Contraction mapping principle

A point $x$ with $f(x) = x$ is called a fixed point for a map $f$.

Lemma 1 (Contraction mapping principle) If $f : M \to M$ is a strict contraction from a complete non-empty metric space to itself, then there exists a point $x \in M$ with $f(x) = x$.

Proof: Since $M$ is non-empty, there exists a point $x_0$ in $M$. Assume we have defined $x_1$ already, then define $x_{n+1} = f(x_n)$.

We claim that for every $n$ we have

$$d(x_{n+1}, x_n) \leq L^n d(x_1, x_0)$$

For $n = 0$ this is clear since $L^0 = 1$. Assume this inequality holds for some $n \geq 0$.

Note that by the contraction property we have

$$d(x_{n+2}, x_{n+1}) = d(f(x_{n+1}), f(x_n)) \leq Ld(x_{n+1}, x_n) \leq L^{n+1} d(x_1, x_0)$$

Now we claim that there is a constant $C$ such that for any $n, m$

$$d(x_n, x_{n+m}) \leq C(2L^n - L^{n+m})$$

This is clear for $m = 0$ as long as $C$ is at least $d(x_0, x_1)$. Assume we have proven this for some $m \geq 0$. Then

$$d(x_n, x_{n+m+1}) \leq d(x_n, x_{n+m}) + d(x_{n+m}, x_{n+m+1}) \leq C(2L^n - L^{n+m}) + L^{n+m}d(x_0, x_1)$$

$$= C(2L^n - L^{n+m+1}) + L^{n+m}(d(x_0, x_1) - C(L^{-1} - 1))$$

The last term is negative as long as $C$ is larger than a number determined by $L < 1$.

This completes the induction.

Now since $f$ is a contraction we have

$$d(f(x), x) \leq d(f(x), f(x_n)) + d(f(x_n), x_n) + d(x_n, x) \leq \epsilon L + 2CL^n + \epsilon$$

Where $\epsilon$ and $L^n$ can both be made arbitrarily small by choosing $n$ large enough. Hence $d(f(x), x) = 0$ and hence $f(x) = x$.  

\qed
Lemma 2 The fix point guaranteed to exists in the above lemma is in unique.

Proof: assume that there are two fixed points $x$ and $y$. Then $d(f(x), f(y)) = d(x, y)$ because both points are fixed points, but als $d(f(x), f(y)) \leq Ld(x, y)$ because of the contraction property. Since $L < 1$, this is impossibly for $d(x, y) \neq 0$. Hence $d(x,y) = 0$ which implies $x = y$. □

The contraction mapping principle is a very ubiquitous method of choice to prove existence for ODE or PDE. We discuss an example

Lemma 3 Let $F$ be a Lipschitz map $\mathbb{R} \to \mathbb{R}$ and let $c \in \mathbb{R}$. Then there exists a solution to the integral equation

$$f(x) = c + \int_0^x F(f(t)) \, dt$$

in $C[0, \infty)$.

Proof. We first find small $\epsilon > 0$ and a solution to this equation in the metric space $C([0, \epsilon])$ with the uniform metric given by $\sup_{x \in [0, \epsilon]} |f(x) - g(x)|$.

Define a map on this space by

$$T(f)(x) = c + \int_0^x F(f(t)) \, dt$$

We claim that this is a contraction mapping with respect to the uniform metric. First of all if $f$ is continuous, then $F \circ f$ is continuous since it is the composition of continuous maps. The primitive of a continuous function is continuous. Hence $T$ maps $M$ to itself.

We have, using that $F$ is Lipschitz

$$d(T(f), T(g)) = \sup_x \left| \int_0^x F(f(t)) \, dt - \int_0^x F(g(t)) \, dt \right|$$

$$\leq d(T(f), T(g)) = \sup_x \int_0^x |F(f(t)) - F(g(t))| \, dt$$

$$\leq \sup_x \int_0^x |f(t) - g(t)| \, dt$$

$$\leq \sup_x \int_0^x Cd(f,g) \, dt$$

$$\epsilon Cd(f,g)$$

Hence this is a contraction mapping provided $\epsilon C < 1$. Hence choosing $\epsilon$ smaller than $1/C$ there is a a continuous function $f : [0, \epsilon] \to \mathbb{R}$ satisfying the equation.

We shall by induction prove that for every $n \geq 0$ there exists a continuous $f$ on $[0, n\epsilon,]$ solving the equation, the case $n = 0$ just having been established.
Assume this is shown for some \( n \geq 0 \). We shall seek an extension of \( f \) on \([0, n\epsilon]\) to \([0, (n+1)\epsilon]\). That extension needs to satisfy for \( x \geq n \)

\[
f(x) = c + \int_0^x F(f(t)) \, dt
\]

\[
f(x) = c + \int_0^{n\epsilon} F(f(t)) \, dt + \int_{n\epsilon}^x F(f(t)) \, dt
\]

This can be done by solving on \([0, \epsilon]\) the equation

\[
g(x) = c + \int_0^{n\epsilon} F(f(t)) \, dt + \int_0^x F(g(t)) \, dt
\]

using the previous arguments and then setting \( f(x) = g(x - n\epsilon) \).

Note that for \( c = 1 \) and \( F \) the identity map, this existence proof can be used to prove existence of the exponential function.

**Exercise 1 (Newton method)** Assume we have a function \( f : [0, 1] \to \mathbb{R} \) that satisfies for any \( x, y \in [0, 1] \) the Taylor expansion estimate

\[
f(y) - f(x) + f'(x)(y - x) \leq Cx^2
\]

for some constant \( C \)

Define the iteration

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

And prove that if \( x_0 \) is within \( \epsilon \) of a point \( x \) with \( f(x) = 0 \), then this iteration converges to \( x \).

This is called the Newton iteration method to find a zero.

**Curves**

We consider mappings \( f : [0, 1] \to M \) with \( M \) a complete metric space.

**Definition 1** An mapping \( f : [0, 1] \to M \) is called

1. geodesic, if it is an isometry or equivalent to an isometry, that is of a becomes an isometry when the distance function on the target space is multiplied by a nonzero constant.

2. a rectifiable curve, if it is Lipschitz

3. a curve if it is continuous (hence uniformly continuous)
Sometimes one restricts attention to injective mappings which are then called non-selfintersecting curves. We will however not need injectivity here.

We say two points \( x \) and \( y \) can be connected by a curve \( f \) if there is a curve with \( f(0) = x \) and \( f(1) = y \).

**Exercise 2** Prove that a subset of \([0, 1]\) which contains the points \( 0 \) and \( 1 \), these two points are connected by a curve if and only if it is the set \([0, 1]\) itself.

**Lemma 4** Let \( M \) be a metric space. Then the following two are equivalent:

1. For every \( x, y \in M \) there is a geodesic \( f : [0, 1] \to M \) such that \( f(0) = x \) and \( f(1) = y \).

2. For every \( x, y \in M \) there is a midpoint of \( x \) and \( y \), that is a \( z \in M \) such that \( d(x, z) = d(z, y) = d(x, y)/2 \).

Proof: If \( x, y \) can be connected by a geodesic curve, then we pick such a geodesic curve and define \( z = f(1/2) \). If \( \lambda \) is such that \( d(f(s), f(t)) = \lambda d(s, t) \) for this geodesic, then

\[
d(x, z) = \lambda/2 = d(x, z)/2
\]

and similarly for \( d(z, y) \).

Conversely, assume there is a midpoint for any two points in \( M \). Assume without loss of generality that \( d(x, y) = 1 \). Define a map \( f : [0, 1] \to M \) as follows: Set \( f(0) = x \) and \( f(1) = y \). Assume we have already defined \( f \) on \( D_n \cap [0, 1] \). Let \( z \in D_{n+1} \setminus D_n \) and let \( u \in D_n \) such that \( u \leq z \leq \nu_n(u) \). Then define \( f(z) \) to be the midpoint of \( f(u) \) and \( f(\nu_n(u)) \). By induction we see that \( d(t, t + \nu_m(t)) = 2^{-m} \) for \( x \in D_m \). Let \( s < t \) be any two points in \( D_m \), then we have by the triangle inequality

\[
1 = d(f(0), f(1)) \leq d(f(0), f(s)) + d(f(s), f(t)) + d(f(t), f(1))
\]

\[
\leq \sum_{r \in D_m \cap [0, 1], r < s} d(f(r), f(\nu_m(r)))+ \sum_{r \in D_m \cap [0, 1], s \leq r < t} d(f(r), f(\nu_m(r)))+ \sum_{t \in D_m \cap [0, 1]} d(f(t), f(\nu_m(t)))
\]

\[
= \sum_{r \in D_m \cap [0, 1], r < s} 2^{-m} \sum_{r \in D_m \cap [0, 1], s \leq r < t} 2^{-m} \sum_{t \in D_m \cap [0, 1]} 2^{-m}
\]

\[
= (s - 0) + (t - s) + (1 - t) = 1
\]

Since both sides of the inequality are 1, all intermediate inequalities in this argument have to be equalities and in particular \( d(f(s), f(t)) = t - s \). Hence \( f \) is an isometry and thus a geodesic. \( \square \)

**Definition 2** We define the length of a curve to be

\[
\sup_n \sum_{x \in D_n, 0 \leq x < 1} d(f(x), f(\nu_n(x)))
\]
Exercise 3 Let $f : [0, 1] \to M$ be a curve. Let $g : [0, 1] \cap D \to [0, 1]$ be a monotone curve with $g(0) = 0$ and $g(1) = 1$ (called a reparameterization). Then the length of $f$ equals the length of $f \circ g$.

Lemma 5 A curve has finite length if and only if there is a curve $g$ onto the range of $f$ which is Lipschitz.

Proof: Assume we have the map $g$ such that $h = f \circ g$ is rectifiable, with Lipschitz constant $L$. Then we have for every $n$

$$\sum_{x \in D_n, 0 \leq x < 1} d(h(x), h(\nu_n(x))) \leq \sum_{x \in D_n, 0 \leq x < 1} L2^{-n} \leq L$$

Conversely, assume the curve has finite length. Assume without loss of generality that this length is 1. Then we can define the arclength for every partial piece $[0, x]$ to be

$$\lambda(x) = \sup_n \sum_{t \in D_n, 0 \leq t < x} d(f(t), f(\nu_n(t)))$$

Then $\lambda$ is obviously a monotone function $[0, 1]$ to $[0, 1)$. We define a function $g[0, 1] \to [0, 1]$ by setting

$$g(\lambda(x)) = f(x)$$

To see that this is well defined, we need to verify that if $\lambda(x) = \lambda(y)$ then $f(x) = f(y)$. However, if $\lambda(x) = \lambda(y)$, then by definition of $\lambda$ in particular $d(f(x), f(y)) = 0$ because that term may appear in a sum that we take supremum over. Similarly we see that $g$ is Lipschitz, since we have for $y > x$

$$\lambda(x) - \lambda(y) \geq d(f(x), f(y))$$

Clearly $g$ is onto the range of $f$. □

Definition 3 A curve $f : D \cap [0, 1] \to M$ is called of finite $r$-variation, if there exists a constant $C$ such that for any partition $x_0 < x_1 < \ldots < x_n$ of $[0, 1]$ (in $D_m$ say) we have

$$\left( \sum_{i=1}^{n} d(x_i, x_{i+1})^r \right)^{1/r} \leq C$$

Exercise 4 A curve $f$ is of bounded $r$ variation if and only if there is a curve $g : [0, 1]$ onto the range of $f$ with which is Hölder with exponent $1/r$.

Definition 4 Let $M$ be a complete metric space and $x \in M$. then the set of all $y$ such that there exists a curve $f$ with $f(0) = x$ and $f(1) = y$ is called the connected component $M_x$ of $x$.  

5
Definition 5 A metric space is called pathwise connected, if for any two points \( x, y \) there exists a curve \( f \) with \( f(0) = x \) and \( f(1) = y \). In other words, it is pathwise connected if \( M_x = M \) for every \( x \in M \).

Definition 6 A metric space is called locally pathwise connected, if for any point \( x \) there exists a \( r > 0 \) so that \( B_r(x) \subset M_x \).

Lemma 6 Assume \( M \) is locally pathwise connected. Then for every \( x \) the set \( M_x \) is open.

Proof: Let \( y \in M_x \). Since \( M \) is locally pathwise connected there exists \( r \) such that \( B_r(y) \subset M_y \). It suffices to show that \( B_r \subset M_x \).

Let \( z \in B_r(y) \). Since also \( z \in M_y \), there exist \( g : [0,1] \to M \) with \( g(0) = y \) and \( g(1) = z \). Since \( y \in M_x \) there exists \( f : [0,1] \to M \) with \( g(0) = y \) and \( g(1) = z \).

Now define \( h : [0,1] \to M \) as follows

1. If \( t \leq 1/2 \) then \( h(t) = f(2t) \)
2. If \( t \geq 1/2 \) then \( h(t) = f(2t - 1) \)

This clearly defines a curve connecting \( x \) with \( z \).

Lemma 7 Assume \( M \) is locally pathwise connected. Then for every \( x \) the set \( M \setminus M_x \) is open.

Proof: If \( y \in M \setminus M_x \), then also \( M_y \subset M \setminus M_x \). We prove this by contradiction, assume \( z \in M_y \cap M_x \). Then we can connect \( z \) both with \( x \) and \( y \), and then we can construct a curve from \( x \) to \( y \) as in the proof of the previous lemma. Since \( M_y \) is open for every \( y \in M \setminus M_x \), then \( M \setminus M_x \) is open. \( \Box \)

Definition 7 A metric space is called connected if it is not the union of two nonempty disjoint open sets.

Lemma 8 A locally pathwise connected and connected set is pathwise connected.

Proof: For every \( x \), since both \( M_x \) and \( M \setminus M_x \) are open, one of them has to be empty. Since \( x \in M_x \), we have \( M \setminus M_x = \emptyset \).

Exercise 5 Present and prove carefully an example for a connected metric space which is not pathwise connected.