MATH 3963
NONLINEAR ODES WITH APPLICATIONS

R MARANGELL

Contents

Date: September 6, 2017.

1. Metric Spaces - The Beginnings

So.... I am from California, and so I fly through L.A. a lot on my way to my parents house. How far is that from Sydney? Well... Google tells me it's 12000 km (give or take), and in fact this agrees with my airline - they very nicely give me 12000 km on my frequent flyer points. But.... well.... are they really 12000 km apart? What if I drilled a hole through the surface of the Earth? Using a little trigonometry, and the fact that the radius of the earth is $R = 6371$ km, we have that the distance 'as the mole digs' so-to-speak from Sydney to Los Angeles CA is given by (see Figure 1):

$$x = 2R\sin\frac{12000}{2R} \approx 10303 \text{ km} \quad \text{(give or take)}$$

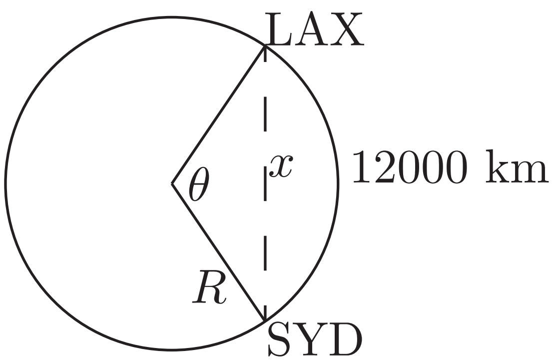

Figure 1. A schematic of the earth showing an 'equatorial' (or great) circle from Sydney to L.A.

So I have two different answers, to the same question, and *both* are correct. Indeed, Sydney is both 12000 km and around 10303 km away from L.A. What's going on here is pretty obvious, but we're going to make it precise. We are using a different definition of *distance* in each case. To make this mathematically precise we are going to define the following:

Definition 1.1. Let X be a (non-empty) set. A *metric* or a *distance function* is a map $d : X \times X \to \mathbb{R}$ such that the following three properties hold.

(1) $d(x, y) \geq 0 \quad \forall x, y \in X \quad$ and $\quad d(x, y) = 0 \Leftrightarrow x = y$ (nonnegativity)

(2) $d(x, y) = d(y, x) \quad \forall x, y \in X$ (symmetry)

(3) $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in X$ (the triangle inequality).

If X is a space, and d is a metric on X, the we call the two together (X, d) a *metric space*.

Metric spaces are pretty useful, in that they abstract the concept of distance, so really, any space you can put a concept of distance on, has a bunch of properties, (some that you might not even have thought about) and these properties, can be characterised in terms of the space, and the distance function on it. We are going to focus on the first three examples outlined in these notes, but I am including more (both here and in the tutorial questions), so that you can get a feel for how broad this topic actually is.

 ©University of Sydney

Example 1.1. Let X be the real vector space $X = \mathbb{R}^n$, and let d be the usual Euclidean distance

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

where the x_is and the y_is are the components of the vectors x and $y \in \mathbb{R}^n$. This is the metric that we used to compute the distance from Sydney to L.A. 'as the mole digs', only our space wasn't all of $\mathbb{R}^3$, but rather the surface of the Earth, viewed as a body in 3-space.

Example 1.2. Let $X = C^0[a,b]$ be the space of continuous real valued functions on the closed interval $[a,b]$, and let d_∞ be the metric defined by

$$d_\infty(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|.$$

That is, the distance between any two functions is the maximum distance the functions are simultaneously apart. By way of example (in this example) let $[a,b]$ be the interval $[0,2\pi]$ and let $f(x) = \sin(x)$ and $g(x) = \cos(x)$. Then

$$d_\infty(f,g) = \sup_{x\in[0,2\pi]} |\sin(x) - \cos(x)|.$$

The way to determine this is to find the critical points of the new function $h(x) = \sin(x) - \cos(x)$ for $x \in [0,2\pi]$. We take the derivative of $h(x)$ and see that $h'(x) = 0 \Rightarrow \sin(x) = -\cos(x)$. Which means that $x = \frac{3\pi}{4}, \frac{7\pi}{4}$. So for $x \in [0,2\pi]$ we have that $d_\infty(f,g) = \left|\sin(\frac{3\pi}{4}) - \cos(\frac{3\pi}{4})\right| = \sqrt{2}$ (see figure 2).

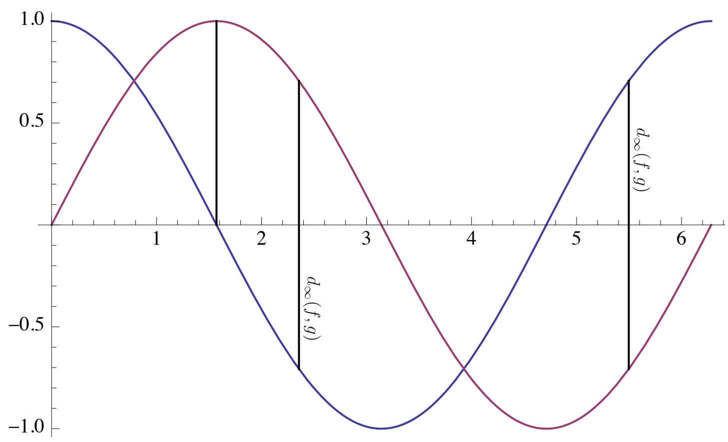

FIGURE 2. The distance in the supremum metric d_∞ between $\sin(x)$ (purple) and $\cos(x)$ (blue) are the longer lines (of length $\sqrt{2}$). The shorter line shows that 1 is not the supremum of $|\sin(x) - \cos(x)|$ for $x \in [0,2\pi]$.

Example 1.3. Here we use the same space $X = C^0[a,b]$ of continuous functions on a closed interval. But this time we define a new metric d_2 as follows. For two functions $f(x)$ and $g(x)$ we let the distance between them be defined as:

$$d_2(f,g) = \sqrt{\int_a^b (f-g)^2 dx}.$$

This is called the L^2 metric. Here again, we'll compute an example. Suppose again that $[a,b] = [0,2\pi]$ and that $f(x) = \sin(x)$ and $g(x) = \cos(x)$. Then you can verify yourself that

$$d_2(f,g) = \left[\int_0^{2\pi} (\sin(x) - \cos(x))^2 dx\right]^{\frac{1}{2}} = \sqrt{2\pi}.$$

Here the thing to notice is that $d_2(f,g) \neq d_\infty(f,g)$.

So these are the three important specific examples of metric spaces that we're going to be considering a lot in this class. However, they all come from a couple of more general ideas, and it is a good idea to familiarise yourself with how to get a whole bunch more examples of metric spaces.

Example 1.4 (Normed Spaces)**.** Suppose that X is a vector space over either $\mathbb{R}$ or $\mathbb{C}$ (denoted by F for convenience). Then 'recall' that a *norm* on X is a function (usually denoted) $||\cdot|| : X \to \mathbb{R}$ such that the following three properties hold

(1) $||x|| \geq 0 \quad \forall x \in X \quad \text{and} \quad ||x|| = 0 \Leftrightarrow x = 0$

(2) $||\lambda x|| = |\lambda|||x|| \quad \forall \lambda \in F(= \mathbb{R} \text{ or } \mathbb{C}) \quad \text{and} \quad \forall x \in X$

(3) $||x+y|| \leq ||x|| + ||y|| \quad \forall x, y \in X.$

If X is a vector space, and $||\cdot||$ is a norm on it, we call $(X, ||\cdot||)$ a *normed space.* Some examples of normed spaces are $\mathbb{R}^n$ with the Euclidean norm.

The point is that if $(X, ||\cdot||)$ is a normed space, then you can turn it into a metric space by defining

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

This is where the metrics for d_2 and d_∞ on $C^0[a,b]$ came from in fact (as well as the Euclidean metric on $\mathbb{R}^n$). The metric d_2 comes from the L^2 norm which is defined on continuous functions $f \in C^0[a,b]$ as

$$||f||_{L^2} := \sqrt{\int_a^b f^2 dx},$$

while d_∞ comes from the 'sup' norm which is defined on continuous functions $f \in C^0[a,b]$ as

$$||f||_\infty := \sup_{x\in[a,b]} |f(x)|\,.$$

Just to check that you are following along, you should verify yourself that if $[a,b] = [0,2\pi]$ then

$$||\sin(x)||_\infty = ||\cos(x)||_\infty = 1 \quad \text{and} \quad ||\sin(x)||_{L^2} = ||\cos(x)||_{L^2} = \sqrt{\pi}.$$

Example 1.5 (Inner Product Spaces)**.** Another host of examples comes from an *inner product* on a vector space. Suppose that X is a real or complex vector space. Then an inner product on X is a mapping $\langle\, ,\rangle : X \times X \to \mathbb{R}$ or $\mathbb{C}$ such that the following hold:

(1) $\langle x, y\rangle = \overline{\langle y, x\rangle} \quad \forall x, y \in X$

(2) $\langle ax+y, z\rangle = a\langle x, z\rangle + \langle y, z\rangle \quad \forall a \in \mathbb{R} \text{ or } \mathbb{C} \quad \text{and} \quad \forall x, y, z \in X$

(3) $\langle x, x\rangle \geq 0 \quad \forall x \in X \quad \text{and} \quad \langle x, x\rangle = 0 \Leftrightarrow x = 0.$

(Think of the dot product on $\mathbb{R}^n$). Now if we have an inner product, we can define a norm and hence a metric on X in the following way. For each $x \in X$ define the *norm induced by the inner product* as

$$||x||_{\langle,\rangle} := \sqrt{\langle x, x\rangle}.$$

To put this all in context, it is fairly clear that the Euclidean norm (and hence the Euclidean metric) comes from an inner product, the one you know and love, the standard dot product on $\mathbb{R}^n$. What might not be so obvious is that the L_2 norm (and hence the d_2 metric) on $C^0[a, b]$ comes from an inner product. Let f and g be two continuous functions on $[a, b]$. Then define the L^2 *inner product* as follows:

$$\langle f, g\rangle_{L^2} := \int_a^b fg dx.$$

It might be interesting to think about what it means for two functions to be *orthogonal* in this sense. This also allows one to define the *angle* between two functions, and really a host of other things that you can do with the standard dot product on $\mathbb{R}^n$. Lastly it is perhaps interesting to note that though the L^2 norm comes from an inner product, the sup norm $|| \cdot ||_\infty$ does not. This isn't too hard to show, and so you should go ahead and see what makes it different. (Hint: Google the *parallelogram identity.*)

2. The Topology of Metric Spaces

One of the points of introducing metric spaces was that metrics are *intrinsic* objects that *change the shape* of your space. In order to better understand this sentence, we begin to explore the concept of *open* and *closed* sets in a metric space. We start with the most basic sets. For what follows, we'll let (X, d) be a metric space.

Definition 2.1. For $x \in X$ and $\varepsilon > 0$ we define the *open ball of radius ε about x*, the *ε-ball* about x, or the *ε-neighbourhood* about x as all the points in the space which are within ε (as measured by our metric) of the element x. That is:

$$B_d(x; \varepsilon) = B(x; \varepsilon) := \{y \in X | d(x, y) < \varepsilon\}$$

Here we will sometimes use the B_d notation to indicate the metric, but if the metric is clear, we will drop the d subscript. Lastly, if $\varepsilon = 1$, then $B(x, 1)$ is called the *unit ball* about x.

In general the shape of an open ball of radius ε depends on the metric.

Example 2.1. Let $(X, d) = (\mathbb{R}, |\cdot|)$ be the usual Euclidean metric (i.e. the absolute value i.e. the metric induced from the Euclidean norm). Then the open ball around a point $a \in \mathbb{R}$ of radius ε is simply the open interval $(a - \varepsilon, a + \varepsilon)$.

Example 2.2. Let $(X, d) = (\mathbb{R}^n, || \cdot ||)$ be the usual Euclidean metric (i.e. the metric induced from the Euclidean norm). Then the open ball around a point $a \in \mathbb{R}^n$ of radius ε is simply the open n ball

$$B(a; \varepsilon) = \left\{x \in \mathbb{R}^n | \sqrt{(a_1 - x_1)^2 + \cdots + (a_n - x_n)^2} < \varepsilon\right\}$$

where the a_is are the components of a and the x_is are the components of the points in $\mathbb{R}^n$ in the $\varepsilon-$ball around a. For $\mathbb{R}^2$ this is the usual open disc of radius ε that you are used to (see Figure 3).

Example 2.3. An example that you may not find so familiar is if we let $(X, d) = (C^0[a, b], d_\infty)$ be the space of continuous functions on a closed interval, equipped with the sup metric (coming from the sup norm). What does a ball of radius ε about a fixed

©University of Sydney

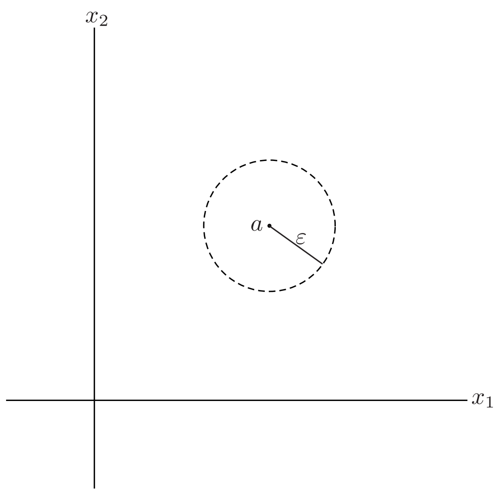

FIGURE 3. A disc of radius ε about a point $a \in \mathbb{R}^2$.

function $f \in (C^0[a,b], d_\infty)$ look like? Well, by chasing down the definitions, we have that a function g will be in $B_{d_\infty}(f;\varepsilon)$ provided that

$$\sup_{x\in[a,b]} |f(x) - g(x)| < \varepsilon.$$

This means that the function can get kind of 'wild' (whatever that means - and provided it stays continuous) but that it can't 'stray' too far from the function $f(x)$. See Figure 4 for a qualitative illustration. Just as a remark - the ε-ball around a function with the d_2 metric is a bit weirder...

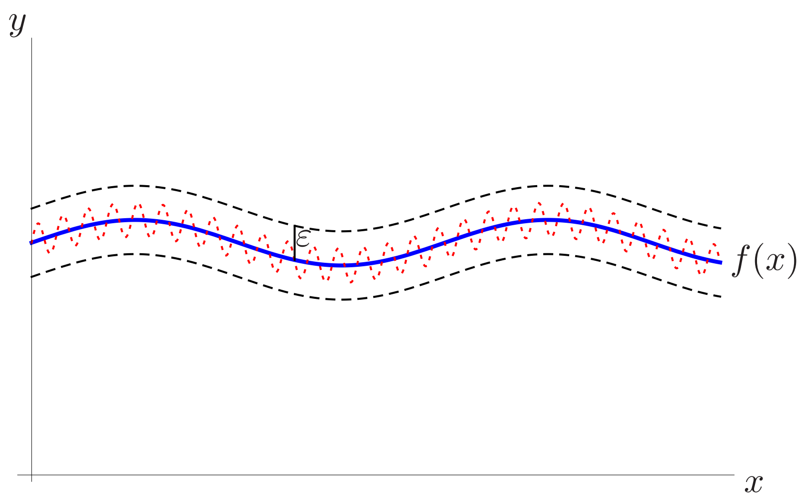

FIGURE 4. The 'ball' of radius ε about the function $f(x)$ in the sup norm. Note how the very oscillatory function in red dots is still in the 'ball'. A continuous function can do whatever it wants around the function $f(x)$ so long as it stays within the tube of radius ε.

Okay, so now for a more general idea than just an 'open' $\varepsilon-$ball around a point.

Definition 2.2. A subset $U \subset (X,d)$ of a metric space is called *open* if around every point $x \in U$ you can find an $\varepsilon-$ball that is entirely contained in U. That is

$$\forall x \in U,\ \exists \varepsilon \ \text{ s.t. } B(x;\varepsilon) \subset U.$$

©University of Sydney

We also have the following, related ideas/definitions

Definition 2.3. A subset $A \subset (X, d)$ of a metric space is called *closed* if its complement in X is open, that is $X \setminus A$ is open.

Definition 2.4. The *interior* of a set U in a metric space X, denoted $\text{Int}\,(U)$ is the largest open set contained in U, in the following sense: if B is any open set and $B \subset U$, then, $B \subset \text{Int}\,(U)$.

Definition 2.5. The *closure* of a set A in a metric space X, denoted $\overline{A}$ is the smallest closed set containing A, in the following sense: if V is a closed set, and $A \subset V$ then $\overline{A} \subset V$.

Definition 2.6. The *boundary* of a set U in a metric space X, denoted ∂U is defined as the closure of U minus the interior $\partial U := \overline{U} \setminus \text{Int}\,(U)$. The following, equivalent definition is also sometimes useful $\partial U = \overline{U} \cap \overline{(X \setminus U)}$.

Example 2.4. As a sort of sanity check with all these definitions, let's look at what these are in $\mathbb{R}$ with the usual Euclidean metric. First, we have that open intervals are open - this is useful, since you've probably been using the term 'open interval' for quite a while, so it would be a pain if somehow they were different things. Also it is sort of important to note that the whole space $\mathbb{R}$ itself is open, as is its 'complement' $\varnothing$ the empty set (there's nothing in it to put an $\varepsilon-$ball around so the definition is trivially satisfied). Further, we have that the union of any number of open intervals is also open. And we have that any finite number of these open intervals can be intersected. Putting this together we have the following

Fact. The open sets in $(\mathbb{R}, |\cdot|)$ are

- Open intervals (a, b),
- The whole set $\mathbb{R}$ and its complement $\varnothing$, the empty set,
- Arbitrary unions of open intervals (as well as arbitrary unions of open sets),
- Finite intersections of open intervals (as well as finite intersections of open sets).

What this is doing is making formal what you've already known since you knew the meaning of an open interval in $\mathbb{R}$. We can do the same thing for closed sets

Fact. The closed sets in $(\mathbb{R}, |\cdot|)$ are

- Closed intervals $[a, b]$,
- The whole set $\mathbb{R}$ and its complement $\varnothing$, the empty set (the complements of both are open),
- Arbitrary intersections of closed intervals (as well as arbitrary intersections of closed sets),
- Finite unions of closed intervals (as well as finite unions of closed sets).

Again, the aim here is to simply make precise what you already know, and to put it in the language of metric spaces. It is easy to see that the 'half-open' interval $[a, b)$ is neither open nor closed, according to our definitions. It is also straightforward to see that the interior of an open interval (a, b) is itself, while the closure of an open interval is the closed interval $\overline{(a, b)} = [a, b]$.... This sort of chasing the definitions carries on and on.

Example 2.5. This example is just the previous one, but replace $\mathbb{R}$ with $\mathbb{R}^n$ and the metric induced from the Euclidean norm in higher dimensions is denoted $\|\cdot\|$. It is straightforward to see that the open $n-$ball, of radius r about a point a - the one which you have been working with for quite a while is open i.e.

$$B(a, r) = \left\{ y \in \mathbb{R}^n | \sqrt{(y_1 - a_1)^2 + \cdots + (y_n - a_n)^2} < r \right\}$$

©University of Sydney

is open (where a_j are the components of a and y_j are the components of y). Likewise, we have that

Fact. The open sets in $(\mathbb{R}^n, \|\cdot\|)$ are
- Open balls of radius r about any point $a \in \mathbb{R}^n, B(a; r)$,
- The whole set $\mathbb{R}^n$ and its complement $\varnothing$, the empty set,
- Arbitrary unions of open balls (as well as arbitrary unions of open sets),
- Finite intersections of open balls (as well as finite intersections of open sets).

And similarly we have that closed sets are the closed disks you've been working with.

$$D(a, r) = \overline{B(a; r)} = \left\{ y \in \mathbb{R}^n | \sqrt{(y_1 - a_1)^2 + \cdots + (y_n - a_n)^2} \leq r \right\}.$$

Fact. The closed sets in $(\mathbb{R}^n, \|\cdot\|)$ are
- Closed disks of radius r about any point $a \in \mathbb{R}^n, D(a; r)$,
- The whole set $\mathbb{R}^n$ and its complement $\varnothing$, the empty set,
- Arbitrary intersections of closed disks (as well as arbitrary intersections of closed sets),
- Finite unions of closed disks (as well as finite unions of closed sets).

It is straightforward to see that the boundary of an open ball of radius r about a point $a \in \mathbb{R}^n$ is the $n-$sphere of radius r centred at a:

$$\partial B(a; r) = S^n(a; r) = \left\{ y \in \mathbb{R}^n | \sqrt{(y_1 - a_1)^2 + \cdots + (y_n - a_n)^2} = r \right\}.$$

You are of course, encouraged to prove these facts rigorously, though it is pretty much the same in both cases and in both cases is just an application of the definitions of open and closed sets.

Before we move on to more complicated things, I would like to point out that this definition of closed that we've been using is a bit 'clunky', and not so intuitive. I would be nice if we had a better, (but equivalent) definition of a set A in a metric space being closed. To that end, we define the following:

Definition 2.7. A point x in a metric space, (X, d) is called a *limit point* or an *accumulation point* of a subset $A \subset (X, d)$ if every $\varepsilon-$neighbourhood $B(x; \varepsilon)$ about x has at least one point, not equal to x in A. That is $x \in (X, d)$ is a limit point of A if and only if the following holds:

$$\forall \varepsilon > 0 \quad B(x; \varepsilon) \setminus \{x\} \cap A \neq \varnothing.$$

One thing to note is that the point x need not necessarily be in A as the next (straightforward) example shows.

Example 2.6. Let $(X, d) = \mathbb{R}, |\cdot|$. It is pretty straightforward to see that the limit points of the open interval (a, b), are the points in the closed interval $[a, b]$.

Example 2.7. A slightly more complicated example is the following. Let $\mathbb{N}^+$ denote the positive integers (i.e. $\{1, 2, 3, \ldots\}$. Define the set $A \subset \mathbb{R}$ as

$$A = \left\{ \frac{1}{n} \,\middle|\, n \in \mathbb{N}^+ \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \ldots \right\}.$$

The claim is that $\{0\}$ is the only limit point of A. First, we show that 0 is a limit point. This is easy enough, for any $\varepsilon > 0$ we have that there is an $N \in \mathbb{N}^+$ such that $\frac{1}{N} < \varepsilon$. Thus we have that for all $n > N$, $\frac{1}{n} \in B(0; \varepsilon) \setminus \{0\}$, so 0 is a limit point. Next we have to show that 0 is the only one. For any $x \in \mathbb{R} \setminus \{0\}$ we have that one of the four cases

(1) $x < 0$. In this case $0 - x = \delta > 0$, and so we have that $B(x; \frac{\delta}{2}) \setminus \{x\} \cap A = \varnothing$.

(2) $x > 1$. Here we have that $x - 1 = \delta > 0$ and so we have that $B(x; \frac{\delta}{2}) \setminus \{x\} \cap A = \varnothing$.
(3) $x = \frac{1}{n}$ for some $n \in \mathbb{N}^+$. Then $B(\frac{1}{n}; \frac{1}{2(n^2+n)}) \setminus \{\frac{1}{n}\} \cap A = \varnothing$.
(4) $\frac{1}{n+1} < x < \frac{1}{n}$ Then choose $\delta = \min\left(x - \frac{1}{n+1}, \frac{1}{n} - x\right)$, and $B(x; \frac{\delta}{2}) \setminus \{x\} \cap A = \varnothing$.
See Figure 5.

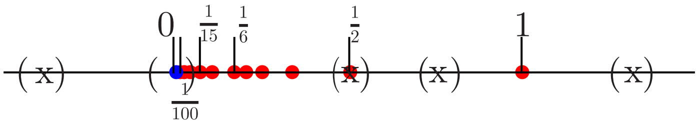

FIGURE 5. A Proof by Picture

The reason for introducing Definition 2.7 was that we wanted a 'better' definition of a closed set. Well, here it is

Theorem 2.1 (Alternate definition of a closed set). *A set $A \subset (X, d)$ of a metric space is closed if and only if it contains all its limit points.*

3. Sequences and Completeness

Related to the idea of limit points are is the notion of a (convergent) sequence. A sequence in a metric space is just a list of points (x_n), indexed by $\mathbb{N}$ or $\mathbb{N}^+$ (depending on whether or not you start counting from 0 or from 1.

Definition 3.1. A sequence (x_n) in a metric space is said to *converge* to a point $x \in X$ if it gets and stays as close to x as we please (using the distance function provided in the metric space of course), provided we choose the index high enough. That is:

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) \text{ s.t. } n > N(\varepsilon) \Rightarrow d(x, x_n) < \varepsilon.$$

If a sequence (x_n) converges to x then we call x the *limit* and write

$$\lim_{n\to\infty} x_n = x \quad \text{or} \quad x_n \to x.$$

If $(X, d) = (\mathbb{R}^n, \|\cdot\|)$ then you already know, and are (should be) reasonably familiar with this. An alternate way to characterise the definition of a limit of a sequence is the following.

Definition 3.2 (Alternate definition of a limit). A sequence (x_n) converges to x in a metric space (X, d) if and only if for every $\varepsilon > 0$ the ball of radius epsilon about x contains all but a finite number of terms of (x_n).

Continuing on in this vein, we say that

Definition 3.3. A sequence (x_n) in a metric space (X, d) is *Cauchy* if the elements of the sequence eventually all get close together. That is:

$$\forall \varepsilon > 0, \exists N = N(\varepsilon) \text{ s.t. } n, m > N \Rightarrow d(x_n, x_m) < \varepsilon.$$

We have the following proposition.

Proposition 3.1. *A convergent sequence is Cauchy.*

Proof. Let $(x_n) \to x$ be a convergent sequence in (X, d). Then we have that for an arbitrary $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that $d(x_n, x) < \frac{\varepsilon}{2}$ for all $n > N$. Thus we have that

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) \text{ (by the } \Delta \text{ inequality) } \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

The next natural question is to ask whether a Cauchy sequence is convergent. Evidently not, since we have the following definition

Definition 3.4. A metric space in which all Cauchy sequences converge is called *complete.*

Another reason for introducing this definition is the following

Axiom of Mathematics. *The real numbers $\mathbb{R}$ are complete.*

This is a basic assumption that you pretty much can't get away from. If you start with a sequence of real numbers, and you know that it is Cauchy, then it will converge to a real number. In some senses this is one of the main reasons that the real numbers were invented. Consider the following:

Example 3.1 (Easy Example)**.** The rational numbers $(\mathbb{Q}, |\cdot|)$ with the usual Euclidean metric are not complete. To see this take any decimal expansion of your favourite irrational number. Consider the expansion of $\pi \approx 3.1415926535897932384626........$ Then write

$$x_n = \sum_{j=0}^{n} \frac{a_j}{10^j} \quad \text{where } a_j \text{ is the } j\text{th decimal digit of } \pi,$$

so that $x_0 = 3, x_1 = 3.1, x_2 = 3.14, \ldots$ etc. These are all rational numbers (by construction), and we have that for any $\varepsilon > 0$ we can choose N so that $|\pi - x_N| < \frac{\varepsilon}{2}$. We then have that $\forall m, n > N, d(x_n, x_m) = |x_n - x_m| < |x_n - \pi| + |x_m - \pi| = \varepsilon$. And so the sequence is Cauchy. However, we *explicitly* chose it so that it would 'converge' (not in $\mathbb{Q}$ mind you, but in $\mathbb{R}$), to the number π which we know is not in $\mathbb{Q}$ (how do we know this?), and so our space $\mathbb{Q}$ is not complete.

Example 3.2. Let $(X, d) = (C^0[0,1], d_2)$. Let $(f_n) = x^n$ be the sequence of monic polynomials of a single degree. We claim that this sequence is Cauchy in (X, d_2). To see this, first suppose without loss of generality that $m > n$. We have

$$\begin{aligned} d_2(x^m, x^n) &= \left[\int_0^1 (x^m - x^n)^2 dx\right]^{\frac{1}{2}} = \left[\frac{1}{2m+1} - \frac{2}{1+m+n} + \frac{1}{2n+1}\right]^{\frac{1}{2}} \\ &\leq \left[\frac{1}{1+2m} + \frac{2}{1+m+n} + \frac{1}{1+2n}\right]^{\frac{1}{2}} \leq \left[\frac{4}{1+2n}\right]^{\frac{1}{2}} \\ &\leq \frac{2}{\sqrt{1+2n}} \end{aligned}$$

Now for any $\varepsilon > 0$ choose $N = N(\varepsilon)$ so that $\frac{2}{\sqrt{1+2N}} < \varepsilon$. Then if $m, n > N$ we have that $d_2(x^n, x^m) < \varepsilon$ and so the sequence is Cauchy in this metric.

Why am I going on about this? Well for starters, complete metric spaces are very nice, and secondly, this notion of being Cauchy and/or converging is quite subtle. For example consider the following.

Example 3.3. Let $(X, d) = (C^0[0,1], d_\infty)$. Let $(f_n) = x^n$. So we have the *same* space X and the *same* sequence in it. The only difference this time is that the metric has changed. We claim that for any fixed n, as $m \to \infty$ we have that $d_\infty(x^n, x^m) \to 1$, the distance in the sup norm tends towards 1 (see Figure 6). You can prove this by taking derivatives of the function $h(x) = x^n - x^m$. Again first, we assume without loss of generality that $m > n$. Then we have

$$h'(x) = nx^{n-1} - mx^{m-1} = x^{n-1}(n - mx^{m-n}) = 0 \Rightarrow x = 0 \text{ or } x = \left(\frac{n}{m}\right)^{\frac{1}{m-n}}.$$

Plugging this back into h and taking the limit for a fixed n as $m \to \infty$ we have

$$\lim_{m\to\infty}\left[\left(\frac{n}{m}\right)^{\frac{n}{m-n}} - \left(\frac{n}{m}\right)^{\frac{m}{m-n}}\right] = \lim_{m\to\infty}\left[\left(\frac{n}{m}\right)^{\frac{n}{m-n}} - 0\right] \overset{\text{L'H}}{=} 1$$

where we've used L'Hôpital's rule to compute the last limit.

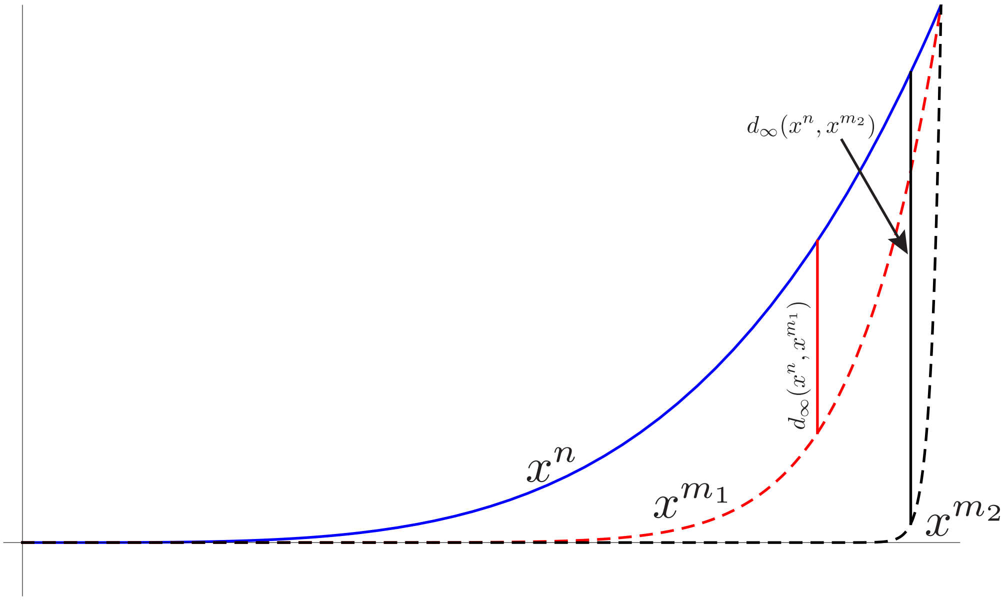

FIGURE 6. It's pretty easy to see from the picture that for any fixed n the supremum of the distance between x^n and x^m on $[0,1]$ tends towards 1 as $m \to \infty$.

Okay, so to recap, we have that the sequence $(f_n) = x^n$ *is* Cauchy in $(C^0[0,1], d_2)$, but is *not* in $(C^0[0,1], d_\infty)$. Now let's define the function $f : [0,1] \to \mathbb{R}$ as follows:

$$f(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{else.} \end{cases}$$

It is easy to see that for any fixed $x_0 \in [0,1]$ we have that the sequence of real numbers $(f_n(x_0)) \to f(x_0)$. It is also apparent that $f(x)$ is *not* in $C^0[0,1]$ (no matter what metric). However, this convergence issue is quite subtle. For example, you can verify for yourself that $(f_n) \to 0$ in $(C^0[0,1], d_2)$. So while we have that our sequence converges *pointwise* to one function, it converges in the d_2 metric to another. It actually will converge to an infinite number of functions which are not the same, however, only one of them is in $C^0[0,1]$. So we have that this sequence $(f_n) = x^n$ is Cauchy in $(C^0[0,1], d_2)$ and it converges to 0, which is in the same space. It also converges in the d_2 metric to a bunch of other functions which aren't in the space. Ao is $(C^0[0,1], d_2)$ a complete metric space? The answer is no, as the following example shows:

Example 3.4. Let f_n for $n \in \mathbb{N}^+$ be the sequence of functions defined as follows

$$f_n(x) = \begin{cases} 0 & 0 \le x \le \frac{1}{2} \\ nx - \frac{n}{2} & \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n} \\ 1 & \frac{1}{2} + \frac{1}{n} \le x \le 1. \end{cases} \tag{3.1}$$

 ©University of Sydney

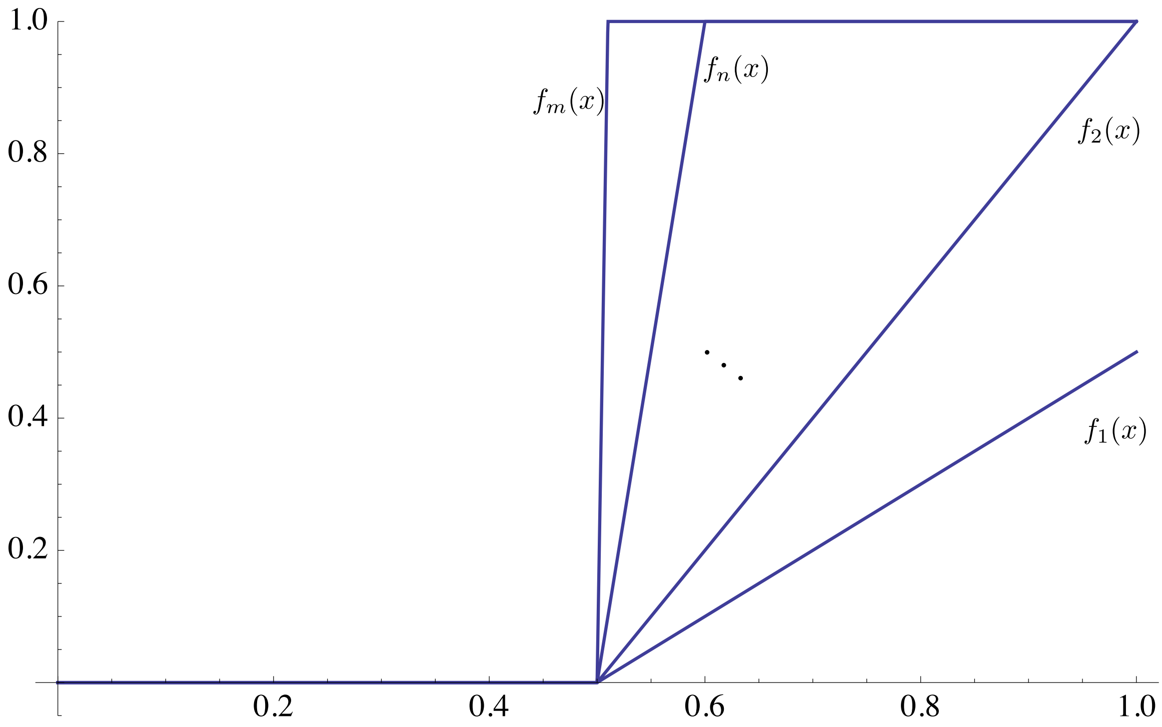

FIGURE 7.

For plots of the graphs of $f_n(x)$ on $[0,1]$ see Figure 7. It is reasonably straightforward to see that (f_n) is a Cauchy sequence in this metric. Assuming without loss of generality that $m > n > 2$ and proceeding by direct calculation we have that

$$\int_0^1 (f_n - f_m)^2 dx = \frac{(m-n)^2}{3m^2n} < \frac{(m+n)^2}{3m^2n} < \frac{4m^2}{3m^2n} < \frac{4}{3n}.$$

So we have that if $\frac{1}{N} < \frac{3}{4}\varepsilon$ then for all $n, m > N$ we have that $d_2(f_n, f_m)^2$ (and hence $d_2(f_n, f_m)$) can be made as small as we like. It should be reasonably clear that (f_n) converges pointwise to the discontinuous function

$$H(x) = \begin{cases} 0 & x \leq \frac{1}{2} \\ 1 & \frac{1}{2} < x. \end{cases}$$

(the Heavyside function shifted to the right by $\frac{1}{2}$). But what if, like in the case of x^n we could find some other function $f(x)$ which was continuous, and for which $d_2(f_n, f) \to 0$ as $n \to \infty$? We claim no such function can exist. To see this, suppose that there was such a function, call it $f(x)$. Then by properties of the definite integral, $f(x) \equiv 0$ on an uncountable number of points in the interval $[0, \frac{1}{2}]$, and $f(x) \equiv 1$ on an uncountable number of points in the interval $[\frac{1}{2}, 1]$, and $f(x)$ could only take a value different from 0 or 1 on a countable number of points in the interval from $[0,1]$ (otherwise the limit $d_2(f_n, f) \not\to 0$). However, we have assumed that $f(x)$ is a continuous function, and it takes the value of 0 somewhere in the interval, and 1 somewhere in the interval, so by the intermediate value theorem, it must take every other value, however, there are an uncountable number of these, and $f(x)$ can only be different from 0 and 1 on at most a countable number of points, a contradiction. So we can conclude that $(C^0[0,1], d_2)$ is not complete.

 ©University of Sydney

So $(C^0[0,1], d_2)$ is not a complete metric space, and we have another example of a space that isn't complete. However, there is still hope for the metric space $(C^0[0,1], d_\infty)$, because the sequences (x^n) and $f_n(x)$ aren't Cauchy in regards to this metric (we showed this for (x^n), can you show it for $f_n(x)$?). In fact, we have the following proposition.

Proposition 3.2. *The metric space $(C^0[a,b], d_\infty)$ is complete.*

Proof. Suppose that $(f_n) \subset (C^0[a,b], d_\infty)$ was a Cauchy sequence. Then for each $x_0 \in [a,b]$ we have that $f_n(x_0)$ is a Cauchy sequence of real numbers. Since $\mathbb{R}$ is complete, we have that this sequence converges to a real number say y_{x_0}. Define the limiting function $f : [a,b] \to \mathbb{R}$ by $f(x_0) = y_{x_0}$. Now we have to show that this function is continuous. That is for any fixed $x_0 \in [a,b]$, and for any $\varepsilon > 0$, we can choose δ so that $|y - x_0| < \delta \Rightarrow |f(y) - f(x_0)| < \varepsilon$. To see this, we add zero and appeal to the triangle inequality, and then the definition of a Cauchy sequence in this metric. Observe that

$$\begin{aligned} |f(y) - f(x_0)| &= |f(y) - f_n(y) + f_n(y) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\leq |f(y) - f_n(y)| + |f_n(y) - f_n(x_0)| + |f_n(x_0) - f(x_0)|. \end{aligned}$$

Now (f_n) is Cauchy in the sup norm which means that there is an N so that for $m, n \geq N$ we have that

$$\sup_{x \in [a,b]} |f_n(x) - f_m(x)| < \frac{\varepsilon}{3}.$$

In particular letting $m \to \infty$ this means that if $n \geq N$

$$\sup_{x \in [a,b]} |f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

But this means that for all $x \in [a,b]$ we can choose N so that if $n \geq N$ we have that

$$|f(x) - f_n(x)| \leq \sup_{x \in [a,b]} |f(x) - f_n(x)| \leq \frac{\varepsilon}{3}.$$

Now we choose δ so that if $|y - x_0| < \delta$ we have that $|f_N(x_0) - f_N(y)| < \frac{\varepsilon}{3}$ (we can do this because each of the f_n's are continuous). But this is all we need, now we put this together and we see that if $|y - x_0| < \delta$ then

$$\begin{aligned} |f(y) - f(x_0)| &\leq |f(y) - f_N(y)| + |f_N(y) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

□

A quick recap of what we've established.

(1) $(\mathbb{R}, |\cdot|)$ is complete (by Axiom of Mathematics). And consequently so is $(\mathbb{R}^n, \|\cdot\|)$.
(2) $(C^0[a,b], d_\infty)$ is complete.
(3) $(\mathbb{Q}, |\cdot|)$ is not complete.
(4) $(C^0[a,b], d_2)$ is not complete.

4. The Contraction Mapping Theorem

'Recall' that a metric space (X, d) is called *complete* if every Cauchy sequence in it converges to an element in the space. Suppose we have a map $f : X \to X$ from a complete metric space (X, d) to itself.

Definition 4.1. We call f a *contraction* if we can find a $c < 1$ such that

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

for all $x, y \in X$.

Example 4.1 (Linear maps)**.** Suppose we consider a linear map (a line) $f : \mathbb{R} \to \mathbb{R}$ (with the usual metric $d(x, y) = |x - y|$) with $f(x) = mx + b$ where m and b are real constants. When is f a contraction? We have that $|f(x) - f(y)| = |mx + b - my - b| = |m||x - y|$ so we can see right away that $f(x)$ will be a contraction if and only if $|m| < 1$.

Example 4.2. Now for a slightly more complicated example. Supose we consider the function $f(x) = \sqrt[3]{x} = x^{\frac{1}{3}}$. Is this function a contraction? The answer is no, and we'll discuss why in two different ways. The first is by direct calculation. Choose a $u, v \in \mathbb{R}$ so that $u^3 = x$ and $v^3 = y$. Since x, y are real numbers this is always possible. We then have

$$|f(x) - f(y)| = |x^{\frac{1}{3}} - y^{\frac{1}{3}}| = |u - v| \quad \text{and}$$
$$|x - y| = |u^3 - v^3| = |u - v||u^2 + uv + v^2|$$

So we just need to choose u and v so that $\dfrac{1}{|u^2 + uv + v^2|} > 1$ and we're in business. For example if we choose $u = 0$ and $v = \dfrac{1}{10}$ then $x = 0$ and $y = \dfrac{1}{1000}$ and we have that

$$|f(x) - f(y)| = \frac{1}{10} = 100|x - y|,$$

so f is not a contraction.

Another way to show that $f(x)$ is not a contraction is slightly more oblique, but will aid us in our discussion of things later. We can rewrite the definition of a contraction for $x \neq y$ as:

$$\frac{|f(x) - f(y)|}{|x - y|} \leq c,$$

and we see that the left hand side is a difference quotient. So if we can bound the slope of the secant lines by 1 then we have a contraction. In particular, if we have a C^1 function and we can bound the derivative by 1, then we have a contraction. Observe that in this case, we have that $f'(x) = \dfrac{1}{3}x^{-\frac{2}{3}}$ which is unbounded on $\mathbb{R}$ and in particular is strictly greater than 1 on the interval $(0, 1)$ and so we don't have a contraction. (We will return to an idea similar to this later).

The reason that we're interested in contraction maps is the following:

Theorem 4.1 (Contraction Mapping Theorem)**.** *Suppose that $f : X \to X$ is a contraction from a complete metric space to itself, with contraction constant $c < 1$. Then f has a unique fixed point. That is there exists a unique $x_* \in X$ such that $f(x_*) = x_*$.*

Prrof 1. The first proof of the contraction mapping theorem (or sometimes it's called the contraction mapping principle) shows how intuitive it really is: Simply pick a map (the english word) of a place that contains the point where you are standing. The place the map on the floor. It is obvious that there is one and only one point on the map that is exactly on top of where it is on the floor. □

Proof 2. While the above proof highlights how intuitive the contraction mapping principle is - it is perhaps not in the style to which mathematicians are accustomed. So we will also prove the contraction mapping principle by iterating the map $f(x)$. Let d denote the metric on the complete space X then we chose *any* initial value x_0 and define a sequence $x_n = f(x_{n-1})$. We will show that x_n defined this way is Cauchy. First we have that

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq cd(x_n, x_{n-1}) \leq c^2 d(x_{n-1}, x_{n-2}) \leq \cdots \leq c^n d(x_1, x_0).$$

 ©University of Sydney

Now the triangle inequality says that for any $m > n$ we have that

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{i=n}^{m-1} d(x_{i+1}, x_i) \leq \sum_{i=n}^{m-1} c^i d(x_1, x_0) \\ &= \frac{1 - c^{m-n}}{1 - c} c^n d(x_1, x_0) \leq K c^n \end{aligned}$$

where the last line follows from the expression for partial sums in the terms of a geometric series, and $K = \dfrac{d(x_1, x_0)}{c - 1}$ is a fixed constant. Now since $c < 1$ we have that for any $\varepsilon > 0$, we have an N such that for all $m, n > N$ we have that $d(x_m, x_n) < Kc^N < \varepsilon$. So our sequence is Cauchy. Now since X is complete, we have that our sequence must converge to an $x_* \in X$. The claim is that x_* is our unique fixed point.

To see this, choose N large enough so that $d(x_n, x_*) < \varepsilon$ for all $n > N = N(\varepsilon)$ for an arbitrary ε. Then we have that

$$\begin{aligned} d(f(x_*), x_*) \leq& d(f(x_*), x_{n+1}) + d(x_{n+1}, x_*) = d(f(x_*), f(x_n)) + d(x_{n+1}, x_*) \\ \leq& c\, d(x_*, x_n) + d(x_{n+1}, x_*) \leq \varepsilon(c + 1) \end{aligned}$$

And since this is true for all $\varepsilon > 0$ we have that $d(f(x_*), x_*) = 0$ and x_* is a fixed point. To see that this is unique, suppose we had another fixed point y_*, then

$$d(x_*, y_*) = d(f(x_*), f(y_*)) \leq c\, d(x_*, y_*)$$

The only way that this is possible is if $d(x_*, y_*) = 0$ and so our fixed point is unique. $\square$

Now we want to consider the implications of the contraction mapping principle on function spaces. We'll cover two examples - one that is fairly straightforward, and one that is more exotic. We'll begin with the straightforward one.

Example 4.3. We let $X = C^0(S^1)$ be the set of continuous periodic functions $f : \mathbb{R} \to \mathbb{R}$ of period 1. That is $f : \mathbb{R} \to \mathbb{R}$ and $f(x + 1) = f(x)$ for all $x \in \mathbb{R}$. You can also view this as the set of functions $f : [0, 1] \to \mathbb{R}$ where $f(0) = f(1)$ and then extending f to all of $\mathbb{R}$ by periodicity. This metric space is complete with respect to the metric induced from the sup norm which in this case reduces to

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|$$

because of periodicity. We define a transformation (map) on X as follows $T : X \to X$ by $T(f)(x) = \dfrac{1}{2} f(2x)$. To see that this is a contraction, observe that

$$\begin{aligned} d(T(f)(x), T(g)(x)) = \sup_{x \in [0,1]} |\frac{1}{2}(f(2x) - g(2x))| \leq \frac{1}{2} \sup_{x \in [0, \frac{1}{2}]} |f(2x) - g(2x)| \\ \leq \frac{1}{2} \sup_{x \in [0,1]} |f(x) - g(x)|. \end{aligned}$$

So we must have a fixed point. According to the contraction mapping principle, we can start with any initial condition and it will converge to the fixed point. Let's pick a function that makes computations relatively straightforward. Define a sequence $u_0(x) = \sin(2\pi x)$,

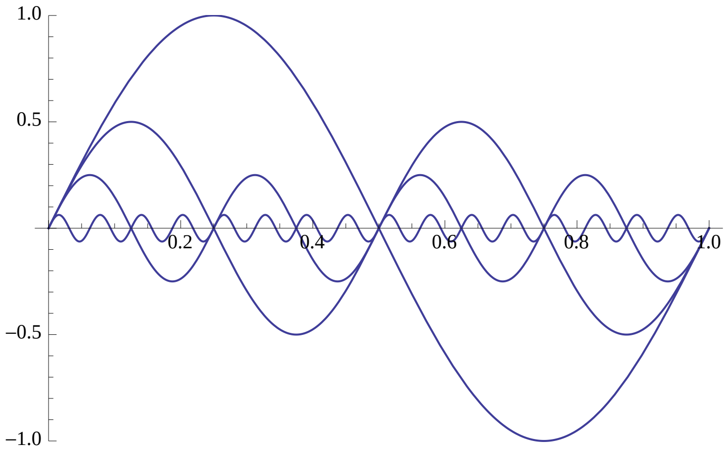

FIGURE 8. Plots of the sequence $u_i(x)$ for $i = 0, 1, 2, 4$. It is pretty clear that this is shrinking (fast) to 0 on all of $[0, 1]$.

and $u_n(x) = T(u_{n-1})(x)$. Then we have

$$
\begin{aligned}
u_0(x) &= \sin(2\pi x) \\
u_1(x) &= \frac{1}{2}\sin(4\pi x) \\
u_2(x) &= \frac{1}{4}\sin(8\pi x) \\
&\vdots \\
u_n(x) &= \frac{1}{2^n}\sin(2^{n+1}\pi x)
\end{aligned}
$$

And we can see that

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |u_n(x)| = 0$$

so that the fixed point of our contraction is 0. For a graphical representation of what's going on, see Figure 8.

Example 4.4. So that example is a bit mundane. For a more exotic example let's define the transformation on the same space $C^0(S^1)$ only this time define the map

$$T(f)(x) = \cos(2\pi x) + \frac{1}{2}f(2x).$$

Just as before we claim that this is a contraction (you should verify this yourself). What is its fixed point? To find it we follow the same iterative procedure. Let $f_0(x) = \sin(2\pi x)$

 ©University of Sydney

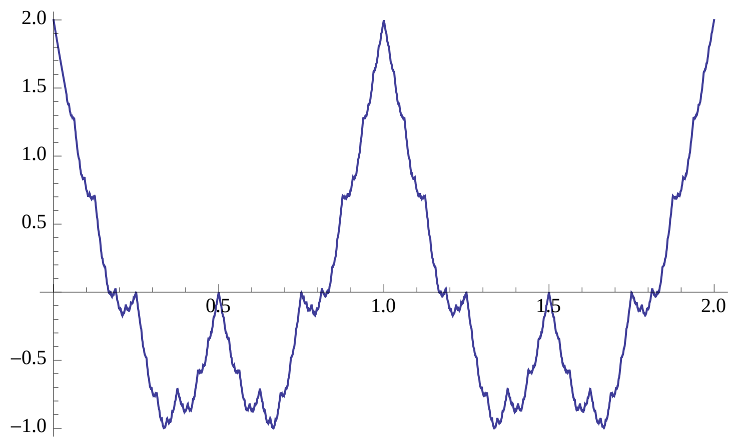

FIGURE 9. A plot of the Weierstrass function. It is (uniformly) continuous on $[0, 1]$, but not differentiable anywhere.

and define a sequence by $f_n(x) = T(f_{n-1})(x)$. Then we have

$$\begin{aligned}
f_0(x) &= \sin(2\pi x)\\
f_1(x) &= \cos(2\pi x) + \frac{1}{2}\sin(4\pi x)\\
f_2(x) &= \cos(2\pi x) + \frac{1}{2}\cos(4\pi x) + \frac{1}{4}\sin(8\pi x)\\
f_3(x) &= \cos(2\pi x) + \frac{1}{2}\cos(4\pi x) + \frac{1}{4}\cos(8\pi x) + \frac{1}{8}\sin(16\pi x)\\
&\vdots\\
f_n &= \sum_{j=0}^{n-1}\frac{1}{2^j}\cos(2^{j+1}\pi x) + \frac{1}{2^n}\sin(2^{n+1}\pi x)
\end{aligned}$$

The last term in the expression goes to 0 by the previous example and we are left with the limit of our sequence (which is the fixed point of our contraction) is

$$f_* = \sum_{j=0}^{\infty}\frac{1}{2^j}\cos(2^{j+1}\pi x).$$

This function is easy enough to plot (see figure 9). What is nice about the contraction mapping principle is that we know that this function must be continuous (and even uniformly continuous) on $[0, 1]$. However, this function is pretty exotic: while being uniformly continuous, it is in fact *nowhere* differentiable. This function was introduced by K. Weierstrass in the 19th century, and serves as a good reminder of the 'space' that exists between C^0 and C^1 functions.

Example 4.5. Here is another exotic example. We consider the space of continuous functions on $[0, 1]$, but with the restriction that $f(0) = 0$ and $f(1) = 1$, again with the supremum metric. Again, we claim that this space is a complete metric space. The proof is the same as for $(C^0[a, b], d_\infty)$, only you need to verify that the limit function also satisfies the end point conditions. Alternatively, you can argue that the condition $f(0) = 0$ is closed (it's complement is open, can you show this?), and then the last tutorial question from week 7 will give you the result. Next we are going to construct a map from

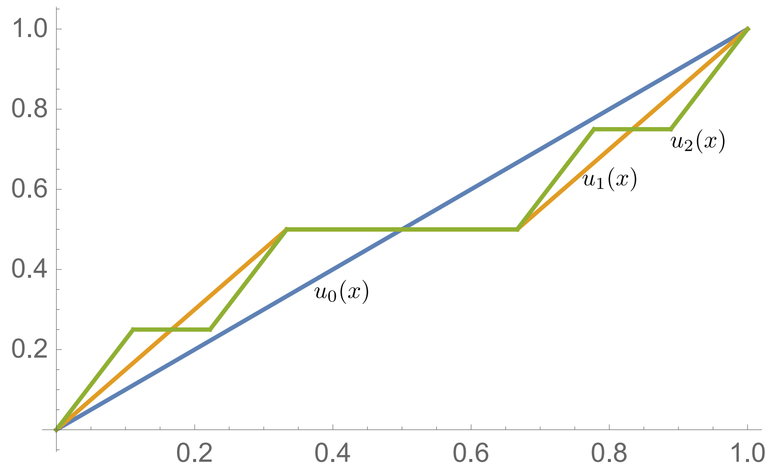

FIGURE 10. The graphs of the first three iterates $u_{0,1,2}(x)$ in the sequence defined by the transformation eq. (4.1)

the space of such functions to itself as follows

$$T(u)(x) = \begin{cases} \frac{1}{2}f(3x) & 0 \leq x \leq \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{1}{2} + \frac{1}{2}f(3x-2) & \frac{2}{3} \leq x \leq 1. \end{cases} \tag{4.1}$$

Next we observe that this map is a contraction. We have

$$\begin{aligned} d(T(f)(x), T(g)(x)) &= \sup_{x\in[0,1]} |T(f)(x) - T(g)(x)| \\ &= \frac{1}{2}\max\left\{ \sup_{x\in[0,1/3]} |f(3x) - g(3x)|, \sup_{x\in[2/3,1]} |f(3x-2) - g(3x-2)| \right\} \\ &= \frac{1}{2} \sup_{x\in[0,1]} |f(x) - g(x)|. \end{aligned}$$

So we must have a fixed point. We follow the same iterative procedure as before to find the fixed point. We let $u_0(x) = x$, and let $u_n(x) = T(u_{n-1})(x)$. The fixed point will then be

$$u_*(x) = \lim_{n\to\infty} u_n(x).$$

Writing down an explicit formula for $u_*(x)$ is a little bit messy, however graphically it is pretty straightforward to see what is going on. See figs. 10 and 11.

Why is this limit function exotic? Well for starters, it is constant *almost everywhere.* This idea can be made precise, but for now, consider the length of the set where the function is constant. $u_1(x)$ is constant (i.e. the derivative $u_1'(x) = 0$) on a set whose length is $1 - \frac{2}{3}$. Similarly, $u_2(x)$ is constant on a set of length $1 - \left(\frac{2}{3}\right)^2$, and indeed $u_n(x)$ is constant on a set of length $1 - \left(\frac{2}{3}\right)^n$. But this means that $u_*(x)$ is constant on a set of length 1. But it is continuous. And increasing from 0 to 1! This means that it takes an uncountable number of values on a set of zero length. You might recognise the set of points where $u_*(x)$ is non-constant is the *Cantor set*, or the Cantor middle thirds set. This is an uncountable set, that has no length. It is also one of the earliest examples of a

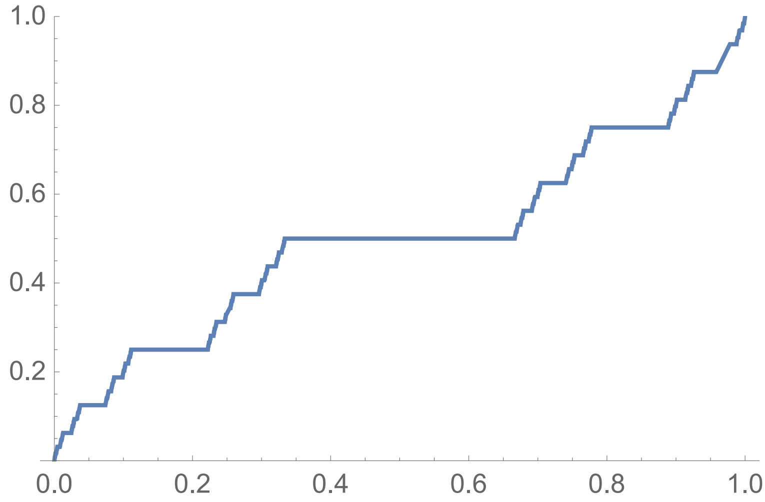

FIGURE 11. The so-called 'Devil's staircase'. This function is monotone increasing from 0 to 1 but is constant on a set of 'length' one.

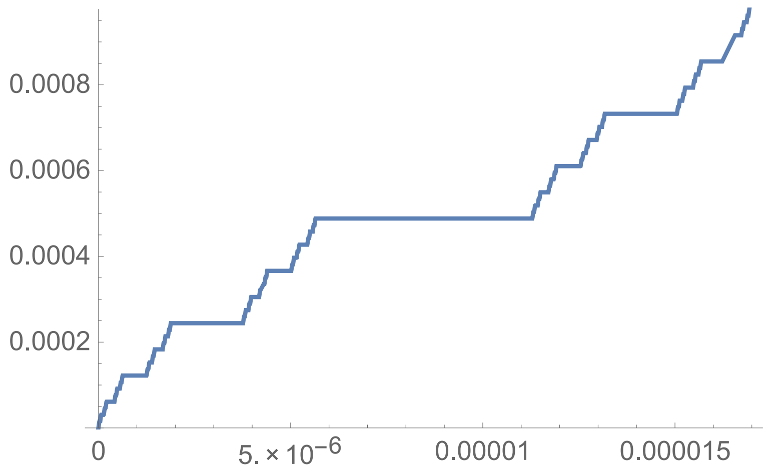

FIGURE 12. A zoom in of the Devil's staircase, notice that the graph looks the same, even though the x and y values are significantly smaller

fractal. Because of the way that we defined this function, its graph is also a fractal, and as such, it exhibits what is called *self-similarity*. That is, the function 'looks' like itself, on very different scales. Compare the axes lables in figs. 11 and 12, to see what I mean.

Lastly, this is an example of a continuous function which is not equal to the integral of its derivative, that is

$$u_*(x) \neq \int u_*'(x)dx.$$

The derivative is not defined on an uncountable number of points, but it is defined on a set of length 1, and on this set $u_*'(x) = 0$ in a perfectly valid way. Lastly,

5. Lipschitz Continuity

Now we're going to push a little bit further into the realm between C^0 and C^1 functions. From Example 4.2, when we were investigating whether or not $f(x) = x^{\frac{1}{3}}$ was a contraction, we wrote out the difference quotient $\frac{|f(x)-f(y)|}{|x-y|}$ and we were investigating whether or not it was less than or equal to a constant $c < 1$. If we are interested in investigating functions which are continuous, but not maybe not differentiable, we can ask if the difference quotient is bounded? To this end we have the following definition:

Definition 5.1. Let $E \subseteq \mathbb{R}^n$ be a subset of Euclidean space. We say a function $f : E \to \mathbb{R}^n$ is *Lipschitz continuous* (or just *Lipschitz*) if there exists a constant $K < \infty$ such that the following holds for all $x, y \in E$

$$|f(x) - f(y)| \leq K|x - y|.$$

The smallest such constant K is called the Lipschitz constant (of f on E).

Lipchitz functions may not be differentiable, but they are 'close'. For example, let's consider the function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = |x|$. This function is not differentiable at $x = 0$, but we still have that

$$|f(x) - f(y)| = ||x| - |y|| \leq |x - y|$$

by what's called the reverse triangle inequality. So our function is Lipschitz continuous with Lipschitz constant $K = 1$. Alternatively, we have that the derivative of $|x|$ apart from at $x = 0$ is ± 1 and so all of our difference quotients

$$\frac{|f(x) - f(y)|}{|x - y|}$$

are bounded. Incidentally, this gives you a sort of straightforward way to determine whether a function which is C^1 except for at a finite number of places is Lipschitz.

Question. Is the Weierstrass function Lipschitz continuous? Why or why not? If it is what's the Lipschitz constant?

A nice geometric interpretation of Lipschitz continuous is that the slope of the line (in $\mathbb{R}^{2n}$) between the points $(x, f(x))$ and $(y, f(y))$ is bounded by K for all $x, y \in E$. One way to think about this is that, a Lipschitz function is 'continuous +'. This means that they have all the properties of continuous functions, 'plus' a little bit more, in the sense that hey share a few properties with differentiable functions. (Though not all). The first part of this sentence is proved in the following proposition:

Proposition 5.1. *If $f : E \to \mathbb{R}^n$ is a Lipschitz continuous function, then f is continuous.*

Proof. The proof is just chasing definitions until we arrive at the result. Fix any $x \in E \subseteq \mathbb{R}^n$ and then for any $\varepsilon > 0$ we choose a y such that $|x - y| < \dfrac{\varepsilon}{K} = \delta$. Then if $y \in B_\delta(x)$ we have $|f(x) - f(y)| < K|x - y| < \varepsilon$. □

Often times Lipschitz continuity is too strong a result to require, so we introduce a weaker notion now:

Definition 5.2. A function $f : E \to \mathbb{R}^n$ on a subset $E \subseteq \mathbb{R}^n$ is said to be *locally Lipschitz* (on E) if for all $x \in E$, there is an $r > 0$ such that f is Lipschitz on $\overline{B_r(x)}$ the closed ball of radius r around x.

The idea is that for each $x \in E$ the Lipschitz constant K can vary with x, and $K(x)$ can get arbitrarily large on E. We also have the following proposition:

Proposition 5.2. *Suppose that f is C^1 on an open set E. Then $f(x)$ is locally Lipschitz.*

Proof. For any $x \in E$ there is an r such that $\overline{B_r(x)} \subseteq E$. Now choose a $y \neq x$ but such that $y \in \overline{B_r(x)}$. Then the line between x and y which is given as $\xi(s) = y + s(x - y)$ will be in $\overline{B_r(x)}$ when $s \in [0, 1]$. By the Fundamental Theorem of Calculus we have

$$f(x) - f(y) = \int_0^1 \frac{d}{ds} f(\xi(s))ds = \int_0^1 Df(\xi(s))(x - y)ds,$$

 ©University of Sydney

where the last equality follows from the chain rule and the definition of $\xi(s)$. Now since f is C^1 on E we have that $||Df||$, the norm of the Jacobian matrix (using the matrix norm from the first tutorial) is continuous and bounded on $\overline{B_r(x)}$. Therefore, if we let $K = \max_{y \in \overline{B_r(x)}} ||Df(y)||$ then we have that $|f(x) - f(y)| \leq K|x - y|$. □

Note that this last proof also gives you a handle on what the Lipschitz constants are. This can be quite useful.

6. Existence and Uniqueness of Solutions to ODEs

After all of these formal asides, analysis and topology, we are finally ready to discuss the existence and uniqueness of solutions to the ODE initial value problem:

$$\dot{x} = f(x) \qquad x(0) = x_0 \tag{6.1}$$

where $x(t) : J \to \mathbb{R}^n$ is some vector of functions on an interval $J := [-a, a]$ of the real line (to be determined later) and $f : E \to \mathbb{R}^n$ is a map from a subset of Euclidean space $E \subseteq \mathbb{R}^n$ to Euclidean space. We have two fundamental questions that we are going to address:

(1) Is there a solution to (6.1)? What sort of conditions do we need to put on $x(t)$ and $f(x)$ in order to ensure this, and
(2) Is such a solution unique? That is is there more than one function which satisfies (6.1) and which passes through the initial value x_0?

In order to answer these questions, we need to reformulate equation (6.1) in an integral form. Formally, integrating this equation and inputting the initial condition gives :

$$x(t) = x_0 + \int_0^t f(x(s))ds. \tag{6.2}$$

Right away we note the following lemma:

Lemma 6.1. *Suppose that $f \in C^k(E, \mathbb{R}^n)$ for some $k \geq 0$ and that $x \in C^0(J, E)$ is a solution to* (6.2)*. Then $x \in C^{k+1}(J, E)$ and is a solution to* (6.1)*.*

Proof. First observe that if $x(t)$ satisfies (6.2), then $x(0) = x_0$. Now if $x(s)$ is continuous then so is $f(x(s))$. Moreover $\int_0^t f(x(s))ds$ is in fact C^1. Thus the right hand side of (6.2) is C^1 and so the left hand side is too, so $x(t)$ is C^1. Then we just lather, rinse and repeat (proof by induction) and we're done. □

The other (cool?) thing about the formulation of (6.1) as (6.2) is that we can look at it as an operator acting on *functions*:

$$T(u)(t) = x_0 + \int_0^t f(u(s))ds.$$

In order that T be well defined, $u(t)$ must be chosen in some suitable function space (say $C^0(J, \mathbb{R}^n)$). Since solutions to eq. (6.1) satisfy eq. (6.2), we have that a continuous function $x_*(t)$ will satisfy the initial value problem (6.1) if and only if it is a fixed point of T. That is if and only if $T(x_*(t)) = x_*(t)$. Like in the second proof of the contraction mapping theorem we can find fixed points of our maps by iteration. In this case this is called *Picard iteration.* The strategy is as follows. Pick an initial function $u_0(t)$. Define a sequence $u_n(t) = T(u_{n-1})(t)$. (The elements of such a sequence are called *Picard iterates* by the way). Write down the sequence of functions and see what you get. Let's do a couple of examples to get a feel for Picard iteration:

 ©University of Sydney

Example 6.1. Let's consider the initial value problem $\dot{x} = rx$ where $x : \mathbb{R} \to \mathbb{R}$ is a scalar function and $x(0) = x_0$. We already know that the solution to this initial value problem is $e^{rt}x_0 = x(t)$. Now let's define an operator on continuous functions

$$T(u)(t) = x_0 + \int_0^t ru(s)ds,$$

chosen as the integral operator whose right hand side is the right hand side of the integral formulation of our initial value problem. Now define a sequence as follows $u_0(t) = x_0$ the constant function, and $u_n(t) = T(u_{n-1})(t)$. We have

$$\begin{aligned}
u_0(t) =& x_0 \\
u_1(t) =& x_0 + \int_0^t rx_0 ds = x_0(1 + rt) \\
u_2(t) =& x_0 + \int_0^t rx_0(1 + rs)ds = x_0 \left(1 + rt + \frac{r^2t^2}{2}\right) \\
&\vdots \\
u_n(t) =& x_0 \left(\sum_{j=0}^{n} \frac{(rt)^j}{j!}\right).
\end{aligned}$$

And so we see that $u_n(t) \to e^{rt}x_0$ as $n \to \infty$ (not so coincidentally). Moreover we can see that $x(t) = e^{rt}x_0$ is in fact a fixed point of the operator T and hence a solution to (6.2) and hence (6.1). What is also cool, and maybe not so obvious is that we don't need to start with $u_0(t) = x_0$ in order for this process of Picard iteration to converge on the solution. If, for example we were to start with $u_0(t) = t$, then we would have

$$\begin{aligned}
u_0(t) =& t \\
u_1(t) =& x_0 + \int_0^t rs ds = x_0 + \frac{rt^2}{2} \\
u_2(t) =& x_0 + \int_0^t r\left(x_0 + \frac{rt^2}{2}\right) ds = x_0 + x_0 rt + \frac{r^2t^3}{3!} \\
&\vdots
\end{aligned}$$

$$u_n(t) = x_0 \sum_{j=0}^{n-1} \frac{(rt)^j}{j!} + \frac{r^n t^{n+1}}{(n+1)!}$$

The right most expression in $u_n(t)$ tends to zero for all $t \in \mathbb{R}$ as $n \to \infty$ and so we have that $u_n(t) \to e^{rt}x_0$ as $n \to \infty$ this time as well.

In the above examples we needed to choose our initial function $u_0(t)$ from a 'suitable space'. If for example, we picked a badly behaved initial function, then we wouldn't have ended up with the solution to (6.1). In this case we want the integrals to all make sense, so we'll say that we want our initial conditions to be continuous functions of t. We also need the right hand side of (6.1) to make sense, so we need to restrict ourselves to regions $E \subseteq \mathbb{R}^n$ where $|f(x)| < \infty$. Neither of these turns out to be that much of a terrible restriction.

We are now ready to state the existence and uniqueness theorem:

 ©University of Sydney

Theorem 6.1 (Picard-Lindelöf). *Suppose that for $x_0 \in \mathbb{R}^n$ there exists an $r > 0$ so that $f : \overline{B_r(x_0)} \to \mathbb{R}^n$ is Lipschitz with Lipschitz constant K. Then the initial value problem (6.1) has a unique solution $x(t)$ for $t \in J = [-a, a]$ where $a = \dfrac{r}{M}$ and $M = \max\limits_{x \in \overline{B_r(x_0)}} |f(x)|$.*

Proof. The main idea is to use the contraction mapping principle. In order to do this we need a complete metric space. We will choose $(X, d) = (C^0([-a, a], \overline{B_r(x_0)}), d_\infty)$. That is the space of continuous functions mapping the interval $[-a, a]$ to the closed ball of radius r about the point x_0, $\overline{B_r(x_0)}$. The next thing we need is a contraction. Define the map T as follows

$$T(u)(t) = x_0 + \int_0^t f(u(s))ds.$$

We claim that T is a contraction from $X \to X$. If the claim is true, then we have a unique fixed point which by construction will be the solution to the initial value problem (6.1). In order to prove the claim, we will need to show that $T : X \to X$ and that T is indeed a contraction on X. First, if $u(t) \in X$ then $T(u)(t)$ is continuous since f is continuous. Now, since f is Lipschitz continuous on the closed set $\overline{B_r(x_0)}$, it is continuous. So it must attain its maximum somewhere on the set. Define $M := \max\limits_{x \in \overline{B_r(x_0)}} |f(x)|$. If $t \in [0, a]$ (the right half of the interval J, then we have the following

$$|T(u)(t) - x_0| = \left| x_0 + \int_0^t f(u(s))ds - x_0 \right| \le \int_0^t |f(u(s))|ds \le M|t| \le Ma.$$

In order to ensure that the image of $T(x)(t)$ is in $\overline{B_r(x_0)}$ (so that T maps X to itself), we must have that the right hand side of the above inequality must not be larger than r. But we chose a so that that $a \le \dfrac{r}{M}$. Next we observe that the same equality (for basically the same reasons) holds for $t \in [-a, 0]$. This shows that $T : X \to X$. To prove that $T(x)$ is a contraction, we have that

$$|T(u) - T(v)| \le \int_0^t |f(u(s)) - f(v(s))|ds \le K \int_0^t |u(s) - v(s)| \le Ka\, d_\infty(u, v).$$

This means that for $t \in J$, $d_\infty(T(u), T(v)) \le Ka\, d_\infty(u, v)$, so T will be a contraction when $c = Ka < 1$. So we have proven the theorem for the case when $a \le \dfrac{r}{M}$ and $a < \dfrac{1}{K}$. □

We are going to move on to a couple of examples, but first a couple of remarks.

Remark. The first thing to note is that we haven't exactly proven the statement of the theorem. In order to remove the second constraint, (i.e. that $a > \frac{1}{K}$), we require some more technology (specifically, we need either the so-called Weierstrass M test, and the notion of uniform continuity and uniform convergence, or the so-called Bielecki norm on a function space). But in the interest of expediency, we are going to skip this. It is just important to know that one can in fact tighten the proof, and remove the second constraint (the $a \le \frac{1}{K}$ part), which gives the theorem, as stated above.

Remark. Another thing to note is that there is nothing particularly special about $t = 0$. That is, we could have chosen *any* initial time t_0 from which to begin our initial value problem, and the result would have held the same. We just need to adjust the interval of existence $J = [t_0 - a, t_0 + a]$ and our map $T(u) = x_0 + \int_{t_0}^t f(u(s))ds$, but the rest of the proof remains the same.

Remark. Finally, it is worth noting, that one can show *existence* under some pretty mild conditions for f (see the Peano and Carathéodory existence theorems), and from an

©University of Sydney

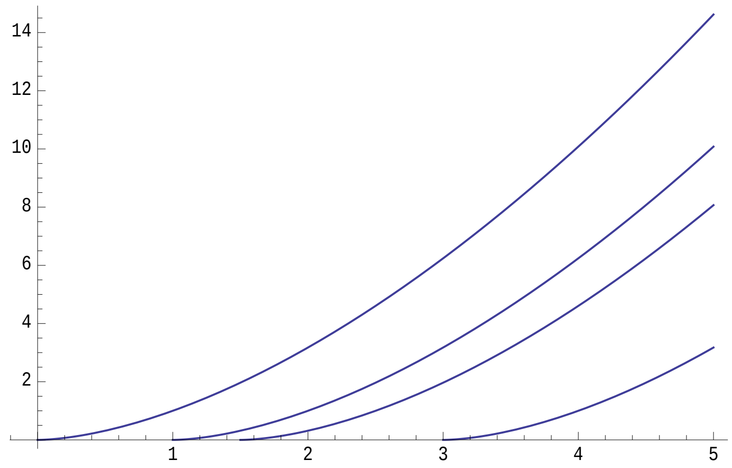

FIGURE 13. Plots of the (multiple) solutions to the IVP in example 1 with $\xi = 0, 1, 1.5$ and 3.

applied/modelling point of view this sort of makes philosophical sense. For uniqueness however, you really really need f to be Lipschitz at your initial condition x_0. As the following example shows.

Example 6.2. Consider the IVP

$$\dot{x} = \frac{5}{3}x^{\frac{2}{5}} \qquad x(0) = 0.$$

Now this function f is not Lipschitz at $x_0 = 0$, and so the theorem will not guarantee a unique solution. In fact we will construct an infinite family of solutions to this IVP. First observe that the constant solution $x = 0$ is a solution to the IVP. Next, use separation of variables to see that $x = t^{\frac{5}{3}}$ is a solution as well. So we have two solutions to this IVP. This is bad. But wait - it gets even worse. Indeed define the functions

$$x_\xi(t) = \begin{cases} 0 & \text{if } t \leq \xi \\ (t-\xi)^{\frac{5}{3}} & \text{if } \xi \leq t \end{cases}$$

for every $\xi > 0$. Each of these is a solution to the IVP. Figure 16 is a plot of a few values of ξ. We have an *infinite* family of solutions to this initial value problem.... Why is this bad? Well, suppose you were modelling something, and this came up. How would you know which member of the family you would choose? How could you predict the state of your future system based on your model, and the fact that you knew what was happening at some initial condition?

Example 6.3. As another example of how the Lipschitz property enters into the behaviour of solutions, it is instructive to consider as the right hand side of eq. (6.1) a function that is Lipschitz, but that doesn't necessarily look like it on all scales. An example is say

$$\dot{x} = \tanh(1000x) \qquad x(0) = x_0. \tag{6.3}$$

This function is perfectly smooth, but if you look at it on a large scale, it appears to be discontinuous (and hence it isn't Lipshitz). See fig. 14. In fact eq. (6.3) has exact solutions. They are given as

$$x(t) = \frac{\text{Arcsinh}\left[e^{1000t}\sinh(1000x_0)\right]}{1000}.$$

 ©University of Sydney

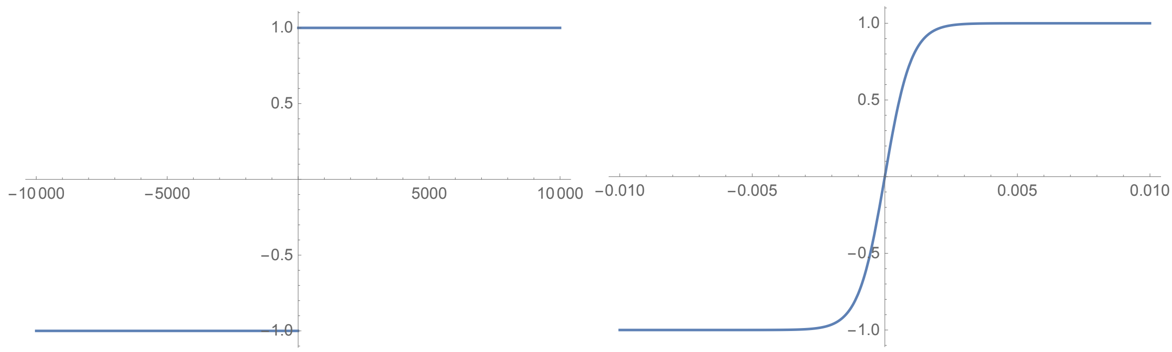

FIGURE 14. Plots of the same function $\tanh(1000x)$, on different scales. On the large scale (left) it appears discontinuous (and hence not Lipschitz).

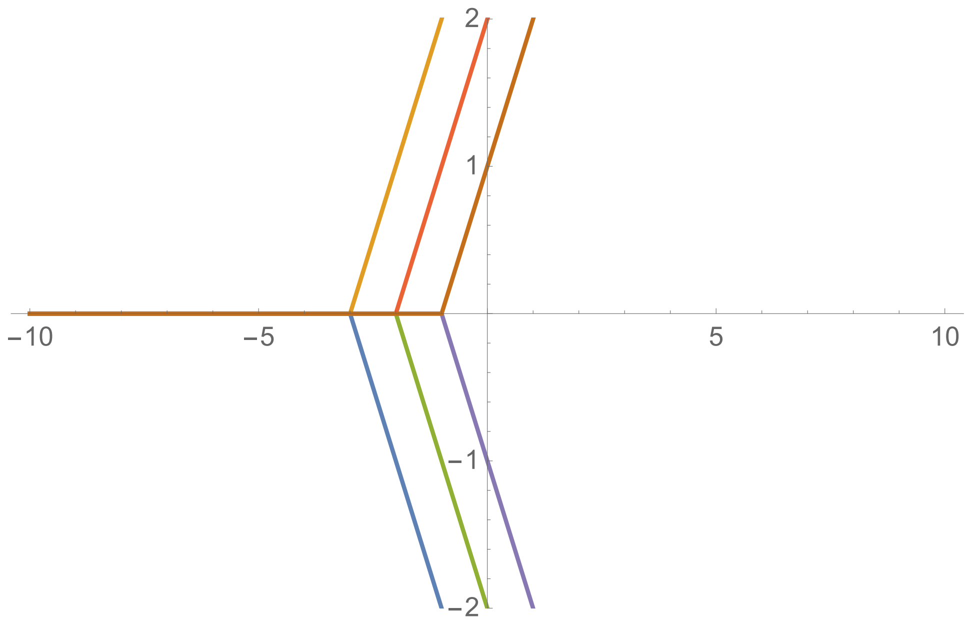

FIGURE 15. Plots of several solutions to eq. (6.3). The solutions are very close to zero, before 'taking off' to satisfy their initial condition.

What do these solutions look like? See fig. 15 for a plot of several initial conditions. What is interesting to note is that for $t < 0$, these solutions are all exponentially small, and they carry on as if they didn't have to satisfy the Picard theorem. If you were just tracking one (say in a model), because of the exponential factor killing off the part with the initial condition, you wouldn't really be able to tell when the function would 'take off' - that is, the non-Lipshcitz nature of the large scale is sort of built into the solutions for times away from $t = 0$. But, then as t nears zero, suddenly the solution remembers that it is in fact a well-behaved solution and needs to satisfy its initial condition, so it 'takes off' and then is recognisably different.

In Theorem 6.1 we have that we are guaranteed a solution on the interval $J = [-a, a]$ but can we do better? The best possible interval of existence would be if we could always find a solution defined on all of $\mathbb{R}$, and we did this for linear equations. However, for *nonlinear* equations, this is generally not possible. But, even though we can't guarantee a solution on all of $\mathbb{R}$, we *can* usually do better than the interval of existence guaranteed by Theorem 6.1. We will discuss this more in Section 8.

©University of Sydney

Example 6.4. Consider the IVP

$$\dot{x} = x^2 \qquad x(0) = x_0 > 0 \tag{6.4}$$

Now our function $f(x) = x^2$ is Lipschitz in the 'closed ball' (in this case just the interval $[x_0 - r, x_0 + r]$) $\overline{B_r(x_0)}$ about x_0 and indeed we can show that $|f(x)| \leq |(x_0 + r)^2| =: M$ for all the $x \in [x_0 - r, x_0 + r]$. So Theorem 6.1 tells us that we have a solution $x(t)$ defined on $[-a, a]$. In this example, $a \leq \frac{r}{M} = \frac{r}{(x_0 + r)^2}$, so we have a solution to (7.1) defined on $[-\frac{r}{(x_0 + r)^2}, \frac{r}{(x_0 + r)^2}]$. We want the largest interval of definition possible, which happens when $r = x_0$ so in particular we have a solution defined for $t \in [-\frac{1}{4x_0}, \frac{1}{4x_0}]$.

On the other hand, we can explicitly solve (7.1) to get that $x(t) = \frac{x_0}{1 - x_0 t}$ which is defined for $t \in (-\infty, \frac{1}{x_0})$. It is worth noting that this is quite a bit larger than what the theorem gave us, but it is *not* all of $\mathbb{R}$.

Before we can move on to the next module of the course, we will need a couple of more technical theorems on the dependence of our solutions on initial conditions and the proof of the existence of a maximal interval on which the solution to eq. (6.1) can be defined. However, I will not be including lectures on this material. You will see it on tutorials and perhaps on the exam, and I am including lecture notes. You should go and read it now before moving on to the next part.

7. Dependence on Initial Conditions and Parameters

Just for completeness, 'recall' The Picard-Lindelhöf theorem of existence and uniqueness (sometimes this is referred to as *local well-posed-ness*) of solutions to ordinary differential equations is as follows

Suppose we have our initial value problem

$$\dot{x} = f(x) \qquad x(t_0) = x_0 \tag{$*$}$$

where $x(t)$ is a vector of functions mapping some (to be determined) interval $x : J \to E \subseteq \mathbb{R}^n$ to Euclidean space, and $f : E \to \mathbb{R}^n$ is some function. We have the following:

Theorem 7.1 (Picard-Lindelhöf)**.** *Suppose that for $x_0 \in \mathbb{R}^n$ there is an $r > 0$ such that on the closed ball of radius r around x_0, $f : \overline{B_r(x_0)} \to \mathbb{R}^n$ is Lipschitz with constant K (i.e. f is locally Lipschitz at x_0). Then the initial value problem $(*)$ has a unique solution $x(t) : J \to \mathbb{R}^n$, on the interval $J = [t_0 - a, t_0 + a]$ where the number $a = \frac{r}{M}$, with $M = \max_{x \in \overline{B_r(x_0)}} |f(x)|$.*

So again, we return to our general IVP

$$\dot{x} = f(x) \qquad x(0) = x_0 \tag{7.1}$$

and suppose that we have that $f(x)$ is Lipschitz on a closed ball of radius r about some point $p \in \mathbb{R}^n$ $\overline{B_r(x_0)}$, then we have a solution for all initial conditions $y \in B_r(x_0)$. That is, for each $y \in B_r(x_0)$ we can write a solution to eq. (7.1) with $x_0 = y$, defined on some interval $[-a(y), a(y)]$ (which might be a function of the initial condition!). We'll write $x(t) = u(t; y)$ as a function of t and of y. Conventionally we'll use a semicolon remind ourselves of u's dependence on the initial condition y but to distinguish it from the independent variable of the ODE. Using this notation our initial value problem eq. (7.1), becomes a *family* of IVPs, one for each $y \in B_r(x_0)$:

$$\frac{d}{dt}u(t;u) = f(u(t;y)) \qquad u(0;y) = y. \tag{7.2}$$

The first question that arises is 'for which values of $y \in B_r(p)$ will we be able to find a common interval of definition for $u(t;y)$ for the variable t?' That is if we are willing to shrink the domain of 'niceness' of f a little bit, can we guarantee that the $u(t;y)$s will all be defined on the same interval $J = [-a, a]$ where a only depends on $f(x)$ and on r (and not on y any more). The answer is given in the following theorem:

Theorem 7.2. *Suppose that for a given $x_0 \in \mathbb{R}^n$ there is an r such that $f : \overline{B_r(x_0)} \to \mathbb{R}^n$ is Lipschitz with Lipschitz constant K. Let $M = \max_{x \in \overline{B(x_0)}} |f(x)|$. Then the family of solutions $u(t;y)$ of eq. (7.1) exists for each $y \in \overline{B_{\frac{r}{2}(x_0)}}$ and is unique for $t \in [-a, a]$ provided $a \leq \dfrac{r}{2M}$.*

Proof. The proof is basically the same as the proof of theorem 6.1. We define the complete metric space $(X, d) = ((C^0[-a, a], \overline{B_r(x_0)}), d_\infty)$ exactly as before, and the contraction for each initial condition

$$T_y(u)(t) = y + \int_0^t f(u(s))ds$$

If $y \in \overline{B_{\frac{r}{2}}(x_0)}$ and $u \in X$ then $T_y(u) \in V$ provided

$$|T_y(u) - x_0| \leq |y - x_0| + \int_0^t |f(u(s))|ds \leq \frac{r}{2} + Ma.$$

In this instance, in order for $T : X \to X$ we must have that $\frac{r}{2} + Ma < r$, which means that $a < \dfrac{r}{2M}$. And in order for T to be a contraction on $\overline{B_r(x_0)}$, we must have $a \leq \dfrac{1}{K}$ as before. Also as in the existence and uniqueness proof, we can eliminate the second condition on a with a little more work. $\square$

Example 7.1. Let's return to the IVP

$$\dot{x} = x^2 \qquad x(0) = x_0 > 0.$$

On the 'closed ball' of radius r around x_0 (the interval $[x_0 - r, x_0 + r]$) we have that $f(x)$ is Lipschitz with Lipschitz constant $K = 2(r + x_0)$. We also have that $f(x)$ achieves it's maximum of $M = (r + x_0)^2$. Thus for all initial conditions y such that $|y - x_0| \leq \dfrac{r}{2}$ we have a solution $u(t;y)$, defined for $t \in [-\dfrac{r}{2(r+x_0)^2}, \dfrac{r}{2(r+x_0)^2}]$. As before, the maximal interval occurs when $r = x_0$ and we have $t \in [-\dfrac{1}{8x_0}, \dfrac{1}{8x_0}]$, which you might notice is an interval of *half* of the length of the interval of definition before.

Notice that the true solution for an initial condition $y \in \overline{B_{\frac{r}{2}}(x_0)}$ is given as

$$u(t;y) = \frac{y}{1 - ty}$$

which is defined for $t \in (-\infty, \frac{1}{y}) \supset (-\frac{1}{y}, \frac{1}{y})$ so the true definition of existence is at least 8 times longer than what Theorem 7.2 guarantees. One last observation is that the previous analysis holds even if we allow $x_0 \to 0$. In that instance (i.e. when $x_0 = 0$), the solution to the IVP is $x(t) = 0$, and subsequently is defined on all of $\mathbb{R}$. We note that the intervals of existence of solutions $u(t;y)$ tend to all of $\mathbb{R}$ as $y \to 0$.

 ©University of Sydney

Now that we know that we have a family of solutions $u(t;y)$ we would like to know how well behaved the solutions are with respect to the y variable (depending of course on how well behaved $f(x)$ is). This is in general kind of a subtle question, but luckily enough we have a few straightforward relationships. For example if $f(x)$ is Lipschitz, then $u(t;y)$ will be a Lipschitz function of y.

Theorem 7.3 (Lipschitz dependence on initial conditions)**.** *Let $x_0 \in \mathbb{R}^n$ and suppose there is an r such that $f : \overline{B_r(x_0)} \to \mathbb{R}^n$ is Lipschitz with constant K, and that $J = [-a, a]$ is the common interval of existence of solutions $u : J \times B_{\frac{r}{2}}(x_0) \to B_r(x_0)$. Then $u(t;y)$ the family of solutions to the IVPs $(*)$ is Lipschitz in y for all $t \in J$ with Lipschitz constant e^{aK}.*

Proof. The proof is a straightforward application of the following useful lemma, called Grönwall's lemma, and is on your tutorial sheet for the week. □

Lemma 7.1 (Grönwall's lemma)**.** *Suppose that $g, k : [0, a] \to \mathbb{R}$ are continuous functions $a > 0$, $k(t) \geq 0$, and suppose that $g(t)$ obeys the functional inequality*

$$g(t) \leq G(t) := c + \int_0^t k(s)g(s)ds$$

for all $t \in [0, a]$. Then for all $t \in [0, a]$,

$$g(t) \leq ce^{\int_0^t k(s)ds}.$$

Proof. First, since $g(t)$ and $k(t)$ are continuous, we have that $G(0) = c$, and that $G(t)$ is C^1. Differentiating the definition of $G(t)$ we have that

$$\dot{G}(t) = k(t)g(t) \leq k(t)G(t) \Rightarrow \dot{G}(t) - k(t)G(t) \leq 0$$

Multiplying through by the integrating factor $e^{-\int_0^t k(s)ds}$ gives

$$\frac{d}{dt}\left(G(t)e^{-\int_0^t k(s)ds}\right) \leq 0$$

Integrating gives

$$G(t)e^{-\int_0^t k(s)ds} \leq G(0) = c$$

and multiplying both sides by $e^{\int_0^t k(s)ds}$ gives the result. □

Alternatively, if $g(t)$ were differentiable on $[0, a]$, then Grönwall's lemma says that if $g(t)$ obeys the differential *inequality* $g'(t) \leq k(t)g(t)$ it is bounded by the solution to the differential equation. In some sense this is a stop-gap measure. It would be nice to say that if $g(t) < G(t)$ and they were both differentiable, then $g'(t) < G'(t)$. This is however not in general true, (though it is true for integration, if $f(x)$ and $g(x)$ are both integrable and $f(x) \leq g(x)$ for all $x \in [a, b]$ then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$). But, we have Grönwall's lemma as the sort of best equivalent true statement.

We also have that if $f(x)$ is C^1, continuously differentiable then $u(t;y)$ is as well. The proof requires a bit more work, and will take us a bit far afield, so I am omitting it. But the theorem is useful.

Theorem 7.4. *Suppose $f : E \to \mathbb{R}^n$ is a C^1 function on some open set E. Then there is an $a > 0$ such that the solution $u(t;y)$ of $(*)$ with initial condition $x_0 = y$ is a C^1 function of y for $t \in J = [-a, a]$.*

 ©University of Sydney

Often we have extra parameters in our ODE's whose exact values may not be precisely known. In this case we can think of our solutions $u = u(t; y, \mu)$ where μ is the parameter. For example if you are considering modelling a simple pendulum, you would use the nonlinear ODE

$$\ddot{\theta} + \frac{g}{l}\sin(\theta) = 0$$

where θ is the displacement of the pendulum, g is the acceleration due to gravity (not likely to change unless you're leaving the earth as you watch your pendulum) and l is the length of the pendulum. Now for any fixed l you can (in theory anyway) write down a solution to this ODE (or initial value problem) $u(t; \theta_0, l)$ and as you vary l, it would be nice to know that your solution will vary nicely. It turns out that since $\frac{g}{l}\sin(\theta)$ is 'nice' for most (i.e. nonzero) values of l, then the solutions will be 'nice' too. In general if you have a system of ODEs

$$\dot{x} = f(x; \mu)$$

which is dependent on parameters, and $f(x; \mu)$ is nice enough with regards to μ at some μ_0, then you're in business.

Theorem 7.5 (Continuous Dependence on Parameters)**.** *Suppose your function $f(x; \mu)$ is Lipschitz on $\overline{B_r(x_0)}$ and is C^1 in your parameters in some closed ball $\overline{B_b(\mu_0)}$ of some radius b around an initial parameter μ_0. Then the family of IVPs*

$$\dot{x} = f(x; \mu) \qquad x(0) = y$$

will have unique solution $u(t; y, \mu)$ for each $y \in \overline{B_{\frac{r}{2}}(x_0)}$ on some interval J. Moreover $u(t; y, \mu)$ will be a C^1 function of μ.

For the time being, don't get too caught up on the interval of existence J for these sections. Typically in examples, this interval makes itself quite clear. Also we have some theorems which allow us to extend our interval of existence a bit further than what is guaranteed in the theorems above.

8. Maximal Intervals of Existence

Definition 8.1. The *maximal interval of existence* $J(t_0, x_0)$ is the largest interval of time that includes t_0 for which the solution to the initial value problem $(*)$ exists.

If the solution can be explicitly found, then we can just compute the maximal interval. Otherwise we can iteratively find it by repeated applications of Theorem 6.1. It turns out that the maximal interval of existence is open.

Theorem 8.1 (Maximal Interval of Existence)**.** *Let $E \subseteq \mathbb{R}^n$ be an open set and $f : E \to \mathbb{R}^n$ be locally Lipschitz. Then there is a maximal, open interval $J = (\alpha, \beta)$ containing t_0 such that the initial value problem*

$$\dot{x} = f(x) \quad x(t_0) = x_0$$

has a unique solution on $x(t) : J \to E$.

Proof. Let's denote by $u(t; t_0, x_0)$, the solution to the initial value problem

$$\dot{x} = f(x) \quad x(t_0) = x_0.$$

We know that for each closed ball $\overline{B_{r_0}(x_0)}$ there is a solution on the interval $J_0 := [t_0 - a_0, t_0 + a_0]$. Indeed the theorem implies that $u(t; t_0, x_0) \in \overline{B_{r_0}(x_0)} \subset E$ and is C^1. Thus we have that $\lim\limits_{t \to a_0} u(t; t_0, x_0) := x_1 \in \overline{B_{r_0}(x_0)} \subset E$. Now as E is open, and as f is locally

Lipschitz on E, we can find another r_1 such that $u(t; t_0 + a_0, x_1)$ i.e. the solution to the initial value problem

$$\dot{x} = f(x) \quad x(t_1) = x_1 \quad \text{where } t_1 = t_0 + a_0 \text{ and } x_1 = u(a_0; t_0, x_0).$$

is unique and maps $J_1 := [t_1 - a_1, t_1 + a_1] \to \overline{B_{r_1}(x_1)} \subset E$. Note that the intersection of J_0 and J_1 is nonempty, and by uniqueness of solutions we have that $u(t; t_0, x_0) = u(t; t_1, x_1)$ on their common interval of definition $J_0 \cap J_1$. From here (see fig. 16), it is just lather, rinse, repeat to extend the solution to obtain a unique solution on larger and larger intervals. Let J be the union of all such intervals, and let $x(t)$ be the unique solution constructed on J. Finally, J must be open, because if it were not, say $J = (\alpha, \beta]$, then we could play the same game again: extend $x(t)$ to β, and note that $x(\beta) \in \overline{B_{r_\beta}(x(\beta))} \subset E$, and we could extend $x(t)$ to a larger interval. $\square$

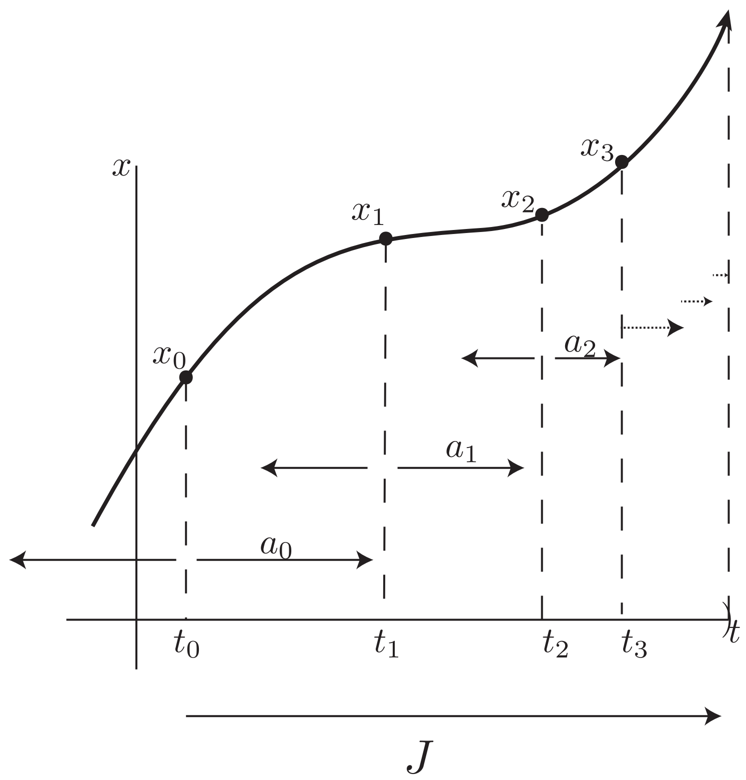

FIGURE 16. The maximal interval of existence is constructed by repeated applications of Theorem 6.1

Example 8.1. This example illustrates how to practically use the ideas in Theorem 8.1 to extend the maximal interval of existence. Consider the IVP

$$\dot{x} = x^3 \quad x(0) = 1.$$

Now Theorem 6.1 guarantees that we have a solution $x(t)$ defined on $[-a_0, a_0]$ where $a = \frac{r}{M}$ where $M = \max_{|x-1| \le r} x^3 = (1 + r)^3$. So we are trying to maximise the function $a = \frac{r}{(1+r)^3}$ where $r > 0$. This has a maximum of $a_0 = \frac{4}{27}$ at $r = \frac{1}{2}$. So we have a solution $u_0(t; 0, 1)$ defined on $[-a_0, a_0] = [-\frac{4}{27}, \frac{4}{27}]$. Now as in the proof of Theorem 8.1, we look at a new IVP

$$\dot{x} = x^3 \quad x(a_0) = u_0(a_0; 0, 1) \equiv x_1.$$

Ordinarily, we would need to numerically approximate $u_0(a_0; 0, 1)$ using MATLAB or some other ODE evolver (you could even build one yourself if you didn't trust MATLAB....) but for the purposes of expediency, here we can exploit the fact that we actually know the solution outright

$$f(t) = \frac{1}{\sqrt{1-2t}}.$$

I might as well point out here that this will give us directly the maximal interval of existence $(-\infty, \frac{1}{2})$, however, the point of this example is to illustrate the algorithm that is described in Theorem 8.1. We now have that we are trying to solve the IVP

$$\dot{x} = x^3 \quad x(a_0) = u_0(a_0; 0, 1) = x_1 = f(\frac{4}{27}) = 3\sqrt{\frac{3}{19}}$$

and we know from Theorem 6.1 that we have a solution $u_1(t; \frac{4}{27}, 3\sqrt{\frac{3}{19}})$, which is defined on the interval $[t_1 - a_1, t_1 + a_1] = [\frac{4}{27} - a_1, \frac{4}{27} + a_1]$. Again, we want to maximise

$$a_1 = \frac{r}{M} = \frac{r}{(3\sqrt{\frac{3}{19}} + r)^3}$$

which happens at $r = \frac{3}{2}\sqrt{\frac{3}{19}}$, giving $a_1 = \frac{76}{729}$. Now it is just lather, rinse, repeat. We solve the new IVP

$$\dot{x} = x^3 \quad x(t_1 + a_1) = x_2 = f(0 + \frac{4}{27} + \frac{76}{729}) = \frac{27}{19}.$$

I iterated this process in Mathematica 9, using the following simple FOR loop

```
For[ii = 1, ii < capN+1, ii++,{x[ii] = f[Sum[a[k], {k, 0, ii - 1}]],
a[ii] = MaxValue[{r/(x[ii] + r)^3, r > 0}, r]}]
```

where 'capN' was set to 1, 3, 5, 10, 50, 100, and 500. After that it becomes pretty clear what the right endpoint of the maximal interval of existence is going to be. Below is a table of the values:

capN	value of $\sum_{j=0}^{\text{capN}} a_j$
1	$4/27 = 0.1481$
3	0.3258
5	0.4137
10	0.4851
50	0.5
100	0.5
500	0.5

So we can see pretty clearly that the right endpoint is going to stop at $\frac{1}{2}$. To get the left endpoint, we just need to change the initial condition to that of the left endpoint of the interval guaranteed by Theorem 6.1. In this case we get that the endpoint tends towards $-\infty$ as the following table shows:

capN	value of $\sum_{j=0}^{\text{capN}} -a_j$
1	-0.1481
3	- 0.5891
5	-1.3302
10	-6.1990
50	- 215865
100	-9.3×10^{10}

This pretty much wraps our study of the existence and uniqueness of solutions to ODEs as well as the second module of this course. We are now going to move on to the next module of the course: 2-D nonlinear systems of ODEs.

©University of Sydney