{\mu}, \frac{\mu - 1}{\mu}).$$

The Jacobian $J(D_\mu, (x_1, x_2))$ of D_μ at a point (x_1, x_2) is given by

$$\begin{bmatrix} \mu - \mu x_2 & -\mu x_1 \\ 1 & 0 \end{bmatrix}.$$

Its eigenvalues are the roots of

$$\begin{aligned}
\det\begin{bmatrix} \mu - \mu x_2 - \lambda & -\mu x_1 \\ 1 & -\lambda \end{bmatrix} &= (\mu - \mu x_2 - \lambda)(-\lambda) + \mu x_1 \\
&= \lambda^2 - \lambda(\mu - \mu x_2) + \mu x_1. \qquad (1.18)
\end{aligned}$$

Let us focus on the fixed point $(0,0)$ first. In this case the eigenvalues of $J(D_\mu,(0,0))$ are the roots of $\lambda^2 - \lambda\mu$, which are 0 and μ. Thus $(0,0)$ is stable and attracting when $|\mu| < 1$ and unstable when $|\mu| > 1$.

Now let us look at the other fixed point, $((\mu-1)/u, (\mu-1)/\mu)$, which we will denote with x^*. Equation (1.18) with $x_1 = x_2 = (\mu-1)/\mu$ becomes

$$\lambda^2 + \lambda + \mu - 1$$

and thus the eigenvalues of $J(D_\mu, x^*)$ are

$$\frac{-1 \pm \sqrt{5-4\mu}}{2}.$$

When $\mu < 1$, we may note that

$$\frac{-1 - \sqrt{5-4\mu}}{2} < -1,$$

so that x^* is unstable.

When $\mu = 1$, the eigenvalues are 0 and -1, so that Theorem 1.3.1 is inconclusive.

When $\mu = 5/4$, we have only one eigenvalue, which is $-1/2$. Since $|-1/2| < 1$, we know that x^* is stable and attracting.

Let us now suppose that $1 < \mu < 5/4$. We note that

$$\begin{aligned} 1 < \mu < 5/4 \quad &\Rightarrow \quad 4 < 4\mu < 1 \\ &\Rightarrow \quad -4 > -4\mu > -5 \\ &\Rightarrow \quad 1 > 5 - 4\mu > 0 \\ &\Rightarrow \quad 1 > \sqrt{5-4\mu} > 0 \quad \text{since } 5 - 4\mu > 0 \\ &\Rightarrow \quad \frac{1}{2} > \frac{\sqrt{5-4\mu}}{2} > 0, \end{aligned}$$

so that

$$\left| \frac{-1 \pm \sqrt{5-4\mu}}{2} \right| < 1,$$

and thus x^* is stable and attracting.

Finally, let us now suppose that $\mu > 5/4$. In this case the eigenvalues are complex, and therefore

$$\left| \frac{-1 \pm \sqrt{5-4\mu}}{2} \right| = \left| \frac{-1 \pm i\sqrt{4\mu-5}}{2} \right| = \sqrt{\frac{1}{4} + \frac{4\mu-5}{4}} = \sqrt{u-1}.$$

Therefore, in this case there is only one distinct eigenvalue and it has magnitude less than 1 only when $\mu < 2$.

In summary, the fixed point $(0,0)$ is stable and attracting when $|\mu| < 1$ and unstable when $|\mu| > 1$. When $|\mu| = 1$ our analysis is inconclusive. As for the other fixed point, $x^* = ((\mu-1)/\mu, (\mu-1)/\mu)$, we have that x^* is unstable when $\mu < 1$, stable and attracting when $1 < \mu < 2$, and unstable again when $\mu > 2$. Our analysis about x^* is inconclusive when $\mu = 1$ or $\mu = 2$. ■

The examples from this section can be also be found in [11], [31] and [10]. The definitions in this section are similar to those found in [10]. Theorem 1.3.1 can be found in most texts, and in fact it is taken as the definition of stability in most ([8], for example). Reference [10], from where we borrowed our definitions, only proves the theorem in the two-dimensional case.

1.3.3 Lyapunov exponents in general.

Let us return to the general setting where $\Phi : X \to X$ is any map (and not necessarily differentiable), and $X \subseteq \mathbf{R}^d$. For every point x in the interior of X, and for every y in $\mathbf{R}^d$, the *local Lyapunov exponent* of Φ at x with respect to direction y is defined as the limit (if it exists)

$$\lim_{\delta \to 0} \ln \frac{||\Phi(x+\delta y) - \Phi(x)||}{||\delta y||};$$

the *global Lyapunov exponent* of Φ at x with respect to y is defined as the limit (if it exists)

$$\lim_{t \to \infty} \frac{1}{t} \lim_{\delta \to 0} \ln \frac{||\Phi^t(x+\delta y) - \Phi^t(x)||}{||\delta y||}.$$

If the global Lyapunov exponent of Φ at x with respect to y equals λ then, for all sufficiently large values of t for all sufficiently small (with respect to the value of t) values of δ, we have

$$\frac{||\Phi^t(x+\delta y) - \Phi^t(x)||}{||\delta y||} \approx e^{\lambda t}.$$

1.3.4 Lyapunov exponents of differentiable maps.

Let $X \subseteq \mathbf{R}^d$ and let $F : X \to X$ be a differentiable map. Recall that $J(F, x)$ is the Jacobian matrix of F evaluated at x. Once again, since

$$||F(x+\delta y) - F(x)|| \; = \; ||J(F,x)\delta y|| + o(||\delta y||) \quad \text{as } \delta \to 0,$$

the local Lyapunov exponent of a differentiable map Φ at a point x with respect to a direction y equals

$$\ln \frac{||J(\Phi, x)y||}{||y||}$$

and the global Lyapunov exponent of Φ at x with respect to y, which we will denote by $\lambda(x, y)$, equals

$$\lim_{n \to \infty} \frac{1}{n} \ln \frac{||J(\Phi^n, x)y||}{||y||}. \tag{1.19}$$

For future reference, let us note that

$$\frac{1}{n}\ln\frac{||J(\Phi^n,x)y||}{||y||} = \frac{1}{n}\sum_{t=0}^{n-1}\ln\frac{\big|\big|J(\Phi^{t+1},x)y\big|\big|}{||J(\Phi^t,x)y||} = \frac{1}{n}\sum_{t=0}^{n-1}\ln\frac{||J(\Phi,\Phi^t(x))J(\Phi^t,x)y||}{||J(\Phi^t,x)y||}. \tag{1.20}$$

This means that the global Lyapunov exponent of Φ at point x with respect to direction y is the average of the local Lyapunov exponents of Φ at points $\Phi^0(x)$, $\Phi^1(x)$, $\Phi^2(x)$, $\ldots$ of the trajectory of x with respect to direction $J(\Phi^t,x)y$ at each point $\Phi^t(x)$.

1.3.5 Computing global Lyapunov exponents.

How can we compute the global Lyapunov exponent of a prescribed differentiable map $\Phi : X \to X$ (such that $X \subseteq \mathbf{R}^d$) at a prescribed point x^0 in the interior of X and with respect to a prescribed direction y^0? Answers to this question depend on the way Φ is prescribed; let us assume that it is prescribed by an oracle that, given any x in X, returns $\Phi(x)$.

In this situation, we can compute iteratively

$$x^t = \Phi(x^{t-1}),\ \ y^t = \delta_t\left(\Phi(x^{t-1}+y^{t-1}) - \Phi(x^{t-1})\right),$$

with each δ_t a (possibly negative) number small enough to ensure that

$$\Phi(x^t+y^t) - \Phi(x^t) \approx J(\Phi,x^t)y^t,$$

until the sequence of averages

$$\frac{1}{n}\sum_{t=0}^{n-1}\ln\frac{||\Phi(x^t+y^t)-\Phi(x^t)||}{||y^t||}$$

shows signs of convergence, at which time we return an estimate of its limit as an estimate of the Lyapunov exponent.

To justify this policy, we use induction on t to show that

$$y^t \approx \delta_1\delta_2\cdots\delta_t J(\Phi^t,x^0)y^0 : \tag{1.21}$$

in the induction step, we argue that

$$\begin{aligned} y^{t+1} =& \delta_{t+1}\left(\Phi(x^t+y^t)-\Phi(x^t)\right)\\ \approx& \delta_{t+1}J(\Phi,x^t)y^t\\ \approx& \delta_{t+1}J(\Phi,x^t)\delta_1\delta_2\cdots\delta_t J(\Phi^t,x^0)y^0\\ =& \delta_1\delta_2\cdots\delta_{t+1}J(\Phi,\Phi^t(x^0))J(\Phi^t,x^0)y^0\\ =& \delta_1\delta_2\cdots\delta_{t+1}J(\Phi^{t+1},x^0)y^0. \end{aligned}$$

From (1.21), it follows that

$$\frac{||\Phi(x^t + y^t) - \Phi(x^t)||}{||y^t||} = \frac{\left|\left|y^{t+1}\right|\right|}{|\delta_{t+1}|\,||y^t||} \approx \frac{\left|\left|J(\Phi^{t+1}, x^0)y^0\right|\right|}{||J(\Phi^t, x^0)y^0||},$$

and so

$$\frac{1}{n}\sum_{t=0}^{n-1} \ln \frac{||\Phi(x^t + y^t) - \Phi(x^t)||}{||y^t||} \approx \frac{1}{n}\sum_{t=0}^{n-1} \ln \frac{\left|\left|J(\Phi^{t+1}, x^0)y^0\right|\right|}{||J(\Phi^t, x^0)y^0||}; \tag{1.22}$$

as n tends to infinity, the right-hand side of (1.22) converges to the global Lyapunov exponent of Φ at point x^0 with respect to direction y^0.

1.3.6 The spectrum of Lyapunov exponents.

We will now consider a specific differentiable map $\Phi : X \to X$ with $X \subseteq \mathbf{R}^d$. In Section 1.3.4, we have noted that for every x such that $x \in X$ and an arbitrary direction y such that $y \in \mathbf{R}^d$,

$$\lambda(x, y) = \lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{||J(\Phi^n, x)y||}{||y||}\right).$$

Note that for every constant, non-zero c we have

$$\lambda(x, cy) = \lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{||J(\Phi^n, x)cy||}{||cy||}\right) = \lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{||J(\Phi^n, x)y||}{||y||}\right) = \lambda(x, y)$$

so that we may assume $||y|| = 1$ without any loss of generality. Thus,

$$\lambda(x, y) = \lim_{n\to\infty} \frac{1}{n} \ln||J(\Phi^n, x)y|| = \lim_{n\to\infty} \frac{1}{2n} \ln\left(y^T J(\Phi^n, x)^T J(\Phi^n, x)y\right).$$

Oseledets ([25]; see also [26] and Section 9.1 of [30]) proved that under certain conditions on Φ, which are only mildly restrictive, there is a $d \times d$ real symmetric matrix A such that

$$\lim_{n\to\infty} \frac{1}{2n} \ln\left(y^T J(\Phi^n, x)^T J(\Phi^n, x)y\right) = \lim_{n\to\infty} \frac{1}{2n} \ln\left(y^T A^{2n} y\right) \text{ for all } y \text{ in } \mathbf{R}^d. \tag{1.23}$$

We should note here that in the one-dimensional case where $d = 1$, this reduces to saying that there exists a number a such that

$$\lim_{n\to\infty} \frac{1}{n} \ln|\Phi'(\Phi^n(x))| = a,$$

which is simply a consequence of the Birkhoff Ergodic Theorem we referred to in Section 1.3.4.

The Spectral Theorem (see Appendix 2.2) guarantees that there are real numbers $\lambda_1, \ldots, \lambda_d$ and an orthonormal basis $y^1, \ldots, y^d$ of $\mathbf{R}^d$ such that $Ay^j = \lambda_j y^j$ for all j. Writing y as $\sum_{i=1}^d c_i y^i$ (with $c_i = y^T y^i$), we find that

$$y^T A^{2n} y = \left(\sum_{i=1}^d c_i y^i\right)^T A^{2n} \left(\sum_{j=1}^d c_j y^j\right) = \sum_{i=1}^d c_i (y^i)^T \sum_{j=1}^d c_j \lambda_j^{2n} y^j = \sum_{i=1}^d c_i^2 \lambda_i^{2n},$$

and so

$$\lambda(x,y) = \ln \max\{|\lambda_i| : c_i \neq 0\}.$$

Suppose we reorder the λ_i's so that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_d|$, with corresponding (orthogonal) eigenvectors $y^1, \ldots, y^k$. Then the *ith Lyapunov exponent* of Φ at x is $\lambda(x, y^i) = \ln|\lambda_i|$. In particular, $\lambda(x, y^1) = \ln|\lambda_1|$ is the *maximal Lyapunov exponent* of Φ at x. Furthermore, for any direction y such that $y \in \mathbf{R}^d$, if

$$i = \min_{i=1,\ldots,d}\{i : y^T y^i \neq 0\},$$

then $\lambda(x,y) = \lambda(x,y^i) = \ln|\lambda_i|$.

For future reference, let us make note that for a given point x, a randomly chosen direction y in $\mathbf{R}^d$ will almost surely have $\lambda(x,y)$ equal to the maximal Lyapunov exponent: all exceptions lie in the $(d-1)$-dimensional space of vectors orthogonal to any of the vectors y^k, for which the eigenvalue λ_k of A has the largest absolute value.

1.3.7 Avoiding Oseledets' Theorem.

The use of Oseledets' Theorem in the previous section is not always necessary. Since the statement of the theorem guarantees the existence of a real symmetric matrix A in (1.23), but gives no way of finding it, we will now consider situations where A can be found directly.

Constant, symmetric Jacobians.

Theorem 1.3.2:
Suppose $\Phi : X \to X$ ($X \subseteq \mathbf{R}^d$) is a differentiable map such that for a certain point $x \in X$, we have

1. *$J(\Phi, \Phi^t(x)) = J(\Phi, x)$ for all positive integers t, and*
2. *$J(\Phi, x)$ is symmetric.*

Furthermore, let $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_d|$ be the (not necessarily distinct) eigenvalues of $J(\Phi, x)$, and let $y^1, y^2, \ldots, y^d$ be their corresponding eigenvectors. For every direction y in $\mathbf{R}^d$ we have

$$\lambda(x,y) = \ln|\lambda_k|,$$

where

$$k = \min_{i=1,\ldots,d}\{i : y^T y^i \neq 0\}.$$

Proof:
In this case we may immediately set $A = J(\Phi, x)$ and see that

$$J(\Phi^n, x) = J(\Phi, \Phi^{n-1}(x)) \cdots J(\Phi, \Phi(x)) J(\Phi, x) = A^n$$

so that

$$J(\Phi^n, x)^T J(\Phi^n, x) = (J(\Phi, x)^T)^n J(\Phi, x)^n = A^{2n}.$$

Now equation (1.23) trivially holds. As before, the Spectral Theorem is applied, and we conclude that the spectrum of Lyapunov exponents is given by the natural logarithm of the absolute values of eigenvalues of $J(\Phi, x)$ (although they are all positive in this case, since $J(\Phi, x)$ is symmetric). □

Example 1.3.9:
Let us recall Arnold's cat map:

$$\Phi(x_1, x_2) = (x_1 + x_2 \bmod 1, x_1 + 2x_2 \bmod 1),$$

in which case

$$J(\Phi, (x_1, x_2)) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

as long as $x_1 + x_2 \notin \mathbf{Z}$, and $x_1 + 2x_2 \notin \mathbf{Z}$, in order to avoid discontinuities. Let us assume that x is a point in $\mathbf{R}^2$ such that the trajectory $x, \Phi(x), \Phi^2(x), \ldots$ never hits any of these discontinuities. Note that $J(\Phi, x)$ is constant and symmetric along any such trajectory. Therefore, the Lyapunov exponents are given by $\ln|\lambda_1|, \ln|\lambda_2|$, where λ_1, λ_2 are the eigenvalues of $J(\Phi, x)$.

The eigenvalues of $J(\Phi, x)$ are the roots of

$$\begin{aligned} \det \begin{bmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} &= (1-\lambda)(2-\lambda) - 1 \\ &= 2 - 2\lambda - \lambda + \lambda^2 - 1 \\ &= \lambda^2 - 3\lambda + 1. \end{aligned}$$

Thus, the eigenvalues are $\lambda_1 = (3 + \sqrt{5})/2$, $\lambda_2 = (3 - \sqrt{5})/2$, with respective eigenvectors $e_1 = (2/(1 + \sqrt{5}), 1)$, $e_2 = (2/(1 - \sqrt{5}), 1)$, which form an orthonormal basis of $\mathbf{R}^2$. Therefore, if y is any direction not parallel to e_2, then $\lambda(x, y) = (3 + \sqrt{5})/2$. If, however, y' is parallel to e_2, then $\lambda(x, y') = (3 - \sqrt{5})/2$. ■

Constant Jacobians.

In fact, we may drop the symmetry condition for a slightly weaker condition.

Theorem 1.3.3:
Suppose $\Phi : X \to X$ ($X \subseteq \mathbf{R}^d$) is a differentiable map such that for a certain point $x \in X$, we have

1. *$J(\Phi, \Phi^t(x)) = J(\Phi, x)$ for all positive integers t, and*
2. *$J(\Phi, x)$ has d distinct, nonzero eigenvalues $\lambda_1, \ldots, \lambda_d$.*

Then the spectrum of Lyapunov exponents of the trajectory of x is given by

$$\ln|\lambda_1|, \ln|\lambda_2|, \ldots, \ln|\lambda_d|.$$

Proof:
Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of $J(\Phi, x)$, and let $v_1, \ldots, v_d$ be the corresponding eigenvectors. We will assume $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_d|$.

Let us begin with an arbitrary direction y, where $y \in \mathbf{R}^d$. Since the d eigenvalues are distinct, the d eigenvectors are linearly independent (see Lemma .0.11 in Appendix), and we can decompose y as $y = \sum_{i=1}^{d} c_i v_i$ for some constant c_i's. Then

$$J(\Phi^n, x)y = J(\Phi, x)^n y = \sum_{i=1}^{d} c_i \lambda_i^n v_i. \tag{1.24}$$

The next step is to consider the action of left-multiplying (1.24) by $J(\Phi^n, x)^T$. Since $J(\Phi, x)$ is not necessarily symmetric, we will do this by decomposing each of the eigenvectors of $J(\Phi, x)$ in terms of the eigenvectors of $J(\Phi, x)^T$. By Lemma .0.8 in the Appendix we know that $J(\Phi, x)$ and $J(\Phi, x)^T$ have the same eigenvalues. Let $w_1, \ldots, w_d$ be the eigenvectors of $J(\Phi, x)^T$ that correspond to the eigenvalues $\lambda_1, \ldots, \lambda_d$, respectively. For each v_i we may write $v_i = \sum_{j=1}^{d} b_{ij} w_j$, for some constants b_{ij}. We now have

$$\begin{aligned}
y^T J(\Phi^n, x)^T J(\Phi^n, x) y &= y^T J(\Phi^n, x)^T \sum_{i=1}^{d} c_i \lambda_i^n v_i \\
&= y^T J(\Phi^n, x)^T \sum_{i=1}^{d} c_i \lambda_i^n (\sum_{j=1}^{d} b_{ij} w_j) \\
&= y^T J(\Phi^n, x)^T \sum_{i=1}^{d} \sum_{j=1}^{d} c_i b_{ij} \lambda_i^n w_j \\
&= y^T \sum_{i=1}^{d} \sum_{j=1}^{d} c_i b_{ij} \lambda_i^n \lambda_j^n w_j \\
&= \left(\sum_{i=1}^{d} c_i {v_i}^T \right) \times \sum_{i=1}^{d} \sum_{j=1}^{d} c_i b_{ij} \lambda_i^n \lambda_j^n w_j \\
&= \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{k=1}^{d} c_i c_k b_{ij} \lambda_i^n \lambda_j^n {v_k}^T w_j.
\end{aligned}$$

By Lemma .0.9, ${v_k}^T w_j = 0$ if $k \neq j$, and so we may proceed with

$$\begin{aligned} y^T J(\Phi^n, x)^T J(\Phi^n, x) y &= \sum_{i=1}^{d} \sum_{j=1}^{d} c_i c_j b_{ij} \lambda_i^n \lambda_j^n {v_j}^T w_j \\ &= \sum_{i=1}^{d} c_i^2 b_{ii} \lambda_i^{2n} {v_i}^T w_i + \sum_{i=1}^{d} \sum_{\substack{j=1,\dots,d,\\ j\neq i}} c_i c_j b_{ij} \lambda_i^n \lambda_j^n {v_j}^T w_j \end{aligned}$$

By Lemma .0.10, $b_{ii} \neq 0$ and $v_i^T w_i \neq 0$ for all i. Therefore

$$\lim_{n\to\infty} \frac{1}{2n} \ln y^T J(\Phi^n, x)^T J(\Phi^n, x) y = \ln |\lambda_1|,$$

unless $c_1 = 0$. Similarly, if $c_1 = 0$, then

$$\lim_{n\to\infty} \frac{1}{2n} \ln y^T J(\Phi^n, x)^T J(\Phi^n, x) y = \ln |\lambda_2|,$$

unless $c_2 = 0$, and so on. Thus, the spectrum of Lyapunov exponents of Φ at x are $\ln |\lambda_1|, \ln |\lambda_2|, \dots, \ln |\lambda_d|$.□

Remark:
The coefficients c_i correspond to the directions of the eigenvectors of $J(\Phi, x)$, which need not be orthogonal. How does this relate to the matrix A in equation (1.23)? We want to know what the d orthogonal directions at x are that would yield the d different Lyapunov exponents of Φ at x. What we are looking for is a matrix A such that the eigenvalues of A are the same as the eigenvalues of $J(\Phi, x)$ and whose corresponding eigenvectors $y_1, \dots, y_d$ are such that

$y_1 = v_1$
y_2 is orthogonal to y_1 and $\text{span}\{y_1, y_2\} = \text{span}\{v_1, v_2\}$
y_3 is orthogonal to y_1 and y_2, and $\text{span}\{y_1, y_2, y_3\} = \text{span}\{v_1, v_2, v_3\}$
$\vdots$
y_d is orthogonal to $y_1, y_2, \dots, y_d$, and $\text{span}\{y_1, y_2, \dots, y_d\} = \text{span}\{v_1, v_2, \dots, v_d\}$.

This can be done via Gram-Schmidt orthonormalization. We now seek a matrix A, such that

a) the eigenvalues of A are the eigenvalues of $J(\Phi, x)$, namely $\lambda_1, \dots, \lambda_d$, and such that

b) the eigenvectors of A are the y_i specified above.

To this end, let S be the matrix whose columns are the orthonormal vectors $y_1^T, \dots, y_d^T$ and let D be the diagonal matrix whose diagonal entries are $\lambda_1, \dots, \lambda_d$. Then $A = SDS^T$ is the matrix we are seeking. An easy check shows that $Ay_i = SDS^T y_i = \lambda_i y_i$.

In fact, the requirement that $J(\Phi, x)$ be constant along a trajectory can be dropped for the slightly weaker assumption that the eigenvectors and eigenvalues be the same for all $J(\Phi, \Phi^t(x))$. All of the above arguments still hold and allow us to reach the same conclusion.

Example 1.3.10:
Consider a modified version of Arnold's cat map:

$$\Phi(x_1, x_2) = (2x_1 + 2x_2 \bmod 1, x_1 + 2x_2 \bmod 1),$$

in which case

$$J(\Phi, (x_1, x_2)) = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}.$$

Noe that $J(\Phi, (x_1, x_2))$ is not symmetric, but that it is the same throughout the trajectory beginning at any point (x_1, x_2) (as long as the trajectory never hits a point with an integer coordinate). Thus, Theorem 1.3.3 may be applied. The eigenvalues are given by the roots of

$$\det \begin{bmatrix} 2-\lambda & 2 \\ 1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 - 2,$$

so that the eigenvalues are $2 \pm \sqrt{2}$. Therefore, the spectrum of Lyapunov exponents is given by $\ln|2 \pm \sqrt{2}|$. ■

A slight improvement.

We may improve slightly on the previous setting as follows. We consider a differentiable map and a trajectory such that the Jacobians have d eigenvalues and $d-1$ eigenvectors in common. More precisely, we have the following.

Theorem 1.3.4:
Consider a differentiable map $\Phi : X \to X$ ($X \subseteq \mathbf{R}^d$) such that there is a specific $x \in X$ for which the following requirements are met:

1. *$J(\Phi, x), J(\Phi, \Phi^2(x)), J(\Phi, \Phi^3(x)), \ldots$ each have the same d distinct, non-zero eigenvalues $\lambda_1, \ldots, \lambda_d$*
2. *There are vectors $v_1, v_2, \ldots, v_{d-1}$ in $\mathbf{R}^d$ such that for all positive integers t we have*

$$J(\Phi, \Phi^t(x))v_i = \lambda_i v_i.$$

Then the spectrum of Lyapunov exponents of the trajectory of x is given by

$$\ln|\lambda_1|, \ln|\lambda_2|, \ldots, \ln|\lambda_d|.$$

Proof:
The proof proceeds similarly to that of the previous section. To start, let

$$\lambda_1, \lambda_2, \ldots, \lambda_d$$

and

$$v_1, v_2, \dots, v_{d-1}$$

be the d eigenvalues and $d-1$ eigenvectors, respectively, of $J(\Phi, x)$ (which are the same as those of $J(\Phi, \Phi^t(x))$ for all t). Let v_d^t be the dth eigenvector of $J(\Phi, \Phi^{t-1}(x))$. The main idea is that the eigenvector v_d^t can be decomposed in terms of the eigenvectors of $J(\Phi, \Phi^t(x))$ as

$$v_d^t = \sum_{i=1}^{d-1} c_i^{t+1} v_i + c_d^{t+1} v_d^{t+1},$$

for some constants $c_1^{t+1}, c_2^{t+1}, \dots, c_d^{t+1}$.

We begin with a direction y and decompose it as $y = \sum_{i=1}^{d-1} c_i^1 v_i + c_d^1 v_d^1$. For simplicity, when $n > m$, let us write

$$J(n, m) = J(\Phi, \Phi^{n-1}(x)) J(\Phi, \Phi^{n-2}(x)) \cdots J(\Phi, \Phi^{m-1}(x)).$$

We now have

$$
\begin{aligned}
J(n,1)y &= J(n,1)\left(\sum_{i=1}^{d-1} c_i^1 v_i + c_d^1 v_d^1\right) \\
&= J(n,2)\left(\sum_{i=1}^{d-1} c_i^1 \lambda_i v_i + c_d^1 \lambda_d v_d^1\right) \\
&= J(n,2)\left(\sum_{i=1}^{d-1} c_i^1 \lambda_i v_i + c_d^1 \lambda_d (\sum_{i=1}^{d-1} c_i^2 v_i + c_d^2 v_d^2)\right) \\
&= J(n,2)\left(\sum_{i=1}^{d-1} (c_i^1 \lambda_i v_i + c_d^1 c_i^2 \lambda_d v_i) + c_d^1 c_d^2 \lambda_d v_d^2\right) \\
&= J(n,3)\left(\sum_{i=1}^{d-1} (c_i^1 \lambda_i^2 v_i + c_d^1 c_i^2 \lambda_d \lambda_i v_i) + c_d^1 c_d^2 \lambda_d^2 v_d^2\right) \\
&= J(n,3)\left(\sum_{i=1}^{d-1} (c_i^1 \lambda_i^2 v_i + c_d^1 c_i^2 \lambda_d \lambda_i v_i) + c_d^1 c_d^2 \lambda_d^2 (\sum_{i=1}^{d-1} c_i^3 v_i + c_d^3 v_d^3)\right) \\
&= J(n,3)\left(\sum_{i=1}^{d-1} (c_i^1 \lambda_i^2 v_i + c_d^1 c_i^2 \lambda_d \lambda_i v_i + c_d^1 c_d^2 c_i^3 \lambda_d^2 v_i) + c_d^1 c_d^2 c_d^3 \lambda_d^2 v_d^3\right) \\
&= J(n,4)\left(\sum_{i=1}^{d-1} (c_i^1 \lambda_i^3 v_i + c_d^1 c_i^2 \lambda_d \lambda_i^2 v_i + c_d^1 c_d^2 c_i^3 \lambda_d^2 \lambda_i v_i) + c_d^1 c_d^2 c_d^3 \lambda_d^3 v_d^3\right) \\
&= \quad \vdots \\
&= \sum_{i=1}^{d-1} \left(c_i^1 \lambda_i^n + c_d^1 c_i^2 \lambda_d \lambda_i^{n-1} + c_d^1 c_d^2 c_i^3 \lambda_d^2 \lambda_i^{n-2} + \cdots + c_d^1 c_d^2 \cdots c_d^{n-1} c_i^n \lambda_d^{n-1} \lambda_i\right) v_i \\
&\quad + c_d^1 c_d^2 c_d^3 \cdots c_d^n \lambda_d^n v_d^n.
\end{aligned}
$$

Similarly, we will let $w_1, w_2, \ldots, w_d^t$ be the eigenvectors of $J(\Phi, \Phi^t(x))^T$. By Lemma .0.9, we may write $v_i = b_i^n w_i$ for $i = 1, \ldots, d-1$, and $v_d^n = b_d^n w_d^n$ for some constants b_i^n. Furthermore, for $t = 2, \ldots, n$ there are constants b_i^{t-1} such that

$$
w_d^t = \sum_{i=1}^{d-1} b_i^{t-1} w_i + b_d^{t-1} w_d^{t-1}.
$$

Before we continue, let us simplify the expressions by writing

$$
C = c_d^1 c_d^2 \cdots c_d^n
$$

and

$$
\beta_i = c_i^1 \lambda_i^n + c_d^1 c_i^2 \lambda_d \lambda_i^{n-1} + c_d^1 c_d^2 c_i^3 \lambda_d^2 \lambda_i^{n-2} + \cdots + c_d^1 c_d^2 \cdots c_d^{n-1} c_i^n \lambda_d^{n-1} \lambda_i. \tag{1.25}
$$

We may now proceed with

$$
\begin{aligned}
y^T J(n,1)^T J(n,1) y &= y^T J(n,1)^T \left[\sum_{i=1}^{d-1} \beta_i v_i + C \lambda_d^n v_d^n \right] \\
&= y^T J(n,1)^T \left[\sum_{i=1}^{d-1} \beta_i b_i^n w_i + C \lambda_d^n b_d^n w_d^n \right] \\
&= y^T J(n-1,1)^T \left[\sum_{i=1}^{d-1} \beta_i b_i^n \lambda_i w_i + C b_d^n \lambda_d^{n+1} w_d^n \right] \\
&= y^T J(n-1,1)^T \left[\sum_{i=1}^{d-1} \beta_i b_i^n \lambda_i w_i + C b_d^n \lambda_d^{n+1} \left(\sum_{i=1}^{d-1} b_i^{n-1} w_i + b_d^{n-1} w_d^{n-1} \right) \right] \\
&= y^T J(n-1,1)^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i + C b_d^n b_i^{n-1} \lambda_d^{n+1}) w_i + C b_d^n b_d^{n-1} \lambda_d^{n+1} w_d^{n-1} \right] \\
&= y^T J(n-2,1)^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^2 + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i) w_i + C b_d^n b_d^{n-1} \lambda_d^{n+2} w_d^{n-1} \right] \\
&= y^T J(n-2,1)^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^2 + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i) w_i \right. \\
&\qquad \left. + C b_d^n b_d^{n-1} \lambda_d^{n+2} \left(\sum_{i=1}^{d-1} b_i^{n-2} w_i + b_d^{n-2} w_d^{n-2} \right) \right] \\
&= y^T J(n-2,1)^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^2 + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i + C b_d^n b_d^{n-1} b_i^{n-2} \lambda_d^{n+2}) w_i \right. \\
&\qquad \left. + C b_d^n b_d^{n-1} b_d^{n-2} \lambda_d^{n+2} w_d^{n-2} \right] \\
&= y^T J(n-3,1)^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^3 + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i^2 + C b_d^n b_d^{n-1} b_i^{n-2} \lambda_d^{n+2} \lambda_i) w_i \right. \\
&\qquad \left. + C b_d^n b_d^{n-1} b_d^{n-2} \lambda_d^{n+3} w_d^{n-2} \right] \\
&\vdots
\end{aligned}
$$

So that

$$y^T J(n,1)^T J(n,1) y = y^T \left[\sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^n + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i^{n-1} + C b_d^n b_d^{n-1} b_i^{n-2} \lambda_d^{n+2} \lambda_i^{n-2} \right.$$
$$+ \cdots + C b_d^n b_d^{n-1} \cdots b_d^2 b_i^1 \lambda_d^{2n-1} \lambda_i) w_i$$
$$\left. + C b_d^n b_d^{n-1} b_d^{n-2} \cdots b_d^1 \lambda_d^{2n} w_d^1 \right].$$

Now, with $y^T = \sum_{i=1}^{d-1} c_i^1 v_i^T + c_d^1 (v_d^1)^T$, and applying Lemma .0.9, we get

$$y^T J(n,1)^T J(n,1) y = \sum_{i=1}^{d-1} (\beta_i b_i^n \lambda_i^n + C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i^{n-1} + C b_d^n b_d^{n-1} b_i^{n-2} \lambda_d^{n+2} \lambda_i^{n-2}$$
$$+ \cdots + C b_d^n b_d^{n-1} \cdots b_d^2 b_i^1 \lambda_d^{2n-1} \lambda_i) c_i^1 v_i^T w_i$$
$$+ C b_d^n b_d^{n-1} b_d^{n-2} \cdots b_d^1 \lambda_d^{2n} c_d^1 (v_d^1)^T w_d^1.$$

Finally, plugging expression (1.25) for β_i, we get

$$y^T J(n,1)^T J(n,1) y = \sum_{i=1}^{d-1} \left[(c_i^1 \lambda_i^{2n} + c_d^1 c_i^2 \lambda_d \lambda_i^{2n-1} + c_d^1 c_d^2 c_i^3 \lambda_d^2 \lambda_i 2n-2 \right.$$
$$+ \cdots + c_d^1 c_d^2 \cdots c_d^{n-1} c_i^n \lambda_d^{n-1} \lambda_i^{n+1}) b_i^n$$
$$+ C b_d^n b_i^{n-1} \lambda_d^{n+1} \lambda_i^{n-1} + C b_d^n b_d^{n-1} b_i^{n-2} \lambda_d^{n+2} \lambda_i^{n-2}$$
$$\left. + \cdots + C b_d^n b_d^{n-1} \cdots b_d^2 b_i^1 \lambda_d^{2n-1} \lambda_i \right] c_i^1 v_i^T w_i$$
$$+ C b_d^n b_d^{n-1} b_d^{n-2} \cdots b_d^1 \lambda_d^{2n} c_d^1 (v_d^1)^T w_d^1.$$

The expression $C b_d^n b_d^{n-1} b_d^{n-2} \cdots b_d^1 c_d^1 (v_d^1)^T w_d^1$ can only be 0 if $c_d^1 = 0$; b_d^t cannot be 0 for any $t \in \{2, 3, \ldots, n\}$, since this would cause $w_1, w_2, \ldots, w_d^{t+1}$ to be linearly dependent, which they are not. Similarly, $c_d^t \neq 0$ for any $t \in \{2, 3, \ldots, n\}$. Furthermore, $(v_d^1)^T w_d^1 \neq 0$ by Lemma .0.10.

Since λ_d always appears with a coefficient c_d^1, this allows us to reach the same conclusion as the previous section. Namely,

$$\lim_{n \to \infty} \frac{1}{n} y^T J(n,1)^T J(n,1) y = \ln \max\{|\lambda_i| : c_i^1 \neq 0\}.$$ □

Example 1.3.11:
Recall the Kaplan-Yorke map:

$$\Phi(x_1, x_2) = (a x_1 \bmod 1, b x_2 + \cos(4\pi x_1) \bmod 1).$$

As long as $ax_1 \notin \mathbf{Z}$ and $bx_2 + \cos(4\pi x_1) \notin \mathbf{Z}$, we have

$$J(\Phi, (x_1, x_2)) = \begin{bmatrix} a & 0 \\ -4\pi \sin(4\pi x_1) & b \end{bmatrix}.$$

The eigenvalues of $J(\Phi, (x_1, x_2))$ are the roots of

$$\begin{aligned} \det \begin{bmatrix} a-\lambda & 0 \\ -4\pi \sin(4\pi x_1) & b-\lambda \end{bmatrix} &= (a-\lambda)(b-\lambda) \\ &= ab - (a+b)\lambda + \lambda^2, \end{aligned}$$

which are a and b. The corresponding eigenvectors are

$$(\frac{a-b}{-4\pi \sin 4\pi x_1}, 1)$$

and

$$(0, 1),$$

respectively. Thus, we are in a situation where the eigenvalues remain constant along a trajectory, and all but one their corresponding eigenvectors remain the same as well. We can apply Theorem 1.3.4 to conclude that the Lyapunov exponents of Φ at x are $\ln |a|$ and $\ln |b|$, as long as the trajectory starting at x never hits any of the discontinuities at integer values. ■

Example 1.3.9 can be found in [16]. The exact Lyapunov exponents calculated in Example 1.3.11 are stated in [36], but without any explanations. Our justifications through the manipulation of eigenvectors as we have presented them in Theorems 1.3.3 and 1.3.4 seem to be absent from the literature.

1.3.8 From a trajectory to its maximal Lyapunov exponent.

Given only the trajectory $x^0, x^1, x^2, \ldots$ of a point in $\mathbf{R}^d$ under some otherwise unknown differentiable map $\Phi : \mathbf{R}^d \to \mathbf{R}^d$, we can estimate the maximal Lyapunov exponent of Φ at x^0 by choosing iteratively suitable nonnegative integers $s(0)$, $s(1)$, $s(2)$, ... until the sequence of averages

$$\frac{1}{n} \sum_{t=0}^{n-1} \ln \frac{||x^{s(t)+1} - x^{t+1}||}{||x^{s(t)} - x^t||} \tag{1.26}$$

shows signs of convergence, at which time we return an estimate of its limit as an estimate of the maximal Lyapunov exponent of the trajectory. (What we are actually estimating here is the global Lyapunov exponent of Φ at x^0 with respect to direction $x^{s(0)} - x^0$; as noted at the end of Section 1.3.6, this Lyapunov exponent is likely to be the maximal Lyapunov exponent of the trajectory.)

This procedure is a close relative of the procedure described in Section 1.3.5: with y^t standing for $x^{s(t)} - x^t$, we have

$$\frac{||x^{s(t)+1} - x^{t+1}||}{||x^{s(t)} - x^t||} = \frac{||\Phi(x^t + y^t) - \Phi(x^t)||}{||y^t||}.$$

With this in mind, we must establish how to choose $s(t)$ in a suitable fashion.

How to choose s(t).

Given a nonnegative threshold z^- and positive threshold z^+, representing distances that are considered to be too close (possibly due to noise) and too far, respectively, we proceed as follows. Choose $s(0)$ to make $\left|\left|x^{s(0)} - x^0\right|\right|$ as small as possible but larger than z^-. The policy of Section 1.3.5 would guide us to choose each subsequent $s(t)$ so that it is different from t and so that

$$x^{s(t)} = x^t + \delta_t \left(x^{s(t-1)+1} - x^t \right)$$

with each δ_t a (possibly negative) number small enough in magnitude to ensure that

$$\Phi(x^{s(t)}) - \Phi(x^t) \approx J(\Phi, x^t)(x^{s(t)} - x^t), \tag{1.27}$$

but big enough to ensure that the distance between $x^{s(t)}$ and x^t is above the prescribed level of noise, z^-. Let us write

$$C_t = \{x^k : z^- < \left|\left|x^k - x^t\right|\right| < z^+\}.$$

If $x^{s(t-1)+1} \in C_t$ then we declare $\delta_t = 1$ to be small enough to satisfy (1.27), and we simply set $s(t) = s(t-1) + 1$. Otherwise we consider all candidates in the set $\{x^{s(t-1)+1}\} \cup C_t$. From this set we choose $x^{s(t)}$ so that $x^{s(t)} - x^t$ approximates a multiple of $x^{s(t-1)+1} - x^t$ that is small in magnitude. In order to do this, we note the following.

All multiples of $x^{s(t-1)+1} - x^t$ lie on a line passing through the origin; all multiples of $x^k - x^t$ lie on another line passing through the origin; the cosine of the angle between these two lines is

$$\frac{(x^{s(t-1)+1} - x^t) \cdot (x^k - x^t)}{\left|\left|x^{s(t-1)+1} - x^t\right|\right| \left|\left|x^k - x^t\right|\right|}. \tag{1.28}$$

This quantity lies between -1 and 1. Choices of k for which $x^k - x^t$ is a good approximation of $x^{s(t-1)+1} - x^t$ will have the magnitude of this quantity lying closer to 1. Consequently, candidates for $s(t)$ are ranked by a partial order $\succeq$: with $c_t(k)$ standing for the magnitude of the value in (1.28), we have

$$\left|\left|x^i - x^t\right|\right| \le \left|\left|x^j - x^t\right|\right| \text{ and } c_t(i) \ge c_t(j) \quad \Rightarrow \quad i \succeq j \tag{1.29}$$

The selected $s(t)$ is a maximal element in this partial order.

The fact that $\succeq$ is only a partial ordering leads to some ambiguity: different rules for breaking ties may lead to different outputs from the algorithm. In order to remove this ambiguity we will assume that we are given a linear order that satisfies (1.29). We will refer to this input as an *operator* (we can imagine someone manually running the algorithm and hand-picking a partner for each x^t).

The following two definitions will make this idea precise.

Definition 1.3.4:
We call a linear order $\rhd$ on $[0,1] \times (0,\infty)$ *monotone* if for all c_1, c_2, d_1, d_2 such that $c_1, c_2 \in [0,1]$ and $d_1, d_2 \in (0,\infty)$ we have

$$c_1 \geq c_2 \text{ and } d_1 \leq d_2 \quad \Rightarrow \quad (c_1, d_1) \rhd (c_2, d_2).$$

Definition 1.3.5:
By an *operator* we mean a triple $(z^-, z^+, \rhd)$ such that

- $z^- \geq 0$,
- $z^+ > z^-$, and
- $\rhd$ is a monotone order on $[0,1] \times (0,\infty)$.

Given a sequence of points $x^0, x^1, x^2, \ldots$ in $\mathbf{R}^d$ and an index t, an operator $(z^-, z^+, \rhd)$ selects $s(t)$ as follows:

if t=0 **then**
    $s(t) = \underset{k}{\operatorname{argmin}}\{\left|\left|x^k - x^0\right|\right| : z^- < \left|\left|x^k - x^0\right|\right|\}$
else if $t > 0$ **then**
    if $z^- < \left|\left|x^{s(t-1)+1} - x^t\right|\right| < z^+$ **then**
        $s(t) = s(t-1) + 1$
    else
        Let $C = \{k : z^- < \left|\left|x^k - x^t\right|\right| < z^+\}$
        Let $D = \left\{ \left(\dfrac{(x^k - x^t) \cdot (x^{s(t-1)+1} - x^t)}{||x^k - x^t|| \left|\left|x^{s(t-1)+1} - x^t\right|\right|}, \left|\left|x^k - x^t\right|\right| \right) : k \in \{s(t-1)+1\} \cup C \right\}$
        s(t) = $\underset{k}{\operatorname{argmax}} D$ with respect to $\rhd$

Algorithm 1.3.1: Selecting $s(t)$.

Note that in the case where the set C is empty (there are no valid candidates) we resort to setting $s(t) = s(t-1) + 1$, even if $\left|\left|x^{s(t-1)+1} - x^t\right|\right| \leq z^-$ or $\left|\left|x^{s(t-1)+1} - x^t\right|\right| \geq z^+$. We are not left with any other choice.

We will now use $\lambda(x^0, (z^-, z^+, \rhd))$ to denote the maximal Lyapunov exponent of the sequence $x^0, x^1, \ldots$ with respect to the operator $(z^-, z^+, \rhd)$, which can be computed via the sum in (1.26).

Distance biased operators.

One particularly simple class of operators use angles only to break ties.

Definition 1.3.6:
A linear order $\rhd$ is *distance biased* if for all c_1, c_2, d_1, d_2 such that $c_1, c_2 \in [0,1]$ and $d_1, d_2 \in (0, \infty)$ we have

$$d_1 < d_2 \quad \Rightarrow \quad (c_1, d_1) \rhd (c_2, d_2).$$

and

$$d_1 = d_2 \text{ and } c_1 \geq c_2 \quad \Rightarrow \quad (c_1, d_1) \rhd (c_2, d_2)$$

with respect to $\rhd$.

We will call $(z^-, z^+, \rhd)$ a *distance biased operator* if $\rhd$ is distance biased.

We will consider distance biased operators in more detail in Section 2.2.

A seminal operator.

The operator described in a seminal paper of Wolf et al. [40] is as follows.

The choice of $s(0)$ is such that $\left|\left|x^{s(0)} - x^0\right|\right|$ is as small as possible, but above the prescribed level of noise, z^-.

When $t > 0$, there are four values $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ such that

$$1 \geq \alpha_1 > \alpha_2 > \alpha_3 > \alpha_4 \geq 0$$

and four distances z_1, z_2, z_3, z_4 such that

$$z_1 < z_2 < z_3 < z_4.$$

Letting $\alpha_t(k)$ stand for the magnitude of the cosine of

$$\frac{(x^{s(t-1)+1} - x^t) \cdot (x^k - x^t)}{\left|\left|x^{s(t-1)+1} - x^t\right|\right| \left|\left|x^k - x^t\right|\right|},$$

if there is no k with $\alpha_t(k) \geq \alpha_4$ and $z < \left|\left|x^k - x^t\right|\right| \leq z_4$, then we set

$$s(t) = s(t-1) + 1. \tag{1.30}$$

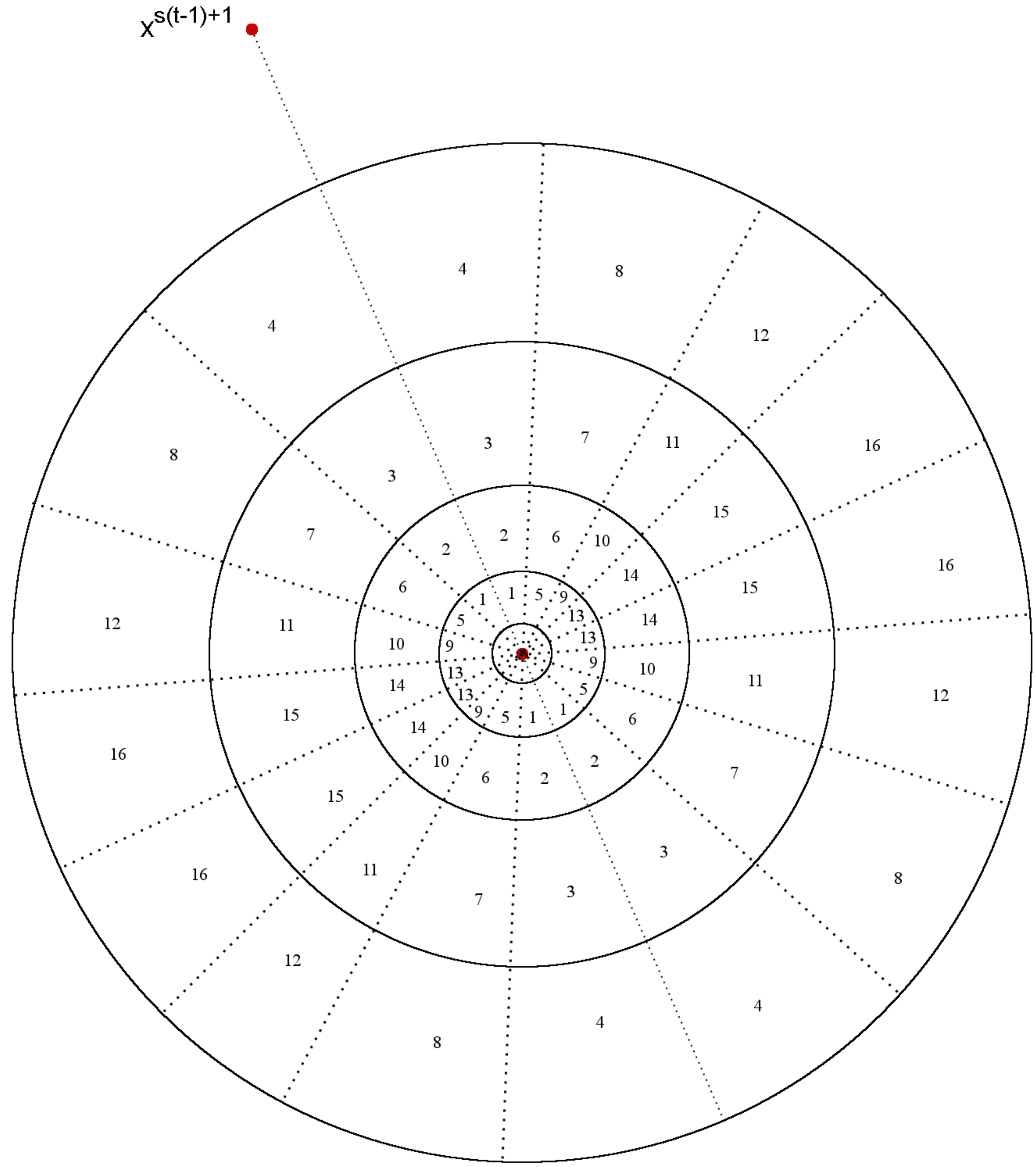
Otherwise we find

- first the smallest m for which there is a k with $\alpha_t(k) \geq \alpha_m$ and $z < \left|\left|x^k - x^t\right|\right| \leq z_4$,

- then the smallest n for which there is a k with $\alpha_t(k) \geq \alpha_m$ and $\left|\left|x^k - x^t\right|\right| \leq z_n$,
- and finally a k that maximizes $\alpha_t(k)$ subject to $z < \left|\left|x^k - x^t\right|\right| \leq z_n$;

then we set $s(t) = k$. (The "fixed evolution time" program of Wolf et al. [40] resorts to this policy only when t is a positive integer multiple of a prescribed parameter EVOLV and uses the default (1.30) for all other values of t.)

Figure 1.15: The ordering given by Wolf et al. The middle point is the current reference point, x^t. The four outer circles correspond to the four distances z_1, z_2, z_3, z_4 and the innermost circle corresponds to the noise threshold, z^-. Dotted lines correspond to the angles associated with α_1, α_2, α_3, α_4. Points from regions with smaller numbers are preferred.



The finite case.

In the case of a finite sequence of points $x^0, x^1, \ldots, x^N$ in $\mathbf{R}^d$ coming from some unknown differentiable map Φ, the procedure we just described remains the same, except that instead of waiting for equation (1.26) to show signs of convergence, we simply take the finite sum

$$\frac{1}{N}\sum_{t=0}^{N-1} \ln \frac{\left|\left|x^{s(t)+1} - x^{t+1}\right|\right|}{\left|\left|x^{s(t)} - x^{t}\right|\right|}$$

as an approximation of the maximal Lyapunov exponent of Φ at x^0.

Let us note here that in this case there is an extra condition imposed on the choice of $s(t)$. Namely that $s(t) < N$, so that $x^{s(t)+1}$ is a point we have access to. Accordingly, we modify the algorithm for selecting $s(t)$ to the following.

if t=0 **then**
  $s(t) = \underset{0<k<N}{\operatorname{argmin}}\{\left|\left|x^k - x^0\right|\right| : z^- < \left|\left|x^k - x^0\right|\right|\}$
else if $t > 0$ and $s(t-1)+1 < N$ **then**
  if $z^- < \left|\left|x^{s(t-1)+1} - x^t\right|\right| < z^+$ **then**
    $s(t) = s(t-1)+1$
  else
    Let $C = \{k : z^- < \left|\left|x^k - x^t\right|\right| < z^+,\ 0 \le k < N\}$
    Let $D = \left\{ \left(\frac{(x^k - x^t)\cdot(x^{s(t-1)+1} - x^t)}{\left|\left|x^k - x^t\right|\right| \left|\left|x^{s(t-1)+1} - x^t\right|\right|}, \left|\left|x^k - x^t\right|\right| \right) : k \in \{s(t-1)+1\} \cup C \right\}$
    s(t) = $\underset{0\le k<N}{\operatorname{argmax}} D$ with respect to $\triangleright$

Algorithm 1.3.2: Selecting $s(t)$ in the finite case.

1.3.9 Trajectory-like sequences.

We began Section 1.3.8 by supposing we were given a sequence of points $x^0, x^1, x^2, \ldots$ in $\mathbf{R}^d$, and that the sequence came from some unknown differentiable map. In practical situations, however, it is likely that we will be given only a sequence of points, with no knowledge of what may have been the generating force behind it. In such instances the method for approximating the maximal Lyapunov exponent of a sequence of points that we have described may still be applied, as long as the sequence obeys the deterministic nature of a dynamical system. The following is really what we are after.

Fact:
Given any sequence of points $x^0, x^1, x^2, \ldots$ in $\mathbf{R}^d$, the following are logically equivalent:

1. There exists a map $\Phi : \mathbf{R}^d \to \mathbf{R}^d$ such that for all nonnegative integers i, we have $\Phi(x^i) = x^{i+1}$.

2. For all nonnegative integers i, j, k we have that

$$x^i = x^j \quad \Rightarrow \quad x^{i+k} = x^{j+k}.$$

Let us give a name to the sequences for which these statements hold.

Definition 1.3.7:
The sequence of points in $\mathbf{R}^d$ given by $x^0, x^1, x^2, \ldots$ is called *trajectory-like* if for all nonnegative integers i, j, k we have that

$$x^i = x^j \quad \Rightarrow \quad x^{i+k} = x^{j+k}.$$

We will also refer to a finite sequence of points as *trajectory-like*, if it is the truncated sequence of some infinite trajectory-like sequence.

Example 1.3.12:
We should note that a sequence can be trajectory-like without there necessarily existing a continuous map that generates it. Consider, for example, the sequence $x^1, x^2, x^3, \ldots$ of points in $\mathbf{R}$ given by

$$x^i = \begin{cases} 0 + \dfrac{1}{i} & \text{if } i = 1 \pmod 4, \\[2ex] \dfrac{1}{3} + \dfrac{1}{i} & \text{if } i = 0 \pmod 2, \\[2ex] \dfrac{2}{3} + \dfrac{1}{i} & \text{if } i = 3 \pmod 4. \end{cases}$$

No number appears more than once in the sequence, so it is trivially trajectory-like. However, there are points arbitrarily close to 1/3 that get mapped arbitrarily close to 0, and points arbitrarily close to 1/3 that get mapped arbitrarily close to 2/3, so no map generating this sequence can be continuous at 1/3. Of course, this means that trajectory-like sequences need not have a differentiable map generating them. ■

With these observations out of the way, how does the algorithm behave on specific trajectory-like sequences? Our first result is about trajectory-like sequences that end where they begin. We start with the following Lemma.

Lemma 1.3.2:
For every trajectory-like sequence $x^0, x^1, x^2, \ldots, x^N$ of points in $\mathbf{R}^d$, for every operator $(z^-, z^+, \triangleright)$, and for all integers t such that $0 < t < N$ we have that $\left|\left|x^{s(t)} - x^t\right|\right| \leq \left|\left|x^{s(t-1)+1} - x^t\right|\right|$.

Proof:
The proof proceeds by examining algorithm 1.3.2. Let t be any integer such that $0 < t < N$. We see that if $z^- < \left|\left|x^{s(t-1)+1} - x^t\right|\right| < z^+$, then $s(t) = s(t-1) + 1$ and the statements holds trivially. If

this is not the case, then we note that $s(t)$ is such that it results in a maximal element in the set D. This means that with v standing for $x^{s(t)} - x^t$, and w standing for $x^{s(t-1)+1} - x^t$, then with respect to $\rhd$ we must have

$$\left(\frac{v \cdot w}{||v||\,||w||}, ||v||\right) \geq \left(\frac{w \cdot w}{||w||\,||w||}, ||w||\right) = (1, ||w||).$$

Since

$$\frac{v \cdot w}{||v||\,||w||} \leq 1,$$

we must have $||v|| \geq ||w||$. □

Theorem 1.3.5:
If $x^0, x^1, x^2, \ldots, x^N$ is a trajectory-like sequence of points in $\mathbf{R}^d$ such that $x^0 = x^N$, then for every operator $(z^-, z^+, \rhd)$ such that $z^- = 0$ we have $\lambda(x^0, (z^-, z^+, \rhd)) \geq 0$.

Proof:
Recall that the maximal Lyapunov exponent of the trajectory is

$$\frac{1}{N}\sum_{t=0}^{N-1} \ln \frac{\left|\left|x^{s(t)+1} - x^{t+1}\right|\right|}{\left|\left|x^{s(t)} - x^t\right|\right|},$$

which is equal to

$$\frac{1}{N}\sum_{t=0}^{N-1} \left[\ln \left|\left|x^{s(t)+1} - x^{t+1}\right|\right| - \ln \left|\left|x^{s(t)} - x^t\right|\right|\right].$$

By the previous Lemma, we must have

$$\left|\left|x^{s(t)+1} - x^{t+1}\right|\right| \geq \left|\left|x^{s(t+1)} - x^{t+1}\right|\right|$$

for all integers t such that $0 \leq t \leq N-2$. Thus, we have

$$\begin{aligned}\sum_{t=0}^{N-1} \left[\ln \left|\left|x^{s(t)+1} - x^{t+1}\right|\right| - \ln \left|\left|x^{s(t)} - x^t\right|\right|\right] &= \sum_{t=0}^{N-2} \left[\ln \left|\left|x^{s(t)+1} - x^{t+1}\right|\right| - \ln \left|\left|x^{s(t+1)} - x^{t+1}\right|\right|\right] \\ &\quad + \ln \left|\left|x^{s(N-1)+1} - x^N\right|\right| - \ln \left|\left|x^{s(0)} - x^0\right|\right| \\ &\geq -\ln \left|\left|x^{s(0)} - x^0\right|\right| + \ln \left|\left|x^{s(N-1)+1} - x^N\right|\right|.\end{aligned}$$

Since $z^- = 0$ and we know that $s(0) = \operatorname{argmin}_k \{\left|\left|x^k - x^0\right|\right| : z^- < \left|\left|x^k - x^0\right|\right|\}$, we must have

$$\left|\left|x^{s(N-1)+1} - x^0\right|\right| \geq \left|\left|x^{s(0)} - x^0\right|\right|, \tag{1.31}$$

unless $\left|\left|x^{s(N-1)+1} - x^0\right|\right| \leq z^- = 0$. This cannot be the case, as we will now show. Suppose first that x^0 appears exactly twice in the sequence (as x^0 and x^N). In order to have $\left|\left|x^{s(N-1)+1} - x^0\right|\right| = 0$, we must have $x^{s(N-1)+1} = x^0$, and therefore we must have that $x^{s(N-1)} = x^{-1}$, which is not allowed, or $x^{s(N-1)} = x^{N-1}$, which is not allowed. Now suppose that x^0 appears more than twice. Let p be the smallest positive integer for which $x^0 = x^p$. Noting that

$$\left|\left|x^{s(N-1)+1} - x^0\right|\right| = 0 \quad \Leftrightarrow \quad x^{s(N-1)+1} = x^0,$$

let k be the positive integer such that $s(N-1)+1 = kp$. Since the sequence in question is trajectory-like, we have

$$x^0 = x^p = x^{2p} = \cdots = x^{kp} = \cdots x^N$$

and therefore,

$$x^{p-1} = x^{2p-1} = \cdots = x^{kp-1} = \cdots x^{N-1}.$$

We cannot have $x^{kp-1} = x^{s(N-1)} = x^{N-1}$, because we require $\left|\left|x^{s(N-1)} - x^{N-1}\right|\right| > z^-$.

We conclude that the maximal Lyapunov exponent of $x^0, x^1, x^2, \ldots, x^{N-1}, x^0$ is nonnegative, with respect to the given operator. □

Note that the proof shows that the statement of the Theorem actually holds for every z^- such that $\left|\left|x^{s(N-1)+1} - x^0\right|\right| \leq z^-$.

This immediately lends itself to the following observation.

Corollary 1.3.1:
Let $x^0, x^1, x^2, \ldots, x^N$ be a trajectory-like sequence of points in $\mathbf{R}^d$ and let $(z^-, z^+, \triangleright)$ be an operator such that $z^- = 0$. If the sequence of points $x^0, x^1, x^2, \ldots, x^N, x^0$ is trajectory-like, then its maximal Lyapunov exponent is nonnegative with respect to $(z^-, z^+, \triangleright)$.

This Corollary is particularly interesting if we apply it to a trajectory-like sequence of points that has a negative maximal Lyapunov exponent, which we will encounter in Section 2.2.

Note also that if

$$x^0, x^1, x^2, \ldots, x^N$$

has no repeating points and is trajectory-like, then

$$x^0, x^1, x^2, \ldots, x^N, x^0$$

is trajectory-like.

1.3.10 Primitive trajectories.

There is a special case in which the sum in equation 1.26 is trivial. Before explicitly stating that case, we will introduce the notion of a *forward-primitive* trajectory.

Definition 1.3.8:
A trajectory-like sequence of points $x^0, x^1, \ldots, x^N$ in $\mathbf{R}^d$ is *forward-primitive* with respect to an operator $(z^-, z^+, \rhd)$ if for every integer t such that $0 \leq t \leq N-2$ we have $s(t) = t+1$.

Here is why these trajectories are special.

Theorem 1.3.6:
If the sequence of trajectory-like points in $\mathbf{R}^d$ given by $x^0, x^1, \ldots, x^N$ is forward-primitive with respect to an operator $(z^-, z^+, \rhd)$, then

$$\lambda(x^0, (z^-, z^+, \rhd)) = \frac{1}{N}\left(-\ln\left|\left|x^1 - x^0\right|\right| + \ln\left|\left|x^N - x^{N-1}\right|\right| + \ln\frac{\left|\left|x^{s(N-1)+1} - x^N\right|\right|}{\left|\left|x^{s(N-1)} - x^{N-1}\right|\right|}\right).$$

Proof:
In this case, equation (1.26) simply telescopes and becomes

$$\begin{aligned}
&\frac{1}{N}\sum_{t=0}^{N-1}\ln\frac{\left|\left|x^{s(t)+1} - x^{t+1}\right|\right|}{\left|\left|x^{s(t)} - x^t\right|\right|} \\
&= \frac{1}{N}\left(\ln\frac{\left|\left|x^{s(0)+1} - x^1\right|\right|}{\left|\left|x^{s(0)} - x^0\right|\right|} + \sum_{t=1}^{N-2}\ln\frac{\left|\left|x^{s(t)+1} - x^{t+1}\right|\right|}{\left|\left|x^{s(t)} - x^t\right|\right|} + \ln\frac{\left|\left|x^{s(N-1)+1} - x^N\right|\right|}{\left|\left|x^{s(N-1)} - x^{N-1}\right|\right|}\right) \\
&= \frac{1}{N}\left(\ln\frac{\left|\left|x^2 - x^1\right|\right|}{\left|\left|x^1 - x^0\right|\right|} + \sum_{t=1}^{N-2}\left(\ln\left|\left|x^{t+2} - x^{t+1}\right|\right| - \ln\left|\left|x^{t+1} - x^t\right|\right|\right) + \ln\frac{\left|\left|x^{s(N-1)+1} - x^N\right|\right|}{\left|\left|x^{s(N-1)} - x^{N-1}\right|\right|}\right) \\
&= \frac{1}{N}\left(-\ln\left|\left|x^1 - x^0\right|\right| + \ln\left|\left|x^N - x^{N-1}\right|\right| + \ln\frac{\left|\left|x^{s(N-1)+1} - x^N\right|\right|}{\left|\left|x^{s(N-1)} - x^{N-1}\right|\right|}\right). \qquad \square
\end{aligned}$$

Similarly, a trajectory may be *backward-primitive*.

Definition 1.3.9:
A trajectory-like sequence of points $x^0, x^1, \ldots, x^N$ in $\mathbf{R}^d$ is *backward-primitive* with respect to an operator $(z^-, z^+, \rhd)$ if for every integer t such that $1 \leq t \leq N-1$ we have $s(t) = t-1$.

Theorem 1.3.7:
If the sequence of trajectory-like points in $\mathbf{R}^d$ given by $x^0, x^1, \ldots, x^N$ is backward-primitive with respect to an operator $(z^-, z^+, \rhd)$, then

$$\lambda(x^0, (z^-, z^+, \rhd)) = \frac{1}{N}\left(\ln\frac{\left|\left|x^{s(0)+1} - x^1\right|\right|}{\left|\left|x^{s(0)} - x^0\right|\right|} + \ln\left|\left|x^{N-1} - x^N\right|\right| - \ln\left|\left|x^0 - x^1\right|\right|\right).$$

Proof:
Once again, equation (1.26) telescopes and we get

$$\begin{aligned}
\lambda &= \frac{1}{N}\sum_{t=0}^{N-1}\ln\frac{\left|\left|x^{s(t)+1}-x^{t+1}\right|\right|}{\left|\left|x^{s(t)}-x^{t}\right|\right|}\\
&= \frac{1}{N}\left(\ln\frac{\left|\left|x^{s(0)+1}-x^{1}\right|\right|}{\left|\left|x^{s(0)}-x^{0}\right|\right|}+\sum_{t=1}^{N-1}\ln\frac{\left|\left|x^{s(t)+1}-x^{t+1}\right|\right|}{\left|\left|x^{s(t)}-x^{t}\right|\right|}\right)\\
&= \frac{1}{N}\left(\ln\frac{\left|\left|x^{s(0)+1}-x^{1}\right|\right|}{\left|\left|x^{s(0)}-x^{0}\right|\right|}+\sum_{t=1}^{N-1}\left(\ln\left|\left|x^{t}-x^{t+1}\right|\right|-\ln\left|\left|x^{t-1}-x^{t}\right|\right|\right)\right)\\
&= \frac{1}{N}\left(\ln\frac{\left|\left|x^{s(0)+1}-x^{1}\right|\right|}{\left|\left|x^{s(0)}-x^{0}\right|\right|}+\ln\left|\left|x^{N-1}-x^{N}\right|\right|-\ln\left|\left|x^{0}-x^{1}\right|\right|\right). \qquad \square
\end{aligned}$$

Chapter 2

Time Series and Lyapunov Exponents

2.1 The maximal Lyapunov exponent of a time series.

Given positive integers d, τ and a sequence $T = \xi_0, \xi_1, \xi_2, \ldots$ of real numbers, let us write

$$E(T, \tau, d) = x^0, x^1, x^2, \ldots$$

where

$$x^t = (\xi_t, \xi_{t+\tau}, \xi_{t+2\tau}, \ldots, \xi_{t+(d-1)\tau})$$

for all nonnegative integers t. This process is usually found in the literature under the term *phase space reconstruction* or *method of delays*.

If $E(T, \tau, d)$ is trajectory-like then we refer to its maximal Lyapunov exponent (computed as in Section 1.3.8) as *the maximal Lyapunov exponent of T in dimension d and with time delay τ with respect to the operator* $(z^-, z^+, \rhd)$, and we will denote this by $\lambda(E(T, \tau, d), (z^-, z^+, \rhd))$. This notation is becoming cumbersome, but we will leave it this way as a reminder that there is no such thing as "the maximal Lyapunov exponent of a time series": many parameters are required.

Takens' Theorem.

Here are a few comments on the notion we have just described: When a map $\Phi : X \to X$ describes deterministic evolution of some physical system, measurements in this system define functions $\pi : X \to \mathbf{R}$. The trajectory $x^0, x^1, x^2, \ldots$ of a state x^0 under Φ can be often reconstructed from the measured values $\pi(x^0), \pi(x^1), \pi(x^2), \ldots$. The celebrated and often-cited Theorem of Takens (proved in 1981, see [38] for the proof or [32] for an in depth discussion) states that under certain conditions (for instance Φ and π must be smooth[1]), the mapping

$$x^t \to (\pi(x^t), \pi(\Phi(x^t)), \pi(\Phi^2(x^t)), \ldots, \pi(\Phi^{d-1}(x^t)))$$

is a diffeomorphism[2], provided d is large enough. Thus, with the notation

$$\widetilde{x}^t = (\pi(x^t), \pi(x^{t+\tau}), \pi(x^{t+2\tau}), \ldots, \pi(x^{t+(d-1)\tau})),$$

the asymptotic behaviour of the trajectory

$$\widetilde{x}^0, \widetilde{x}^1, \widetilde{x}^2, \ldots$$

can be thought of as a model of the trajectory

$$x^0, x^1, x^2, \ldots.$$

The maximal Lyapunov exponent of the trajectory $x^0, x^1, x^2, \ \ldots$ is a quantity that has a physical meaning; the maximal Lyapunov exponent of the trajectory $\widetilde{x}^0, \widetilde{x}^1, \widetilde{x}^2, \ldots$ (also known as the maximal Lyapunov exponent of the time series $\pi(x^0), \pi(x^1), \pi(x^2), \ \ldots$) is a quantity that we can compute.

Example 2.1.1:
The Hénon map $H : \mathbf{R}^2 \to \mathbf{R}^2$ defined by

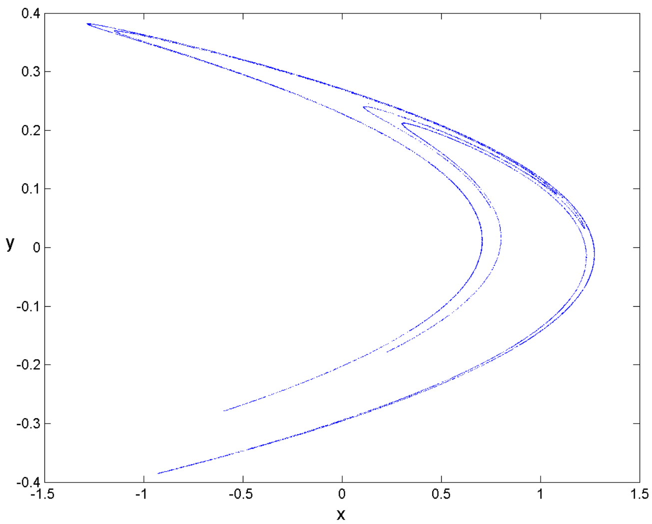
$$H(x, y) = (y + 1 - ax^2, bx)$$

was introduced by Michel Hénon in 1976 (see [13]) and will help clarify the meaning of Takens' Theorem. In particular, the values of a and b that make the map interesting are $a = 1.4$ and $b = 0.3$.

A typical trajectory results in the following set of points in the plane.

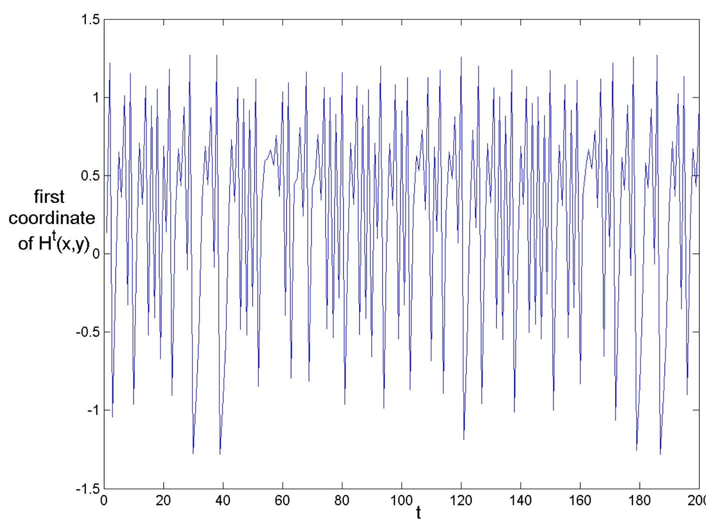
[1] A function is smooth if it has derivatives of all orders.
[2] A bijective map between manifolds that is differentiable and has a differentiable inverse.

Figure 2.1: Hénon attractor with initial condition $(0.13, 0.24)$.



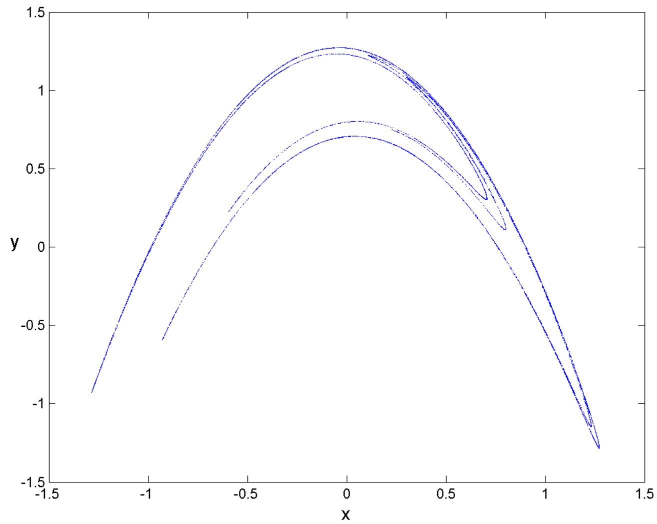
If we ignore the second coordinate of each point along the trajectory, we get a one-dimensional trajectory. More precisely, we are taking the measurement function to be $\pi(x, y) = x$, for every point (x, y) on the trajectory. This results in a sequence of real numbers, which we may view as a time series.

Figure 2.2: First coordinates of a trajectory of the Hénon map as a time series.



We may now embeded the time series in $\mathbf{R}^2$ by taking $x^t = \left(\pi\left(H^t(x,y)\right), \pi\left(H^{t+1}(x,y)\right)\right)$ for all nonnegative integers t.

Figure 2.3: Embeded trajectory.



What results is the familiar shape of the Hénon attractor, although oriented in a different manner.■

Finite case.

Before proceeding, let us make note that in the case of a finite time series $T = \xi_0, \xi_1, \ldots, \xi_N$ and embedding dimension d, the remark about the restriction on $s(t)$ at the end of Section 1.3.8 now becomes that $s(t)$ cannot be larger than $N - d$.

2.2 The maximal Lyapunov exponent of strictly monotonic time series.

In order to help build some intuition about the connection between what a time series looks like and what its maximal Lyapunov exponent is, we should begin by taking a look at some simple time series. If we can find situations in which the procedure for choosing $s(t)$ from Section 1.3.8 is straightforward, then the calculation is vastly simplified.

We shall begin by dealing with strictly monotonic time series.

Definition 2.2.1:
A time series $T = \xi_0, \xi_1, \xi_2, \ldots$ is *strictly monotonic* if it is strictly increasing or strictly decreasing.

Note that if T is strictly monotonic then $E(T, \tau, d)$ is trivially trajectory-like.

For the rest of this section we will assume that we are given a distance biased operator (operators that only consider angles in order to break distance ties). Furthermore, since we are in a pure setting we will assume that our time series are noiseless, so that we may take $z^- = 0$. For this same reason, it will be with no loss of generality that we may set $\tau = 1$.

Lemma 2.2.1:
If $T = \xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a strictly monotonic time series then for every distance biased operator $(0, z^+, \rhd)$ we have $s(0) = 1$.

Proof:
Recall that

$$s(0) = \underset{0<k<N-d+1}{\text{argmin}} \{ \left|\left|x^k - x^0\right|\right| : z^- < \left|\left|x^k - x^0\right|\right| \}.$$

Suppose that T is strictly increasing. Then for every positive integer k such that $k < N - d + 1$ we have

$$\xi_0 < \xi_1 < \xi_2 < \ldots < \xi_k$$

and therefore

$$\xi_k - \xi_0 > \xi_{k-1} - \xi_0 > \xi_{k-2} - \xi_0 > \ldots > \xi_1 - \xi_0.$$

Since each of these differences is positive, we have

$$(\xi_k - \xi_0)^2 > (\xi_{k-1} - \xi_0)^2 > (\xi_{k-2} - \xi_0)^2 > \ldots > (\xi_1 - \xi_0)^2.$$

Thus, for every positive integer k such that $1 < k < N - d + 1$, we have

$$0 < \left|\left|x^1 - x^0\right|\right| = \sqrt{\sum_{i=0}^{d-1} (\xi_{i+1} - \xi_i)^2} < \sqrt{\sum_{i=0}^{d-1} (\xi_{i+k} - \xi_i)^2} = \left|\left|x^k - x^0\right|\right|.$$

Therefore $s(0) = 1$. The proof for the case where T is strictly decreasing is similar. □

Lemma 2.2.2:
If $T = \xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a strictly monotonic time series, then

1. *for all integers k such that $t + 1 < k \leq N - d + 1$ we have $\left|\left|x^{t+1} - x^t\right|\right| < \left|\left|x^k - x^t\right|\right|$, and*

2. *for all integers k such that $0 \leq k < t-1$ we have $\left|\left|x^{t-1} - x^t\right|\right| < \left|\left|x^k - x^t\right|\right|$.*

Proof:
Suppose that T is strictly increasing. To prove the first claim, note that for all integers k such that $t+1 < k \leq N-d+1$, we have

$$\left|\left|x^{t+1} - x^t\right|\right| = \sqrt{\sum_{i=0}^{d-1}(\xi_{t+1+i} - \xi_{t+i})^2} < \sqrt{\sum_{i=0}^{d-1}(\xi_{k+i} - \xi_{t+i})^2} = \left|\left|x^k - x^t\right|\right|,$$

since $k > t+1$ implies $\xi_{t+1+i} < \xi_{k+i}$ for all integers k such that $t+1 < k \leq N-d+1$.

Similarly, the second claim follows by noting that for all integers k such that $0 \leq k < t-1$, we have

$$\left|\left|x^{t-1} - x^t\right|\right| = \sqrt{\sum_{i=0}^{d-1}(\xi_{t-1+i} - \xi_{t+i})^2} < \sqrt{\sum_{i=0}^{d-1}(\xi_{k+i} - \xi_{t+i})^2} = \left|\left|x^k - x^t\right|\right|,$$

since $k < t-1$ implies $\xi_{t-1+i} < \xi_{k+i}$ for all integers k such that $0 \leq k < t-1$. The case where T is strictly decreasing is handled similarly. □

Lemma 2.2.3:
If $T = \xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a strictly monotonic time series, then for every positive integer t such that $1 \leq t \leq N-d$ and for every distance biased operator $(0, z^+, \triangleright)$ we have that either $s(t) = t+1$ or $s(t) = t-1$.

Proof:
We will prove this by induction on t. When $t = 0$, the statement holds by Lemma 2.2.1. Suppose now that the statement is true for $t-1$, which means that $s(t-1) = t-2$ or $s(t-1) = t$.

If

$$0 < \left|\left|x^{s(t-1)+1} - x^t\right|\right| < z^+$$

then our algorithm for choosing $s(t)$ simply sets $s(t) = s(t-1)+1$, which is either $t-1$ or $t+1$, so the statement holds.

If

$$\left|\left|x^{s(t-1)+1} - x^t\right|\right| \geq z^+$$

then if there is no k for which $0 < \left|\left|x^k - x^t\right|\right| < z^+$, our algorithm sets $s(t) = s(t-1)+1$ and we are done by the arguments in the previous case. Otherwise, by Lemma 2.2.2 we set $s(t) = t-1$ or $s(t) = t+1$.

Finally, it cannot be the case that

$$\left|\left|x^{s(t-1)+1} - x^t\right|\right| = 0,$$

since if $s(t-1) = t-2$ we have $\left|\left|x^{t-1} - x^t\right|\right| = 0$, which is impossible since T is strictly monotonic. Similarly, if $s(t-1) = t$ then $\left|\left|x^{t+1} - x^t\right|\right| = 0$ which is impossible since T is strictly monotonic. □

Our plan from here is to get our hands on a family of time series that will turn out to be primitive, in either the forward or backward sense, as defined in Section 1.3.10. In order to arrive at this, let us introduce the notion of *convex* and *concave* sequences.

Definition 2.2.2:
If $\xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a sequence of real numbers such that for all integers t such that $0 \leq t \leq N-2$ we have $\xi_{t+1} < (\xi_t + \xi_{t+2})/2$, then the sequence is *strictly convex.*

Definition 2.2.3:
If $\xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a sequence of real numbers such that for all integers t such that $0 \leq t \leq N-2$ we have $\xi_{t+1} > (\xi_t + \xi_{t+2})/2$, then the sequence is *strictly concave.*

Lemma 2.2.4:
Let $T = \xi_0, \xi_1, \xi_2, \ldots, \xi_N$ be a strictly monotonic time series. Then the sequence of points $E(T,1,d) = x^0, x^1, \ldots, x^{N-d+1}$ in $\mathbf{R}^d$ is such that

1. *If T is strictly convex then for all positive integers t and k such that $t < k \leq N-d+1$ we have $\left|\left|x^t - x^{t-1}\right|\right| < \left|\left|x^k - x^{k-1}\right|\right|$.*
2. *If T is strictly concave then for all positive integers t and k such that $t < k \leq N-d+1$ we have $\left|\left|x^t - x^{t-1}\right|\right| > \left|\left|x^k - x^{k-1}\right|\right|$.*

Proof:
Let us fix an integer t, where $0 \leq t \leq N-d$. First let us show that the first statement is true for $k = t+1$. For every integer i such that $0 \leq i+t \leq N-2$ we have $2\xi_{i+t} < \xi_{i+t-1} + \xi_{i+t+1}$ by the definition of a strictly convex time series.

Thus, if T is strictly increasing then $\xi_{i+t} - \xi_{i+t-1} < \xi_{i+t+1} - \xi_{i+t}$, which implies $(\xi_{i+t} - \xi_{i+t-1})^2 < (\xi_{i+t+1} - \xi_{i+t})^2$.

If T is strictly decreasing then $\xi_{i+t} - \xi_{i+t+1} < \xi_{i+t-1} - \xi_{i+t}$, which implies $(\xi_{i+t} - \xi_{i+t+1})^2 < (\xi_{i+t-1} - \xi_{i+t})^2$.

In either case, this implies

$$\sqrt{\sum_{i=0}^{d-1} (\xi_{t-1+i} - \xi_{t+i})^2} < \sqrt{\sum_{i=0}^{d-1} (\xi_{t+1+i} - \xi_{t+i})^2}$$

and thus $\left|\left|x^{t-1} - x^t\right|\right| < \left|\left|x^{t+1} - x^t\right|\right|$. Iterating this inequality gives us the desired result. The proof in the concave case follows along the same lines. $\square$

We are now ready to return to the notion of a primitive trajectory.

Theorem 2.2.1:
If $T = \xi_0, \xi_1, \xi_2, \ldots, \xi_N$ is a strictly monotonic time series, then

1. *If T is strictly convex then for every distance biased operator $(0, z^+, \rhd)$ such that $\left|\left|x^0 - x^1\right|\right| < z^+ \leq \left|\left|x^1 - x^2\right|\right|$ we have that $E(T, 1, d)$ is backward-primitive.*
2. *If T is strictly concave then for every distance biased operator $(0, z^+, \rhd)$ we have that $E(T, 1, d)$ is forward-primitive.*

Proof:
Let us first prove the second statement by induction. We know that $s(0) = 1$ by Lemma 2.2.1. Suppose now that $s(t-1) = t$ for some t such that $t < N - d - 1$. If $0 < \left|\left|x^{s(t-1)+1} - x^t\right|\right| < z^+$ then we are done, since our algorithm sets $s(t) = s(t-1) + 1 = t + 1$. Otherwise, by Lemma 2.2.3 there is only one candidate to consider: x^{t-1}. However, by the previous Lemma we know that $\left|\left|x^{t-1} - x^t\right|\right| > \left|\left|x^{t+1} - x^t\right|\right| > z^+$ so that x^{t-1} is not a candidate either. Thus, the algorithm choose $s(t) = t + 1$ by default, and we are done.

Let us now prove the first statement. First, we know that $s(0) = 1$. We will prove the statement by induction on t. Note that we have chosen z^+ conveniently enough so that $s(1) = 0$. Now take any t such that $t < N - d$. We inductively assume $s(t-1) = t - 2$. We have two candidates to consider for $s(t)$: $t - 1$ and $t + 1$. By the previous Lemma, $\left|\left|x^{t+1} - x^t\right|\right| > \left|\left|x^2 - x^1\right|\right| \geq z^+$, so that our algorithm resorts to the default choice of $s(t) = s(t-1) + 1 = t - 1$. □

Corollary 2.2.1:
Let $T = \xi_0, \xi_1, \ldots, \xi_N$ be a strictly monotonic time series. We have the following.

1. *If T is strictly convex then for every distance biased operator $(0, z^+, \rhd)$ such that $\left|\left|x^0 - x^1\right|\right| < z^+ \leq \left|\left|x^1 - x^2\right|\right|$ we have*

$$\lambda(E(T,1,d),(0,z^+,\rhd)) = \frac{-2\ln\left|\left|x^1 - x^0\right|\right| + \ln\left|\left|x^2 - x^1\right|\right| + \ln\left|\left|x^{N-d+1} - x^{N-d}\right|\right|}{N-d+1}. \quad (2.1)$$

2. *If T is strictly concave then for every distance biased operator $(0, z^+, \rhd)$ we have*

$$\lambda(E(T,1,d),(0,z^+,\rhd)) = \frac{-\ln\left|\left|x^1 - x^0\right|\right| - \ln\left|\left|x^{N-d} - x^{N-d-1}\right|\right| + 2\ln\left|\left|x^{N-d+1} - x^{N-d}\right|\right|}{N-d+1}. \quad (2.2)$$

Proof:
In either case, by Lemma 2.2.1 we know that $s(0) = 1$. Let us suppose that T is strictly convex. Then by Theorem 2.2.1 we know that $E(T, 1, d)$ is backward-primitive. Thus we can apply Theorem 1.3.7.

Suppose now that T is strictly concave. Then by Theorem 2.2.1 we know that $E(T, 1, d)$ is forward-primitive and we can apply Theorem 1.3.6. □

Corollary 2.2.2:
Let $T = \xi_0, \xi_1, \ldots, \xi_N$ be a strictly monotonic time series that is strictly convex. If we let $\bar{T}$ denote T in reverse order ($\bar{T} = \omega_0, \omega_1, \ldots, \omega_N$ where $\omega_i = \xi_{N-i}$) then for every distance biased operator $(0, z^+, \triangleright)$ such that $\left|\left|x^0 - x^1\right|\right| < z^+ \leq \left|\left|x^1 - x^2\right|\right|$ we have

$$\lambda(E(T, 1, d), (0, z^+, \triangleright)) = -\lambda(E(\bar{T}, 1, d), (0, z^+, \triangleright)).$$

Proof:
For every d and for all integers i such that $1 \leq i \leq N - d + 1$ let us write

$$y^i = (\xi_{N-i}, \xi_{N-i-1}, \ldots, \xi_{N-i-d+1}).$$

Now we may note that for every positive integer i such that $1 \leq i \leq N - d + 1$ we have

$$\begin{aligned} \left|\left|y^i - y^{i-1}\right|\right| &= ||(\xi_{N-i} - \xi_{N-i+1}, \xi_{N-i-1} - \xi_{N-i}, \ldots, \xi_{N-i-d+1} - \xi_{N-i-d+2})|| \\ &= ||(\xi_{N-i-d+1} - \xi_{N-i-d+2}, \ldots, \xi_{N-i-1} - \xi_{N-i}, \ldots, \xi_{N-i} - \xi_{N-i+1})|| \\ &= \left|\left|x^{N-i-d+1} - x^{N-i-d+2}\right|\right|. \end{aligned}$$

If T is strictly convex and strictly monotonic then $\bar{T}$ is strictly concave and strictly monotonic. Thus we may apply equation (2.2) to get that

$$\begin{aligned} \lambda(E(T, 1, d), (0, z^+, \triangleright)) &= \frac{-\ln\left|\left|y^1 - y^0\right|\right| - \ln\left|\left|y^{N-d} - y^{N-d-1}\right|\right| + 2\ln\left|\left|y^{N-d+1} - y^{N-d}\right|\right|}{N - d + 1} \\ &= \frac{-\ln\left|\left|x^{N-d} - x^{N-d+1}\right|\right| - \ln\left|\left|x^1 - x^2\right|\right| + 2\ln\left|\left|x^0 - x^1\right|\right|}{N - d + 1} \\ &= -\lambda(E(\bar{T}, 1, d), (0, z^+, \triangleright)). \end{aligned}$$

□

Similarly, we have the following.

Corollary 2.2.3:
Let $T = \xi_0, \xi_1, \ldots, \xi_N$ be a strictly monotonic time series that is strictly concave. If we let $\bar{T}$ denote T in reverse order ($\bar{T} = \omega_0, \omega_1, \ldots, \omega_N$ where $\omega_i = \xi_{N-i}$) then for every distance biased operator $(0, z^+, \triangleright)$ such that $\left|\left|x^{N-d+1} - x^{N-d}\right|\right| < z^+ \leq \left|\left|x^{N-d} - x^{N-d-1}\right|\right|$ we have

$$\lambda(E(T, 1, d), (0, z^+, \triangleright)) = -\lambda(E(\bar{T}, 1, d), (0, z^+, \triangleright)).$$

As we have mentioned before, there is often much emphasis placed on whether the maximal Lyapunov exponent is positive or negative. The following Corollary tells that this is easily deduced for the types of time series we have been considering.

Corollary 2.2.4:
If $T = \xi_0, \xi_1, \ldots, \xi_N$ is a strictly monotonic time series, then

1. *If T is strictly convex then for every distance biased operator $(0, z^+, \triangleright)$ such that $\left|\left|x^0 - x^1\right|\right| < z^+ \leq \left|\left|x^1 - x^2\right|\right|$ we have*

$$\lambda(E(T, 1, d), (0, z^+, \triangleright)) > 0.$$

2. *If T is strictly concave then for every distance biased operator $(0, z^+, \triangleright)$ we have*

$$\lambda(E(T,1,d),(0,z^+,\triangleright)) < 0.$$

Proof:
Suppose T is strictly convex. Then by the equation (2.2) we have,

$$\lambda(E(T,1,d)) = \frac{-2\ln\left|\left|x^1 - x^0\right|\right| + \ln\left|\left|x^2 - x^1\right|\right| + \ln\left|\left|x^{N-d+1} - x^{N-d}\right|\right|}{N-d+1}.$$

By Lemma 2.2.4 we know that $\left|\left|x^2 - x^1\right|\right| > \left|\left|x^1 - x^0\right|\right|$ and $\left|\left|x^{N-d} - x^{N-d+1}\right|\right| > \left|\left|x^1 - x^0\right|\right|$, so that $\lambda(E(T,1,d)) > 0$. The strictly concave case is dealt with similarly. □

Now that we have an example of a time series that has a negative Lyapunov exponent, we can return to Corollary 1.3.1. Combining it with the second part of Corollary 2.2.4, it tells us that if we take a strictly monotonic strictly concave time series, its Lyapunov exponent becomes nonnegative by simply appending the first point to the end of the time series.

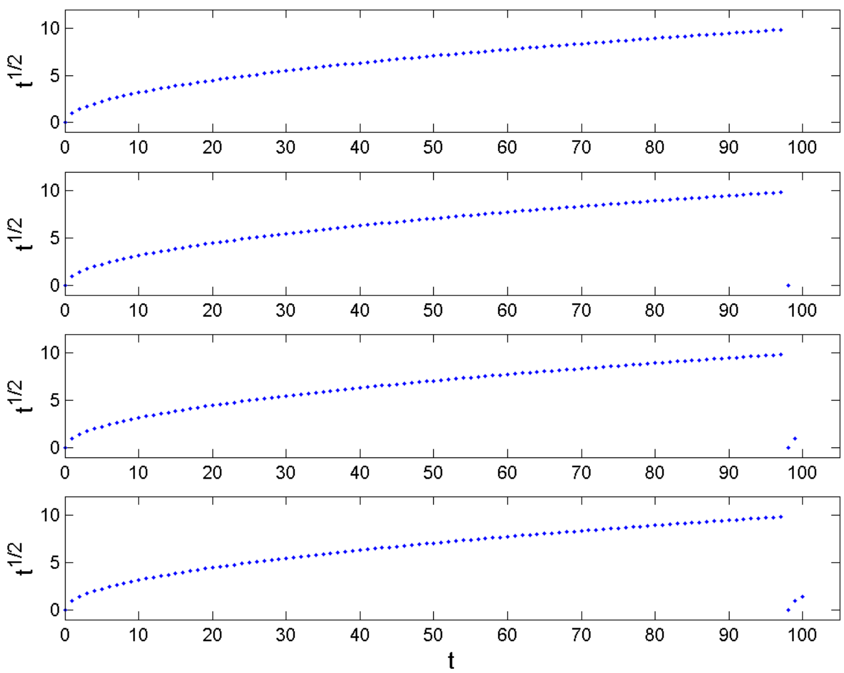
In other words, given a time series $T = \xi_0, \xi_1, \ldots, \xi_N$ that is strictly monotonic and strictly concave, for every distance biased operator $(0, z^+, \triangleright)$ we have that

$\lambda(E(T,1,d),(z^-,z^+,\triangleright)) < 0$, where $T = \xi_0, \xi_1, \ldots, \xi_N$
$\lambda(E(T_0,1,1),(z^-,z^+,\triangleright)) \geq 0$, where $T_0 = \xi_0, \xi_1, \ldots, \xi_N, \xi_0$
$\lambda(E(T_1,1,2),(z^-,z^+,\triangleright)) \geq 0$, where $T_1 = \xi_0, \xi_1, \ldots, \xi_N, \xi_0, \xi_1$
$\lambda(E(T_2,1,3),(z^-,z^+,\triangleright)) \geq 0$, where $T_2 = \xi_0, \xi_1, \ldots, \xi_N, \xi_0, \xi_1, \xi_2$

and so on.

This unusual result demonstrates that the Lyapunov exponent of a time series may be quite sensitive to the addition of a single point. In pictures, we have the following.

Figure 2.4: Time series $T = \xi_0, \xi_1, \ldots, \xi_{97}$, where $\xi_t = \sqrt{t}$. The first time series has a negative Lyapunov exponent in every embedding dimension. The second, third and fourth time series have nonnegative Lyapunov exponents in embedding dimensions 1, 2 and 3, respectively.



Before concluding, let us note that most of the results from this section actually hold for all operators, and not just distance biased operators. In particular, all of the results for strictly monotonic strictly concave time series will hold. The fact that consecutive distances shrink means that we will never be in a situation where we need to consider new candidates; the default is always used. The case of strictly monotonic strictly convex time series is almost identical, except that the default is not used for $s(1)$ (the default is 2 in this case). In order to keep the arguments for showing that $s(1) = 0$ simple, we restricted our attention to distance biased operators only.

Concluding Remarks and Further Work.

We have introduced the necessary background for understanding what Lyapunov exponents are with respect to pure mathematics. In doing so, we have given a treatment of many of the notions typically found in the theory of dynamical systems. In particular, those of fixed points, stability, and periodic trajectories. Once a foundation was in place, we described what the spectrum of Lyapunov exponents of a trajectory of a map are, and how they are calculated. We then gave a rigorous exposition of the algorithm given in [40] for estimating the maximal Lyapunov exponent of trajectory. Finally, we ended Chapter 1 with some definitions which will facilitate future arguments in the theory of Lyapunov exponents.

Chapter 2 explained how the maximal Lyapunov exponent of a time series (with respect to various parameters) can be calculated, via the algorithm found in [40]. Results on strictly monotonic and strictly concave/convex time series were then presented, and displayed how the previously laid foundation can be used to rigorously prove statements about particular time series, without the use of numerical estimation.

The application of the theory of Lyapunov exponents to time series is complicated and still seems to require some leaps of faith. We have developed what we feel to be a starting point for a theory that was in desperate need of clarification. With it in place, the door is open not only to give precise answers, but to ask precise questions. For example, is there a bound on the maximal Lyapunov exponent of all time series that take on values within a specified range?

Intimately connected to the theory we have just presented is that of Takens' Theorem. In a sense, it seems to be abused in the same way Lyapunov exponents are, and often by the same people. We hope to provide a similar treatment of this issue in the future, so that it is not blindly appealed to when it happens to sound convenient.

Bibliography

[1] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos*, Springer, 2000.

[2] V. I. Arnold and A. Avez, *Ergodic problems in classical mechanics*, New York: Benjamin, 1968.

[3] M. Barnsley, *Fractals everywhere*, Morgan Kaufmann, 2000.

[4] G. D. Birkhoff, *Proof of the ergodic theorem*, Proc. Natl. Acad. Sci. USA **17** (1931), 656660.

[5] L. S. Block and W. A. Coppel, *Dynamics in one dimension*, Lecture notes in mathematics, vol. 1513, Springer-Verlag, 1992.

[6] V. Chvátal, *Notes on the maximal Lyapunov exponent of a time series*, http://users.encs.concordia.ca/ chvatal/neuro/lyapunov.pdf, 2009.

[7] I.P. Cornfeld, S.V. Fomin, and Y. G. Sinai, *Ergodic theory*, Springer-Verlag, 1982.

[8] R. L. Devaney, *An introduction to chaotic dynamical systems*, Addison-Wesley Publishing Company, 1989.

[9] P. G. Drazin, *Nonlinear systems*, Cambridge University Press, 1992.

[10] S. N. Elaydi, *Discrete chaos*, Chapman & Hall/CRC, 2008.

[11] R. H. Enns and G. McGuire, *Nonlinear physics with mathematica for scientists and engineers*, Birkhäuser, 2001.

[12] J. Gleick, *Chaos: Making a new science*, Penguin, 1988.

[13] M. Hénon, *A two-dimensional mapping with a strange attractor*, Communications in Mathematical Physics **50** (1976), 69–77.

[14] R. C. Hilborn, *Sea gulls, butterflies, and grasshoppers: A brief history of the butterfly effect in nonlinear dynamics*, American Journal of Physics **72** (2004), 425–427.

[15] K. Hoffman and R. A. Kunze, *Linear algebra*, Prentice Hall, 1971.

[16] A. V. Holden, *Chaos*, John Wiley & Sons Australia, 1986.

[17] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

[18] L. D. Iasemidis, D.-S. Shiau, J. C. Sackellares, P. M. Pardalos, and A. Prasad, *Dynamical resetting of the human brain at epileptic seizures: Application of nonlinear dynamics and global optimization techniques*, IEEE Transactions On Biomedical Engineering **51** (2004), no. 3.

[19] Leonidas Iasemidis, *On the dynamics of the human brain in temporal lobe epilepsy*, Ph.D. thesis, University of Michigan, 1991.

[20] J. L. Kaplan and J. A. Yorke, *Functional differential equations and approximations of fixed points*, Lecture notes in Mathematics **730** (1979).

[21] E. N. Lorenz, *The predictability of hydrodynamic flows*, Trans. N. Y Acad. Sci. **25** (1963), 409–432.

[22] S. Lynch, *Dynamical systems with applications using matlab*, Birkhäuser, 2004.

[23] R. M. May, *Simple mathematical models with very complicated dynamics*, Nature **261** (1976), 459.

[24] S. P. Nair, D.-C. Shiau, J. Principe, C. Leonidas, D. Iasemidis, P. M. Pardalos, W. M. Norman, P. R. Carney, K. M. Kelly, and J. C. Sackellares, *An investigation of EEG dynamics in an animal model of temporal lobe epilepsy using the maximum Lyapunov exponent*, Experimental Neurology **216** (2009).

[25] V. Oseledets, *A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems*, Trudy Moskov. Mat. Obšč **19** (1968), 179–210 (Russian).

[26] V. Oseledets, *Oseledets' theorem*, http://www.scholarpedia.org/article/Oseledets_theorem, 2008.

[27] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Geometry from a time series*, Physical Review Letters **45** (1980), 712–716.

[28] Y. Pesin and V. Climenhagam, *Lectures on fractal geometry and dynamical systems*, American Mathematical Society, 2009.

[29] C. Robinson, *Dynamical systems: stability, symbolic dynamics, and chaos*, CRC, 1998.

[30] D. Ruelle, *Chaotic evolution and strange attractors*, Lincei Lectures, Cambridge University Press, 1989.

[31] S. Sastry, *Nonlinear systems: analysis, stability, and control*, Springer, 1999.

[32] T. Sauer, J. A. Yorke, and M. Casdagli, *Embedology*, Journal of Statistical Physics **65** (1991).

[33] E. R. Scheinerman, *Invitation to dynamical systems*, Prentice-Hall, 1996.

[34] H. G. Schuster, *Deterministic chaos: an introduction*, Wiley-VCH, 1988.

[35] L. Smith, *Linear algebra*, Springer, 1998.

[36] J. C. Sprott, *Chaos and time-series analysis*, Oxford University Press, 2003.

[37] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry and engineering*, Westview Press, 1994.

[38] F. Takens, *Detecting strange attractors in turbulence*, Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898, 1981, pp. 366–381.

[39] P. F. Verhulst, *Recherches mathématiques sur la loi d'accroissement de la population*, Nouv. mém. de l'Academie Royale des Sci. et Belles-Lettres de Bruxelles **18** (1845), 1–41.

[40] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, *Determining Lyapunov exponents from a time series*, Physica D: Nonlinear Phenomena **16** (1985), 285–317.

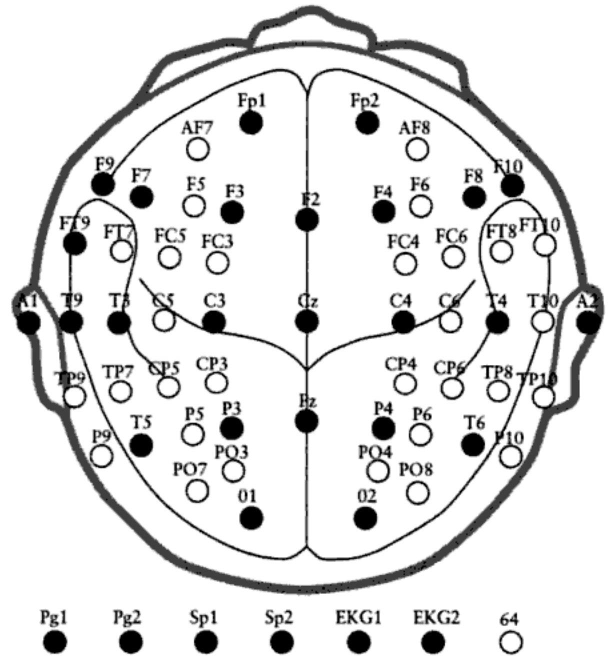
Lyapunov Exponents and Epilepsy.

In this section we will briefly describe some of the claims in the literature ([10], [2]) about Lyapunov exponents and their relation to the prediction of epileptic seizures.

Introduction.

Sufferers of epilepsy must currently live with the fact that seizures tend to occur spontaneously, often with little or no warning signs. In attempts to study the disease, patients with severe cases of epilepsy are continuously watched by medical staff for weeks at a time. During these periods EEG (electroencephalogram) or ECoG (electrocorticography) recordings are made. The former of these recordings results from the placement of multiple electrodes placed on the scalp of the patient. The latter is the result of intracranial measurements, which involves invasive surgery so that a grid of electrodes may be placed directly on the cortex of the patient. Thus, ECoG recordings are far more robust in terms of analyzable data. This process is reserved for patients with the most severe forms of the disease.

Figure 5: Montage of electrodes for EEG. (Image borrowed from reference [11]).



To the untrained eye, an EEG/ECoG recording looks like nothing more than an indecipherable random series of data. However, specialized neurologists working in the field of epilepsy are trained to read these recordings, and seem to develop some intuition as to what seizure activity looks like. Mere visual inspection allows them to distinguish between preictal, ictal and postictal stages of a seizure (corresponding to before the seizure, during the seizure, and after the seizure, respectively).

Figure 6: Temporal seizure. (Image borrowed from reference [11]).

Until recently, visual inspection of EEG data was a time consuming but unfortunately necessary task for neurologists; reading through days of EEG data had no shortcuts. This was the problem of seizure *detection*: given an EEG recording, can we tell when a seizure might have occured to a reliable degree of accuracy? Although the solution to this problem may only offer a small improvement, if any, to the quality of life for the patient, it is a great time saver for the medical staff. This problem was essentially solved by Jean Gotman in the 80's ([1]). Neurologists now have software at their disposal that reads through EEG data and indicates the points in time that might be of interest, allowing the neurologist to ignore much of the reading.

In contrast, the problem of seizure *prediction* asks: given some continuous EEG data, is a seizure about to occur? If yes, then when? In 1991, the Ph.D. thesis written by Leonidas Iasemidis [6] was one of the first of many publications written in an attempt to use techniques from dynamical systems to solve this problem. In particular, emphasis was placed on the calculation of Lyapunov exponents from a time series via the algorithm developed by Wolf et al. in 1985 [12]. We will now say a few words on how this is done.

How they do it.

A patient is hooked up to the necessary equipment while EEG and/or ECoG recordings are made for hours at a time. These measurements constitute of a sequence of real numbers and thus is the time series under our consideration. In fact, each electrode gives rise to its own time series. We mainly focus on just one particular electrode, and hence one particular time series. During the monitoring period, an epileptic seizure occurs, typically lasting a few minutes in length. Given an EEG/ECoG time series, a time delay τ and an embedding dimension d is chosen. We will not defend or justify the choices here, but for the sake of having some numbers to look at, [2] and [10] each use $\tau \approx 14$msec and an embedding dimension of $d = 7$. With this is place, the time series is embedded using these values in the same way we described in the beginning of Section 2.1. The time series is now considered as the trajectory of some d-dimensional point. This trajectory is divided into small windows or epochs, of around 10 to 12 seconds (again, see [10]). Next, a modified[3] version of the algorithm found in [12] is ran on each of these epochs, resulting in a sequence of values, each of which is referred to as STL_{max} or sometimes L_{max}, depending on the paper ([2] and [10], respectively), and is claimed to be an estimate of the maximal Lyapunov exponent of the epoch, or as "a reflexion of the chaoticity of the signal" ([2]).

What they find.

The general findings in [10] are that L_{max} takes on a mean value[4] of approximately 6bits/sec during the preictal stage, with occasional drops to 5bits/sec. Once a clinical seizure appears to be taking place, a drastic drop of L_{max} occurs, reaching as low as 2bits/sec. Finally, the seizure enters the postictal stage and L_{max} jumps back up to approximately 8bits/sec. The authors claim that not

[3]The algorithm is essentially the same, but a slightly different operator is used. See [2] for the details.
[4]The maximal Lyapunov exponent is calculated in base 2 instead of base e as we have presented it.

only do their algorithms detect these seizures (by drops in L_{max} during the ictal stage), but that the drops in the preictal stage, which can occur more than 20 minutes before the onset of the seizure, can be used to predict that a seizure is on the way. The authors refer to these drops as *entrainment*.

Whether or not these methods have validity or meaning in terms of pure mathematics is difficult to argue. However, the fact that interesting and useful results seem to be produced beg for as much justification as possible, and we hope to continue developing and contributing to this process.

Bibliography for predicting epileptic seizures

[1] J. Gotman, *Automatic recognition of interictal spikes*, Electroencephalography Clinical Neurophysiology Supplement (1985), 93–114.

[2] L. D. Iasemidis, J. C. Principe, and J. C. Sackellares, *Measurement and quantification of spatio-temporal dynamics of human epileptic seizures*, Nonlinear Biomedical Signal Processing (2000), 294–318.

[3] L. D. Iasemidis and J. C. Sackellares, *The evolution with time of the spatial distribution of the largest Lyapunov exponent on the human epileptic cortex*, Measuring Chaos in the human brain, World Scientific, 1991.

[4] L. D. Iasemidis, D.-S. Shiau, J. C. Sackellares, P. M. Pardalos, and A. Prasad, *Dynamical resetting of the human brain at epileptic seizures: Application of nonlinear dynamics and global optimization techniques*, IEEE Transactions On Biomedical Engineering **51** (2004), no. 3.

[5] L. D. Iasemidis, K. Tsakalis, J. C. Sackellares, and P. M. Pardalos, *Comment on "Inability of Lyapunov exponents to predict epileptic seizures"*, Physical Review Letters **94** (2005).

[6] Leonidas Iasemidis, *On the dynamics of the human brain in temporal lobe epilepsy*, Ph.D. thesis, University of Michigan, 1991.

[7] A. D. Krystal, C. Zaidmana, H. S. Greenside, R. D. Weinera, and C. E. Coffey, *The largest Lyapunov exponent of the EEG during ECT seizures as a measure of ECT seizure adequacy*, Electroencephalography and Clinical Neurophysiology **103** (1997).

[8] Y.-C. Lai, Harrison M. A. F., M. G. Frei, and I. Osorio, *Inability of Lyapunov exponents to predict epileptic seizures*, Physical Review Letters **91** (2003).

[9] S. P. Nair, D.-C. Shiau, J. Principe, C. Leonidas, D. Iasemidis, P. M. Pardalos, W. M. Norman, P. R. Carney, K. M. Kelly, and J. C. Sackellares, *An investigation of EEG dynamics in an animal model of temporal lobe epilepsy using the maximum Lyapunov exponent*, Experimental Neurology **216** (2009).

[10] J. C. Sackerelles, L. D. Iasemidis, R. L. Golmore, and S. N. Rober, *Epilepsy - when chaos fails*, Chaos in the Brain?, pp. 112–133, World Scientific, 2000.

[11] W. O. Tatum, *Handbook of EEG interpretation*, Demos Medical Publishing, 2007.

[12] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, *Determining Lyapunov exponents from a time series*, Physica D: Nonlinear Phenomena **16** (1985), 285–317.

Linear Algebra

Norms

Definition .0.4:
Let V be a vector space over $\mathbf{R}$. A function $||\cdot|| : V \to \mathbf{R}$ is a *norm* if for all x and y such that $x, y \in V$,

1. $||x|| \geq 0$
2. $||x|| = 0 \quad \Leftrightarrow \quad x = 0$
3. $||cx|| = |c|\,||x||$ for all c such that $c \in \mathbf{R}$
4. $||x + y|| \leq ||x|| + ||y||$

In particular, the only norm we use in this thesis is the Euclidean norm.

Definition .0.5:
When $v = (v_1, v_2, \ldots, v_n)$ is a vector in $\mathbf{R}^n$, we use $||v||$ to denote the usual Euclidean norm,

$$||v|| = \sqrt{\sum_{i=1}^{n} |v_i|^2}.$$

Definition .0.6:
When M is an $m \times n$ matrix, the Euclidean norm of M is

$$||M|| = \max\{||Mv|| : v \in \mathbf{R}^n \text{ and } ||v|| \leq 1\}.$$

Any standard text on linear algebra should contain proofs that these are indeed norms. For instance see [2] or [1].

We also make use of the following fact.

Lemma .0.5:
If M is an $m \times n$ matrix and $v \in \mathbf{R}^n$, then $||Mv|| \leq ||M||\,||v||$.

Proof:
Let v be any vector such that $v \in \mathbf{R}^n$ and let $w = v/\,||v||$. Note that $||w|| = 1$, so that $||Mw|| \leq ||M||$ by definition of the matrix norm. Now

$$||Mv|| = ||v||\,||Mw|| \leq ||M||\,||v||$$

. □

Jordan Canonical Form.

Here we will introduce some of the definitions and claims about Jordan normal form. Material from this section can be found in Section 3.1 of [2], with the exception of Lemmas .0.6 and .0.7.

Definition .0.7:
A *Jordan block* $J_k(\lambda)$ is a $k \times k$ upper triangular matrix of the form

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

The k diagonal entries are all λ, the $k-1$ upper diagonal entries all 1, and all other entries are 0.

Definition .0.8:
A *Jordan matrix* J is an $n \times n$ matrix that is the direct sum of Jordan blocks:

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & & & \\ & J_{n_2}(\lambda_2) & & & \\ & & J_{n_3}(\lambda_3) & & \\ & & & \ddots & \\ & & & & J_{n_k}(\lambda_k) \end{bmatrix},$$

where $n_1 + n_2 + \cdots + n_k = n$. The diagonal entries take the form of Jordan blocks, all other entries are 0.

Theorem .0.2:
Let A be an $n \times n$ matrix with entries in $\mathbf{C}$. There is an invertible $n \times n$ matrix S such that

$$A = SJS^{-1},$$

where

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & & & \\ & J_{n_2}(\lambda_2) & & & \\ & & J_{n_3}(\lambda_3) & & \\ & & & \ddots & \\ & & & & J_{n_k}(\lambda_k) \end{bmatrix}$$

and $n_1 + n_2 + \cdots + n_k = n$. Furthermore, $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the eigenvalues of A, which may not be distinct.

Proof:
See [2] □

Lemma .0.6:
Let $J_k(\lambda)$ be a Jordan block, let t be a positive integer, and let $a_{ij}(t)$ denote the the entry in the ith row and jth column of $J_k(\lambda)^t$. Then

$$a_{i,j}(t) = \binom{t}{j-i} \lambda^{t-(j-i)},$$

with the convention that $\binom{i}{t} = 0$ if $t > i$.

Proof:
We will prove the statement by induction on t. When $t = 1$ the statement is trivial. Now suppose the statement holds for $t - 1$, so that

$$a_{i,j}(t-1) = \binom{t-1}{j-i} \lambda^{t-1-(j-i)}$$

for all integers i and j such that $1 \le i \le k$ and $1 \le j \le k$. Note that the entry $a_{i,j}(t)$ of $J_k(\lambda)^t$ is calculated by multiplying $J_k(\lambda)^{t-1}$ on the right by the relatively simple matrix $J_k(\lambda)$, so that

$$\begin{aligned} a_{i,j}(t) &= a_{i,j-1} + a_{i,j}\lambda \\ &= \binom{t-1}{j-1-i}\lambda^{t-1-(j-1-i)} + \binom{t-1}{j-i}\lambda^{t-1-(j-i)} \\ &= \binom{t}{j-i}\lambda^{t-(j-i)}. \end{aligned}$$

□

Lemma .0.7:
If J is an $n \times n$ Jordan matrix, then

$$J^t = \begin{bmatrix} J_{n_1}(\lambda_1)^t & & & & \\ & J_{n_2}(\lambda_2)^t & & & \\ & & J_{n_3}(\lambda_3)^t & & \\ & & & \ddots & \\ & & & & J_{n_k}(\lambda_k)^t \end{bmatrix},$$

for every positive integer t. (All non-diagonal entries are 0).

Proof:
By induction on t. When $t = 1$ the statement is trivial. Suppose it holds for $t - 1$, then $J^t = J^{t-1}J$ clearly has the correct form by inspecting

$$\begin{bmatrix} J_{n_1}(\lambda_1)^{t-1} & & & \\ & J_{n_2}(\lambda_2)^{t-1} & & \\ & & \ddots & \\ & & & J_{n_k}(\lambda_k)^{t-1} \end{bmatrix} \times \begin{bmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_k}(\lambda_k) \end{bmatrix}.$$

□

The Spectral Theorem.

Here we state the theorem as it is found on p. 316 of [3]. Its proof can be found there as well.

Theorem .0.3 (The Spectral Theorem):
Let $T : V \to V$ be a self-adjoint[5] linear transformation in the finite-dimensional inner product space V. Then there exists an orthonormal basis $\{v_1, \ldots, v_n\}$ and number $\lambda_1, \ldots, \lambda_n$ such that

$$T(v_i) = \lambda_i v_i$$

for all integers i such that $1 \le i \le n$.

[5] $T^* : V \to V$ is the adjoint of T if $< Tx, y >=< x, T^*y >$ for all x, y such that $x, y \in V$. Furthermore, T is self-adjoint if $T = T^*$.

In particular, we note that in the case of a matrix with entries in $\mathbf{R}$, all matrices M for which $M = M^T$ are self-adjoint. Thus, if M is a real symmetric matrix then it is self-adjoint.

Other Lemmas.

The following Lemmas are used in Section 1.3.7.

Lemma .0.8:
Let M be an $n \times n$ matrix having n distinct, non-zero eigenvalues. Then M and M^T have the same eigenvalues.

Proof:
Suppose λ is an eigenvalue of M, then $Mv = \lambda v$ for some non-zero v. Since $(M - \lambda I)v = 0$ we know that $(M - \lambda I)$ is not invertible. To complete the proof we will show that $(M - \lambda I)^T = (M^T - \lambda I)$ is not invertible either, which allows us to conclude that λ is an eigenvalue of M^T. To this end, let $N = M - \lambda I$ and suppose N^T is invertible. Then $I = N^T(N^T)^{-1} = ((N^T)^{-1})^T N = N^{-1}N$ so that N is invertible, which is contrary to our assumption. $\square$

Lemma .0.9:
Let M be an $n \times n$ matrix with entries in $\mathbf{R}$ having n distinct, non-zero eigenvalues $\lambda_1, \ldots, \lambda_n$. Let v_i and w_i $(i = 1, \ldots, n)$ be the corresponding eigenvectors of M and M^T, respectively. Then v_i and w_j are orthogonal if $i \neq j$.

Proof:
For every i, $Mv_i = \lambda_i v_i$, and so $v_i^T M^T = \lambda_i v_i^T$. Now for every $j \neq i$, $v_i^T M^T w_j = \lambda_i v_i^T w_j$. Since w_j is an eigenvector of M^T, it follows that $v_i^T \lambda_j w_j = \lambda_i v_i^T w_j$. Since $\lambda_i \neq \lambda_j$, we must have $v_i^T w_j = 0$. $\square$

Lemma .0.10:
If $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ are eigenvectors of M and M^T, respectively, and $v_j = \sum_{i=1}^n a_i w_i$, then $a_j \neq 0$ and $v_j^T w_j \neq 0$.

Proof:
We have $0 \neq v_i^T v_i = v_i^T \sum_{i=1}^n a_i w_i = a_i v_i^T w_i$ by Lemma .0.9. $\square$

Lemma .0.11:
Let M be an $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the corresponding eigenvectors $v_1, \ldots, v_n$ are linearly independent.

Proof:
Suppose the eigenvectors are linearly dependent. Then there exists a smallest positive integer k such that $k < n$ and constants $c_1, \ldots, c_k$ such that

$$v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1}.$$

Multiplying by M, we get

$$M v_k = c_1 M v_1 + c_2 M v_2 + \cdots + c_{k-1} M v_{k-1},$$

which yields

$$\lambda_k v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_{k-1} \lambda_{k-1} v_{k-1}.$$

However, we also have

$$\lambda_k v_k = c_1 \lambda_k v_1 + c_2 \lambda_k v_2 + \cdots + c_{k-1} \lambda_k v_{k-1},$$

so that

$$c_1(\lambda_k - \lambda_1) v_1 + c_2(\lambda_k - \lambda_2) v_2 + \cdots + c_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1} = 0.$$

This contradicts the minimality of k, since we now have a smaller linear dependent set of vectors.□

References

[1] K. Hoffman and R. A. Kunze, *Linear algebra*, Prentice Hall, 1971.

[2] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

[3] L. Smith, *Linear algebra*, Springer, 1998.