MATH255 - Analysis 2

Basic point-set topology; metric spaces; Hölder-Minkowski Inequalities; compactness.

Based on lectures from Winter, 2024 by Prof. Dimitry Jakobson

Notes by Louis Meunier

Contents

1 Introduction

1.1 Metric Spaces

↪ **Definition 1.1: Metric Space**

A set X is a *metric space* with distance d if

1. (symmetric) $d(x,y) = d(y,x) \geqslant 0$
2. $d(x,y) = 0 \iff x = y$
3. (triangle inequality) $d(x,y) + d(y,z) \geqslant d(x,z)$

Remark 1.1. *If 1., 3. are satisfied but not 2., d can be called a "pseudo-distance".*

↪ **Definition 1.2: Open Metric Space**

Let (X,d) be a metric space. A subset $A \subseteq X$ is open $\iff \forall x \in A, \exists r = r(x) > 0$ s.t. $B(x, r(x)) \subseteq A$.

↪ **Definition 1.3: Normed Space**

Let X be a vector space over $\mathbb{R}$. The norm on X, denoted $||x|| \in \mathbb{R}$, is a function that satisfies

1. $||x|| \geqslant 0$
2. $||x|| = 0 \iff x = 0$
3. $||c \cdot x|| = |c| \cdot ||x||$
4. $||x + y|| \leqslant ||x|| + ||y||$

If X is a normed vector space over $\mathbb{R}$, we can define a distance d on X by $d(x,y) = ||x - y||$.

↪ **Proposition 1.1**

If X is a normed vector space over $\mathbb{R}$, a distance d on X by $d(x,y) = ||x-y||$ makes (X,d) a metric space.

Proof.

1. $d(x,y) = ||x-y|| \geqslant 0$
2. $d(x,y) = 0 \iff ||x-y|| = 0 \iff x - y = 0 \iff x = y$
3. $d(x,y) + d(y,z) = ||x-y|| + ||y-z|| \geqslant ||(x-y) + (y-z)|| = ||x-z|| := d(x,z)$

■

⊛ **Example 1.1: L^p distance in $\mathbb{R}^n$**

Let $\overline{x} \in \mathbb{R}^n, x = (x_1, x_2, \ldots, x_n)$. The L^p norm is defined

$$||x||_p := (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}.$$

In the case $p = 2, n = 2$, we simply have the standard Euclidean distance over $\mathbb{R}^2$.

Unit Balls: consider when $||x||_p \leqslant 1$, over $\mathbb{R}^2$.

- $p = 1 : |x_1| + |x_2| \leqslant 1$; this forms a "diamond ball" in the plane.
- $p = 2 : \sqrt{|x_1|^2 + |x_2|^2} \leqslant 1$; this forms a circle of radius 1. Clearly, this surrounds a larger area than in $p = 2$.

A natural question that follows is what happens as $p \to \infty$? Assuming $|x_1| \geqslant |x_2|$:

$$\begin{aligned} ||x||_p &= \left(|x_1|^p + |x_2|^p\right)^{\frac{1}{p}} \\ &= \left[|x_1|^p \left(1 + \left|\frac{x_2}{x_1}\right|^p\right)\right]^{\frac{1}{p}} \\ &= |x_1| \left(1 + \left|\frac{x_2}{x_1}\right|^p\right)^{\frac{1}{p}} \end{aligned}$$

If $|x_1| > |x_2|$, this goes to $|x_1|$. If they are instead equal, then $||x||_p = |x_1| \cdot 2^{\frac{1}{p}} \to |x_1| \cdot 1$ as well. Hence, $\lim_{p\to\infty} ||x||_p = \max\{|x_1|, |x_2|\}$. Thus, the unit ball will approach $\max\{|x_1|, |x_2|\} \leqslant 1$, that is, the unit square.

↪ **Proposition 1.2**

Let $x \in \mathbb{R}^n$. Then, $||x||_p \to \max\{|x_1|, \ldots, |x_n|\}$ as $p \to \infty$.

Remark 1.2. *This is an extension of the previous example to arbitrary real space; the proof follows nearly identically.*

↪ **Definition 1.4: Convex Set**

Let X be a normed space, and take $x, y \in X$. The line segment from x to y is the set

$$\{t \cdot x + (1 - t) \cdot y : 0 \leqslant t \leqslant 1\}.$$

Let $A \subseteq X$. A is *convex* $\iff \forall\, x, y \in A$, we have that

$$(t \cdot x + (1 - t) \cdot y) \in A \,\forall\, 0 \leqslant t \leqslant 1.$$

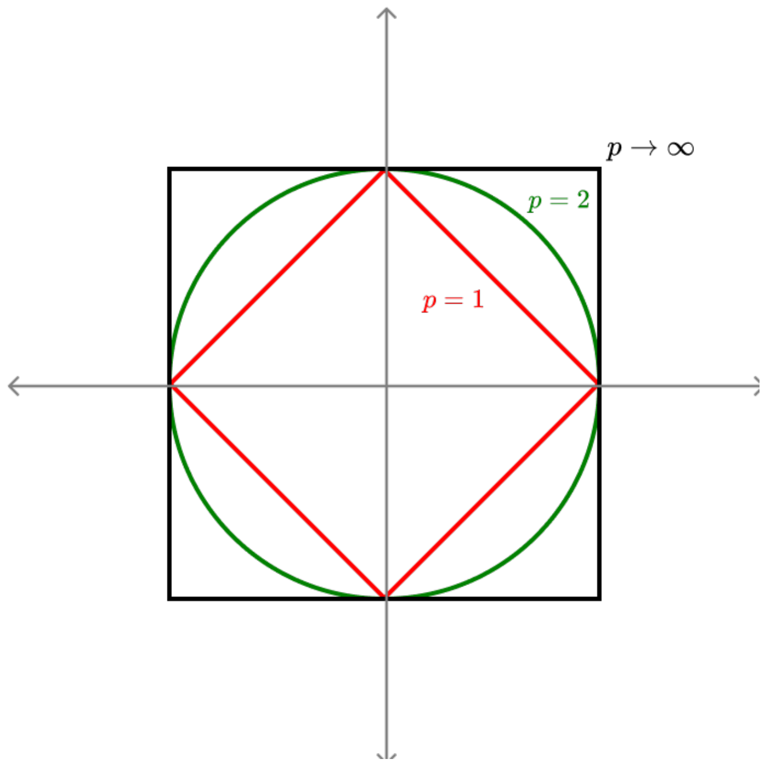

Figure 1: Regions of $\mathbb{R}^2$ where $||x||_p \leqslant 1$ for various values of p.

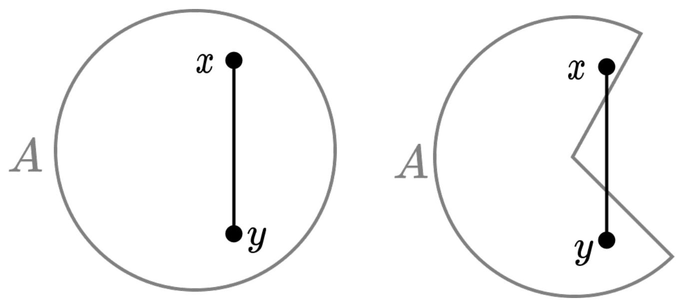

Figure 2: Convex (left) versus not convex (right) sets.

Remark 1.3. *Think of this as saying "a set is convex iff every point on a line segment connected any two points is in the set".*

$\hookrightarrow$ **Definition 1.5:** ℓ_p

The space ℓ_p of sequences is defined as

$$\{x = (x_1, x_2, \ldots, x_n, \ldots) : \sum_{n=1}^{\infty} |x_n|^p < +\infty\} \quad * .$$

Then, $*$ defines the ℓ^p norm on the space of sequences; that is, $||x||_p := \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}}$.

⊛ **Example 1.2:** ℓ_p, $x_n = \frac{1}{n}$

. Let $x_n = \frac{1}{n}$. For which p is $x \in \ell_p$? We have, raising the norm to the power of p for ease:

$$\begin{aligned} ||x||_p^p &= |x_1|^p + |x_2|^p + \cdots + |x_n|^p + \cdots \\ &= 1^p + \left(\frac{1}{2}\right)^p + \cdots < \infty \iff p > 1. \end{aligned}$$

In the case that $p = 1$, this becomes a harmonic sum, which diverges.

⊛ **Example 1.3:** L^p **space of functions**

Let $f(x)$ be a continuous function. We define the norm of f over an interval $[a, b]$

$$||f||_p = \left[\int_a^b |f(x)|^p \, dx\right]^{\frac{1}{p}}.$$

Remark 1.4. *Triangle inequality for $||x||_p$ or $||f||_p$ is called Minkowski inequality; $||x||_p + ||y||_p \geqslant ||x + y||_p$. This will be discussed further.*

⊛ **Example 1.4: Distances between sets in $\mathbb{R}^2$**

Let A, B be bounded, closed, "nice" sets in $\mathbb{R}^2$. We define

$$d(A, B) := \text{Area}(A \triangle B),$$

where

$$A \triangle B : (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

It can be shown that this is a "valid" distance.

Remark 1.5. *$\triangle$ denotes the "symmetric difference" of two sets.*

⊛ **Example 1.5:** p**-adic distance**

Let p be a prime number. Let $x = \frac{a}{b} \in \mathbb{Q}$, and write $x = p^k \cdot \left(\frac{c}{d}\right)$, where c, d are not divisible by p. Then, the p-adic norm is defined $||x||_p := p^{-k}$. It can be shown that this is a norm.

Suppose $p = 2$, $x = 28 = 4 \cdot 7 = 2^2 \cdot 7$. Then, $||28||_2 = 2^{-2} = \frac{1}{4}$; similarly, $||1024||_2 = ||2^{10}||_2 = 2^{-10}$.

More generally, we have that $||2^k||_2 = 2^{-k}$; coversely, $||2^{-k}|| = 2^k$. That is, the closer to 0, the larger the distance, and vice versa, contrary to our notion of Euclidean distance.

↪ **Proposition 1.3**

$||x||_p$ as defined above is a well-defined norm over $\mathbb{Q}$.

Proof. Left as a (homework) exercise. ■

2 Point-Set Topology

2.1 Definitions

↪ **Definition 2.1: Topological space**

A set X is a topological space if we have a collection of subsets τ of X called *open sets* s.t.

1. $\varnothing \in \tau, X \in \tau$
2. Consider $\{A_\alpha\}_{\alpha \in I}$ where A_α an open set for any α; then, $\bigcup_{\alpha \in I} A_\alpha \in \tau$, that is, it is also an open set.
3. If J is a finite set, and A_β open for all $\beta \in J$, then $\bigcap_{\beta \in J} A_\beta \in \tau$ is also open.

In other words, 2.: arbitrary unions of open sets are open, and 3.: finite intersections of open sets are open.

↪ **Definition 2.2: Closed sets**

Closed sets are complements of open sets; hence, axioms for closed sets follow appropriately;

1.* $X, \varnothing$ closed;

2.* B_α closed $\forall \alpha \in I \implies \bigcap_{\alpha \in I} B_\alpha$ closed.

3.* B_β closed $\forall \beta \in J$, J finite, then $\bigcup_{\beta \in J} B_\beta$ also closed.

↪ Lecture 01; Last Updated: Tue Feb 13 09:24:32 EST 2024

↪ **Definition 2.3: Equivalence of Metrics**

Suppose we have a metric space X with two distances d_1, d_2; will these necessarily admit the same topology?

A sufficient condition is that, if $\forall x \neq y \in X, \exists 1 < C < +\infty$ s.t.

$$\frac{1}{C} < \frac{d_1(x,y)}{d_2(x,y)} < C.$$

That is, the distances are equivalent, up to multiplication by a constant.

Indeed, this condition gives that $d_2 < Cd_1$ and $d_2 > \frac{d_1}{C}$; this gives

$$B_{d_1}(x, \frac{r}{c}) \subseteq B_{d_2}(x, r) \subseteq B_{d_1}(x, C \cdot r).$$

Hence, d_1, d_2 define the same open/closed sets on X thus admitting the same topologies. We write $d_1 \asymp d_2$.

Remark 2.1. *If $d_1 \asymp d_2$ and $d_2 \asymp d_3$, then also $d_1 \asymp d_3$. Moreover, clearly, $d_1 \asymp d_1$ and $d_1 \asymp d_2 \implies d_2 \asymp d_1$, hence this is a well-defined equivalence relation.*

Hence, its enough to show that $\forall 1 < p < +\infty$, we have $||x||_p \asymp ||x||_\infty$ to show that any $||x||_q$ norm are equivalent for all q on $\mathbb{R}^n$.

↪ **Definition 2.4: Interior, Boundary of a Topological Set**

Let X be a topological space, $A \subseteq X$ and let $x \in X$. We have the following possibilities

1. $\exists U$-open : $x \in U \subseteq A$. In this case, we say $x \in$ the *interior* of A, denoted

$$x \in \text{Int}(A).$$

2. $\exists V$-open : $x \in V \subseteq X \setminus A = A^C$. In this case, we write

$$x \in \text{Int}(X^C).$$

3. $\forall U$-open : $x \in U, U \cap A \neq \varnothing$ AND $U \cap A^C \neq \varnothing$. In this case, we say x is in the *boundary* of A, and denote

$$x \in \partial A.$$

↪ **Definition 2.5: Closure**

$x \in \text{Int}(A)$ or $x \in \partial A$ (that is, $x \in \text{Int}(A) \cup \partial A$) $\iff$ every open set U that contains x intersects A.[1]Such points are called *limit points* of A. The set of all limits points of A is called the *closure* of A, denoted $\overline{A}$.

Remark 2.2. *We have that*

$$\mathrm{Int}(A) \subseteq A \subseteq \overline{A} = \mathrm{Int}(A) \cup \partial A.$$

↪ **Proposition 2.1: Properties of** $\mathrm{Int}(A)$

$\mathrm{Int}(A)$ is *open*, and it is the largest open set contained in A. It is the union of all U-open s.t. $U \subseteq A$. Moreover, we have that

$$\mathrm{Int}(\mathrm{Int}(A)) = \mathrm{Int}(A).$$

↪ **Proposition 2.2: Properties of** $\overline{A}$

$\overline{A}$ is *closed*; $\overline{A}$ is the smallest closed set that contains A, that is, $\overline{A} = \bigcap B$ where B closed and $A \subseteq B$. We have too that

$$\overline{(\overline{A})} = \overline{A}.$$

↪ **Proposition 2.3**

1. A is open $\iff A = \mathrm{Int}(A)$
2. A is closed $\iff A = \overline{A}$

2.2 Basis

↪ **Definition 2.6: Basis for a Toplogy**

Let τ be a topology on X. Let $\mathcal{B} \subseteq \tau$ be a collection of open sets in X such that every open set is a union of open sets in $\mathcal{B}$.

⊛ **Example 2.1: Example Basis**

$X = \mathbb{R}$, and $\mathcal{B} = \{\text{all open intervals } (a, b) : -\infty < a < b < +\infty\}$.

↪ **Proposition 2.4**

Let $\mathcal{B}$ be a collection of open sets in X. Then, $\mathcal{B}$ is a basis $\iff$

1. $\forall x \in X, \exists U\text{-open} \in \mathcal{B}$ s.t. $x \in U$.
2. If $U_1 \in \mathcal{B}$ and $U_2 \in \mathcal{B}$, and $x \in U_1 \cap U_2$, then $\exists U_3 \in \mathcal{B}$ s.t. $x \in U_3 \subseteq U_1 \cap U_2$.

[1] "Requires" proof.

⊛ **Example 2.2**

Consider $X = \mathbb{R}$. Requirement 1. follows from taking $U = (x-\varepsilon, x+\varepsilon)$ for any $\varepsilon > 0$. For 2., suppose $x \in (a,b) \cap (c,d) =: U_1 \cap U_2$. Let $U_3 = (\max\{a,c\}, \min\{b,d\})$; then, we have that $U_3 \subseteq U_1 \cap U_2$, while clearly $x \in U_3$.

↪ **Proposition 2.5**

In a metric space, a basis for a topology is a collection of open balls,

$$\{B(x,r) : x \in X, r > 0\} = \{\{y \in X : d(x,y) < r\} : x \in X, r > 0\}.$$

Proof. We prove via proposition 2.4. Property 1. holds clearly; $x \in B(x,\varepsilon)$-open $\subseteq \mathcal{B}$.

For property 2., let $x \in B(y_1,r_1) \cap B(y_2,r_2)$, that is, $d(x,y_1) < r_1$ and $d(x,y_2) < r_2$. Let

$$\delta := \min\{r_1 - d(x,y_1), r_2 - d(x,y_2)\}.$$

We claim that $B(x,\delta) \subseteq U_1 \cap U_2$.

Let $z \in B(x,\delta)$. Then,

$$d(z,y_1) \overset{\triangle\neq}{\leqslant} d(z,x) + d(x,y_1) < \delta + d(x,y_1) \leqslant r_1 - d(x,y_1) + d(x,y_1) = r_1,$$

hence, as $d(z,y_1) < r_1 \implies z \in B(y_1,r_1) = U_1$. Replacing each occurrence of y_1, r_1 with y_2, r_2 respectively gives identically that $z \in B(y_2,r_2) = U_2$. Hence, we have that $B(x,\delta) \subseteq U_1 \cap U_2$ and 2. holds. ■

2.3 Subspaces

↪ **Definition 2.7**

Let X be a topological space and let $Y \subseteq X$. We define the subspace topology on Y:

1. Open sets in $Y = \{Y \cap \text{ open sets in } X\}$

↪ **Proposition 2.6: Consequences of Subspace Topologies**

Suppose $\mathcal{B}$ is a basis for a topology in X. Then, $\{U \cap Y : U \in \mathcal{B}\}$ forms a basis for the subspace $Y \subseteq X$.

Suppose X a metric space. Then, Y is also a metric space, with the same distance.

↪ **Proposition 2.7**

Let $Y \subseteq X$- a metric space. Then, the metric space topology for (Y,d) is the same as the subspace topology.

Proof. (Sketch) A basis for the open sets in X can be written $\bigcup_{\alpha \in I} B(x_\alpha, r_\alpha)$; hence

$$Y \cap (\bigcup_{\alpha \in I} B(x_\alpha, r_\alpha)) = \bigcup_{\alpha \in I} (Y \cap B(x_\alpha, r_\alpha))$$

is an open set topology for Y. ■

↪ **Lemma 2.1**

Let $A \subseteq X$-open, $B \subseteq A$; B-open in subspace topology for $A \iff B$-open in X.

↪ **Lemma 2.2**

Let $Y \subseteq X$, $A \subseteq Y$. Then, $\overline{A}$ in $Y = Y \cap \overline{A}$ in X. We can denote this

$$\overline{A}_Y = \overline{A}_X \cap Y.$$

2.4 Continuous Functions

↪ **Definition 2.8: Continuous Function**

Let X, Y be topological spaces. Let $f : X \to Y$. f is *continuous* $\iff \forall$ open $V \in Y$, $f^{-1}(V)$-open in X.

↪ **Proposition 2.8**

This definition is consistent with the normal ε-δ definition on the real line.

Proof. Let $f : \mathbb{R} \to \mathbb{R}$, continuous; that is, $\forall \varepsilon > 0, \forall x \in \mathbb{R} \exists \delta > 0$ s.t. $|x_1 - x| < \delta$, then $|f(x_1) - f(x)| < \varepsilon$.

Let $V \subseteq \mathbb{R}$ open. Let $y \in V$. Then, $\exists \varepsilon : (y - \varepsilon, y + \varepsilon) \subseteq V$. Let $y = f(x)$, hence $y \in f^{-1}(V)$. Now, if $d(x, x_1) < \delta$, we have that $d(f(x_1), f(x)) < \varepsilon$ (by continuity of f), hence $f(x_1) \in (y - \varepsilon, y + \varepsilon) \subseteq V$; moreover, $(x - \delta, x + \delta) \subseteq f{-1}(V)$, thus $f^{-1}(V)$ is open as required.

The inverse of this proof follows identically. ■

↪ *Lecture 02; Last Updated: Thu Jan 11 08:52:09 EST 2024*

↪ **Proposition 2.9**

Suppose $\mathcal{B}$ forms a basis of topology for Y. Then, $f : X \to Y$ is continuous if $f^{-1}(U)$ open $\forall U \in \mathcal{B}$.

Proof. If U-open set in Y, then $\exists I$-index set and a collection of open sets $\{A_\alpha\}_{\alpha \in I}, A_\alpha \in \mathcal{B}$, s.t. $U = \cup_{\alpha \in I} A_\alpha$. Then, we have

$$f^{-1}(U) = f^{-1}(\cup_{\alpha \in I}(A_\alpha)) = \cup_{\alpha \in I} \underbrace{f^{-1}(A_\alpha)}$$

Hence, if each $f^{-1}(A_\alpha)$ open, then $\cup_{\alpha\in I} f^{-1}(A_\alpha)$ open; hence it suffices to check if $f^{-1}(U)\,\forall\, U$-open in V is open to see if f continuous. ∎

↪ Theorem 2.1: Continuity of Composition

If $f : X \to Y$ continuous and $g : Y \to Z$ continuous, then $g \circ f$ continuous as well.

Proof. Let U-open in Z. Then

$$(g \circ f)^{-1}(U) = \underbrace{f^{-1}(\underbrace{g^{-1}(U)}_{\text{open in } Y})}_{\text{open in } X}$$

∎

↪ Proposition 2.10

If $f : X \to Y$ continuous and $A \subseteq X$, A has subspace topology, then $f|_A : A \to Y$ is also continuous.[2]

Proof. Let U-open in Y. Then

$$(f|_A)^{-1}(U) = \underbrace{f^{-1}(U)}_{\text{open}} \cap \underbrace{A}_{\text{open}}$$

By the definition of subspace topology, this is an open set and hence $f|_A$ is continuous. ∎

2.5 Product Spaces

↪ Definition 2.9: Finite Product Spaces

Let $X_1, \ldots, X_n$ be topological spaces. We define

$$(X_1 \times X_2 \times \cdots \times X_n),$$

and aim to define a *product topology*; a basis of which consists of cylinder sets.

[2] We denote $f|_A$ as the restriction of the domain of f to A.

↪ **Definition 2.10: Cylinder Set**

A *cylinder set* has the form

$$A_1 \times A_2 \times \cdots \times A_n$$

where each A_j-open in X_j.

⊛ **Example 2.3**

Given an open interval $(a_1, b_1), (a_2, b_2) \subset \mathbb{R}$, the set $(a_1, b_1) \times (a_2, b_2) \subset \mathbb{R}^2$ is a basis for the topology on $\mathbb{R}^2$.

↪ **Definition 2.11: Projection**

Let $X_1 \times X_2 \times \cdots X_n =: X$. The *projection* $\pi_j : X \to X_j$ maps $(x_1, \ldots, x_n) \to x_j \in X_j$.

Remark 2.3. *One can show π_j continuous.*

↪ **Definition 2.12: Coordinate Function**

Given a function $f : Y \to X_1 \times \cdots X_n = (x_1(y), x_2(y), \ldots, x_n(y))$. The *coordinate function* is

$$f_j = \pi_j \circ f; \quad f_j = x_j(y).$$

↪ **Proposition 2.11**

$f : Y \to X = X_1 \times \cdots \times X_n$ continuous $\iff$ $f_j : Y \to X_j$ continuous.

Proof. Its enough to show that $\forall\, U \in \mathcal{B}$-basis for X-product space, $f^{-1}(U)$-open in Y. Take $U = A_1 \times \cdots A_n$-open. Then, we claim that

$$f^{-1}(U) = f^{-1}(A_1 \times \cdots \times A_n) = f_1^{-1}(A_1) \cap f_2^{-1}(A_2) \cap \cdots \cap f_n^{-1}(A_n). \quad \star$$

If this holds, then as each f_i continuous (being a composition of continuous functions) and each A_i open in X_i, then each $f_i^{-1}(A_i)$ open in Y and hence $\star$, being the finite intersection of open sets in Y, is itself open in Y.

■

⊛ **Example 2.4: Fourier Transform: Motivation for Infinite Product Toplogies**

Let $f \in C([0, 2\pi])$ is real-valued. We write the nth Fourier coefficients

$$\begin{aligned}\hat{f}(n) &= \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x - i\frac{1}{2\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x\,.\end{aligned}$$

And the Fourier transform of f as the infinite product

$$f(x) \mapsto (\dots, \hat{f}(-n), \hat{f}(-n+1), \dots, \hat{f}(-1), \hat{f}(0), \hat{f}(1), \cdots, \hat{f}(n), \dots) \in \prod_{n\in\mathbb{Z}} (\mathbb{C})_n.$$

Hence, this is an (countably, as indexed by integers) infinite product space.

Now, let $f : \mathbb{R} \to \mathbb{R}$. Suppose $f(x) \to 0$ "fast enough" as $|x| \to \infty$ and f continuous. Then, we can define the Fourier coefficients

$$\hat{f}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-itx}\,\mathrm{d}x\,,$$

where $t \in \mathbb{R}$. We then have the transform

$$f \mapsto \{\hat{f}(t)\}_{t\in\mathbb{R}}.$$

In this case, our index set is $\mathbb{R}$ is (uncountably) infinite.

↪ **Definition 2.13: Product Topology/Cylinder Sets for ∞ Products**

Let $X = \prod_{\alpha\in I} X_\alpha$. Then, a basis for X is given by cylinder sets of the form $A = \prod_{\alpha\in I} A_\alpha$ where A_α-open in X_α, AND $A_\alpha = X_\alpha$ except for finitely many indices α.

That is, there exists a finite set $J = (\alpha_1, \dots, \alpha_k) \subseteq I$, such that we can write $A = \prod_{\alpha\in J} A_\alpha \times \prod_{\alpha\notin J} X_\alpha$ (where A_α open in X_α).

↪ **Proposition 2.12**

Given $f : Y \to \prod_{\alpha\in I} X_\alpha = X$, then (taking $f_\alpha = \pi_\alpha \circ f$ as before) we have that f is continuous in $X \iff f_\alpha : Y \to X_\alpha$ continuous in $X_\alpha \,\forall\, \alpha \in I$.

Remark 2.4. *Extension of proposition 2.11 to infinite product space.*

Proof. Write $U = \prod_{\alpha \in J} A_\alpha \times \prod_{\alpha \notin J} X_\alpha$. Then,

$$f^{-1}(U) = \bigcap_{\alpha \in J} f_\alpha^{-1}(A_\alpha)$$

which is open in Y, hence f continuous. ∎

Remark 2.5. *The intersection of the entire spaces give no restriction.*

↪ *Lecture 03; Last Updated: Fri Jan 19 11:49:27 EST 2024*

2.6 Metrizability

↪ **Proposition 2.13**

Different metrics can define the same topology.

⊛ **Example 2.5**

1. Different ℓ_p metrics in $\mathbb{R}^n$ (PSET 1)

2. Let (X, d) be a metric space. Then,

$$\tilde{d}(x, y) := \frac{d(x, y)}{d(x, y) + 1}$$

is also a metric (the first two axioms are trivial), and defines the same topology. Note, moreover, that $\tilde{d}(x, y) \leqslant 1 \, \forall \, x, y$; this distance is bounded, and can often be more convenient to work with in particular contexts.

↪ **Question 2.1**

Suppose (X_k, d_k) are metric spaces $\forall \, k \geqslant 1$. Then, we can define the product topology τ on

$$X := \prod_{k=1}^{\infty} X_k.$$

Does the product topology τ come from a metric? That is, is τ *metrizable*?

Remark 2.6. *There do indeed exist examples of non-metrizable topological spaces; this question is indeed well-founded.*

Answer. Let $\underline{x} = (x_1, x_2, \ldots, x_n, \ldots), \underline{y} = (y_1, y_2, \ldots, y_n, \ldots) \in \prod_{k=1}^{\infty}$ (where $x_i, y_i \in X_i$) be infinite sequences of elements. Then, for each metric space X_k take the metric

$$\tilde{d}_k(x_k, y_k) = \frac{d_k(x_k, y_k)}{1 + d_k(x_k, y_k)}$$

(as in the example above). Then, we define

$$D(\underline{x}, \underline{y}) = \sum_{k=1}^{\infty} \frac{\tilde{d}_k(x_k, y_k)}{2^k},$$

noting that $D(\underline{x}, \underline{y}) \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$ (by our construction, "normalizing" each metric), hence this is a valid, *converging* metric (which wouldn't otherwise be guaranteed if we didn't normalize the metrics). It remains to show whether this metric omits the same topology as τ. ■

2.7 Compactness, Connectedness

↪ **Definition 2.14: Compact**

A set A in a topological space is said to be *compact* if every cover has a finite subcover. That is, if

$$A \subseteq \bigcup_{\alpha \in I} U_\alpha - \text{open},$$

then $\exists \{\alpha_1, \ldots, \alpha_n \in I\}$ such that $A \subseteq \bigcup_{i=1}^n U_{\alpha_i}$.

↪ **Proposition 2.14**

A closed interval $[a, b]$ is compact.

Proof. If[3] $a = b$, this is clear. Suppose $a < b$, and let $[a, b] \subseteq \bigcup_{i \in I} U_i =: \mathcal{U}$ be an arbitrary cover. Then, we proceed in the following steps:

1. **Claim:** Given $x \in [a, b], x \neq b, \exists y \in [a, b]$ s.t. $[x, y]$ has a finite subcover.

   Let $x \in [a, b]$, $x \neq b$. Then, $\exists U_\alpha \in \mathcal{U} : x \in U_\alpha$. Since U_α open, and $x \neq b$, we further have that $\exists c \in [a, b]$ s.t. $[x, c) \subseteq U_\alpha$.

   Now, let $y \in (x, c)$; then, the interval $[x, y] \subseteq [x, c) \subseteq U_\alpha$, that is, $[x, y]$ has a finite subcover.

2. Define $C := \{y \in [a, b] : y > a, [a, y] \text{ has a finite subcover}\}$. We note that

   - $C \neq \varnothing$; taking $x = a$ in Step 1. above, we have that $\exists y \in [a, b]$ such that $[a, y]$ has a finite step cover, so this $y \in C$.
   - C bounded; by construction, $\forall\, y \in C, a < y \leqslant c$.

   Thus, we can validly define $c := \sup C$, noting that $a < c \leqslant b$. Ultimately, we wish to prove that $c = b$, completing the proof that $[a, b]$ has a finite subcover.

3. **Claim:** $c \in C$.

   Let $U_\beta \in \mathcal{U} : c \in U_\beta$. Then, by the openness of U_β, $\exists d \in [a, b]$ s.t. $(d, c] \subseteq U_\beta$.

[3] This proof is adapted from that of Theorem 27.1 in Munkre's Topology, an identical theorem but applied to more general ordered topologies.

Supposing $c \notin C$, then $\exists z \in C$ such that $z \in (d, c)$; if one did not exist, then this would imply that d was a smaller upper bound that c, a contradiction. Thus, $[z, c] \subseteq (d, c] \subseteq U_\beta$.

Moreover, we have that, given $z \in C$, $[a, z]$ has a finite subcover; call it $U_z \subseteq \mathcal{U}$. This gives, then:

$$[a, c] = [a, z] \cup [z, c] \subseteq U_z \cup U_\beta.$$

But this is a finite subcover of $[a, c]$, contradicting the fact that $c \notin C$. We conclude, then, that $c \in C$ after all.

4. **Claim:** $c = b$.

   Suppose not; then, since we have $c \leqslant b$, then assume $c < b$. Then, applying Step 1. with $x = c$ (which we can do, by our assumption of $c \neq b$!), then we have that $\exists y > c$ s.t. $[c, y]$ has a finite subcover, call this $U_y \subseteq \mathcal{U}$.

   Moreover, we had $c \in C$, hence $[a, c]$ has a finite subcover, call this $U_c \subseteq \mathcal{U}$.

   Then, this gives us that

$$[a, y] = [a, c] \cup [c, y] \subseteq U_c \cup U_y,$$

   that is, $[a, y]$ has a finite subcover, and so $y \in C$. But recall that $y > c$; hence, this a contradiction to c being the least upper bound of C. We conclude that $c = b$, and thus $[a, b]$ has a finite subcover, and is thus compact.

■

Remark 2.7. *A similar proof shows that $[a, b]$ is* connected*; we cannot cover it by two disjoint open sets.*

↪ **Theorem 2.2: On Compactness**

Let $A \subseteq \mathbb{R}^n$. Then, A compact $\iff$ A closed and bounded.

↪ **Proposition 2.15**

If X, Y are compact topological spaces, then $X \times Y$ is compact.

Remark 2.8. *By induction, if $X_1, \ldots, X_n$ compact, so is $\prod_{i=1}^n X_i$.*

↪ **Proposition 2.16**

A closed subset of a compact topological space is compact in the subspace topology.

Proof. (Of theorem 2.2)

($\impliedby$) If $A \subseteq \mathbb{R}^n$ closed and bounded, then $A \subseteq [-R, +R]^n$ for some $R > 0$ (it is contained in some "n-cube"). Then, we have that $[-R, R]$ is compact, by proposition 2.14, proposition 2.15, and proposition 2.16, A itself compact.

($\implies$) Suppose $A \subseteq \mathbb{R}^n$ is compact. Then, $\bigcup_{x \in A} B(x, \varepsilon)$ for some $\varepsilon > 0$ is an open cover of A. As A compact, there must exist a finite subcover of this cover, $A \subseteq \bigcup_{i=1}^N B(x_i, r_i)$. Let $R := \max_{i=1}^N (||x_i|| + r_i)$. Then, $A \subseteq \overline{B(0, R)}$, that is, A is bounded.

Now, suppose x is a limit point of A. Then, any neighborhood of x contains a point in A, so $\forall r > 0, B(x,r) \cap A \neq \varnothing$, and so $\overline{B}(x,r)$ also contains a point of A for any $r > 0$.

Now, suppose $x \notin A$ (looking for a contradiction). Then,

$$U := \bigcup_{r>0} U_r := \bigcup_{r>0} (\mathbb{R}^n \setminus \overline{B(x,r)}) = \mathbb{R}^n \setminus \{x\}$$

is an open cover for the set A. A being compact implies that U has an finite subcover such that $A \subset U_{r_1} \cup U_{r_2} \cup \cdots \cup U_{r_N}$. Let $r_0 = \min_{i=1}^N r_i$. Then, $A \subseteq U_{r_0}$, and $A \cap B(x, r_0) = \varnothing$; but this is a contradiction to the definition of a limit point, hence any limit point x is contained in A and A is thus closed by definition. ■

↪ **Proposition 2.17**

Compact $\implies$ sequentially compact; that is, every sequence in a compact set has a convergent subsequence.

↪ Lecture 04; Last Updated: Wed Jan 24 21:27:59 EST 2024

↪ **Definition 2.15: Connected**

A topological space X is *not connected* if $X = U \cup V$ for two open, nonempty, disjoint sets U, V.

If this does not hold, X is said to be *connected*.

A set $A \subseteq X$ is not connected if A is not connected in the subspace topology $\iff A =\subseteq U \cup V$, for U, V-open in X, $(U \cap A) \neq \varnothing$, $(V \cap A) \neq \varnothing$ and $U \cap V = \varnothing$.

↪ **Theorem 2.3**

Let X be a connected topological space. Let $f : X \to Y$ be a continuous function. Then, $f(X)$ is also connected.

Proof. Suppose, seeking a contradiction, that X is connected, but $f(X)$ is not. Then, we can write $f(X) \subseteq Y$ as $\overline{f(X)} \subseteq U \cup V$, such that U, V open in Y and $U \cap V = \varnothing$. Then,

$$(U \cap f(X)) \cap (V \cap f(X)) = \varnothing.$$

We also have that

$$X \subseteq \underbrace{f^{-1}(U)}_{\text{open in } X, \neq \varnothing} \cup \underbrace{f^{-1}(V)}_{\text{open in } X, \neq \varnothing}.$$

$f^{-1}(U) \cap f^{-1}(V) = \varnothing$ (that is, they are disjoint) by our assumption; this is a contradiction to the connectedness of X, as we are able to write it as a subset of two disjoint open sets. Hence, $f(X)$ is indeed connected. ■

↪ **Lemma 2.3**

Any interval $(a,b),[a,b],[a,b),\dots,\subseteq \mathbb{R}$ is connected.

Proof. ■

↪ **Theorem 2.4: "Intermediate Value Theorem"**

Suppose X is connected and $f : X \to \mathbb{R}$ is a continuous function. Then, f takes intermediate values.

More precisely, let $a = f(x), b = f(y)$ for $x, y \in X$. Assume $a < b$. Then, $\forall a < c < b$, $\exists z \in X$ s.t. $f(z) = c$.

Proof. Suppose, seeking a contradiction, that $\exists c : a < c < b$ s.t. $c \notin f(X)$ (that is, there exists an intermediate value that is "not reached" by the function).

Let $U = (-\infty, c)$ and $V = (c, +\infty)$; note that these are disjoint open sets. Then, we have that

$$X = f^{-1}(U) \cup f^{-1}(V),$$

by our assumption of $c \notin f(X)$. But this gives that X is not connected, as the union of two open (by continuity), disjoint, nonempty ($f(x) = a \in U \implies x \in f^{-1}(U)$, and $f(y) = b \in V \implies y \in f^{-1}(V)$) sets, a contradiction. ■

↪ **Theorem 2.5**

Suppose X is compact, Y-topological space, $f : X \to Y$ is a continuous function. Then, $f(X)$ is also compact.

Proof. Let $\{U_\alpha\}_{\alpha \in I}$ be an open cover of $f(X) \subseteq Y$, that is,

$$f(X) \subseteq \bigcup_{\alpha \in I} U_\alpha \implies X \subseteq f^{-1}(\bigcup_{\alpha \in I} U_\alpha) = \bigcup_{\alpha \in I} f^{-1}(U_\alpha) =: \bigcup_{\alpha \in I} V_\alpha - \text{open}.$$

Then, this is an open cover of X; X is compact, thus there exists a finite subcover, that is, indices $\{\alpha_1, \dots, \alpha_n\} \subseteq I$ such that $X = \bigcup_{i=1}^n V_{\alpha_i}$. Thus,

$$f(X) \subseteq \bigcup_{i=1}^n U_{\alpha_i},$$

which is a finite subcover of $f(X)$. Thus, $f(X)$ is compact. ■

Remark 2.9. *Recall the "extreme value theorem": let $f : [a,b] \to \mathbb{R}$ a continuous function; then, a minimum and maximum is obtained for $f(x)$ on this interval for values in this interval.*

↪ **Theorem 2.6**

Let X compact, and $f : X \to \mathbb{R}$ a continuous function. Then,

$$\max_{x \in X} f(x) \text{ and } \min_{x \in X} f(x)$$

are both attained.

Proof. $f(X) \subseteq \mathbb{R}$ is compact by theorem 2.5, and so by theorem 2.2, $f(X)$ is closed and bounded. Let, then, $m = \inf f(X)$ and $M = \sup f(X)$; these necessarily exist, since $f(X)$ is bounded. Both m and M are limit points of $f(X)$. But $f(X)$ is closed, and hence contains all of its limit points, and thus $m \in f(X)$ and $M \in f(X)$, and thus $\exists y_m : f(y_m) = m$ and $y_M : f(y_M) = M$. ∎

↪ **Definition 2.16: Path Connected**

A set $A \subseteq X$ is called *path connected* if $\forall x, y \in A$, $\exists f : [a, b] \to X$, continuous, s.t. $f(a) = x, f(b) = y$ and $f([a, b]) \subseteq A$.

The set $\{f(t) : a \leqslant t \leqslant b\}$ is called a *path* from x to y.

↪ **Theorem 2.7: Path connected $\implies$ connected**

If $A \subseteq X$ is path connected, then A is connected.

Proof. Suppose, seeking a contradiction, that A is path connected, but not connected. Then, we can write $A \subseteq U \cup V$, for open, disjoint, nonempty subsets $U, V \subseteq X$.

Let $x \in U \cap A$ and $y \in V \cap A$. Then, $\exists f : [a, b] \to A$ s.t. $f(a) = x, f(b) = y$, and $f([a, b]) \subseteq A$, by the path connectedness of A. Then,

$$[a, b] \subseteq f^{-1}(A) \subseteq \underbrace{f^{-1}(U \cap A)}_{\text{open}} \cup \underbrace{f^{-1}(V \cap A)}_{\text{open}} =: \underbrace{U_1}_{a \in} \cup \underbrace{U_2}_{b \in},$$

that is, $[a, b]$ is contained in a union of open, nonempty, disjoint sets, contradicting $[a, b]$ the connectedness of $[a, b]$ by lemma 2.3. Thus, A is connected. ∎

Remark 2.10. *A counterexample to the opposite side of the implication is the Topologist's sine curve, the set*

$$\{(x, \sin\left(\frac{1}{x}\right)) : x \in (0, 1]\} \cup \{0\} \times [-1, 1].$$

This set is connected in $\mathbb{R}^2$, but is not path connected.

↪ **Proposition 2.18**

For open sets in $\mathbb{R}^n$, path connected $\iff$ connected.

2.8 Path Components, Connected Components

Remark 2.11. *Remark that if a metric space X is not connected, then we can write $X = U \cup V$ where U, V are open, nonempty and disjoint. It follows, then, that $U = V^C$ (and vice versa) and hence U, V are both open and closed.*

↪ **Definition 2.17: Connected Component**

A connected component of $x \in X$ is the largest connected subset of X that contains x.

⊛ **Example 2.6**

Let $X = (0,1) \cup (1,2)$. Here, we have two connected components, $(0,1)$ and $(1,2)$

⊛ **Example 2.7: Middle Thirds Cantor Set**

Let $C_0 := [0,1]$, and given C_n, define $C_{n+1} := \frac{1}{3}(C_n \cup (2 + C_n))$ for $n \geqslant 0$. C_∞ is totally disconnected.

↪ **Definition 2.18: Path Component**

A *path component* $P(x)$ of $x \in X$ is the largest path connected subset of X that contains x.

↪ **Proposition 2.19**

$P(x) = \{x \in X : \exists \text{ conintuous path } \gamma : [0,1] \to X : \gamma(0) = x, \gamma(1) = y\}$.

Remark 2.12. *Where we "start" a path does not matter. We write $x \sim y$ if $\exists \gamma$ from x to y; this is an equivalence relation on the elements of X.*

Remark 2.13. *The choice of $[0,1]$ here is arbitrary; any closed interval is homeomorphic.*

↪ **Lemma 2.4**

If $P(x) \cap P(y) \neq \emptyset$, then $P(x) = P(y)$.

Proof. $P(x) \cap P(y) \neq \emptyset \implies \exists z : x \sim z \wedge y \sim z \implies x \sim y$. ∎

↪ **Lemma 2.5**

If $A \subseteq X$ is connected, then $\overline{A}$ is also connected.

↪ **Lemma 2.6**

Suppose $A \subseteq X$ is both open and closed. Then, if $C \subseteq X$ is connected and $C \cap A \neq \emptyset$, then $C \subseteq A$.

Proof. If A is both open and closed, then $C \cap A$ is both open and closed in C. If $C \cap A^C \neq \emptyset$, then this is also open and closed in C. Hence, we can write $C = (C \cap A) \cup (C \cap A^C)$, that is, a disjoint union of two nonempty open sets, contradicting the connectedness of C. Hence, $C \cap A^C = \emptyset$, and so $C \subseteq A$. ■

↪ **Proposition 2.20**

Let $\{C_\alpha\}_{\alpha \in I}$ be a collection of nonempty connected subspaces of X s.t. $\forall \alpha, \beta \in I, C_\alpha \cap C_\beta \neq \emptyset$. Then, $\cup_{\alpha \in I} C_\alpha$ is connected.

↪ **Proposition 2.21**

Suppose each $x \in X$ has a path-connected neighborhood. Then, the path components in X are the same as the connected components in X.

2.8.1 Cantor Staircase Function

↪ **Definition 2.19: An Explicit Definition**

Let $x \in \mathrm{C} : x = 0.a_1 a_2 a_3 \ldots$ (base 3), ie $a_j = \begin{cases} 0 \\ 2 \end{cases}$. Define

$$f(x) = \begin{cases} \sum \frac{a_j/2}{2^j} & x \in C \\ \text{extend by continuity} & x \notin C. \end{cases}$$

That is, if $x \notin C$, set $f(y) = \sup_{x \in C, x < y} f(x) = \inf_{x \in C, x > y} f(x)$.

↪ **Definition 2.20: Complement Definition**

To construct the complement of the Cantor set, begin with $[0,1]$ and at a step n, we remove 2^n open intervals from this interval. $f(x)$ will be constant on each of these intervals with values $\frac{k}{2^n}$ where k odd and $0 < k < 2^n$. Extend by continuity to all $x \in C$.

Remark 2.14. *Wikipedia's explanation of this is far better than whatever this definition is trying to say.*

↪ Lecture 06; Last Updated: Tue Jan 23 11:03:35 EST 2024

3 L^p Spaces

3.1 Review of ℓ^p Norms

Remark 3.1. *Recall that for* $1 \leqslant p \leqslant +\infty$*, we define for* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ *the norm*

$$||x||_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}, \quad ||x||_\infty = \max_{i=1}^{n} |x_i|.$$

Similarly, for infinite vector spaces, we had, for $x = (x_1, \ldots, x_n, \ldots)$*, the norm*

$$||x||_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}, \quad ||x||_\infty = \sup_{i \geqslant 1} |x_i|.$$

Here, we define

$$\ell_p := \{x = (x_1, \ldots, x_n) : ||x||_p < +\infty\}.$$

3.2 ℓ^p Norms, Hölder-Minkowski Inequalities

↪ **Definition 3.1: Hölder Conjugates**

For $1 \leqslant p, q \leqslant +\infty$, we say that p, q are said to be *Hölder conjugates* if

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Remark 3.2. *We refer to these simply as "conjugates" throughout as no other conception of conjugate numbers will be discussed.*

Further, we take by convention $\frac{1}{\infty} = 0$.

↪ **Proposition 3.1: Hölder's Inequality**

Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Suppose $p, q : 1 \leqslant p, q \leqslant +\infty$ are conjugate. Then,

$$\langle x, y \rangle_{\mathbb{R}^n} := \left|\sum_{i=1}^{n} x_i y_i\right| \leqslant ||x||_p \cdot ||y||_q$$

⊛ **Example 3.1**

For the case $p = 1$ or ∞ (functionally, the same case):

↪ **Lemma 3.1**

Let p, q be conjugates, and $x, y \geqslant 0$. Then,

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}.$$

Remark 3.3. *If the inequality holds, then, for some $t > 0$, let $\tilde{x} = t^{\frac{1}{p}} \cdot x$, $\tilde{y} = t^{\frac{1}{q}} y$. Substituting x for $\tilde{x}$ and y for $\tilde{y}$, we have*

$$\text{LHS: } \tilde{x}\tilde{y} = t^{\frac{1}{p}} x \cdot t^{\frac{1}{q}} y = t^{\frac{1}{p}+\frac{1}{q}} \cdot xy = xy$$
$$\text{RHS: } \cdots = t(\frac{x^p}{p} + \frac{y^q}{q}.)$$

That is, we have

$$t \cdot xy \leqslant t\left(\frac{x^p}{p} + \frac{y^q}{q}\right),$$

hence, the inequality is preserved under multiplication by a positive scalar; moreover, the original inequality holds iff this "scaled" version holds. Hence, choosing t such that $\tilde{y} = 1$ (let $t = \left(\frac{1}{y}\right)^q$), it suffices to prove the lemma for $y = 1$.

Proof. If $x = 0$ or $y = 0$, then the entire LHS becomes 0 and we are done; assume $x, y > 0$; by the previous remark, assume wlog $y = 1$. Then, we have

$$x \cdot y \leqslant \frac{x^p}{p} + \frac{y^q}{q} \iff x \cdot 1 \leqslant \frac{x^p}{p} + \frac{1}{q}$$
$$\iff \frac{x^p}{p} - x + \frac{1}{q} =: f(x) \geqslant 0.$$

Taking the derivative, we have

$$f'(x) = \frac{\cancel{p}x^{p-1}}{\cancel{p}} - 1 = x^{p-1} - 1$$
$$p > 1 \implies p - 1 > 0 \implies \begin{cases} f'(x) > 0 & \forall\, x > 1 \\ f'(x) = 0 & x = 0 \\ f'(x) < 0 & \forall\, 0 < x < 1 \end{cases}$$

Hence, $x = 1$ is a local minimum of the function, and thus $f(x) \geqslant f(1) \,\forall\, 0 < x \leqslant 1$. But $f(1) = \frac{1^p}{p} - 1 + \frac{1}{q} = 1 - 1 = 0$, hence $f(x) \geqslant 0 \,\forall\, x \geqslant 0$, as desired, and the inequality holds. ■

Proof. Assume $||x||_p = ||y||_q = 1$. Then,

$$\begin{aligned}
\left|\sum_{i=1}^n x_i y_i\right| &\leqslant \sum_{i=1}^n |x_i y_i| && \textit{(by triangle inequality)} \\
&\leqslant \sum_{i=1}^n \left|\frac{x_i^p}{p} + \frac{y_i^q}{q}\right| && \textit{(by lemma 3.1)} \\
&= \frac{1}{p}\left(\sum_{i=1}^n |x_i|^p\right) + \frac{1}{q}\left(\sum_{i=1}^n |y_i|^q\right) \\
&= \frac{1}{p}||x||_p^p + \frac{1}{q}||y||_q^q && \textit{(by staring)} \\
&= \frac{1}{p} \cdot 1^p + \frac{1}{q} \cdot 1^1 = \frac{1}{p} + \frac{1}{q} = 1 && \textit{(by assumption)} \\
&= ||x||_p \cdot ||y||_q,
\end{aligned}$$

and the proposition holds, in the special case $||x||_p = ||y||_q = 1$.

If $||x||_p = 0$ or $||y||_q = 0$, then $x_1 = \cdots = x_n = 0$ or $y_1 = \cdots = y_n = 0$, resp, then we'd have ($||x||_p = 0$ case)

$$0 \cdot y_1 + \cdots + 0 \cdot y_n \leqslant 0,$$

which clearly holds.

Assume, then, $||x||_p > 0, ||y||_q > 0$. Let $\tilde{x} := \frac{x}{||x||_p}, \tilde{y} := \frac{y}{||y||_q}$. Then,

$$||\tilde{x}||_p^p = \frac{\left(\sum_{i=1}^n |x_i|^p\right)}{||x||_p^p} = \frac{||x||_p^p}{||x||_p^p} = 1 \implies ||\tilde{x}||_p = 1.$$

The same case holds for $\tilde{y}$, hence $||\tilde{y}||_q = 1$; that is, we have "rescaled" both vectors. Hence, we can use the case we proved above for when the norms were identically 1 on $\tilde{x}, \tilde{y}$. We have:

$$\left|\sum_{i=1}^n \tilde{x}_i \tilde{y}_i\right| \leqslant 1$$

But by definition of $\tilde{x}, \tilde{y}$, we have

$$\left|\sum_{i=1}^n \tilde{x}_i \tilde{y}_i\right| = \left|\frac{1}{||x||_p ||y||_q} \sum_{i=1}^n x_i y_i\right| \leqslant 1 \implies \left|\sum_{i=1}^n x_i y_i\right| \leqslant ||x||_p \cdot ||y||_q,$$

and the proof is complete. ■

↪ **Proposition 3.2: Minkowski Inequality**

Let $1 \leqslant p \leqslant \infty$, $x, y \in \mathbb{R}^n$. Then,

$$||x + y||_p \leqslant ||x||_p + ||y||_p.$$

Remark 3.4. *This is just the triangle inequality for ℓ_p norms.*

Proof. The cases $p = 1, \infty$ are left as an exercise.

Assume $1 < p < \infty$. Then,

$$\begin{aligned}
||x + y||_p^p = \sum_{j=1}^{n} \left|x_j + y_j\right|^p &= \sum_{j=1}^{n} \left|x_j + y_j\right| \left|x_j + y_j\right|^{p-1} \\
&\leqslant \sum_{j=1}^{\infty} \left(\left|x_j\right| + \left|y_j\right|\right) \cdot \left|x_j + y_j\right|^{p-1} \\
&= \underbrace{\sum_{j=1}^{n} \left|x_j\right| \cdot \left|x_j + y_j\right|^{p-1}}_{:=A} + \overbrace{\sum_{j=1}^{n} \left|y_j\right| \cdot \left|x_j + y_j\right|^{p-1}}^{:=B} \quad \circledast
\end{aligned}$$

Let $\vec{u} = (|x_1|, \cdots, |x_n|)$ and $\vec{v} = (|x_1 + y_1|^{p-1}, \cdots, |x_n + y_n|^{p-1})$, then, $A = \vec{u} \cdot \vec{v} = \langle \vec{u}, \vec{v} \rangle_{\mathbb{R}^n}$. We have

$$\begin{aligned}
||\vec{u}||_p &= \left(\sum_{i=1}^{n} (|x_i|^p)\right)^{\frac{1}{p}} = ||x||_p \\
||\vec{v}||_q &= \left(\sum_{i=1}^{n} \left(|x_i + y_i|^{p-1}\right)^q\right)^{\frac{1}{q}} \\
&= \left[\sum_{i=1}^{n} \left(|x_i + y_i|^{p-1}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \\
&= \left[\sum_{i=1}^{n} |x_i + y_i|^p\right]^{\frac{p-1}{p}} \\
&= ||x + y||_p^{p-1}
\end{aligned}$$

where the second-to-last line follows from p, q being conjugate, hence $q = \frac{p}{p-1}$. Thus, by Hölder's Inequality, we have that

$$A = \langle \vec{u}, \vec{v} \rangle \leqslant ||u||_p \cdot ||v||_q = ||x||_p \cdot ||x + y||_p^{p-1}.$$

By a similar construction, we can show that

$$B \leqslant ||y||_p \cdot ||x + y||_p^{p-1}.$$

Thus, returning to our original inequality in ⊛, we have

$$\begin{aligned} ||x+y||_p^p &\leqslant A+B \\ &\leqslant ||x||_p \cdot ||x+y||_p^{p-1} + ||y||_p \cdot ||x+y||_p^{p-1} \\ &\implies ||x+y||_p \leqslant ||x||_p + ||y||_p, \end{aligned}$$

and the proof is complete. ■

↪ *Lecture 07; Last Updated: Tue Jan 30 12:54:59 EST 2024*

3.3 An Aside on Complete Metric Spaces

↪ **Theorem 3.1**

The sequence of centers of balls with monotonically decreasing radii is a Cauchy sequence in X.

Proof. Let $\varepsilon > 0$ and let $N : \forall j > N, r_j < \varepsilon$. Then,

$$d(x_j, x_k) < r_{\min(j,k)} = r_j$$

■

↪ **Definition 3.2: Complete Metric Space**

A metric space is complete if every Cauchy sequence converges to a limit in that space.

⊛ **Example 3.2: Examples of Complete Metric Spaces**

1. $\mathbb{R}$, p-adic integers $(\mathbb{Z}_p)$/rationals$(\mathbb{Q}_p)$.
2. $\ell_p = \{x = (x_i)_{i=1}^{\infty} : \sum_{i=1}^{\infty} |x_i|^p < +\infty\}, 1 \leqslant p \leqslant +\infty$
3. $\ell_\infty = \{x = (x_i) : \sup_{i=1}^{\infty} |x_i| < +\infty\}$.

↪ **Proposition 3.3**

Hölder's Inequality and Minkowski Inequality inequalities hold for infinite sequences. that is,

1. if $x = (x_i) \in \ell_p$ and $y = (y_i) \in \ell_q$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\left|\sum_{i=1}^{\infty} x_i y_i\right| \leqslant ||x_i||_{\ell_p} ||y_i||_{\ell_q}.$$

2. if $x, y \in \ell_p$, then

$$||x+y||_p \leqslant ||x||_p + ||y||_p.$$

Remark 3.5. *2. gives the triangle inequality for the $||x||_p$ norm on ℓ_p.*

Moreover,

$$\begin{aligned} ||c \cdot x||_p &= ||(c_1 x_1, \ldots, c_n x_n, \ldots)||_p \\ &= \left(\sum_{i=1}^{\infty} |c x_i|^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^{\infty} |c|^p \, |x_i|^p \right)^{\frac{1}{p}} \\ &= \left(|c|^p \right)^{\frac{1}{p}} ||x||_p = c \cdot ||x||_p \end{aligned}$$

Proof. (of 2.) If $x, y \in \ell_p$, we have that $\sum_{i=1}^{\infty} |x_i|^p < +\infty, \sum_{i=1}^{\infty} |y_i|^p < +\infty$, so $\exists N > 0 : \sum_{i=N+1}^{\infty} |x_i|^p < \varepsilon, \sum_{i=N+1}^{\infty} |y_i|^p < \varepsilon$. Let $x_i^{(n)} = (x_1, \ldots, x_n, 0, 0, \ldots)$ be (x) truncated after n (finite) coordinates. This gives

$$||(x_i + y_i)^{(n)}||_p \leqslant ||x_i^{(n)}||_p + ||y_i^{(n)}||_p \leqslant ||x||_p + ||y||_p$$

by Minkowski on finite spaces. Taking $n \to \infty$ (ie, "detruncating"), we have $(x + y) \in \ell_p$, and thus $||x + y||_p \leqslant ||x||_p + ||y||_p$.

1. left as an exercise. ■

↪ **Proposition 3.4**

Let $1 \leqslant p \leqslant +\infty$, and $||x||_\infty = \sup_{i=1}^{\infty} |x_i| = A < +\infty, ||y||_\infty = \sup_{i=1}^{\infty} |y_i| = B < +\infty$. Then, the triangle inequality $||x + y||_\infty \leqslant ||x||_\infty + ||y||_\infty$ holds.

Proof. We have

$$\sup_{i=1}^{\infty} |x_i + y_i| \leqslant \sup_{i=1}^{\infty} (|x_i| + |y_i|) \leqslant \sup_{i=1}^{\infty} |x_i| + \sup_{i=1}^{\infty} |y_i| = ||x||_\infty + ||y||_\infty.$$

■

↪ **Proposition 3.5**

$||x||_\infty := \sup_{i=1}^{\infty} |x_i|$ is a well-defined norm on ℓ_∞.

Proof. The triangle inequality is prove in proposition 3.4. The remainder of the requirements are left as an exercise. ■

↪ **Proposition 3.6**

$\ell_p \subseteq \ell_q$ if $p < q$.

Proof. Let $x \in \ell_p$. If $\sum_{i=1}^{\infty} |x_i|^p < +\infty$, then $\exists N : \forall i \geqslant N, |x_i| \leqslant 1$. Then,

$$\begin{aligned} \sum_{i \geqslant N} |x_i|^q &\leqslant \sum_{i \geqslant N} |x_i|^p < \infty \\ \implies \sum_{i=1}^{\infty} |x_i|^q < +\infty &\implies x \in \ell_q \\ &\implies \ell_p \subseteq \ell_q \end{aligned}$$

■

↪ Lecture 08; Last Updated: Thu Feb 1 09:10:36 EST 2024

3.4 Contraction Mapping Theorem

↪ Definition 3.3: Contraction Mapping

Let (X, d) be a metric space. A *contraction mapping* on X is a function $f : X \to X$ for which $\exists$ a constant $0 < c < 1$ such that

$$d(f(x), f(y)) \leqslant c \cdot d(x, y) \quad \forall x, y \in X.$$

↪ Theorem 3.2: Contraction Mapping Theorem

Let (X, d) be a complete metric space, and let $f : X \to X$ be a contraction. Then, there exists a unique fixed point z of f such that $f(z) = z$.

Moreover, $f^{[n]}(x) := f \circ f \circ \cdots \circ f(x) \to z$ as $n \to \infty$ for any $x \in X$.

Remark 3.6. *The "functional construction" of the Cantor set is an example of a contraction mapping, with* $f_1(x) = \frac{x}{3}, f_2(x) = \frac{x+2}{3}$. *The first has a fixed point of* 0, *and the second a fixed point of* 1.

Remark 3.7. *This is a generalization of this proof done in Analysis I, an equivalent claim over the reals.*

Proof. Fix $x \in X$. Consider the sequence $\{x_0, x_1, x_2, \ldots, x_n, \ldots\} := \{x, f(x), f \circ f(x), \cdots, f^{[n]}(x), \ldots\}$ (we call $f^{[n]}$ the *orbit* of x under iterations of f). We claim that this is a Cauchy sequence. Let $n \in \mathbb{N}$ arbitrary, then we have, by the property of the contraction mapping,

$$d(f^{[n+1]}(x) - f^{[n]}(x)) \leqslant c \cdot d(f^{[n]}(x) - f^{[n-1]}(x)) \leqslant c^2 d(f^{[n-1]}(x) - f^{[n-2]}(x)).$$

Arguing inductively, it follows that

$$d(f^{[n+1]}(x) - f^{[n]}(x)) \leqslant c^n d(f(x), x). \quad \star$$

Let now $m, k \in \mathbb{N}, m, k > 0$. It follows that

$$\begin{aligned} d(f^{[m]}, f^{[m+k]}(x) &\leqslant d(f^{[m]})(x), f^{[m+1]}(x)) + d(f^{[m+1]}(x), f^{[m]}(x)) + \cdots + d(f^{[m+k-1]}(x), f^{m+k}(x)) \\ &\overset{\star}{\leqslant} d(x, f(x))[c^m + c^{m+1} + \cdots + c^{m+k-1}] \\ &\leqslant c^m d(x, f(x))[1 + c + \cdots + c^k + c^{k+1} + \cdots] = \frac{c^m d(x, f(x))}{1-c} \end{aligned}$$

Now, given $\varepsilon > 0$, choose N such that $\frac{c^N d(x,f(x))}{1-c} < \varepsilon$. It follows, then, that $\{f^{[n]}(x)\}_{n\in\mathbb{N}}$ a Cauchy sequence, and thus converges, $f^{[n]}(x) \to z$ as $n \to \infty$ for some z.

We further have to show that $f(z) = z$. It is easy to show that f continuous due to the contraction mapping (it is clearly Lipschitz with constant c), and it thus follows that

$$\lim_{n\to\infty} f(f^{[n]}(x)) = \lim_{n\to\infty} f^{[n]}(x) \implies f(z) = z,$$

by sequential characterization of continuous functions.

Finally, we need to show that this limit is unique. Suppose $\exists y_1 \neq y_2$, ie two fixed points with $f(y_1) = y_1$ and $f(y_2) = y_2$. Then, by the property of the contraction mapping,

$$d(f(y_1), f(y_2)) \leqslant c \cdot d(y_1, y_2),$$

but by assumption of being fixed points,

$$d(f(y_1), f(y_2)) = d(y_1, y_2),$$

implying $d(y_1, y_2) \leqslant c \cdot d(y_1, y_2)$. This is only possible if $d(y_1, y_2) = 0$, and thus $y_1 = y_2$ and the fixed point is indeed unique. ■

↪ **Theorem 3.3: ℓ_p complete**

The space ℓ_p is complete for all $1 \leqslant p \leqslant +\infty$.

Equivalently, if $(x^1), (x^2), \ldots, (x^n)$ is a Cauchy sequence in ℓ^p, $\exists y \in \ell^p$ s.t. $x^n \to y$ as $n \to \infty$.

Proof. (Sketch) We suppose first $p < +\infty$. Consider an arbitrary number of Cauchy sequences in ℓ_p:

$$\begin{aligned} x^{(1)} &= (x_1^{(1)}, \ldots, x_n^{(1)}, \ldots) \\ x^{(2)} &= (x_1^{(2)}, \ldots, x_n^{(2)}, \ldots) \\ \vdots \quad & \quad \vdots \quad \quad \vdots \\ x^{(k)} &= (x_1^{(k)}, \ldots, x_n^{(k)}, \ldots) \in \ell_p \end{aligned}$$

We claim that, for any $k \in \mathbb{N}$, the $(x_k^{(n)})_{n\in\mathbb{N}}$ is a Cauchy sequence; note that in this definition we are taking a *fixed-index* (namely, the kth) element from different sequences (namely, the nth sequence).

Since $x^{(1)}, x^{(2)}, \dots, x^{(n)}, \dots$ are Cauchy sequences in ℓ^p, we have for a fixed $\varepsilon > 0$, $\exists N \in \mathbb{N} : \forall m, n > N$, $d_p(x^{(m)}, x^{(n)}) < \varepsilon$:

$$d_p(x^{(m)}, x^{(n)})^p = ||x^{(m)} - x^{(n)}||_p^p = \sum_{i=1}^{\infty} \left| x_i^{(m)} - x_i^{(n)} \right|^p < \varepsilon^p$$

$$\left| x_k^{(m)} - x_k^n \right|^p \leqslant \sum_{i=1}^{\infty} \left| x_i^{(m)} - x_i^{(n)} \right|^p \implies \left| x_k^{(m)} - x_k^n \right|^p < \varepsilon^p$$

$$\implies \left| x_k^{(m)} - x_k^{(n)} \right| < \varepsilon,$$

since we are taking "less of the summands in the second line". It follows, then, that for each k, $\exists z_k : x_k^{(n)} \to z_k$ as $n \to \infty$. Let $z = (z_1, \dots, z_n, \dots)$. We claim that $x^{(n)} \to z \in \ell_p$ as $n \to \infty$.

First, we show that $d_p(x^{(n)}, z) \to 0$ as $n \to 0$ (that is, $x^{(n)} \to z$ as $n \to \infty$). Fix $\varepsilon > 0$, and choose $N \in \mathbb{N}$ for which $d_p(x^{(m)}, x^{(n)}) < \varepsilon \ \forall m, n \geqslant N$ (by Cauchy). Fix $K \in \mathbb{N}, K > 0$.

$$d_p^p(x^{(n)}, z) = ||x^{(n)} - z||_p^p = \sum_{i=1}^{\infty} \left| x_i^{(n)} - z_i \right|^p$$

$$||x^{(m)} - x^{(n)}||_p^p < \varepsilon^p \implies \sum_{i=1}^{K} \left| x_i^{(m)} - x_i^{(n)} \right|^p \leqslant \varepsilon^p$$

Let $m \to \infty$; then $x_i^{(m)} \to z_i$ (note that i fixed!), and we have

$$\sum_{i=1}^{K} \left| z_i - x_i^{(n)} \right|^p \leqslant \varepsilon^p.$$

Let $K \to \infty$; then,

$$\sum_{i=1}^{\infty} \left| z_i - x_i^{(n)} \right|^p \leqslant \varepsilon^p \implies ||z - x||_p \leqslant \varepsilon \implies d_p(z, x^n) \leqslant \varepsilon,$$

and thus $x^n \to z$ as $n \to \infty$.

It remains to show that $z \in \ell_p$, ie $||z||_p < +\infty$. We have:

$$||z||_p \leqslant \underbrace{||z - x^{(n)}||_p}_{\to 0} + ||x^{(n)}||_p.$$

For sufficiently large n, $||z - x^{(n)}|| \leqslant 1$ (for instance); $x^{(n)} \in \ell_p$, hence $||x^{(n)}||_p < +\infty$ (say, $||x^{(n)}||_p \leqslant M$). Thus:

$$||z||_p \leqslant 1 + M < +\infty \implies z \in \ell_p,$$

and the proof is complete. ∎

3.5 Compactness in Metric Spaces

↪ **Definition 3.4: Totally Bounded**

Let (X, d) be a metric space. If for every $\varepsilon > 0, \exists x_1, \ldots, x_n \in X, n = n(\varepsilon) : \bigcup_{i=1}^{n} B(x_1, \varepsilon) = X$, we say X is *totally bounded*.

↪ Lecture 09; Last Updated: Mon Feb 19 22:32:57 EST 2024

↪ **Theorem 3.4**

Let (X, d) be a metric space. TFAE:

1. X is complete and totally bounded;
2. X is compact;
3. X is sequentially compact (every sequence has a convergent subsequence).

Proof. (1. $\implies$ 2.) Suppose X complete and totally bounded. Assume towards a contradiction that X not compact, ie there exists an open cover $\{U_\alpha\}_{\alpha \in I}$ of X with no finite subcover.

X being totally bounded gives that it can be covered by finitely many open balls of radius $\frac{1}{2}$. It must be that at least one of these open balls cannot be finitely covered, otherwise we would have a finite subcover. Let F_1 be the closure of this ball. F_1 closed, with diameter $\mathrm{diam}(F_1) \leqslant 1$. X.

We also have that X can be covered by finitely many balls of radius $\frac{1}{4}$; again, there must be at least one ball B_1 such that $B_1 \cap F_1$ cannot be covered by finitely many open sets from the cover. Let $F_2 = \overline{B_1} \cap F_1$-closed, with $\mathrm{diam}(F_2) \leqslant \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.[4]

Arguing inductively, at some step n, X can can be covered by finitely many balls of radius $\frac{1}{2^n}$; at least one of these balls B cannot be covered by a finite subcover hence $B \cap F_{n-1}$ cannot be covered by finitely many U_α's. Let $F_n = \overline{B} \cap F_{n-1}$ -closed, with $\mathrm{diam}(F_n) \leqslant \frac{1}{2^{n-1}}$.

As such, we have a nested sequence $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$ of closed sets, where $\mathrm{diam}(F_k) \leqslant \frac{1}{2^{k-1}} \to 0$ as $k \to \infty$.

↪ **Lemma 3.1** (Cantor Intersection Theorem). $\bigcap_{k=1}^{\infty} F_k \neq \emptyset$.

Proof. (Of Lemma) Let $x_k \in F_k$. Then, $\{x_k\}_{k \in \mathbb{N}}$ is a Cauchy sequence, since

$$d(x_n, x_{n+k}) \leqslant \mathrm{diam}(F_n) + \cdots + \mathrm{diam}(F_{n+k}) \leqslant \frac{1}{2^{n-1}},$$

by the nested property, which can be made arbitrarily small for sufficiently large n, k. Hence, $x_n \to y \in X$ for some y, as X complete. The tail of x_n lies in F_n for all sufficiently large n, and as each F_n closed, the limit must lie in F_n for all sufficiently large n. We conclude the intersection nonempty. ■

[4] B_1 has radius $\frac{1}{4}$ and hence diameter $\frac{1}{2}$. The intersection of B_1 with a set with a larger diameter must have diameter leq $\frac{1}{2}$

This y from the lemma is covered by some U_{α_0}-open for some $\alpha_0 \in I$. Being open, $\exists \varepsilon > 0 : B(y, \varepsilon) \subseteq U_{\alpha_0}$. Let $n : \frac{1}{2^n - 1} < \varepsilon$. Then, $y \in F_n$, and as $\operatorname{diam}(F_n) \leqslant \frac{1}{2^{n-1}}$, we have that $F_n \subseteq B(y, \frac{1}{2^{n-1}}) \subseteq B(y, \varepsilon) \subseteq U_{\alpha_0}$. But then, we have that F_n covered by a single open set U_{α_0}, a contradiction to our inductive construction of F_n. We conclude X compact.

(2. $\implies$ 3.) Suppose X compact. Let $\{x_n\}_{n \in \mathbb{N}} \in X$. Let $F_n = \overline{\bigcup_{k \geqslant n} \{x_k\}}$-closed; we have too that $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$.

↪ **Definition 3.5: Finite Intersection Property**

$\mathcal{F}$ has finite intersection property provided any finite subcollection of sets in $\mathcal{F}$ has a non-empty intersection.

↪ **Lemma 3.2** (Finite Interesection Formulation of Compactness). *X-compact $\iff$ every collection $\mathcal{F}$ of closed subsets of X with finite intersection property has non-empty intersection.*

Proof. ■

This lemma directly gives that $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$, $\{F_n\}_{n \in \mathbb{N}}$ being a collection of closed subsets with any subset having nonempty intersection (by the nestedness). Let $y \in \cap_{n=1}^{\infty} F_n$. Take $B(y, \frac{1}{k})$, which thus has nonempty intersection with $\{x_k\}_{k \geqslant n} \, \forall \, n$, ie $\exists n_1 : d(y, x_{n_1}) < 1$ and $\exists n_2 > n_1 : d(y, x_{n_2}) < \frac{1}{2}$. Arguing inductively, $\exists n_j > n_{j-1} : d(y, x_{n_j}) < \frac{1}{j}$ for any given n_{j-1}. It follows that $\lim_{j \to \infty} x_{n_j} = y$, and thus $\{x_{n_j}\}$ is a convergent subsequence of $\{x_n\}$ that converges within X, and thus X is sequentially compact.

(3. $\implies$ 1.) Suppose X sequentially compact. Let $\{x_n\} \in X$ be a Cauchy sequence in X, which thus have a convergent subsequence $\{x_{n_k}\} \to y$.

↪ **Lemma 3.3.** *Let $\{x_n\}$ be a Cauchy sequence in X where X sequentially compact. Then, if $\{x_{n_k}\} \to y$, so does $\{x_n\} \to y$*

Proof. ■

Then, $\{x_n\}_n \to y$ and so X complete.

Suppose X not totally bounded, ie $\exists \varepsilon > 0 : X$ cannot be covered by a finite union of balls of $B(x_j, \varepsilon)$. Let $x_1 \in X$ s.t. $B(x_1, \varepsilon) \not\supseteq X$; $\exists x_2 \in X \setminus B(x_1, \varepsilon)$, and so $X \not\subseteq B(x_1, \varepsilon) \cup B(x_2, \varepsilon)$ by assumption. Then, choose $x_3 \in X \setminus (B(x_1, \varepsilon) \cup B(x_2, \varepsilon))$. Arguing inductively, we have that $\exists x_n \in X \setminus (\bigcup_{i=1}^{n} B(x_i, \varepsilon))$, noting that $d(x_n, x_j) \geqslant \varepsilon \, \forall \, 1 \leqslant j \leqslant n$.

Consider the sequence $\{x_j\}_{j \in \mathbb{N}}$:

↪ **Lemma 3.4.** *$\{x_j\}$ cannot have a convergent subsequence.*

Proof. Follows by $d(x_m, x_n) \geqslant \varepsilon \, \forall \, m, n$. ■

This contradicts our assumption that X sequentially compact, and we conclude X must be totally bounded. ■

↪ Lecture 10; Last Updated: Tue Feb 6 09:50:59 EST 2024

⊛ **Example 3.3: Complete Metric Space Example: L^p norm**

Let $f \in C([a,b])$. We define the norm

$$||f||_p := \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}.$$

As desired, $||f||_p \geqslant 0$; $||f||_p = 0 \iff f \equiv 0$; $||c \cdot f||_p = c \cdot ||f||_p$.

Hölder's and Minkowski's inequalities for functions also hold; for $\frac{1}{p} + \frac{1}{q} = 1, 1 \leqslant p, q \leqslant \infty$,

$$\int |fg| \leqslant ||f||_p \cdot ||g||_q; \quad ||f+g||_p \leqslant ||f||_p + ||g||_q,$$

respectively.

We similarly have the L^∞ norm, namely, for a function $f : [a,b] \to \mathbb{R}$,

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

which obeys all the necessary properties as well.

Let $f_n \to f$ in $C([a,b])$, wrt $|| \cdots ||_\infty$, where $\{f_n\}_{n \in \mathbb{N}}$ a sequence of functions. Namely, we say that

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall n \geqslant N, \sup_{x \in [a,b]} |f_n(x) - f(x)| < \varepsilon.$$

If this holds, we say that f_n *uniformly converges*.

We say that $f_n(x) \to f(x)$ *pointwise* on $[a,b]$ if $\forall x \in [a,b], f_n(x) \to f(x)$. Note that uniform convergence implies pointwise convergence, but not the converse.

↪ **Theorem 3.5**

Suppose $f_n(x)$ continuous, and $f_n(x) \to f(x)$ uniformly on $[a,b]$. Then, $f(x)$ also continuous on $[a,b]$.

Proof. Fix $\varepsilon > 0$, $x_0 \in [a,b]$. We have that $\exists N : n \geqslant N, |f_n(x) - f(x)| < \frac{\varepsilon}{3}, \forall x \in [a,b]$.

Let $n \geqslant N$. $f_n(x)$ continuous at x_0, hence $\exists \delta(x_0) > 0 : |y - x_0| \implies |f_n(y) - f_n(x_0)| < \frac{\varepsilon}{3}$. We have

$$\begin{aligned}
|f(x_0) - f(y)| &\leqslant |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(y)| + |f_n(y) - f(y)| \\
&\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,
\end{aligned}$$

completing the proof. ■

Remark 3.8. *This does not hold with pointwise convergence.*

Remark 3.9. *We will prove later that $C([a,b])$ is complete for $||f||_\infty$, but not for arbitrary $||f||_p$, $1 \leqslant p < +\infty$. To "complete" $C([a,b])$ for $p \neq \infty$, we will need to consider measurable functions and redefine our notion of integration.*

↪ *Lecture 11; Last Updated: Thu Feb 8 09:51:13 EST 2024*

4 Derivatives

4.1 Introduction

↪ **Definition 4.1: Differentiable**

We say $f(x)$ differentiable at c if $\exists \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$. If so, we denote the limit $f'(c)$.

Remark 4.1. *For x close to c, then $f(x) \approx f(c) + f'(c)(x-c)$; this is a linear approximation of f at c.*

⊛ **Example 4.1: Weierstrass**

$f(x) = \sum_{n=1}^{\infty} \frac{\cos(3^n)x}{2^n}$ is continuous in $\mathbb{R}$, but nowhere differentiable.

↪ **Definition 4.2**

The derivative, dx, is a linear map $C^{(}[a,b]) \longrightarrow C^0([a,b])$.

4.2 Chain Rule

Remark 4.2. *See Analysis I notes as well.*

↪ **Theorem 4.1: Caratheodory's Theorem**

Let $f : I \longrightarrow \mathbb{R}$, $c \in I$. f is differentiable at $x = c$ iff $\exists \varphi(x) : I \longrightarrow \mathbb{R}$ s.t. φ continuous at c and $f(x) - f(c) = \varphi(x)(x-c)$.[5]

Proof. If $f'(c)$ exists, let

$$\varphi(x) = \begin{cases} \frac{f(x)-f(c)}{x-c} & x \neq c \\ f'(c) & x = c. \end{cases},$$

which is well defined. Moreover, for $x \neq c$, $\varphi(x)(x-c) = \frac{f(x)-f(c)}{x-c}(x-c) = f(x) - f(c)$ as desired; the case for $x = c$ is clear. Continuity at c:

$$\lim_{x\to c} \varphi(x) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} = f'(c) = \varphi(c).$$

[5]If not stated otherwise, sets named I or J are intervals.

Conversely, suppose such a φ exists. Then, by continuity,

$$\exists \varphi(c) = \lim_{x \to c} \varphi(x) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

which gives directly that f differentiable at c. ■

↪ **Theorem 4.2: Chain Rule**

Let $f : J \to \mathbb{R}$, $g : I \to \mathbb{R}$, $f(J) \subseteq I$. If $f(x)$ differentiable at c and $g(y)$ is differentiable at $y = f(c)$, then $g \circ f(x)$ is also differentiable at c, and

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

Proof. Using Caratheodory's Theorem, $\exists \varphi : f(x) - f(c) = \varphi(x)(x - c)$ with $\varphi(c) = f'(c)$. Let $d = f(c)$, then similarly $\exists \psi : g(y) - g(d) = \psi(y)(y - d)$ with $\psi(d) = g'(d)$, with φ, ψ continuous at c, d resp. Then,

$$g(f(x)) - g(f(c)) = \psi(f(x))(f(x) - f(c)) = (\psi \circ f)(x) \cdot (\phi(x)(x - c))$$

$\psi \circ f$ is continuous at c, as a composition of continuous functions (ψ, ϕ continuous by construction, f differentiable and thus continuous). It follows, then, that

$$\lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} = \lim_{x \to c} (\psi \circ f)(x) \cdot \varphi(x) = \psi(f(c))\varphi(c) = g'(f(c)) \cdot f'(c),$$

by construction. ■

4.3 Critical Points

↪ **Definition 4.3**

$f : I \to \mathbb{R}$ has a max/min c if $\exists J \subseteq I : x \in J$ s.t. $\max_{x \in J} f(x)/\min_{x \in J} f(x) = f(c)$.

↪ **Theorem 4.3: Rolle's**

Let $f : [a, b] \to \mathbb{R}$ continuous. Suppose $f'(x)$ exists for all $x \in (a, b)$ and $f(a) = f(b) = 0$. Then, $\exists c \in (a, b) : f'(c) = 0$.

Remark 4.3. *A "complex-version" of Rolle's:*

↪ **Theorem 4.4: Gauss-Lucas**

Let $P(z)$ be a complex-valued polynomial. Then, the roots of $P'(z)$ lie inside the convex hull of roots of $P(z)$, where a convex hull is the smallest polygon with vertices at the roots of $P(z)$.

↪ **Definition 4.4**

Consider $P(z) = z^n - 1$ for some $n \in \mathbb{N}$. If z a root, we can show that $(|z|)^n = 1$, hence all roots lie on the unit circle in the complex plane at multiples of the same angle. This gives us a regular n-gon in the complex plane. We then have that $P'(z) = nz^{n-1}$, with has root $z = 0$, which clearly lies within the n-gon hull.

↪ **Theorem 4.5: Mean Value**

Let f be continuous on $[a, b]$ and differentiable on (a, b). Then, $\exists c \in (a, b)$ s.t. $f(b) - f(a) = f'(c)(b - a)$.

Proof. Let $\varphi(x) = f(x) - f(a) = \frac{f(b)-f(a)}{(b-a)}(x - a)$, where $\varphi(a) = 0 = \varphi(b)$. By Rolle's theorem, $\exists c \in (a, b) : \varphi'(c) = 0 = f'(c) - \frac{f(b)-f(a)}{(b-a)}$, as desired. ∎

↪ *Lecture 12; Last Updated: Thu Feb 15 09:47:18 EST 2024*

4.4 Aside: Continued Fractions

We have that, for any $x \in \mathbb{R}$, $x = \lfloor x \rfloor + \{x\}$, with $\{x\} \in (0, 1)$; $\lfloor x \rfloor$ and $\{x\}$ are the integral and fractional parts of x respectively.

Fix $x \in \mathbb{R}$, assuming $x \neq 0$.

Let $x_1 := \frac{1}{\{x\}}$. We can write

$$x = \lfloor x \rfloor + \frac{1}{x_1}.$$

If $\{x_1\} \neq 0$, let $x_2 := \frac{1}{\{x_1\}}$ and write

$$x = \lfloor x \rfloor + \frac{1}{\lfloor x_1 \rfloor + \frac{1}{x_2}}.$$

Continuing in this manner, this process stops if $\{x_i\} = 0$ for some i; if $x \in \mathbb{Q}$, this process will stop, else, it will continue infinitely. For instance, the Golden Ratio $x = \frac{\sqrt{5}\pm 1}{2}$ has continued fraction expansion

$$x = \frac{1}{1 + \frac{1}{1+\frac{1}{1+\cdots}}}.$$

More succinctly, we can denote $a_0 := \lfloor x \rfloor$ and $a_i = \lfloor x_i \rfloor, i \geqslant 1$, and write

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2+\frac{1}{a_3+\ddots}}}.$$

We notate, accordingly, $x := (a_1, a_2, a_3, \dots)$; in this case, the Golden Ratio can be notated $(1, 1, 1, \dots)$.

We denote $\frac{p_n}{q_n}$ as the nth continued fraction of a given x. It turns out that this is the best possible rational approximation for $x \notin \mathbb{Q}$.

4.5 Back To Derivatives

↪ **Theorem 4.6**

$f : I \to \mathbb{R}$, differentiable. f is increasing (resp decreasing) iff $f'(x) \geqslant 0 \,\forall\, x \in I$ (resp $f'(x) \leqslant 0 \,\forall\, x \in I$).

↪ **Proposition 4.1**

Let f continuous on $I = [a, b]$. Let $a < c < b$ and suppose f differentiable on (a, c) and (c, b). Suppose $f'(x) \geqslant 0$ on $(c - \delta, c)$ and $f'(x) \leqslant 0$ on $(c, c + \delta)$ for some $\delta > 0$. Then, f has local max at $x = c$.

↪ **Lemma 4.1**

Let $I \subseteq \mathbb{R}$, and assume $f : I \to \mathbb{R}$ is differentiable at $x = c \in I$.

1. If $f'(c) > 0$, then $\exists \delta > 0 : f(x) > f(c) \,\forall\, x \in I, x \in (c, c + \delta)$.
2. (Reverse statement for $f'(c) < 0$)

↪ **Theorem 4.7: Darboux**

Suppose f differentiable on $I := [a, b]$ and $f'(a) < k < f'(b)$. Then, $\exists c \in (a, b)$ such that $f'(c) = k$.

↪ Lecture 13; Last Updated: Thu Feb 15 09:49:55 EST 2024

4.6 L'Hopital's Rules

↪ **Proposition 4.2**

Suppose $f(x), g(x) : [a, b] \to \mathbb{R}$ with $f(a) = g(a) = 0$, and $g(x) \neq 0 \,\forall\, a < x < b$. Suppose f, g are differentiable at $x = a$ and $g'(a) \neq 0$. Then, $\lim_{x \to a^+} \frac{f(x)}{g(x)}$ exists, and moreover, it is equal to $f'(a)g'(x)$.

Proof.

$$\lim_{x \to a^+} \left(\frac{f(x) - f(a)}{x - a}\right) / \left(\frac{g(x) - g(a)}{x - a}\right) = \lim_{x \to a^+} \frac{f(x)}{x - a} \frac{x - a}{g(x)} = \lim_{x \to a^+} \frac{f(x)}{g(x)},$$

but the original line is simply $\frac{f'(a)}{g'(a)}$. ∎

⊛ **Example 4.2**

$\lim_{x \to 0} \frac{\sin x}{x} = \frac{\cos(0)}{1} = 1.$

↪ **Theorem 4.8: Cauchy Mean Value**

Let $f(x), g(x) : [a,b] \to \mathbb{R}$ where f, g continuous on $[a,b]$ and differentiable on (a,b). Assuming $g'(x) \neq 0$, $\forall x \in (a,b)$, then $\exists c \in (a,b)$ such that

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

↪ **Proposition 4.3: More General L'Hopital**

let $-\infty \leqslant a < b \leqslant +\infty$ and f, g differentiable on (a,b). Suppose $\lim_{x\to a^+} f(x) = 0 = \lim_{x\to a^+} g(x)$.

1. If $\exists L := \lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ where L some real number, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$ as well.
2. If $\exists L := \lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ where $L = +\infty$ or $-\infty$, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$ as well.

↪ **Proposition 4.4**

Let $-\infty \leqslant a < b \leqslant +\infty$, f, g differentiable on (a,b) and $g'(x) \neq 0 \, \forall x \in (a,b)$. Suppose $\lim_{x\to a^+} g(x) = \pm\infty$.

1. If $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} =: L$ exists and is some finite real number, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$ as well.
2. If $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} =: L$ exists and is $\pm\infty$, then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$ as well.

4.7 Taylor's Theorem

↪ **Theorem 4.9: Taylor's Theorem**

Let $I = [a,b] \subseteq \mathbb{R}$, $f : I \to \mathbb{R}$, $f \in C^n(I)$ and suppose $f^{(n+1)}(x)$ exists on (a,b). Let $x_0 \in [a,b]$. Then, for any $x \in [a,b]$, $\exists c$ between x, x_0 such that

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f'(x_0)}{2}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

↪ Lecture 14; Last Updated: Thu Feb 22 08:37:23 EST 2024

↪ **Theorem 4.10: Relative Extrema**[6]

Let $I \subseteq \mathbb{R}$ be an open interval, $x_0 \in I$, and $n \geqslant 2$. Suppose $f', f'', \ldots, f^{(n)}$ are continuous in a neighborhood of x_0, and $f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0$ and $f^{(n)}(x_0) \neq 0$. Then:

1. if n is even and $f^{(n)}(x_0) > 0$, then f has a local minimum at x_0;
2. if n is even and $f^{(n)}(x_0) < 0$, then f has a local maximum at x_0;
3. if n is odd, then f has neither a local minimum nor maximum at x_0.

Proof. If $n := 2m$-even and $f^{(2m)}(x_0) > 0$, then $f^{(n)}(c) > 0$ so $f(x) - f(x_0) = f^{(2m)}(c)(x - x_0) > 0$. ■

4.8 Convex Sets

↪ **Definition 4.5: Convex Set**

$A \subseteq V$-vector space over $\mathbb{R}$ is *convex* if for any $x, y \in A$ and any $0 \leqslant t \leqslant 1$, $t \cdot x + (1-t) \cdot y \in A$.

↪ **Definition 4.6: Convex Function**

Let $f : I \to \mathbb{R}$. f is *convex* if $\forall\, x_1, x_2 \in I$ and $0 \leqslant t \leqslant 1$,

$$f((1-t)x_1 + tx_2) \leqslant (1-t)f(x_1) + tf(x_2).$$

↪ *Lecture 15; Last Updated: Thu Feb 22 09:37:31 EST 2024*

[6] Bartle-Sherbert, Theorem 6.4.4

5 Appendix

5.1 Notes from Tutorials

↪ **Theorem 5.1**

Let (X, d) be a compact metric space.[7] Let $C(X) := \{f : X \to \mathbb{R} : f \text{ continuous}\}$ be a vector space. Take the uniform norm $||f|| := \sup_{x\in X} |f(x)|$ on $C(x)$. Then, $(C(x), || \bullet ||)$ is complete.[8]

Proof. Denote the "canonical norm" $\rho(f, g) := ||f - g||$.

Let $(f_n) \in C(X)$ be a Cauchy sequence. Then, $\forall\, \varepsilon > 0, \exists N \in \mathbb{N} : \forall\, m, n \geqslant N, \rho(f_n, f_m) < \varepsilon$.

Fix $x \in X$, noting that

$$|f_n(x) - f_m(x)| \leqslant \sup_{y\in X} |f_n(y) - f_m(y)| = \rho(f_n, f_m) < \varepsilon. \quad *^1$$

Define, for this fixed x, a sequence *in* $\mathbb{R}$ $\{f_n(x)\}_{n\in\mathbb{N}}$. By $*^1$, we have that this sequence is Cauchy in $\mathbb{R}$, but as $\mathbb{R}$ complete, $f_n(x)$ hence converges, to some limit we call $f(x) := \lim_{n\to\infty} f_n(x)$. Note that x is still fixed at this point; these are but real numbers we are working with here.

Now, as x was completely arbitrary, we can repeat this process for all of X, and define a function $f : X \to \mathbb{R}$ where $f(x) := \lim_{n\to\infty} f_n(x)$.

For a fixed x, we have that $f_m(x) \to f(x)$ as $m \to \infty$. This implies:

$$\begin{aligned} 0 \leqslant \lim_{m\to\infty} |f_n(x) - f_m(x)| \leqslant \lim_{m\to\infty} \varepsilon &= \varepsilon \\ &\implies |f_n(x) - f(x)| \leqslant \varepsilon\, \forall\, n \geqslant N \\ &\implies \rho(f_n, f) = \sup_{x\in X} |f_n(x) - f(x)| \leqslant \varepsilon \implies f_n \to f \end{aligned}$$

It remains to show that $f \in C(X)$. Let $c \in X$ and $\varepsilon > 0$, and the corresponding $N \in \mathbb{N} : \rho(f_n, f) < \frac{\varepsilon}{3}\, \forall\, n \geqslant N$. By construction, $f_N \in C(X)$, and is thus continuous at c. This gives that $\exists \delta > 0 : |f_N(x) - f_N(c)| < \frac{\varepsilon}{3}$ whenever $d(x, c) < \delta$. [9]

Hence, if $d(x, c) < \delta$, we have

$$\begin{aligned} |f(x) - f(c)| &\leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(c)| + |f_N(c) - f(c)| \\ &\leqslant \rho(f, f_N) + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

[8] In this proof, the compactness is necessary for the norm to be well-defined.

[8] In this way, this becomes a Banach Space: a complete, normed vector space.

[9] Be careful here, there are three different metrics going on; ρ from the vector space, d from the underlying metric space, and $|\cdots|$ from $\mathbb{R}$.

hence f continuous at c, which was completely arbitrary, and thus $f \in C(X)$. ■

↪ **Theorem 5.2**

Let (X, d)-complete. Let $\{F_n\}$ be a decreasing family of non-empty closed sets with $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$. Then, $\exists z : \bigcap_{n\in\mathbb{N}} F_n = \{z\}$.

↪ **Theorem 5.3**

Let (X, d)-complete, and $f : X \to X$ an "expanding map", such that $d(x, y) \leqslant d(f(x), f(y)) \,\forall\, x, y \in X$. Then, f is a surjective isometry, ie, $f(X) = X$ and $d(f(x), f(y)) = d(x, y) \,\forall\, x, y \in X$.

↪ **Lemma 5.1**

Differentiable $\implies$ Continuous.

Proof. Let $f : I \to \mathbb{R}$, and $c \in I$ arbitrary. Notice that $\forall\, x \neq c \in I$, $f(x) - f(c) = (x - c)\frac{f(x)-f(c)}{x-c}$. Hence,

$$\begin{aligned}\lim_{x\to c}(f(x) - f(c)) &= \lim_{x\to c}(x - c)\frac{f(x) - f(c)}{x - c} \\ &= \lim_{x\to c}(x - c) \cdot \lim_{x\to c} \frac{f(x) - f(c)}{x - c} \\ &= 0 \cdot f'(x) = 0 \\ &\implies \lim_{x\to c} f(x) = f(c),\end{aligned}$$

hence f continuous, noting that the splitting of the limits is valid as both are defined. ■

⊛ **Example 5.1**

Let $f : \mathbb{R} \to \mathbb{R}, f(x) := \begin{cases} x^2 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$

Claim: f discontinuous at all $x \neq 0$.

Proof. Let $x \neq 0 \in \mathbb{R}$. By density of $\mathbb{Q} \subseteq \mathbb{R}$, there exist sequences $(r_n) \in \mathbb{Q}$ s.t. $r_n \to x$ and $(z_n) \in \mathbb{R} \setminus \mathbb{Q}$ s.t. $z_n \to x$. Then:

$$\lim_{n\to\infty} f(r_n) = \lim r_n^2 = x^2$$
$$\lim_{n\to\infty} f(z_n) = \lim 0,$$

hence f discontinuous by the sequential criterion at $x \neq 0$. ■

Claim: $f'(0) = 0$.

Proof. Let $\varepsilon > 0$ and $\delta = \varepsilon$. Notice that $f(x) \leqslant x^2 \,\forall\, x$. Then, we have that $\forall\ |x| < \delta$,

$$\begin{aligned} \left| \frac{f(x) - f(0)}{x - 0} - 0 \right| &= \left| \frac{f(x)}{x} \right| \\ &\leqslant \left| \frac{x^2}{x} \right| = |x| < \delta = \varepsilon \end{aligned}$$

■

↪ **Definition 5.1**

Let $f : I \to \mathbb{R}$. A point $c \in I$ is a local max (resp min) if $\exists \delta > 0$ s.t. $f(x) \leqslant f(c)$ (resp $f(x) \geqslant f(c)$) $\forall\, x \in (c - \delta, c + \delta) \cap I$.

↪ **Lemma 5.2**

Let $f : I \to \mathbb{R}$ be differentiable at $c \in I^\circ$. If c a local extrema of f, then $f'(c) = 0$.

Proof. Assume wlog that c a local max; if a local min, take $\tilde{f} := -f$ and continue.

Since I° open, $\exists \delta_1 > 0 : (c - \delta_1, c + \delta_1) \subseteq I^\circ \subseteq I$. We also have that $\exists \delta_2 > 0 : f(x) \leqslant f(c) \,\forall\, x \in (c - \delta_2, c + \delta_2) \cap I$, by c an extrema.

Let $\delta := \min\{\delta_1, \delta_2\}$. Then, we have both $(c - \delta, c + \delta) \subseteq I$ and $f(x) \leqslant f(c) \,\forall\, x \in (c - \delta, c + \delta)$.

Since $f'(c)$ exists, $\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c}$. But we have from the property of being a maximum that

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \geqslant 0, \qquad \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \leqslant 0,$$

hence, as these two limits must agree, they must equal 0 and thus $f'(c) = 0$. ■

5.2 Miscellaneous

⊛ **Example 5.2: Rudin, Chapter 7: Differentiability**

1. Let f be defined $\forall\, x \in \mathbb{R}$, and suppose that $|f(x) - f(y)| \leqslant (x-y)^2$, $\forall\, x, y \in \mathbb{R}$. Prove that f is constant.[10]

   Proof. Let $x > y \in \mathbb{R}$. Then, as $|x - y| = x - y$, we have

   $$|f(x) - f(y)| \leqslant (x-y)^2 \implies \left|\frac{f(x) - f(y)}{x - y}\right| \leqslant x - y = |x - y| \to 0 \text{ as } y \to x$$
   $$\implies \left|\frac{f(x) - f(y)}{x - y}\right| \to 0$$

   This implies, then, that $f'(x)$ is defined $\forall\, x \in \mathbb{R}$, and moreover, that $f'(x) = 0 \,\forall\, x \in \mathbb{R}$. We conclude, then, that $f(x)$ constant $\forall\, x \in \mathbb{R}$. ■

2. Suppose $f'(x) > 0$ in (a, b). Prove that f is strictly increasing in (a, b), and let g be its inverse function. Prove that g is differentiable, and that

   $$g'(f(x)) = \frac{1}{f'(x)} \quad (a < x < b).$$

   Proof. Fix $x > y \in (a, b)$. Then, by the mean value theorem, $\exists z \in (x, y) : f'(z) = \frac{f(x) - f(y)}{x - y}$. Since $f'(z) > 0$, it follows that

   $$\frac{f(x) - f(y)}{x - y} > 0 \implies f(x) - f(y) > x - y > 0 \implies f(x) > f(y),$$

   hence, f increasing, as $x > y$ arbitrary.

   Let now $g := f^{-1}$. ■

[10] Note that this means that f *Hölder continuous* with constant $\alpha = 2$. Indeed, Hölder continuous functions with $\alpha > 1$ are always constant by a similar proof. For $0 < \alpha \leqslant 1$, we have the inclusion continuously differentiable $\implies$ Lipschitz $\implies$ $\alpha-$Hölder $\implies$ uniformly continuous $\implies$ continuous.