Lecture 9
Metric spaces.
The contraction fixed point theorem.
The implicit function theorem.
The existence of solutions to differential equations.

Shlomo Sternberg
1. Metric spaces

2. Completeness and completion.

3. The contraction fixed point theorem.

4. Dependence on a parameter.

5. The Lipschitz inverse function theorem

6. The implicit function theorem.

7. The local existence theorem for solutions of differential equations.
Cartesian or direct product

Until now we have used the notion of metric quite informally. It is time for a formal definition. For any set $X$, we let $X \times X$ (called the Cartesian product of $X$ with itself) denote the set of all ordered pairs of elements of $X$. (More generally, if $X$ and $Y$ are sets, we let $X \times Y$ denote the set of all pairs $(x, y)$ with $x \in X$ and $y \in Y$, and is called the Cartesian product of $X$ with $Y$.)
A **metric** for a set $X$ is a function $d$ from $X \times X$ to the real numbers $\mathbb{R}$,

$$d : X \times X \to \mathbb{R}$$

such that for all $x, y, z \in X$

1. $d(x, y) = d(y, x)$
2. $d(x, z) \leq d(x, y) + d(y, z)$
3. $d(x, x) = 0$
4. If $d(x, y) = 0$ then $x = y$.

The inequality in 2) is known as the **triangle inequality** since if $X$ is the plane and $d$ the usual notion of distance, it says that the length of an edge of a triangle is at most the sum of the lengths of the two other edges. (In the plane, the inequality is strict unless the three points lie on a line.)
Condition 4):

If $d(x, y) = 0$ then $x = y$,

is in many ways inessential, and it is often convenient to drop it, especially for the purposes of some proofs. For example, we might want to consider the decimal expansions $0.49999\ldots$ and $0.50000\ldots$ as different, but as having zero distance from one another. Or we might want to “identify” these two decimal expansions as representing the same point.

A function $d$ which satisfies only conditions 1) - 3) is called a **pseudo-metric**.
A **metric space** is a pair \((X, d)\) where \(X\) is a set and \(d\) is a metric on \(X\). Almost always, when \(d\) is understood, we engage in the abuse of language and speak of “the metric space \(X\”).

Similarly for the notion of a **pseudo-metric space**.

In like fashion, we call \(d(x, y)\) the **distance** between \(x\) and \(y\), the function \(d\) being understood.
Open balls.

If $r$ is a positive number and $x \in X$, the (open) **ball of radius** $r$ about $x$ is defined to be the set of points at distance less than $r$ from $x$ and is denoted by $B_r(x)$. In symbols,

$$B_r(x) := \{ y \mid d(x, y) < r \}.$$
The intersection of two balls.

If \( r \) and \( s \) are positive real numbers and if \( x \) and \( z \) are points of a pseudometric space \( X \), it is possible that \( B_r(x) \cap B_s(z) = \emptyset \). This will certainly be the case if \( d(x, z) > r + s \) by virtue of the triangle inequality. Suppose that this intersection is not empty and that

\[ w \in B_r(x) \cap B_s(z). \]

If \( y \in X \) is such that \( d(y, w) < \min[r - d(x, w), s - d(z, w)] \) then the triangle inequality implies that \( y \in B_r(x) \cap B_s(z) \). Put another way, if we set \( t := \min[r - d(x, w), s - d(z, w)] \) then

\[ B_t(w) \subset B_r(x) \cap B_s(z). \]
Put still another way, this says that the intersection of two (open) balls is either empty or is a union of open balls. So if we call a set in $X$ **open** if either it is empty, or is a union of open balls, we conclude that the intersection of any finite number of open sets is open, as is the union of any number of open sets. In technical language, we say that the open balls form a base for a topology on $X$. 
Continuous maps.

A map $f : X \rightarrow Y$ from one pseudo-metric space to another is called \textbf{continuous} if the inverse image under $f$ of any open set in $Y$ is an open set in $X$. Since an open set is a union of balls, this amounts to the condition that the inverse image of an open ball in $Y$ is a union of open balls in $X$, or, to use the familiar $\epsilon, \delta$ language, that if $f(x) = y$ then for every $\epsilon > 0$ there exists a $\delta = \delta(x, \epsilon) > 0$ such that

$$f(B_\delta(x)) \subseteq B_\epsilon(y).$$

Notice that in this definition $\delta$ is allowed to depend both on $x$ and on $\epsilon$. The map is called \textbf{uniformly continuous} if we can choose the $\delta$ independently of $x$. 
Lipschitz maps.

An even stronger condition on a map from one pseudo-metric space to another is the **Lipschitz condition**. A map $f : X \to Y$ from a pseudo-metric space $(X, d_X)$ to a pseudo-metric space $(Y, d_Y)$ is called a **Lipschitz map** with **Lipschitz constant** $C$ if

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) \quad \forall x_1, x_2 \in X.$$ 

Clearly a Lipschitz map is uniformly continuous.
Example: the distance to a set is a Lipschitz map.

For example, suppose that $A$ is a fixed subset of a pseudo-metric space $X$. Define the function $d(A, \cdot)$ from $X$ to $\mathbb{R}$ by

$$d(A, x) := \inf \{d(x, w), \ w \in A\}.$$  

The triangle inequality says that

$$d(x, w) \leq d(x, y) + d(y, w)$$

for all $w$, in particular for $w \in A$, and hence taking lower bounds we conclude that

$$d(A, x) \leq d(x, y) + d(A, y).$$

or

$$d(A, x) - d(A, y) \leq d(x, y).$$
We know that
\[ d(A, x) - d(A, y) \leq d(x, y). \]
Reversing the roles of \( x \) and \( y \) then gives
\[ |d(A, x) - d(A, y)| \leq d(x, y). \]
Using the standard metric on the real numbers where the distance between \( a \) and \( b \) is \( |a - b| \) this last inequality says that \( d(A, \cdot) \) is a Lipschitz map from \( X \) to \( \mathbb{R} \) with \( C = 1 \).
\[ d(A, x) - d(A, y) \leq d(x, y). \]
Closed sets.

A **closed set** is defined to be a set whose complement is open. Since the inverse image of the complement of a set (under a map $f$) is the complement of the inverse image, we conclude that the inverse image of a closed set under a continuous map is again closed.
The closure of a set.

For example, the set consisting of a single point in $\mathbb{R}$ is closed. Since the map $d(A, \cdot)$ is continuous, we conclude that the set

$$\{x | d(A, x) = 0\}$$

consisting of all point at zero distance from $A$ is a closed set. It clearly is a closed set which contains $A$. Suppose that $S$ is some closed set containing $A$, and $y \notin S$. Then there is some $r > 0$ such that $B_r(y)$ is contained in the complement of $C$, which implies that $d(y, w) \geq r$ for all $w \in S$. Thus $\{x | d(A, x) = 0\} \subset S$. 
In short \( \{ x | d(A, x) = 0 \} \) is a closed set containing \( A \) which is contained in all closed sets containing \( A \). This is the definition of the **closure** of a set, which is denoted by \( \overline{A} \). We have proved that

\[
\overline{A} = \{ x | d(A, x) = 0 \}.
\]

In particular, the closure of the one point set \( \{ x \} \) consists of all points \( u \) such that \( d(u, x) = 0 \).
Now the relation $d(x, y) = 0$ is an equivalence relation, call it $R$. (Transitivity being a consequence of the triangle inequality.) This then divides the space $X$ into equivalence classes, where each equivalence class is of the form $\{x\}$, the closure of a one point set. If $u \in \{x\}$ and $v \in \{y\}$ then

$$d(u, v) \leq d(u, x) + d(x, y) + d(y, v) = d(x, y).$$

since $x \in \{u\}$ and $y \in \{v\}$ we obtain the reverse inequality, and so

$$d(u, v) = d(x, y).$$
The metric on the quotient.

In other words, we may define the distance function on the quotient space $X/R$, i.e. on the space of equivalence classes by

$$d(\{x\}, \{y\}) := d(u, v), \quad u \in \{x\}, v \in \{y\}$$

and this does not depend on the choice of $u$ and $v$. Axioms 1)-3) for a metric space continue to hold, but now

$$d(\{x\}, \{y\}) = 0 \Rightarrow \{x\} = \{y\}.$$  

In other words, $X/R$ is a metric space. Clearly the projection map $x \mapsto \{x\}$ is an isometry of $X$ onto $X/R$. (An isometry is a map which preserves distances.) In particular it is continuous. It is also open.
In short, we have provided a canonical way of passing (via an isometry) from a pseudo-metric space to a metric space by identifying points which are at zero distance from one another.
Dense subsets.

A subset $A$ of a pseudo-metric space $X$ is called *dense* if its closure is the whole space. From the above construction, the image $A/R$ of $A$ in the quotient space $X/R$ is again dense. We will use this fact in the next section in the following form:

*If $f : Y \to X$ is an isometry of $Y$ such that $f(Y)$ is a dense set of $X$, then $f$ descends to a map $F$ of $Y$ onto a dense set in the metric space $X/R$.***
Cauchy sequences.

The usual notion of convergence and Cauchy sequence go over unchanged to metric spaces or pseudo-metric spaces $Y$: A sequence $\{y_n\}$ is said to converge to the point $y$ if for every $\epsilon > 0$ there exists an $N = N(\epsilon)$ such that

$$d(y_n, y) < \epsilon \quad \forall \ n > N.$$
Cauchy sequences.

A sequence $\{y_n\}$ is said to be **Cauchy** if for any $\epsilon > 0$ there exists an $N = N(\epsilon)$ such that

$$d(y_n, y_m) < \epsilon \quad \forall \ m, n > N.$$

The triangle inequality implies that every convergent sequence is Cauchy. But not every Cauchy sequence is convergent. For example, we can have a sequence of rational numbers which converge to an irrational number, as in the approximation to the square root of 2.
So if we look at the set of rational numbers as a metric space $\mathbb{R}$ in its own right, not every Cauchy sequence of rational numbers converges in $\mathbb{R}$. We must “complete” the rational numbers to obtain $\mathbb{R}$, the set of real numbers. We want to discuss this phenomenon in general.
Complete metric spaces.

So we say that a (pseudo-)metric space is **complete** if every Cauchy sequence converges. The key result of this section is that we can always “complete” a metric or pseudo-metric space. More precisely, we claim that

Any metric (or pseudo-metric) space can be mapped by a one to one isometry onto a dense subset of a complete metric (or pseudo-metric) space.

By what we have already proved, it is enough to prove this for a pseudo-metric spaces.
Construction of the completion as a sequence space.

Let $X_{seq}$ denote the set of Cauchy sequences in $X$, and define the distance between the Cauchy sequences $\{x_n\}$ and $\{y_n\}$ to be

$$d(\{x_n\}, \{y_n\}) := \lim_{n \to \infty} d(x_n, y_n).$$

It is easy to check that $d$ defines a pseudo-metric on $X_{seq}$. Let $f : X \to X_{seq}$ be the map sending $x$ to the sequence all of whose elements are $x$;

$$f(x) = (x, x, x, x, \cdots).$$

It is clear that $f$ is one to one and is an isometry. The image is dense since by definition

$$\lim d(f(x_n), \{x_n\}) = 0.$$
Using Cantor’s argument.

Now since $f(X)$ is dense in $X_{seq}$, it suffices to show that any Cauchy sequence of points of the form $f(x_n)$ converges to a limit. But such a sequence converges to the element $\{x_n\}$. □
Normed vector spaces.

Of special interest are vector spaces which have a metric which is compatible with the vector space properties and which is complete: Let $V$ be a vector space over the real numbers. A norm is a real valued function $v \mapsto \|v\|$ on $V$ which satisfies

1. $\|v\| \geq 0$ and $> 0$ if $v \neq 0$,
2. $\|rv\| = |r|\|v\|$ for any real number $r$, and
3. $\|v + w\| \leq \|v\| + \|w\| \ \forall \ v, w \in V$.

Then $d(v, w) := \|v - w\|$ is a metric on $V$, which satisfies $d(v + u, w + u) = d(v, w)$ for all $v, w, u \in V$. The ball of radius $r$ about the origin is then the set of all $v$ such that $\|v\| < r$. A vector space equipped with a norm is called a normed vector space and if it is complete relative to the metric it is called a Banach space.

Shlomo Sternberg
Lecture 9 Metric spaces. The contraction fixed point theorem. The implicit function theorem. The existence of solutions to differential equations.
Let $X$ and $Y$ be metric spaces. Recall that a map $f : X \to Y$ is called a **Lipschitz map** or is said to be “Lipschitz continuous”, if there is a constant $C$ such that

$$d_Y(f(x_1), f(x_2)) \leq Cd_X(x_1, x_2), \quad \forall x_1, x_2 \in X.$$ 

If $f$ is a Lipschitz map, we may take the greatest lower bound of the set of all $C$ for which the previous inequality holds. The inequality will continue to hold for this value of $C$ which is known as the Lipschitz constant of $f$ and denoted by $\text{Lip}(f)$. 
Contractions.

A map $K : X \rightarrow Y$ is called a **contraction** if it is Lipschitz, and its Lipschitz constant satisfies $\text{Lip}(K) < 1$.

Suppose $K : X \rightarrow X$ is a contraction, and suppose that $Kx_1 = x_1$ and $Kx_2 = x_2$. Then

$$d(x_1, x_2) = d(Kx_1, Kx_2) \leq \text{Lip}(K)d(x_1, x_2)$$

which is only possible if $d(x_1, x_2) = 0$, i.e. $x_1 = x_2$. So a contraction can have at most one fixed point. The contraction fixed point theorem asserts that if the metric space $X$ is complete (and non-empty) then such a fixed point exists.
The contraction fixed point theorem.

**Theorem**

Let $X$ be a non-empty complete metric space and $K : X \to X$ a contraction. Then $K$ has a unique fixed point.
Proof of the contraction fixed point theorem.

Choose any point \( x_0 \in X \) and define

\[ x_n := K^n x_0 \]

so that

\[ x_{n+1} = Kx_n, \quad x_n = Kx_{n-1} \]

and therefore

\[ d(x_{n+1}, x_n) \leq Cd(x_n, x_{n-1}), \quad 0 \leq C < 1 \]

implying that

\[ d(x_{n+1}, x_n) \leq C^n d(x_1, x_0). \]
\[ d(x_{n+1}, x_n) \leq C^n d(x_1, x_0). \]

Thus for any \( m > n \) we have

\[ d(x_m, x_n) \leq \sum_{n}^{m-1} d(x_{i+1}, x_i) \leq (C^n + C^{n+1} + \cdots + C^{m-1}) d(x_1, x_0) \]

\[ \leq C^n \frac{d(x_1, x_0)}{1 - C}. \]

This says that the sequence \( \{x_n\} \) is Cauchy. Since \( X \) is complete, it must converge to a limit \( x \), and \( Kx = \lim Kx_n = \lim x_{n+1} = x \) so \( x \) is a fixed point. We already know that this fixed point is unique. \( \square \)
Local contractions.

We often encounter mappings which are contractions only near a particular point $p$. If $K$ does not move $p$ too much we can still conclude the existence of a fixed point, as in the following:

**Theorem**

Let $D$ be a closed ball of radius $r$ centered at a point $p$ in a complete metric space $X$, and suppose $K : D \rightarrow X$ is a contraction with Lipschitz constant $C < 1$. Suppose that

$$d(p, Kp) \leq (1 - C)r.$$

Then $K$ has a unique fixed point in $D$. 
Proof.

We simply check that $K : D \to D$ and then apply the preceding theorem with $X$ replaced by $D$: For any $x \in D$, we have

$$d(Kx, p) \leq d(Kx, Kp) + d(Kp, p) \leq Cd(x, p) + (1 - C)r \leq Cr + (1 - C)r = r \quad \square.$$
Another version.

**Theorem**

Let $B$ be an open ball or radius $r$ centered at $p$ in a complete metric space $X$ and let $K : B \to X$ be a contraction with Lipschitz constant $C < 1$. Suppose that

$$d(p, Kp) < (1 - C)r.$$ 

Then $K$ has a unique fixed point in $B$.

**Proof.**

Restrict $K$ to any slightly smaller closed ball centered at $p$ and apply the preceding theorem.
Estimating the distance to the fixed point.

**Corollary**

Let $K : X \rightarrow X$ be a contraction with Lipschitz constant $C$ of a complete metric space. Let $x$ be its (unique) fixed point. Then for any $y \in X$ we have

$$d(y, x) \leq \frac{d(y, Ky)}{1 - C}.$$  

**Proof.**

We may take $x_0 = y$ and follow the proof of Theorem 1. Alternatively, we may apply Prop. 10 to the closed ball of radius $d(y, Ky)/(1 - C)$ centered at $y$. Prop. 10 implies that the fixed point lies in the ball of radius $r$ centered at $y$. 

Shlomo Sternberg  
Lecture 9 Metric spaces. The contraction fixed point theorem. The implicit function theorem. The existence of solutions to differential equations.
Future applications of the Corollary.

The Corollary we just proved will be of use to us in proving continuous dependence on a parameter in the next section. Later, when we study iterative function systems for the construction of fractal images, the corollary becomes the “collage theorem”. We might call our corollary the “abstract collage theorem”.

Shlomo Sternberg
Suppose that the contraction “depends on a parameter $s$”. More precisely, suppose that $S$ is some other metric space and that

$$K : S \times X \to X$$

with

$$d_X(K(s, x_1), K(s, x_2)) \leq Cd_X(x_1, x_2), \quad 0 \leq C < 1, \quad \forall s \in S, \ x_1, x_2 \in X.$$  \hspace{1cm} (1)

(We are assuming that the $C$ in this inequality does not depend on $s$.)
Holding $s$ fixed.

If we hold $s \in S$ fixed, we get a contraction

$$K_s : X \to X, \quad K_s(x) := K(s, x).$$

This contraction has a unique fixed point, call it $p_s$. We thus obtain a map

$$S \to X, \quad s \mapsto p_s$$

sending each $s \in S$ into the fixed point of $K_s$. 
**Theorem**

Suppose that for each fixed \( x \in X \), the map \( s \mapsto K(s, x) \)

of \( S \rightarrow X \) is continuous. Then the map \( s \mapsto p_s \)

is continuous.
Proof.

Fix a $t \in S$ and an $\epsilon > 0$. We must find a $\delta > 0$ such that $d_X(p_s, p_t) < \epsilon$ if $d_S(s, t) < \delta$. Our continuity assumption says that we can find a $\delta > 0$ such that

$$d_X(K(s, p_t), p_t) = d_X(K(s, p_t), K(t, p_t)) \leq (1 - C)\epsilon$$

if $d_S(s, t) < \delta$. This says that $K_s$ moves $p_t$ a distance at most $(1 - C)\epsilon$. But then the "abstract collage theorem", Prop. 4, says that

$$d_X(p_t, p_s) \leq \epsilon.$$
Combining previous results.

It is useful to combine two of our preceding results into a single theorem:

**Theorem**

Let $B$ be an open ball of radius $r$ centered at a point $q$ in a complete metric space. Suppose that $K : S \times B \to X$ (where $S$ is some other metric space) is continuous, satisfies

$$d_X(K(s, x_1), K(s, x_2)) \leq Cd_X(x_1, x_2), \quad 0 \leq C < 1, \quad \forall s \in S,$$

and

$$d_X(K(s, q), q) < (1 - C)r, \quad \forall s \in S.$$

Then for each $s \in S$ there is a unique $p_s \in B$ such that $K(s, p_s) = p_s$, and the map $s \mapsto p_s$ is continuous.
The inverse function theorem

Consider a map $F : B_r(0) \to E$ where $B_r(0)$ is the open ball of radius $r$ about the origin in a Banach space, $E$, and where $F(0) = 0$. Under suitable conditions on $F$, wish to conclude the existence of an inverse to $F$, defined on a possible smaller ball by means of the contraction fixed point theorem.

For example, suppose that $F$ is continuously differentiable with derivative $A$ at the origin which is invertible. Replacing $F$ by $A^{-1}F$ we may assume that the derivative of $F$ at 0 is the identity map, id.

So $F - \text{id}$ vanishes at the origin together with its derivative. Hence the mean value theorem implies that we can arrange that $F - \text{id}$ has Lipschitz constant as small as we like. This justifies the hypothesis of the following theorem:
Theorem

Let \( F : B_r(0) \to E \) satisfy \( F(0) = 0 \) and

\[
\text{Lip}[F - \text{id}] = \lambda < 1. \tag{2}
\]

Then the ball \( B_s(0) \) is contained in the image of \( F \) where

\[
s = (1 - \lambda)r \tag{3}
\]

and \( F \) has an inverse, \( G \) defined on \( B_s(0) \) with

\[
\text{Lip}[G - \text{id}] \leq \frac{\lambda}{1 - \lambda}. \tag{4}
\]
Proof.

Let us set $F = \text{id} + v$ so

$$\text{id} + v : B_r(0) \to E, \quad v(0) = 0, \quad \text{Lip}[v] < \lambda < 1.$$ 

We want to find a $w : B_s(0) \to E$ with

$$w(0) = 0$$

and

$$(\text{id} + v) \circ (\text{id} + w) = \text{id}.$$ 

This equation is the same as

$$w = -v \circ (\text{id} + w).$$
Let $X$ be the space of continuous maps of $\overline{B_s(0)} \to E$ satisfying

$$u(0) = 0$$

and

$$\text{Lip}[u] \leq \frac{\lambda}{1 - \lambda}.$$ 

Then $X$ is a complete metric space relative to the sup norm, and, for $x \in \overline{B_s(0)}$ and $u \in X$ we have

$$\|u(x)\| = \|u(x) - u(0)\| \leq \frac{\lambda}{1 - \lambda} \|x\| \leq r.$$ 

Thus, if $u \in X$ then

$$u : \overline{B_s} \to \overline{B_r}.$$ 

If $w_1, w_2 \in X$,

$$\| - \nu \circ (\text{id} + w_1) + \nu \circ (\text{id} + w_2) \| \leq \lambda \| (\text{id} + w_1) - (\text{id} + w_2) \| = \lambda \| w_1 - w_2 \|.$$
If \( w_1, w_2 \in X \),

\[
\| -v \circ (\text{id} + w_1) + v \circ (\text{id} + w_2) \| \leq \lambda \| w_1 - w_2 \| .
\]

So the map \( K : X \to X \)

\[
K(u) = -v \circ (\text{id} + u)
\]

is a contraction. Hence there is a unique fixed point. This proves the inverse function theorem.
The setup.

We want to solve the equation $F(x, y) = 0$ for $y$ as a function of $x$. In other words, we are looking for a function $y = g(x)$ such that $F(x, g(x)) \equiv 0$.

Here $x$ and $y$ are vector variables, say $x$ ranges over an open ball $A$ in a Banach space $X$ and $y$ ranges over and open ball $B$ in some other Banach space $Y$.

Here $F : A \times B \to Z$ where $Z$ is some third vector space.

To keep the notation simple, we will assume that $A$ and $B$ are open balls about the origin(s) and that

$$F(0, 0) = 0.$$
The assumptions about $F$.

- $F$ is continuous as a function of $(x, y)$.
- $\frac{\partial F}{\partial y}$ exists and is continuous as a function of $(x, y)$. Remember that $\frac{\partial F}{\partial y}$ is a linear transformation. For example, it is a matrix if $B$ is some ball in $\mathbb{R}^n$.
- $\frac{\partial F}{\partial y}(0, 0)$ is invertible.

We set $T := \frac{\partial F}{\partial y}(0, 0)$ and define

$$K(x, y) := y - T^{-1}F(x, y).$$

So $K$ is a continuous map from $A \times B \to Y$ and

$$K(x, y)y = y \iff F(x, y) = 0.$$
\[ K(x, y) := y - T^{-1}F(x, y). \]

\[ K(x, y)y = y \iff F(x, y) = 0. \]

\[ K \text{ is a continuous map from } A \times B \to Y \text{ and } \frac{\partial K}{\partial y} \text{ is a continuous } \]

\[ Y \text{-valued function of } (x, y) \text{ with } \]

\[ \left\| \frac{\partial K}{\partial y}(0, 0) \right\| = 0. \]

So by choosing smaller balls (which we now rename as \( A \) and \( B \))
we can arrange that

\[ \left\| \frac{\partial K}{\partial y}(x, y) \right\| < \frac{1}{2} \]

for all \( (x, y) \in A \times B. \)
\[ \left\| \frac{\partial K}{\partial y}(x, y) \right\| < \frac{1}{2} \]

for all \((x, y) \in A \times B\). The mean value theorem implies that

*K is a contraction in its second variable with Lipschitz constant \(\frac{1}{2}\).*

Let \(r\) denote the radius of the ball \(B\). Since \(K(x, y) - y\) is a continuous function of \(x\) and \(K(0, 0) - 0 = 0\), we can shrink the ball \(A\) still further (and reuse \(A\) as the name of the shrunken ball) so that

\[ \|K(x, y) - y\| < \frac{r}{2} \]

for all \((x, y) \in A \times B\).
Recalling a version of the contraction fixed point theorem with a slight change in notation suitable for our case.

<table>
<thead>
<tr>
<th>Outline Metric spaces</th>
<th>Completeness and completion. The contraction fixed point theorem. Dependence on a parameter. The Linear function theorem. The implicit function theorem. The local existence theorem for solutions of differential equations.</th>
</tr>
</thead>
</table>

Suppose that $K : A \times B \to Y$ is continuous, satisfies

$$\|K(x, y_1) - K(x, y_2)\| \leq \|y_1 - y_2\| C, \quad 0 \leq C < 1, \quad \forall s \in S,$$

and

$$\|K(x, y) - y\| < (1 - C)r, \quad \forall x \in A.$$ 

Then for each $x \in A$ there is a unique $y_x \in B$ such that $K(x, y) = y_x$, and the map $x \mapsto y_x$ is continuous.

We have verified the hypotheses of the theorem with $C = \frac{1}{2}$. So we have proved:
The implicit function theorem.

Theorem

Let \((x, y) \mapsto F(x, y) \in Z\) be a continuous map defined on an open set of \(X \times Y\) where \(X, Y,\) and \(Z\) are Banach spaces and such that \(\frac{\partial F}{\partial y}\) is continuous. Suppose that \(F(x_0, y_0) = 0\) at some point \((x_0, y_0)\) and \(\frac{\partial F}{\partial y}(x_0, y_0)\) is invertible. Then there are open balls \(A\) and \(B\) about \(x_0\) and \(y_0\) such that for each \(x \in A\) there is a unique \(y = g(x) \in B\) such that \(F(x, y) = 0\). The map \(g\) so defined is continuous.

By the arguments we gave in Lecture 2, we know that if \(F\) is also differentiable in both variables then \(g\) is differentiable near \(x_0\) and we know how to compute its derivative.
The set up.

$A$ is an open subset of a Banach space $X$, $I \subset \mathbb{R}$ is an open interval and

$$F : I \times A \to X$$

is continuous. We want to study the differential equation

$$\frac{dx}{dt} = F(t, x).$$

A solution of this equation is a map $f : J \to A$, where $J$ is an open subinterval of $I$ such that $f'(t)$ exists for all $t \in J$ and

$$f'(t) = F(t, f(t)).$$
A solution of this equation is a map $f : J \rightarrow A$, where $J$ is an open subinterval of $I$ such that $f'(t)$ exists for all $t \in J$ and

$$f'(t) = F(t, f(t)).$$

If $f'$ exists then $f$ must be continuous, and the the right hand side of the above equation is then continuous. So any solution must be continuously differentiable.
The function $F$ is uniformly Lipschitz in the second variable. That is, there is a constant $c$ independent of $t$ such that

$$\|F(t, x_1) - F(t, x_2)\| \leq c\|x_1 - x_2\| \quad \forall \ x_1, x_2 \in A.$$
The conclusion.

**Theorem**

For any \((t_0, x_0) \in I \times A\) there is a neighborhood \(U\) of \(x_0\) such that for any sufficiently small interval \(J\) containing \(t_0\) there is a unique map \(f: J \to U\) such that \(f\) is a solution to the differential equation and

\[ f(t_0) = x_0. \]
The idea of the proof.

If $f$ is a solution to our differential equation defined on the interval $J$, then

$$f(t) - f(t_0) = \int_{t_0}^{t} F(s, f(s))ds.$$  

If $f(t_0) = x_0$ then we get

$$f(t) = x_0 + \int_{t_0}^{t} F(s, f(s))ds.$$  

Conversely, if $f$ satisfies this last equation, then $f(t_0) = x_0$ and also $f$ is differentiable and is a solution to our differential equation.
The idea of the proof, continued.

So for any interval $J$ about $t_0$ let $\mathbb{B}(J)$ denote the space of **bounded, continuous** maps from $J$ to $X$ and try to define the map

$$K : \mathbb{B}(J) \rightarrow \mathbb{B}(J)$$

by

$$K(g)(t) = x_0 + \int_{t_0}^{t} F(s, g(s)) ds.$$  

So if we can arrange by suitable choice of $U$ that for small enough $J$ the map $K$ is defined and is a contraction, then the fixed point theorem gives us a unique solution to our differential equation with initial condition.
The proof.

Choose $U$ to be a ball of radius $r$ about $x_0$ and an interval $L$ about $t_0$ so that $F$ is bounded on $L \times U$ with bound $m$. Recall that $c$ is the Lipschitz constant of $F$ in its second variable. Let $x_0$ denote the constant function of $t$ with value $x_0$ i.e.

$$x_0(t) \equiv x_0.$$

Let $B_r$ be the ball of radius $r$ in $\mathbb{B}(J)$ about $x_0$ relative to the sup norm

$$\|g - h\|_\infty := \text{l.u.b.}_{t \in J} \|g(t) - h(t)\|.$$

So for any $g \in B_r$, $g(t) \in U$ for all $t \in J$ and so $F(t, g(t))$ is defined.

Let $\delta$ denote the length of $J$. 

Shlomo Sternberg
Lecture 9 Metric spaces. The contraction fixed point theorem. The implicit function theorem. The existence of solutions to differential equations.
The proof, continued.

If $g_1, g_2 \in B_r$ then $K(g_1)$ and $K(g_2)$ are defined, and for $t \in J$ we have

$$\|K(g_1)(t) - K(g_2)(t)\| = \left\| \int_{t_0}^{t} (F(s, g_1(s)) - F(s, g_2(s))) \, ds \right\| \leq c\delta \|g_1 - g_2\|_{\infty}.$$ 

Taking the least upper bound with respect to $t$ gives

$$\|Kg_1 - Kg_2\|_{\infty} \leq c\delta \|g_1 - g_2\|_{\infty}.$$ 

In other words, $K$ is Lipschitz with Lipschitz constant $C = c\delta$. We need to choose $\delta$ so that $C = c\delta < 1$ if we want $K$ to be a contraction.
The proof, continued.

Now let’s see how far $K$ moves the center, $x_0$ of $B_r$: We have

$$\|K(x_0)(t) - x_0(t)\| = \left\| \int_{t_0}^{t} F(s, x_0) ds \right\| \leq \delta m$$

so

$$\|K(x_0) - x_0\|_{\infty} \leq \delta m.$$
The proof, continued, recall a theorem proved earlier using a slight change in notation:

**Theorem**

Let $D$ be a closed ball of radius $r$ centered at a point $p$ in a complete metric space $Y$, and suppose $K : D \to X$ is a contraction with Lipschitz constant $C < 1$. Suppose that

$$d(p, Kp) \leq (1 - C)r.$$ 

Then $K$ has a unique fixed point in $D$.

So we want to choose $\delta$ small enough that

$$\delta m \leq (1 - C)r = (1 - \delta c)r.$$
The proof, concluded.

The condition \( \delta m \leq (1 - \delta c)r \) translates into \( \delta(m + cr) \leq 1 \) so if we choose

\[
\delta < \frac{r}{m + cr}
\]

then \( C = \delta c < 1 \) and \( \|K(x_0) - x_0\|_{\infty} < (1 - C)r \) so \( K \) satisfies the conditions of the theorem and there is a unique fixed point. \( \square \)