Metric Spaces

A metric space is a set $X$ with a distance $d: X \times X \to \mathbb{R}^+$ such that

$$d(x, z) \leq d(x, y) + d(y, z) \quad \text{triangle inequality}$$

$$d(x, y) = d(y, x)$$

$$d(x, y) = 0 \iff x = y$$

We'll always assume that $X$ is given the standard topology generated by open balls.

**def:** A sequence $x_0, x_1, x_2, \ldots$ in $X$ is a Cauchy sequence if, for any $\varepsilon > 0$, there is an $N$ such that $d(x_i, x_j) < \varepsilon$ for $i, j > N$.

**def:** A metric space is **complete** if every Cauchy sequence converges to a limit.

**def:** $X \xrightarrow{f} Y$ is uniformly continuous if it is continuous and if, given any $\varepsilon > 0$, there is a $\delta > 0$ such that $d(x, x') < \delta \Rightarrow d(f(x), f(x')) < \varepsilon$.

**def:** $X \xrightarrow{f} X$ is a contraction mapping if $d(f(x), f(x')) \leq d(x, x') \cdot k$ for some fixed $0 < k < 1$.

**Theorem:** A continuous function on a compact metric space is uniformly continuous. Proof: Homework.

**Theorem:** A contraction mapping has a unique fixed point. Proof: Homework.
Typical Applications

Riemann Integration $[a, b] \to \mathbb{R}$ continuous

"Partitions" $\Pi = r_0, r_1, r_2, \ldots, r_n \quad r_0 = a, \quad r_n = b \quad r_i \leq r_{i+1}$

Ordered by inclusion

$\rho(r_i, r_{i+1}) = \max_i (r_{i+1} - r_i) \quad \text{the radius of } \Pi = r_0, r_1, \ldots, r_n$

$I(r_0, r_1, \ldots) = \sum_{i=0}^{n-1} (r_{i+1} - r_i) f(r_i)$

$\Pi_0, \Pi_1, \Pi_2, \ldots \leftarrow \text{sequence with } \rho(\Pi_i) \to 0$

$I_{\Pi_0}, I_{\Pi_1}, I_{\Pi_2}, \ldots \text{ is a Cauchy sequence in } \mathbb{R} \to \text{has a limit }$

$\equiv \int_a^b f(x) \, dx$

Evolution Equation $\frac{df}{dt} = F(f(t)) \quad \text{Picard's theorem}$

$f_0(t) \equiv f(0)$

$f_1(t) \equiv f(0) + \int_0^t F(f_0(s)) \, ds$ \quad i

$f_{i+1}(t) \equiv f(0) + \int_0^t F(f_i(s)) \, ds$

$\Rightarrow \text{Cauchy sequence, showing that } f \text{ has a unique solution}$

with $f(0)$. Solving Laplace Equation by smoothing

Convergence of Markov Processes

$\ldots$ Typically very algorithm friendly.
As a warm-up, let's consider calculus in the light of what we have done in the course. What does
\[
    df_x(h) = \lim_{h \to 0} \frac{f(x + \lambda h) - f(x)}{\lambda}
\]
mean?

Setting: If \( X \xrightarrow{f} Y \), evidently \( X \) and \( Y \) must be both topological spaces and real finite dimensional vector spaces. Groth shows that if \( + : X \times X \to X \) and \( \cdot : \mathbb{R} \times X \to X \) are required to be continuous, then \( \mathbb{R}^n \) must have the standard topology.

Objects: \( \mathbb{R}^n \)  
monic, epic, iso as in top

Morphisms: Continuous functions

\[
\begin{align*}
    X & \leftarrow X \times Y \rightarrow Y \\
    f & \uparrow \quad \uparrow \\
    Z & \quad g \\
    \Rightarrow & \Rightarrow \quad \Rightarrow \\
    Z \to (f(z), g(z)) & \text{is continuous} \quad \Rightarrow \quad f \circ g : X \to Y \text{ is continuous}
\end{align*}
\]
iff \( f, g \) are both continuous

\( f : X \times Y \to Z \), even if \( f_x : y \to f(x, y) \) and \( f_y : x \to f(x, y) \) are continuous for all \( x, y \), it may be that \( f \) is not continuous.
What are limits?

\[
\lim_{x \to a} f(x) = y \iff \text{For any } \epsilon > 0, \text{ there is a } \delta > 0 \text{ s.t. } f([B_a^\delta]) \subseteq B_y^\epsilon.
\]

\[
\iff x \mapsto \begin{cases} f(x) & \text{if } x \neq a \\ y & \text{if } x = a \end{cases}
\]

This has the advantage that you don't need "\( \delta, \epsilon \)" arguments anymore. e.g.,

Limits are unique: Suppose \( \lim_{x \to a} f(x) = y \), \( \lim_{x \to a} f(x) = y' \)

\[
\begin{align*}
\{ f(x) & \text{ if } x \neq a \\
y & \text{ if } x = a \end{align*}
\]

\[
\begin{align*}
\{ f(x) & \text{ if } x \neq a \\
y' & \text{ if } x = a \end{align*}
\]

\[
\begin{cases} 0 & \text{if } x \neq a \\
y - y' & \text{if } x = a \end{cases}
\]

\( \Rightarrow y = y' \).

Proof. If \( y \neq y' \), since \( Y \) is Hausdorff, choose \( \delta \in Y \setminus \{ y, y' \} \), \( \delta \neq 0 \).

\( \Rightarrow f^{-1}(\delta) = \{ a \} \) is both open and closed (because \( X \) is Hausdorff).

\( \Rightarrow \).

Fact. If \( \lim_{x \to a} f(x) = y \), \( \lim_{x \to a} g(x) = z \), then \( \lim_{x \to a} (f + g)(x) = y + z \).

Proof:

\[
\begin{cases} f(x) & \text{if } x \neq a \\
y & \text{if } x = a \end{cases} + 
\begin{cases} g(x) & \text{if } x \neq a \\
z & \text{if } x = a \end{cases} = 
\begin{cases} f(x) + g(x) & \text{if } x \neq a \\
y + z & \text{if } x = a \end{cases}
\]

is continuous. \( \blacksquare \)
Fact: If $X \xrightarrow{f} Y \xrightarrow{g} Z$ are continuous and $\lim_{x \to a} f(x) = y$, then $\lim_{x \to a} g(f(x)) = g(\lim_{x \to a} f(x))$.

Proof. $g \circ \{ f(x) \text{ if } x \neq a \}
\left\{ \begin{array}{ll}
y & \text{if } x = a
\end{array} \right\} \xrightarrow{g \circ f} \left\{ g(y) \text{ if } x = a \right\}$ is continuous. $\Rightarrow \lim_{x \to a} g(f(x)) = g(y)$.

.. etc. Let the nice properties of continuous functions do all the work for you.

Back to $f : X \to Y$

$$d f_x(h) \equiv \lim_{\lambda \to 0} \frac{f(x + \lambda h) - f(x)}{\lambda}$$

is now well defined.

$\Leftrightarrow \lambda \mapsto F_x(\lambda, h)$ is continuous for all $h \in X$ where

$$F_x(\lambda, h) \equiv \left\{ \begin{array}{ll}
\frac{f(x + \lambda h) - f(x)}{\lambda} & \text{if } \lambda > 0 \\
d f_x(h) & \text{if } \lambda = 0
\end{array} \right.$$ 

(a) $d f_x(a \cdot h) = a \cdot d f_x(h)$.

Proof. $a \neq 0 \Rightarrow 
F_x(\lambda/a, a \cdot h) = \left\{ \begin{array}{ll}
\frac{f(x + \lambda h) - f(x)}{\lambda/a} & \text{if } \lambda > 0 \\
d f_x(a \cdot h) & \text{if } \lambda = 0
\end{array} \right.$

$\lambda \mapsto \frac{a}{\lambda} F_x(\lambda/a, a \cdot h)$ is continuous. $\Rightarrow$ by the uniqueness of limits, $d f_x(a \cdot h) = a \cdot d f_x(h)$. \(\blacksquare\)
(b) \( \Delta x (f + g) = \Delta f_x + \Delta g_x \)

**Proof.** \( \lambda \mapsto F_x (\lambda, h) + G_x (\lambda, h) \) is continuous where

\[
G_x (\lambda, h) = \begin{cases} 
\frac{g(x + \lambda h) - g(x)}{\lambda} & \text{if } \lambda > 0 \\
0 & \text{if } \lambda = 0
\end{cases}
\]

(c) If \( f : X \to \mathbb{R} \) has a minimum at \( m \in X \), \( df_m = 0 \).

**Proof:**

\[
\lambda \mapsto \begin{cases} 
\frac{f(m + \lambda h) - f(m)}{\lambda} & \text{if } \lambda > 0 \\
0 & \text{if } \lambda = 0
\end{cases}
\]

\( \Rightarrow df_m (h) \geq 0 \) for all \( h \in X \). \( \Rightarrow df_m = 0 \).

(d) The chain rule: Let \( X \xrightarrow{f} Y \xrightarrow{g} Z \),

\[
G_y (\lambda, k) = \begin{cases} 
\frac{g(y + \lambda k) - g(y)}{\lambda} & \text{if } \lambda > 0 \\
0 & \text{if } \lambda = 0
\end{cases}
\]

\( \lambda, h \mapsto \lambda, F_x (\lambda, h) \) \{ \text{continuous for any } x \in X, y \in Y \}

\( \lambda, k \mapsto \lambda, G_y (\lambda, k) \)

\( \Rightarrow \lambda, h \mapsto G_{f(x)} (\lambda, F_x (\lambda, h)) \) is continuous.

\( \Rightarrow \lambda \mapsto \begin{cases} 
\frac{g(f(x) + \lambda F_x (\lambda, h)) - g(f(x))}{\lambda} & \text{if } \lambda > 0 \\
0 & \text{if } \lambda = 0
\end{cases}
\]

\( \Rightarrow d (g \circ f)_x = d g_{f(x)} \circ df_x \). \( \Box \)
The chain rule is the thing that makes the differential into a functor. 

\[ (X, x) \xrightarrow{f} (Y, y) \xrightarrow{g} (Z, z) \]

\[ X \xrightarrow{d f_x} Y \xrightarrow{d g_y} Z \]

\[ d (g \circ f)_x \]

This is called the "tangent space functor" for reasons which will become apparent soon.

**Definitions:** Suppose \( \mathbb{R}^3 \xrightarrow{f} \mathbb{R} \) is differentiable at \( x \in \mathbb{R}^3 \).

Then \( df_x \in (\mathbb{R}^3)^* \). Rememeber that

\[
\begin{align*}
    dx &: (a, b, c) \mapsto a \\
    dy &: (a, b, c) \mapsto b \\
    dz &: (a, b, c) \mapsto c
\end{align*}

\]

is a basis of \( (\mathbb{R}^3)^* \)

\[
 df_x = \left( \frac{\partial f}{\partial x} \right)_x \, dx + \left( \frac{\partial f}{\partial y} \right)_x \, dy + \left( \frac{\partial f}{\partial z} \right)_x \, dz
\]

defines partial derivatives.
The Inverse function theorem:

Smooth $X \xrightarrow{f} Y$ is a local diffeomorphism at $x \in X$ iff $df_x$ is an isomorphism.

This is a "deep theorem" according to Guillemin & Pollack, and it certainly is crucial in the theory of manifolds. The proofs that I have seen are all pretty complicated, but at least let me do:

\[ \text{Lemma: If smooth } X \xrightarrow{f} Y \text{ has } df_x \text{ iso, then } f \text{ is locally a set isomorphism.} \]

\[ \text{Proof. Since } x \mapsto \det(df_x) \text{ is continuous (because } f \text{ is smooth and } \det \text{ is a polynomial), } df_x \text{ is iso on some open } U \ni x. \]

Suppose that $f$ fails to be 1-1 on every $B_x^\varepsilon \subset U \Rightarrow$ for each $B_x^\varepsilon$, $f(y) = f(z)$ for some $y, z \in B_x^\varepsilon$, $y \neq z$.

$\Rightarrow$ $f$ has a maximum $m$ in the line $z \rightarrow y$.

$\Rightarrow$ $df_m(y - z) = 0$

$\Rightarrow$ $df_m$ is not iso $\Rightarrow$ $f$ is a local bijective function.
**def:** A function $\mathbb{R}^n \to \mathbb{R}^n$ is **smooth** if it has continuous partial derivatives of all orders.

**def:** Same for $\mathcal{C}^0 \to \mathbb{R}^n$.

**def:** If $A$ is a subset of $\mathbb{R}^n$, $A \to \mathbb{R}^n$ is **smooth** if, for each $x \in A$, there is an open $\mathcal{O}_x \subset A$ and a smooth map $F: \mathcal{O}_x \to \mathbb{R}^n$ which coincides with $f$ on $A$.

**def:** A map $X \to Y$, $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ is a **diffeomorphism** if $f$ is a continuous isomorphism whose both $f$ and $f^{-1}$ are smooth.

**def:** A subset $M \subset \mathbb{R}^k$ is an $m$-dimensional **manifold** if each $x$ in $M$ has an open neighborhood $\mathcal{O}_x \subset M$ which is diffeomorphic to an open subset of $\mathbb{R}^m$.

**Aut(\mathbb{R}^n)\to\mathcal{C}(3)$

**not a manifold**

**Why?**
Example:

\[ S' = \{(x,y) : x^2 + y^2 = 1\} \subset \mathbb{R}^2 \]

\[
\begin{align*}
&y > 0: (x,y) \mapsto x, \quad x \mapsto (x, \sqrt{1 - x^2'}) \\
&y < 0: (x,y) \mapsto x, \quad x \mapsto (x, -\sqrt{1 - x^2'}) \\
&x > 0: (x,y) \mapsto y, \quad y \mapsto (\sqrt{1 - y^2}, y) \\
&x < 0: (x,y) \mapsto y, \quad y \mapsto (-\sqrt{1 - y^2}, y)
\end{align*}
\]

Some we'll have better ways of constructing manifolds.

Tangent Spaces

\[ M \subset \mathbb{R}^n \quad T_x M \equiv \text{Im}(d\phi_0) \]

This is independent of which diffeomorphism \( \psi \) you choose because

\[ \Rightarrow \text{Im} \, d\phi_0 = \text{Im} \, d\beta_0 \equiv T_x M \]
As usual, we have a category

**Objects:** Manifolds

**Morphisms:** Smooth functions \( M \to N \)

Given \( M \to N \), \( M \subset \mathbb{R}^k \), \( N \subset \mathbb{R}^l \) are manifolds, we wish to define

\[
"d_f : \text{T}M \to \text{T}N_y \quad y = f(x)."
\]

Let \( df_x = dF_x \mid_{\text{T}M} \) where \( F : W \to \mathbb{R}^l \) is some function that extends \( f \) into \( W \subset \mathbb{R}^k \).

\[
\begin{array}{ccc}
W & \xrightarrow{F} & \mathbb{R}^l \\
\uparrow{\alpha} & & \uparrow{\beta} \\
U & \xrightarrow{\varphi} & \mathbb{R}^m
\end{array}
\]

\[
\begin{array}{ccc}
\mathbb{R}^k & \xrightarrow{dF_x} & \mathbb{R}^l \\
\uparrow{d\alpha} & & \uparrow{d\beta} \\
\mathbb{R}^n & \xrightarrow{d\varphi} & \mathbb{R}^n
\end{array}
\]

\[
\Rightarrow \text{Im}(dF_x) \subset \text{Im}(d\beta) \]

\[
\Rightarrow dF_x \text{ is independent of the choice of extension of } f.
\]
The chain rule on manifolds

\[ M \xrightarrow{f} N \xrightarrow{g} P \]

\[ \mathcal{O}_x \subset M \subset \mathbb{R}^k \]
\[ \mathcal{O}_y \subset N \subset \mathbb{R}^l \]
\[ \mathcal{O}_z \subset P \subset \mathbb{R}^m \]

\[ \mathbb{R}^k \xrightarrow{\text{d}F_x} \mathbb{R}^l \xrightarrow{\text{d}G_y} \mathbb{R}^m \]

\[ \text{d}(G \circ F)_x \]

\[ TM_x \xrightarrow{\text{d}F_x} TN_y \xrightarrow{\text{d}G_y} TP_z \]

\[ \text{commutes.} \]
Thus, we have a tangent space function

\[
(M, x) \xrightarrow{f} (N, y) \xrightarrow{g} (P, z)
\]

\[
\downarrow \quad \downarrow \quad \downarrow
\]

\[
TM_x \quad \xrightarrow{df_x} \quad TN_y \quad \xrightarrow{dg_y} \quad TP_z
\]

which gives classification results.

\[\Rightarrow\] Manifolds with different dimensions cannot be isomorphic.

The inverse function theorem for manifolds:

\[
M \xrightarrow{f} N \quad df_x \text{ is} \Rightarrow \quad df_z \text{ is}\]

\[
x \uparrow \quad \uparrow \beta
\]

\[
\rightarrow \quad \Rightarrow f \text{ locally is} \Rightarrow
\]

\[
\emptyset \xrightarrow{f} U \quad \Rightarrow f \text{ locally is}\]

Regular Values

Given \( M \xrightarrow{f} N \),

\[x \in M \text{ is a regular point if } df_x \text{ is epi}\]

\[y \in N \text{ is a regular value if } f^{-1} [y] \text{ are all regular points}\]

\[y \in N \text{ is a critical value if } y \text{ is not regular}\]
Example:

\[ f: (x, y, z) \mapsto x^2 + y^2 - z^2 \quad \mathbb{R}^3 \to \mathbb{R} \]

\[ df = 2x \, dx + 2y \, dy - 2z \, dz \]

\( \Rightarrow 0 \in \mathbb{R} \) is the only critical value.

\( (0,0,0) \in \mathbb{R}^3 \) is the only non-regular point.

\[ f^{-1}(c) \quad c > 0 \]

\[ f^{-1}(c) \quad c < 0 \]

\[ f^{-1}(c) \] changes topology as \( c \) goes through its critical values. \( f^{-1}(c) \) is another manifold unless \( c = 0 \).

Is this generally true? ...
Yes. Suppose $M$ is an $m$-dimensional manifold, $N$ is $n$-dimensional, and smooth $M \xrightarrow{f} N$ has $df_x$ epic.

\[ TM_x \cong \ker(df_x) \oplus U \]

\[ \begin{array}{c}
\text{F : } x \mapsto (f(x), x) \in N \times \mathbb{R}^{n-m} \text{ has } df_x \text{ eso.}
\end{array} \]

By the inverse function theorem, $F$ is locally a diffeomorphism.

\[ \Rightarrow F : f^{-1}[y] \to y \times \mathbb{R}^{n-m} \text{ is a local diffeomorphism} \]

\[ \Rightarrow f^{-1}[y] \text{ is an } (n-m)\text{-dimensional submanifold of } M. \]

**Example:** $\rho : (x, y, z) \mapsto x^2 + y^2 + z^2$

\[ \Rightarrow S^2 \cong \rho^{-1}[1] \text{ is a manifold.} \]

**Example:** $\text{End}(\mathbb{R}^n)$, the set of $n \times n$ matrices is the manifold $\mathbb{R}^{n^2}$. The subgroup of isometries is given by $\Phi : L \to L^*L$

$\mathfrak{O}(n) = \Phi^{-1}[1]$ is a submanifold of $\text{End}(\mathbb{R}^n)$.

This is also a Lie Group.