Lecture 2-3: Linear Maps and Contractions in Euclidean Space

September 19, 2023

- The primitive discrete-time population model $x_{i+1} = f(x_i) = kx_i$ (with $k > 0$) introduced in lecture 1 has simple dynamics.
- Starting with any $x_0 \neq 0$, the sequence $(x_i)_{i \in \mathbb{N}}$ diverges if $k > 1$ and goes to 0 if $k < 1$.
- Part of the simplicity is that the asymptotic behavior is independent of the initial condition.
- Scaling x_0 by a factor a scales all x_i by the same factor.
- Furthermore, the allowed asymptotic behaviors are quite simple.

- So long as $k \neq 1$, this changes little if we replace $f(x) = kx$ by $g(x) := kx + b$.
- Indeed, changing variables to $y = x - \frac{b}{1-k}$ leads to the recursion $y_{i+1} = ky_i$.
- Therefore we have by now fully described the dynamical possibilities for linear maps.

- The central feature of differentiability is that it guarantees a good linear approximation of a map near any given point.
- A simple example is the approximation of $f(x) := \sqrt{x}$ near 16 by $L(x) = f(16) + f'(16)(x - 16) = 4 + \frac{1}{8}(x - 16)$,
- Such linear approximation can sometimes be useful for dynamics when the orbits of a nonlinear map stay near enough to the reference point for the linear approximation to be relevant.

Proposition 1.1

Suppose F is a differentiable map of the line and $F(b) = b$. If the orbits of the linearization of F at b are asymptotic to b, then all orbits of F that start near enough to b are asymptotic to b as well.

Now we define contracting maps with respect to the Euclidean distance $d(x, y) := \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$.

Definition 2.1

A map f of a subset X of Euclidean space is said to be *Lipschitz-continuous* with Lipschitz constant λ, or λ-Lipschitz if

$$d(f(x), f(y)) \leq \lambda d(x, y)$$

for any $x, y \in X$. If $\lambda < 1$ then f is said to be a *contraction* or a λ-contraction. If a map f is Lipschitz-continuous, then we define $\text{Lip}(f) := \sup_{x \neq y} d(f(x), f(y))/d(x, y)$.

Example 2.2

The function $f(x) = \sqrt{x}$ defines a contraction on $[1, \infty)$. To prove this, we show that for $x \geq 1$ and $t \geq 0$ we hjave $\sqrt{x+t} \leq \sqrt{x} + (1/2)t$. This is most easily seen by squaring:

$$(\sqrt{x} + \frac{t}{2})^2 = x + xt + \frac{t^2}{4} \geq x + xt \geq x + t.$$

Proposition 2.3

Let I be an interval and $f : I \to \mathbb{R}$ a differentiable function with $|f'(x)| \leq \lambda$ for all $x \in I$. Then f is λ-Lipschitz.

Proof.

By the mean Value Theorem, for any two points $x, y \in I$ there exists a point c between x and y such that

$$d(f(x), f(y)) = |f(x) - f(y)| = |f'(c)(x - y)| = |f'(c)|d(x, y) \leq \lambda d(x, y).$$

So f is λ-Lipschitz. □

Example 2.4

This criterion makes it easier to check that $f(x) = \sqrt{x}$ defines a contradiction on $I = [1, \infty)$ because $f'(x) = 1/2\sqrt{x} \leq 1/2$ for $x \geq 1$.

The weaker condition $|f'(x)| < 1$ does not suffice to make f a contraction. However, sometimes it does.

Proposition 2.5

Let I be a closed bounded interval and $f : I \to I$ a continuously differentiable function with $|f'(x)| < 1$ for all $x \in I$. Then f is a contraction.

Proof.

The maximum λ of $|f'(x)|$ is attained at some point x_0 because f' is continuous. It is less than 1 because $|f'(x_0)| < 1$. □

The difference is that the real line is not closed and bounded.

- In calculus, a favorite example of a recursively defined sequence is of the form $a_{n+1} = f(a_n)$, with a_0 given and f a function with $|f'| \leq \lambda < 1$.
- This is a simple dynamical system given by the map f. For each initial value a_0 a sequence is uniquely defined by $a_{n+1} = f(a_n)$.
- If f is invertible, then this sequence is defined for all $n \in \mathbb{Z}$.

Definition 2.6

For a map f and a point x, the sequence $x, f(x), f(f(x)), \cdots, f^n(x), \cdots$ (if f is not invertible) or the sequence $\cdots, f^{-1}(x), x, f(x), \cdots$ is called the *orbit* of x under f. A *fixed point* is a point such that $f(x) = x$. The set of fixed points is denoted by Fix(f). A *periodic point* is a point x such that $f^n(x) = x$ for some $n \in \mathbb{N}$, that is, a point in Fix(f^n). Such an n is said to be a period of x. The smallest such n is called the *prime period* of x.

Example 2.7

If $f(x) = -x^3$ on $\mathbb{R}$, then 0 is the only fixed point and ± 1 is a periodic orbit, that is, 1 and -1 are periodic points with prime period 2.

Proposition 2.8: Contraction Principle

Let $I \subset \mathbb{R}$ be a closed interval, possibly infinite on one or both sides, and $f : I \to I$ a λ-contradiction. Then f has a unique fixed point x_0 and $|f^n(x) - x_0| \leq \lambda^n |x - x_0|$ for every $x \in \mathbb{R}$, that is, every orbit of f converges to x_0 exponentially.

Proof.

By iterating $|f(x) - f(y)| \leq \lambda|x - y|$, one sees that

$$|f^n(x) - f^n(y)| \leq \lambda^n|x - y|$$

for $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$; so for $x \in I$ and $m \geq n$ we can use the triangle inequality to show

$$\begin{aligned} |f^m(x) - f^n(x)| &\leq \sum_{k=0}^{m-n-1} |f^{n+k+1}(x) - f^{n+k}(x)| \\ &\leq \sum_{k=0}^{m-n-1} \lambda^{n+k}|f(x) - x| \leq \frac{\lambda^n}{1-\lambda}|f(x) - x|. \end{aligned}$$

Since the right-hand side of the above inequality becomes arbitrarily small as n get large, we have that $(f^n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence.

Proof.

Thus for any $x \in I$ the limit of $f^n(x)$ as $n \to \infty$ exists because Cauchy sequences converge. The limit is in I because I is closed. Since $|f^n(x) - f^n(y)| \leq \lambda^n |x - y|$, this limit is the same for all x. We denote this limit by x_0 and we shall show that x_0 is a fixed point for f. If $x \in I$ and $n \in \mathbb{N}$, then

$$\begin{aligned} |x_0 - f(x_0)| &\leq |x_0 - f^n(x)| + |f^n(x) - f^{n+1}(x)| + |f^{n+1}(x) - f(x_0)| \\ &\leq (1+\lambda)|x_0 - f^n(x)| + \lambda^n |x - f(x)|. \end{aligned}$$

Since $|x_0 - f^n(x)| \to 0$ and $\lambda^n \to 0$ as $n \to \infty$, we have $f(x_0) = x_0$. □

Example 2.9

In contemplating his rabbits, Leonardo of Pisa, also known as Fibonacci, came up with a model according to which the number of rabbit pairs in the nth month is given by the number b_n, defined by recursive relation $b_0 = 1$, $b_1 = 2$, $b_n = b_{n-1} + b_{n-2}$ for $n \geq 2$. Expecting that the growth of these numbers should be exponential, we would like to see how fast these numbers grow by finding the limit of $a_n := b_{n+1}/b_n$ as $n \to \infty$. To that end we use Contraction Principle. Since

$$a_{n+1} := \frac{b_{n+2}}{b_{n+1}} = \frac{b_{n+1} + b_n}{b_{n+1}} = \frac{1}{b_{n+1}/b_n} + 1 = \frac{1}{a_n} + 1,$$

$(a_n)_{n=1}^{\infty}$ is the orbit of 1 under iteration of the map $g(x) := (1/x) + 1$. Since $g(1) = 2$, we are in fact considering the orbit of 2 under g. Now $g'(x) = -x^{-2}$. This tells us that g is not a contraction on $[0, \infty)$. Therefore we need to find a suitable (closed) interval where this is the case and that is mapped inside itself.

Solution.

Since $g' < 0$, g is decreasing on $(0, \infty)$. This implies that $g([3/2, 2]) \subset [3/2, 2]$ because $3/2 < g(3/2) = 5/3 < 2$ and $g(2) = 3/2$. Furthermore, $|g'(x)| = 1/x^2 \leq 4/9 < 1$ on $[3/2, 2]$, so g is a contraction on $[3/2, 2]$, so g is a contraction on $[3/2, 2]$. By the Contraction Principle, the orbit of 2 and hence that of 1 is asymptotic to the unique fixed point x of g in $[3/2, 2]$. Thus $\lim_{n\to\infty} b_{n+1}/b_n = \lim_{n\to\infty} a_n$ exists. To find the limit we solve the equation $x = g(x) = 1 + 1/x = (x+1)/x$, which is equivalent to $x^2 - x - 1 = 0$. There is only one positive solution: $x = (1 + \sqrt{5})/2$. □

Proposition 2.10

Let $X \subset \mathbb{R}^n$ be closed and $f : X \to X$ a λ-contraction. Then f has a unique fixed point x_0 and $d(f^n(x), x_0) \leq \lambda^n d(x, x_0)$ for every $x \in X$.

Proof.

Iterating $d(f(x), f(y)) \leq \lambda d(x, y)$ shows

$$d(f^n(x), f^n(y)) \leq \lambda^n d(x, y) \tag{2.1}$$

for $x, y \in X$ and $n \in \mathbb{N}$. Thus $(f^n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence because

$$\begin{aligned} d(f^m(x), f^n(x)) &\leq \sum_{k=0}^{m-n-1} d(f^{n+k+1}(x), f^{n+k}(x)) \\ &\leq \sum_{k=0}^{m-n-1} \lambda^{m+n} d(f(x), x) \leq \frac{\lambda^n}{1-\lambda} d(f(x), x) \end{aligned}$$

for $m \geq n$, and $\lambda^n \to 0$ as $n \to \infty$. Thus $\lim_{n\to\infty} f^n(x)$ exists and is in X since X is closed. By (2.1) this limit is the same for all x. Denote the limit by x_0. Then

$$\begin{aligned} d(x_0, f(x_0)) &\leq d(x_0, f^n(x)) + d(f^n(x), f^{n+1}(x)) + d(f^{n+1}(x), f(x_0)) \\ &\leq (1+\lambda) d(x_0, f^n(x)) + \lambda^n d(x, f(x)) \end{aligned}$$

for $x \in X$ and $n \in \mathbb{N}$. Now $f(x_0) = x_0$ because $d(x_0, f^n(x)) \to 0$ as $n \to \infty$. □

Taking the limit in

$$d(f^m(x), f^n(x)) \leq \frac{\lambda^n}{1-\lambda} d(f(x), x)$$

as $m \to \infty$ we obtain $d(f^n(x), x_0) \leq (\lambda^n/(1-\lambda))d(f(x), x)$.

This means that, after n iterations, we can say with certainty that the fixed point is in the $(\lambda^n/(1-\lambda))d(f(x), x)$-ball around $f^n(x)$.

Definition 2.11

We say that two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ of points in $\mathbb{R}^n$ *converge exponentially* (or *with exponential speed*) to each other if $d(x_n, y_n) < cd^n$ for some $c > 0$, $d < 1$. In particular, if one of the sequences is constant, that is, $y_n = y$, we say that x_n *converge exponetially* to y.

The differential

- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a map with continuous partial derivatives.
- Then at each point one can define the derivative or differential of $f = (f_1, \cdots, f_m)$ as the linear map defined by the matrix of partial derivatives

$$Df := \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

- We say that the map is regular at x_0 if this map is invertible.

- We define the norm of the differential by the norm of the matrix Df. In linear algebra the norm of a matrix A is defined by looking at its action as a linear map:

$$\|A\| := \max_{v \neq 0} \frac{\|A(v)\|}{\|v\|} = \max_{\|v\|=1} \|A(v)\|.$$

- Geometrically, this is easy to visualize by considering the second of these expressions: Consider the unit sphere $\{v \in \mathbb{R}^n | \|v\| = 1\}$ and notice that the second maximum is just the size of the largest vectors in the image of this unit sphere.

The mean value theorem

Theorem 2.12: Mean Value Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous and f is differentiable on (a, b), then there is a point $x \in (a, b)$ such that $f(b) - f(a) = (b - a)f'(x)$.

Proof.

Note that $g(t) := (t-a)(f(b)-f(a)) - (f(t)-f(a))(b-a)$ is continuous on $[a,b]$ and differentiable on (a,b) and $g(a) = 0 = g(b)$. If g is constant, then we are done. Otherwise, g has an extremum $g(x)$ at some $x \in (a,b)$ by continuity. g is differentiable at x, hence $0 = g'(x) = f(b) - f(a) - f'(x)(b-a)$. □

Lemma 2.13

If $g : [a, b] \to \mathbb{R}^m$ is continuous and differentiable on (a, b), then there exists $t \in [a, b]$ such that

$$\|g(b) - g(a)\| \leq \|\frac{d}{dt}g(t)\|(b - a).$$

Proof.

Let $v = g(b) - g(a)$, $\varphi(t) = \langle v, g(t)\rangle$. By the Mean Value Theorem for one variable there exists a $t \in (a, b)$ such that $\varphi(b) - \varphi(a) = \varphi'(t)(b - a)$, and so

$$(b-a)\|v\|\|\frac{d}{dt}g(t)\| \geq (b-a)\langle v, \frac{d}{dt}g(t)\rangle \frac{d}{dt}\varphi(t)(b-a) = \varphi(b) - \varphi(a)$$
$$= \langle v, g(b)\rangle - \langle v, g(a)\rangle = \langle v, v\rangle = \|v\|^2.$$

Divide by $\|v\|$ to finish the proof. □

Convexity

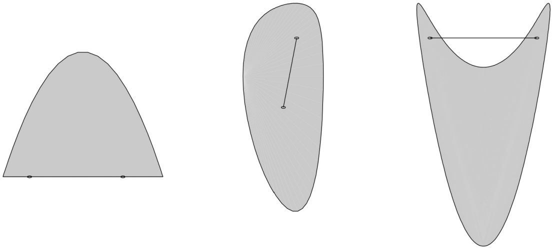
Definition 2.14

A *convex set* in $\mathbb{R}^n$ is set C such that for all $a, b \in C$ the line segment with endpoints a, b is entirely contained in C. It is said to be *strictly convex* if for any points a, b in the closure of C the segment from a to b is contained in C, expect possibly for one or both endpoints.

Example 2.15

The disk $\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 1\}$ is strictly convex. The open upper half-plane $\{(x, y) \in \mathbb{R}^2 | y > 0\}$ is convex. A kidney shape $\{(r, \theta) | 0 \leq r \leq 1 + (1/2) \sin \theta\}$ (in polar coordinates) is not convex. Neither is the annulus $\{(x, y) \in \mathbb{R}^2 | 1 < x^2 + y^2 < 2\}$.

The follow figure gives examples of a convex, a strictly convex and a nonconvex set.



The derivative test

Theorem 2.16

If $C \subset \mathbb{R}^n$ is convex and open and $f : C \to \mathbb{R}^m$ is differentiable with $\|Df(x)\| \leq M$ for all $x \in C$, then $\|f(x) - f(y)\| \leq M\|x - y\|$ for $x, y \in C$.

Proof.

The line segment connecting x and y is given by $c(t) = x + t(y - x)$ for $t \in [0, 1]$, and it is contained in C by convexity. Let $g(t) := f(c(t))$. Then by the chain rule

$$\|\frac{d}{dt} g(t)\| = \|Df(c(t))\frac{d}{dt}c(t)\| = \|Df(c(t))(y - x)\| \leq M\|y - x\|.$$

This implies that $\|f(y) - f(x)\| = \|g(1) - g(0)\| \leq M\|y - x\|$. □

Corollary 2.17

If $C \subset \mathbb{R}^n$ is a convex open set, $f : C \to C$ a map with continuous partial derivatives, and $\|Df\| \leq \lambda < 1$ at every point $x \in \mathbb{R}^n$, then f is a λ-contraction.

Theorem 2.18

If $C \subset \mathbb{R}^n$ is open strictly convex set, $\overline{C}$ its closure, $f : \overline{C} \to \overline{C}$ differentiable on C and continuous on $\overline{C}$ with $\|Df\| \leq \lambda < 1$ on C, then f has a unique fixed point $x_0 \in \overline{C}$ and

$$d(f^n(x), x_0) \leq \lambda^n d(x, x_0)$$

for every $x \in \overline{C}$.

Proof.

For $x, y \in \overline{C}$ we parameterize the line segment connecting x and y by $c(t) = x + t(y - x)$ for $t \in [0,1]$ and let $g(t) := f(c(t))$. Then $c((0,1))$ is contained in C by strict convexity and

$$\|\frac{d}{dt}g(t)\| = \|Df(c(t))\frac{d}{dt}c(t)\| = \|Df(c(t))(y-x)\| \leq \lambda\|y-x\|.$$

This implies $\|f(y) - f(x)\| \leq \lambda|y - x\|$. Thus f is a λ-contraction and has a unique fixed point x_0. Furthermore $d(f^n(x), x_0) \leq \lambda^n d(x, x_0)$ for every $x \in \overline{C}$. □

Now we discuss maps that are not contracting on their entire domain but on a part of it.

Definition 2.19

By a closed neighborhood of X we mean the closure of an open set containing x.

Proposition 2.20

Let f be a continuously differentiable map with a fixed point x_0 where $\|Df_{x_0}\| \leq 1$. Then there is a closed neighborhood U of x_0 such that $f(U) \subset U$ and f is a contraction on U.

Proof.

Since Df is continuous, there is a small closed ball $U = \overline{B(x_0, \eta)}$ around x_0 on which $\|Df_x\| \leq \lambda < 1$. If $x, y \in U$, then $d(f(x), f(y)) \leq \lambda d(x, y)$; so f is a contraction on U. Furthermore, taking $y = x_0$ shows that if $x \in U$, then $d(f(x), x_0) = d(f(x), f(x_0)) \leq \lambda d(x, x_0) \leq \lambda\eta < \eta$ and hence $f(x) \in U$. □

Proposition 2.21

Let f be ba continuously differentiable map with a fixed point x_0 such that all eigenvalues of Df_{x_0} have absolute value less than 1. Then there is a closed neighborhood U of x_0 such that $f(U) \subset U$ and f is a contraction on U with respect to an adapted norm.

Proof.

Later we will show that the assumption on the eigenvalues implies that one can choose a norm that we denote by $\|\cdot\|'$ for which $\|Df\|' < 1$. Now Proposition 2.20 applies. In other words, a sufficiently small closed "ball" around x_0 with respect to the norm $\|\cdot\|'$ can be chosen as the set U. □

Proposition 2.22

Let f be a continuously differentiable map with a fixed x_0 where $\|Df_{x_0}\| < 1$, and let U be a sufficiently small closed neighborhood of x_0 such that $f(U) \subset U$. Then any map g sufficiently close to f is a contraction on U.
Specifically, if $\epsilon > 0$, then there is a $\delta > 0$ and a closed neighborhood U of x_0 such that any map g with $\|g(x) - f(x)\| \leq \delta$ and $\|Dg(x) - Df(x)\| \leq \delta$ on U maps U into U ans is a contraction on U with its unique fixed point y_0 in $B(x_0, \epsilon)$.

Proof.

Since the linear map Df_x depends continuously on the point x, there is a small closed ball $U = \overline{B(x_0, \eta)}$ around x_0 on which $\|Df_x\| \leq \lambda < 1$. Assume $\eta, \epsilon < 1$ and take $\delta = \epsilon\eta(1 - \lambda)/2$. Then

$$\|Dg\| \leq \|Dg - Df\| + \|Df\| \leq \delta + \lambda \leq \lambda_(1 - \lambda)/2 = (1 + \lambda)/2 =: \mu < 1$$

on U, so g is a contraction on U. If $x \in U$, then $d(x, x_0) \leq \eta$ and

$$\begin{aligned} d(g(x), x_0) &\leq d(g(x), g(x_0)) + d(g(x_0), f(x_0)) + d(f(x_0), x_0) \\ &\leq \mu d(x, x_0) + \delta + 0 \\ &\leq \mu\eta + \delta \\ &\leq \eta(1 + \lambda) + \eta(1 - \lambda)/2 = \eta, \end{aligned}$$

so $g(x) \in U$ also, that is, $g(U) \subset U$.

Proof.

Finally, since $g^n(x_0) \to y_0$, we have

$$d(x_0, y_0) \leq \sum_{n=0}^{\infty} d(g^n(x_0), g^{n+1}(x_0)) \leq d(g(x_0), x_0) \sum_{n=0}^{\infty} \mu^n \leq \frac{\delta}{1-\mu} = \frac{\epsilon\eta(1-\lambda)}{1-\lambda},$$

which is less than ϵ. □

Proposition 2.23

If $f : \mathbb{R}\times(a,b) \to \mathbb{R}$ is continuous and $f_y := f(\cdot, y)$ satisfies $|f_y(x_1) - f_y(x_2)| \leq \lambda|x_1 - x_2|$ for all $x_1, x_2 \in \mathbb{R}$ and all $y \in (a,b)$, then the fixed point $g(y)$ of f_y depends continuously on y.

Proof.

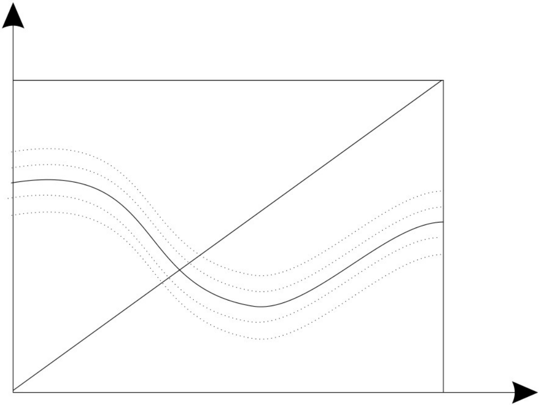
Since

$$|x - g(y)| \leq \sum_{i=0}^{\infty} |f_y^i(x) - f_y^{i+1}(x)| \leq \frac{1}{1-\lambda}|x - f_y(x)|,$$

we take $x = g(y') = f_{y'}(g(y))$ to get

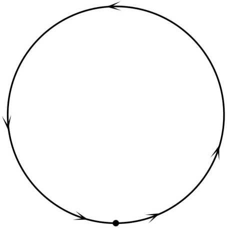
$$|g(y') - g(y)| \leq \frac{1}{1-\lambda}|f_{y'}(g(y')) - f_y(g(y'))|.$$

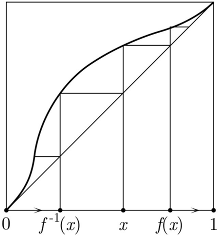
□



- At this point we have encountered two kinds of stability: Given a contraction, each individual orbit exhibits stable behavior in that every nearby orbit (actually, every orbit) has precisely the same asymptotics.
- Put differently, a little perturbation of the initial point has no effect on the asymptotic behavior.
- This constitutes the stability of orbits.
- On the other hand, the above two propositions tell us that contractions are stable as a system; that is, when we perturb the contracting map itself, then the qualitative behavior of all orbits remains the same, and the fixed point changes only slightly.

- This is a good time to make precise what we mean by a stable fixed point. As we said, we want every nearby orbit to be asymptotic to it.
- However, this is not all we want, as the follow figure shows, where we have a semistable fixed point.





Definition 2.24

A fixed point p is said to be *Poisson stable* if, for every $\epsilon > 0$, there is a $\delta > 0$ such that if a point is within δ of p then its positive semiorbit is within ϵ of p. The point p is said to be *asymptotically stable* or an *attracting fixed point* if it is Poisson stable and there is an $a > 0$ such that every point within a of p is asymptotic to p

- Such a map can be given, for example, as $f(x) = x + (1/4)\sin^2 2\pi x$ if the circle is represented as $\mathbb{R}/\mathbb{Z}$.
- We need to make sure that no nearby points ever stray far.
- But, as the example

$$f(x) = \begin{cases} -2x & x \leq 0 \\ -x/4 & x > 0 \end{cases}$$

showa, we must allow points to go a little further for a while.

- Consider a function f on the real line and suppose that we have a reasonable guess x_0 for a root.
- Unless the graph intersects the x-axis at x_0, that is, $f(x_0) = 0$, we need to improve our guess.
- To that end we take the tangent line and see at which points x_1 it intersects the axis by setting $f(x_0) + f'(x_0)(x_1 - x_0) = 0$.
- Thus the improved guess is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Definition 2.25

A fixed point x of a differentiable map F is said to be *superattracting* if $F'(x) = 0$.

Proposition 2.26

If $|f'(x)| > \delta$ and $|f''(x)| < M$ on a neighborhood of the root r, then r is a superattracting fixed point of $F(x) := x - f(x)/f'(x)$.

Remark 2.27

A small first derivative might cause the intersection of the tangent line with the x-axis to go quite far from x_0. The hypothesis $|f''(x)| < M$ holds whenever f'' is continuous.

Proposition 2.28

Approximating $\sqrt{z}$ by the Newton method with initial guess 1 is the same as using the first components of the Greek root extraction method.

Proof.

With initial guess 1 the Newton method gives the recursion

$$x_0 = 1,\ x_{n+1} = x_n - \frac{x_n^2 - z}{2x_n} = \frac{1}{2}(x_n + \frac{z}{x_n}).$$

The Greek method starts with $(x_0, y_0) = (1, z)$, and the recursion

$(x_{n+1}, y_{n+1}) = f(x_n, y_n)$ has the property that $y_n = z/x_n$. Therefore we have

$$x_{n+1} = \frac{x_n + y_n}{2} = \frac{1}{2}(x_n + \frac{z}{x_n}).$$

□