Real Analysis Lecture Notes

MA224: Real Analysis, January –April, 2026

Rajesh Srivastava
Department of Mathematics, IIT Guwahati

Contents

Introduction

Real Analysis provides the rigorous foundations of calculus and, more broadly, of modern mathematical analysis. The guiding theme of the course is the systematic study of limiting processes: convergence of sequences and functions, continuity and differentiability defined through limits, and integration built upon measurable structures. Throughout, emphasis is placed on precise definitions, correct quantifiers, and logically complete proofs, together with carefully chosen examples and counterexamples that clarify the necessity of hypotheses and the sharpness of conclusions.

We begin with metric spaces (X, d), the natural setting in which convergence and continuity can be formulated beyond $\mathbb{R}$. We study open and closed sets, interior and closure, limit points, compactness, and the topological characterization of continuity. We then move to normed linear spaces $(V, \|\cdot\|)$, where algebraic and topological structures interact. A central concept is completeness, which ensures that every Cauchy sequence converges and underlies fundamental existence results such as the contraction mapping principle. Uniform convergence is treated as a key mode of convergence for sequences of functions, since it provides control strong enough to justify passing limits through continuous operations under appropriate assumptions. Classical inequalities, including Young's, Hölder's, and Minkowski's inequalities, are developed as essential tools for norm estimates and convergence arguments.

The second part focuses on functions on $\mathbb{R}^n$. After formalizing limits and continuity in Euclidean space, we study partial derivatives, directional derivatives, and differentiability in the Fréchet sense, where differentiability at a point means approximation by a linear map with a remainder term that is small compared with $\|h\|$. From this viewpoint we develop the multivariable chain rule and Taylor's theorem with remainder, which describe the local structure of smooth functions and provide quantitative error estimates. These results culminate in the inverse mapping theorem and the implicit function theorem, which give precise conditions for local invertibility of maps and for representing solution sets of equations $F(x, y) = 0$ as graphs of functions.

In the final part, we develop Lebesgue measure and integration to address the limitations of Riemann integration. We construct outer measure, define measurable sets using Carathéodory's criterion, and obtain Lebesgue measure on $\mathbb{R}$. Measurable functions are introduced via approximation by simple functions, leading to the definition of the Lebesgue integral for nonnegative functions and then for integrable functions. The principal convergence theorems—the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem—are proved and used to justify the interchange of limits and integrals in a principled way. Classical examples, including the Cantor set, illustrate null sets, non-measurable phenomena, and the distinction between pointwise and uniform convergences.

By the end of the course, students should be able to analyze convergence and continuity in metric and normed spaces, apply the main structural theorems of multivariable differentiability, and use Lebesgue measure and integration as foundational tools for further study in analysis, probability, and partial differential equations.

Chapter 1

Metric and Normed Linear Spaces

This chapter develops the basic language of analysis in abstract spaces. We introduce metrics and norms, discuss sequences and their convergence, and study the topology induced by a metric through open and closed sets, interior and closure. Completeness and Cauchy sequences lead to the key notion of a complete metric space, while density and continuity clarify how analytic structure behaves under mappings. Finally, uniform convergence and the contraction mapping principle (Banach fixed point theorem) provide powerful tools used repeatedly later; Young's, Hölder's, and Minkowski's inequalities are included as essential estimates connecting normed spaces to L^p-type analysis.

1.1 Syllabus map

This chapter is organized into three thematic parts:

(1) **Metric spaces and topology:** open and closed sets, interior and closure, dense subsets, continuity, compactness, and completeness.

(2) **Normed vector spaces:** norms, norm-induced metrics, and standard examples, together with basic inequalities.

(3) **Uniform convergence:** uniform convergence of sequences of functions and differentiation under the limit.

1.2 Metric spaces

Let X be a non-empty set. A map $d : X \times X \to \mathbb{R}^+ = [0, \infty)$ such that

(i) $d(x, y) = 0$ if and only if $x = y$, $\quad x, y \in X$.

(ii) $d(x, y) = d(y, x) \quad$ (symmetric).

(iii) $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality).

is called a *metric* on X, and the pair (X,d) is called a *metric space*.

Example 1.2.1. If $X = \mathbb{R}^n$, then for $x, y \in \mathbb{R}^n$,

1. $d_1(x,y) = \sum_{i=1}^n |x_i - y_i|$;
2. $d_2(x,y) = \left(\sum_{i=1}^n |x_i - y_i|^2\right)^{\frac{1}{2}}$;
3. $d_\infty(x,y) = \sup_{1 \leq i \leq n} |x_i - y_i|$;

   define metrics on $\mathbb{R}^n$.

Example 1.2.2. Let (X,d) be a metric space. Prove that $d'(x,y) = \min\{1, d(x,y)\}$ defines a metric.

Example 1.2.3. If $X = C[0,1]$, the space of continuous functions on $[0,1]$, then for $f, g \in X$,

$$d_\infty(f,g) = \sup_{0 \leq t \leq 1} |f(t) - g(t)|$$

defines a metric on X.

(*Hint:* f is continuous on $[0,1]$, so f is bounded and $|f(t) - h(t)| \leq |f(t) - g(t)| + |g(t) - h(t)|$.)

Example 1.2.4. If $X \neq \emptyset$, then for $x, y \in X$,

$$d_0(x,y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

defines a metric on X. This is called the *discrete metric* on X and (X, d_0) is called *discrete metric space*. Thus, every non-empty set has a metric.

Note that for $d(x,z) \leq d(x,y) + d(y,z)$ to hold, we need to verify three cases:

1. $x = y$, $y \neq z$.
2. $x \neq y$, $y = z$.
3. all of x, y, z are distinct.

Example 1.2.5. Let (X,d) be a metric space, then $\left(X, \dfrac{d}{1+d}\right)$ is also a metric space.

For this, consider, $f(t) = \dfrac{t}{1+t}$, $t \in [0, \infty)$. Then $f'(t) = \dfrac{1}{(1+t)^2} > 0$. Hence, f is a strictly increasing function and $f(0) = 0$. On the other hand

$$\frac{t+s}{1+t+s} < \frac{t}{1+t} + \frac{s}{1+s}$$

Put $t = d(x, y)$, $s = d(y, z)$. Then

$$t + s \geq d(x, z) \quad \text{and} \quad f \quad \text{is strictly increasing}$$

$$\implies f \circ d(x, z) \leq f(t + s) < \frac{t}{1+t} + \frac{s}{1+s} = f \circ d(x, y) + f \circ d(y, z).$$

Example 1.2.6. Let (X, d) be a metric space. Suppose and $f : [0, \infty) \to [0, \infty)$ be an increasing function such that $f(s + t) \leq f(s) + f(t)$ and $f(t) = 0$ if and only if $t = 0$. Then $f \circ d$ is a metric on X.

Example 1.2.7. Let H^∞ (*Hilbert cube*) be the space of sequences $x = (x_n) = (x_1, x_2, \ldots, x_n, \ldots)$ such that $|x_n| \leq 1$. Then

$$d(x, y) = \sum_{n=1}^{\infty} \frac{|x_n - y_n|}{2^n}$$

defines a *metric* on H^∞.

(i) $d(x, y) \leq \sum \frac{2}{2^n} < \infty$.

(ii) $|x_n - z_n| \leq |x_n - y_n| + |y_n - z_n|$

$$\implies \sum_{n=1}^{k} \frac{|x_n - z_n|}{2^n} \leq \sum_{n=1}^{k} \frac{|x_n - y_n|}{2^n} + \sum_{n=1}^{k} \frac{|y_n - z_n|}{2^n}$$

$$\leq d(x, y) + d(y, z) < \infty.$$

Since the left-hand side is an increasing sequence which is bounded above, it follows that

$$\lim_{k \to \infty} \sum_{n=1}^{k} \frac{|x_n - z_n|}{2^n} \leq d(x, y) + d(y, z)$$

$\implies d(x, z) \leq d(x, y) + d(y, z).$

Exercise 1.2.8. Prove that $d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$ defines a metric on $(0, \infty)$.

1.3 Normed linear spaces and fundamental inequalities

1.3.1 Normed linear spaces

Let X be a vector space over the field $\mathbb{R}$ or $\mathbb{C}$. A map $\|\cdot\| : X \to [0, \infty)$ is called a *norm* if.

(i) $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|\alpha x\| = |\alpha| \|x\|$, for all $x \in X$, for all $\alpha \in \mathbb{R}$ or $\mathbb{C}$.

(iii) $\|x + y\| \leq \|x\| + \|y\|$, for all $x, y \in X$.

If we write $d(x, y) = \|x - y\|$, then d is a metric on the vector space X. But all metric on a vector space cannot be obtained by norm.

Example 1.3.1. Let X be a vector space. Then the discrete metric cannot be induced by any norm on X.

For this, if so then $d_0(x, y) = \|x - y\|$. Then for $x \neq 0$,

$$\|x\| = d_0(x, 0) = 1 = d_0(\alpha x, 0) = \|\alpha x\| = |\alpha|\|x\|, \quad \forall \alpha.$$

However, if d is a metric on a vector space X such that $d(x, y) = d(x - y, 0)$ and $d(\alpha x, \alpha y) = |\alpha| d(x, y)$. Then $d(x, 0) = \|x\|$ defines a norm on X. That is,

(i) $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|\alpha x\| = |\alpha|\|x\|$.

(iii) $\|x + y\| = d(x + y, 0) = d(x, -y) \leq d(x, 0) + d(-y, 0)$.

Example 1.3.2. Let ℓ^1 denote the space of all the sequences of real (or complex) numbers such that $\sum_{n=1}^{\infty} |x_n| < \infty$. Then,

$$\|x\|_1 = \sum_{n=1}^{\infty} |x_n|$$

defines a norm on ℓ^1. The pair $(\ell^1, \|\cdot\|_1)$ is a normed linear space. For simplicity, we write ℓ^1 for $(\ell^1, \|\cdot\|_1)$.

(*Hint:* $\sum_{n=1}^{k} |x_n + y_n| \leq \sum_{n=1}^{k} |x_n| + \sum_{n=1}^{k} |y_n| \leq \|x\|_1 + \|y\|_1$.)

Example 1.3.3. ℓ^2 denotes the space of all sequences on $\mathbb{R}$ (or $\mathbb{C}$) such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$. Define

$$\|x\|_2 := \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{\frac{1}{2}}$$

defines a norm on ℓ^2.

(*Hint:* $\sum_{n=1}^{k} |x_n + y_n|^2 \leq ((\sum_{n=1}^{k} |x_n|)^{\frac{1}{2}} + (\sum_{n=1}^{k} |y_n|)^{\frac{1}{2}})^2$.)

Example 1.3.4. ℓ^∞ = space of all sequences on $\mathbb{R}$ (or $\mathbb{C}$) such that $\sup_{n \in \mathbb{N}} |x_n| < \infty$. The function

$$\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$$

defines a norm on ℓ^∞.

Example 1.3.5. c_0 = space of all sequences on $\mathbb{R}$ (or $\mathbb{C}$) such that $\lim_{n \to \infty} x_n = 0$ Then (x_n) must be bounded. Hence

$$\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n| < \infty.$$

Thus, $(c_0, \|\cdot\|_\infty)$ is a normed linear space.

Exercise 1.3.6. If $x = (x_1, x_2, \dots, x_n) \subseteq \mathbb{R}^n$ (or $\mathbb{C}^n$), then

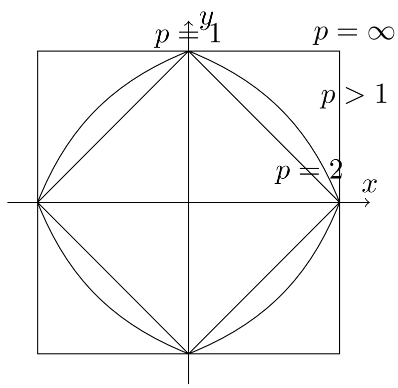
$$\|x\|_\infty \le \|x\|_1 \le \sqrt{n}\|x\|_2 \le n\|x\|_\infty.$$

1.3.2 Geometry of Spheres in $(\mathbb{R}^n,\ \|\cdot\|_p)$

For $0 \le p \le \infty$ and $x \in \mathbb{R}^n$, write

$$\|x\|_p = \left(\sum |x_i|^p\right)^{1/p}.$$

Then $\|\cdot\|_p$ is a norm for $1 \le p < \infty$, and for $0 < p < 1$, $\|x\|_p^p = d_p(0, x)$ with $d_p(x, y) = \|x - y\|_p^p$ is a metric. (We see later).
Let $S_1^p(0) = \{x : d_p(0, x) = 1\}$. Then the following figure can be plotted for different values of $p; 0 < p < \infty; p = \infty$.



Shapes for $0 < p < 1$ would look like star-shaped curves (not shown).

Exercise 1.3.7. If $x = (x_n) \in \ell^1$, then $x \in \ell^\infty$.

$$\sum_{n=1}^{\infty} |x_n|^2 < \sum_{n=1}^{\infty} \|x\|_\infty |x_n| \implies \|x\|_2 \le \|x\|_\infty \|x\|_1.$$

Thus, $\ell^1 \subsetneq \ell^2 \subsetneq c_0 \subsetneq \ell^\infty$.

Exercise 1.3.8. If $1 < p < \infty$, then for $\sum_{n=1}^{\infty} |x_n|^p < \infty$, one can define a norm $\|\cdot\|_p$ on ℓ^p via

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}.$$

To prove this, we need some inequalities.

1.3.3 Young's inequality

Let $1 < p < \infty$ and $a, b > 0$. Then for $\frac{1}{p} + \frac{1}{q} = 1$, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. *Proof:* Let $y = x^{p-1}$,then $x = y^{q-1}$ (since $p - 1 = \frac{1}{q-1}$ by $\frac{1}{p} + \frac{1}{q} = 1$). Now, it is clear that

$$ab \leq \int_0^a x^{p-1}\,dx + \int_0^b y^{q-1}\,dy = \frac{a^p}{p} + \frac{b^q}{q}.$$

Note that equality in $(*)$ holds if and only if $a^p = b^q$ (or $a = b^{q-1}$). For this, consider

$$ab = \frac{a^p}{p} + \frac{b^q}{q}, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

Replace $a \to a^{\frac{1}{p}}$, $b \to b^{\frac{1}{q}}$ and $\frac{1}{p} = \alpha$. Then, we get

$$a^\alpha b^{1-\alpha} = \alpha a + (1 - \alpha) b$$

or

$$t^\alpha - \alpha t - (1 - \alpha) = 0 \quad \text{if} \quad t = a/b.$$

Let

$$f(t) = t^\alpha - \alpha t - (1 - \alpha), \quad t \in (0, \infty).$$

Then $f(1) = 0$ and

$$f'(t) = \alpha t^{\alpha-1} - \alpha = \alpha(t^{\alpha-1} - 1) = 0 \iff t = 1.$$

Since $f'(t) < 0$ if $t > 1$ and $f'(t) > 0$ for $0 < t < 1$. Hence, f is strictly increasing in $(0, 1)$ and strictly decreasing in $(1, \infty)$. Thus, $t = 1$ is the point of absolute maximum of f. Therefore, $f(t) \leq f(1) = 0$, which is another proof of the inequality. On the other hand, $f(t) = 0$ if and only if $t = 1$. This completes the proof.

1.3.4 Hölder's inequality

Let $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then for $x \in \ell^p$ and $y \in \ell^q$, it follows that

$$x \cdot y (= x_1 y_1 + \ldots + x_n y_n + \ldots) \in \ell^1,$$

and

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q \qquad \cdots (*)$$

(where $\frac{1}{\infty} = 0$ adopted.) When $p = 1$, $q = \infty$. In this case $(*)$,

$$\|x \cdot y\|_1 = \sum_{i=1}^{\infty} |x_i y_i| \leq \sum |x_i| \cdot \sup |y_i| = \|x\|_1 \|y\|_\infty.$$

Now, let $1 < p < \infty$, then $1 < q < \infty$. Substitute $a = a_j = \frac{|x_j|}{\|x\|_p}$ and $b = b_j = \frac{|y_j|}{\|y\|_q}$ in the Young's Inequality. Then

$$\sum_{j=1}^{n} \frac{|x_j y_j|}{\|x\|_p \, \|y\|_q} \leq \sum_{j=1}^{n} \left(\frac{|x_j|^p}{p \, \|x\|_p^p} + \frac{|y_j|^q}{q \, \|y\|_q^q} \right) = \frac{1}{p} \sum_{j=1}^{n} \frac{|x_j|^p}{\|x\|_p^p} + \frac{1}{q} \sum_{j=1}^{n} \frac{|y_j|^q}{\|y\|_q^q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

That is,

$$\sum_{j=1}^{n} |x_j y_j| \leq \|x\|_p \|y\|_q, \quad \text{for all} \quad n \geq 1$$

Since the left-hand side is an increasing sequence which is bounded above, hence

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q.$$

Notice that if $\|x\|_p = 1 = \|y\|_q$, then $\|x \cdot y\|_1 \leq 1$, and equality holds if and only if $|x_j|^p / \|x\|_p^p = |y_j|^q / \|y\|_q^q$ for all j (equivalently, $a_j^p = b_j^q$).
This follows from Young's equality. For

$$ab = \frac{a^p}{p} + \frac{b^q}{q},$$

we must have $a^p = b^q$.

1.3.5 Minkowski's inequality

Let $1 \leq p \leq \infty$. Then for $x, y \in \ell^p$, $x + y \in \ell^p$, and $\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad (*)$

Proof. For $p = 1$ or ∞, the proof is trivial. Let $1 < p < \infty$. Then

$$\begin{aligned} \|x + y\|_p &= \left(\sum_{i=1}^{\infty} |x_i + y_i|^p \right)^{1/p} \\ &\leq \left(\sum_{i=1}^{\infty} (|x_i| + |y_i|)^p \right)^{1/p}. \end{aligned} \tag{1}$$

Since

$$(|x_i| + |y_i|)^p = (|x_i| + |y_i|)(|x_i| + |y_i|)^{p-1}.$$

By Hölder's inequality,

$$\sum (|x_i| + |y_i|)^{p-1} |x_i| \leq \left(\sum (|x_i| + |y_i|)^{(p-1)q} \right)^{1/q} \left(\sum |x_i|^p \right)^{1/p}.$$

Thus,

$$\sum (|x_i| + |y_i|)^p \leq \left(\sum (|x_i| + |y_i|)^p \right)^{1/q} (\|x\|_p + \|y\|_p).$$

That is

$$\left(\sum(|x_i|+|y_i|)^p\right)^{1-\frac{1}{q}} \leq \|x\|_p + \|y\|_p.$$

From (1), we get

$$\|x+y\|_p \leq \left(\sum(|x_i|+|y_i|)^p\right)^{1/p} \leq \|x\|_p + \|y\|_p.$$

□

Note that as similar to above cases, it can be shown that equality in (∗) holds if and only if

$$x = \frac{\|x\|_p}{\|y\|_p} y.$$

Now, if $x, y \in \ell^p$, then $x+y \in \ell^p$. Because $a, b > 0$, $(a+b)^p \leq \{2\max\{a,b\}\}^p$ that is, $(a+b)^p \leq 2^p(a^p+b^p)$, and so,

$$\sum |x_j+y_j|^p \leq 2^p(\sum |x_j|^p + \sum |y_j|^p) < \infty.$$

Thus, ℓ^p is closed under $\|\cdot\|_p$. Hence $(\ell^p, \|\cdot\|_p)$ is a normed linear space.

Theorem 1.3.9. *If $f, g \in \mathcal{R}[a,b]$, then for $\|f\|_p = (\int |f|^p)^{\frac{1}{p}}$, we get*

$$\begin{aligned} &(i) \quad \|fg\|_1 \leq \|f\|_p \|g\|_q \quad \textit{where} \quad \frac{1}{p} + \frac{1}{q} = 1 \\ &(ii) \quad \|f+g\|_p \leq \|f\|_p + \|g\|_p, \quad 1 \leq p < \infty \end{aligned}$$

For $p = \infty$,

$$\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|, \ \textit{where } f \in \mathcal{R}[a,b].$$

Then $(\mathcal{R}[a,b], \|\cdot\|_\infty)$ is a normed linear space.

Definition 1.3.10. (Open and Closed balls):

(i) $B_r(x_0) = \{y \in X : d(x_0, y) < r\}$ is called open ball.

(ii) $\overline{B_r(x_0)} = \{y \in X : d(x_0, y) \leq r\}$ is called closed ball.

1.3.6 Open sets in metric spaces

Definition 1.3.11. A set $O \subset (X, d)$ is said to be open if for all $x \in O$, there exists $r > 0$ such that $B_r(x) \subset O$.

Proposition 1.3.12. *If $\{O_i : i \in I\}$, I is any index set. Then*

(i) $\bigcup_{i\in I} O_i$ is open (arbitrary union of open sets is open).

(ii) $\bigcap_{i=1}^n O_i$ is open (finite intersection of open sets is open).

Remark. Arbitrary intersection of open sets need not *be open.*

Example 1.3.13. $X = \mathbb{R}$, $u(x,y) = |x-y|$. $\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$ is not open.

Example 1.3.14. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then $A = \{x \in \mathbb{R} : f(x) > 0\}$ is open.

Proof. Let $x \in A \implies f(x) > 0$. For $\varepsilon = f(x) > 0$, there exists $\delta > 0$ such that for all $y \in (-\delta, \delta) + x = (x-\delta, x+\delta)$,

$$|f(y) - f(x)| < f(x).$$

$$\implies 0 < f(y) < 2f(x), \quad \forall\, y \in (x-\delta, x+\delta).$$

Hence $(x-\delta, x+\delta) \subset A \implies A$ is open. □

Open Sets in $\mathbb{R}$:

A countable union of open intervals is an open set.On the other hand, any open set in $\mathbb{R}$ can be written as a countable union of disjoint open intervals.

Theorem 1.3.15. *Let O be an open set in $\mathbb{R}$, then there exists a unique disjoint family of countably many open intervals I_n such that*

$$O = \bigcup_{n=1}^{\infty} I_n$$

Proof. Since O is open, for $x \in O$, there exists an open interval (a,b) such that $x \in (a,b) \subset O$. Now, we extract the largest open interval containing x and contained in O. Let $a_x = \inf\{a : (a,x] \subset O\}$, and $b_x = \sup\{b : [x,b) \subset O\}$. Then $I_x = (a_x, b_x)$ will be the largest open interval containing x and contained in O.

Note that $I_x = (a_x, b_x) \subset O$. For this, let $a_x < z < b_x$, then $a_x < z - \epsilon$ for small $\epsilon > 0$ $\implies a_x + \epsilon < z$. But by definition of infimum, $\exists\, a < a_x + \epsilon$ and $(a,x] \subset O \implies (a_x + \epsilon, x] \subset O$.

Similarly, $[x, b_x - \epsilon) \subset O \implies (a_x + \epsilon, b_x - \epsilon) \subset O$ for small $\epsilon > 0 \implies (a_x, b_x) \subset O$.

Now, if $x, y \in O$ and $x \neq y$ then either $I_x \cap I_y = \emptyset$ or $I_x = I_y$.

If $I_x \cap I_y \neq \emptyset$, then $I_x \cup I_y$ is an open interval containing x and y.

Therefore, by maximality of I_x for x and I_y for y, it follows that $I_x \cup I_y \subseteq I_x \implies I_y \subseteq I_x$ Since $y \in I_y \implies I_y = I_x$ ($\because I_y$ is maximal).

Now, $O = \bigcup_{x \in O} I_x$. Since I_x and I_y are disjoint (if $x \neq y$), we can assign a distinct rational to each of them. That is, choose $r_x \in I_x$ and $r_y \in I_y$. Then $r_x \neq r_y$.

Thus,

$$\{I_x : x \in O\} \xrightarrow{1-1} \mathbb{Q} \text{ (set of rationals) via } I_x \mapsto r_x$$

Hence,

$$O = \bigcup_{i=1}^{\infty} I_{r_i} \tag{1}$$

The representation (1) is unique. Let $O = \bigcup_{n=1}^{\infty} I_n = \bigcup_{m=1}^{\infty} J_m$.
Then $I_n = I_n \cap O = \bigcup_{m=1}^{\infty}(I_n \cap J_m)$. Since $\{I_n \cap J_m : m \in \mathbb{N}\}$ is a disjoint family and I_n is an open interval, $I_n \subset I_n \cap J_{m_0}$ for some m_0. But then $I_n \subset J_{m_0}$, and given I_n is maximal, $\implies I_n = J_{m_0}$. Thus, the representation (1) is unique upto change in order of union. □

Definition 1.3.16. (Convergent Sequence):

A sequence $(x_n) \in (X, d)$ is said to be convergent if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ and $x_0 \in X$ such that $n \geq N \implies d(x_n, x_0) < \epsilon \iff x_n \in B_\epsilon(x_0)$, $\forall n \geq N$.

Definition 1.3.17. (Cauchy Sequence):

A sequence $(x_n) \in (X, d)$ is said to be a Cauchy sequence if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $m, n \geq N \implies d(x_n, x_m) < \epsilon$

Example 1.3.18. Let $X = (0, 1)$ and $d(x, y) = |x - y|$. Then $\{\frac{1}{n}\}$ is a Cauchy sequence because

$$|x_n - x_m| = \left|\frac{1}{n} - \frac{1}{m}\right| \to 0 \text{ as } n, m \to \infty.$$

But $\lim x_n = 0 \notin X$. Hence not convergent.

However, every convergent sequence is a Cauchy sequence.

Definition 1.3.19. A set $A \subseteq (X, d)$ is said to be bounded if $\exists x_0 \in X$ and $M > 0$ such that $d(a, x_0) \leq M, \quad \forall a \in A \iff a \in \overline{B_M(x_0)}, \forall a \in A$. that is, A is bounded if and only if A is contained in a ball.

Example 1.3.20. The set $\{(x, y) : y = \sin(1/x),\ x \neq 0\} \cup (\{0\} \times [-1, 1])$ is unbounded, since it contains points $(n, \sin(1/n))$ whose Euclidean norm tends to infinity as $n \to \infty$.

Proposition 1.3.21. *Every Cauchy sequence is bounded.*

Proof. Since $(x_n) \subset (X, d)$ is a Cauchy sequence, for $\epsilon = 1$, $\exists N \in \mathbb{N}$ such that

$$d(x_m, x_n) < 1, \quad \forall m, n \geq N.$$

So $d(x_n, x_N) < 1, \quad \forall n \geq N$. Let $M = \max\{1, d(x_i, x_N) : i = 1, 2, \ldots, N-1\}$. Then $d(x_n, x_N) \leq M, \forall n \geq 1 \implies x_n \in \overline{B_M(x_N)}$. □

But converse need not be true. For $(\mathbb{R}, u)$, u = usual metric. $x_n = \{-1, 1, -1, 1, \ldots\}$ is bounded but not Cauchy sequence.

Proposition 1.3.22. *Let (x_n) be a Cauchy sequence in (X, d). If (x_{n_k}) is a subsequence which converges to x. Then $x_n \to x$.*

Proof. For $\varepsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that

$$d(x_n, x_m) < \frac{\varepsilon}{2}, \quad \forall\, n, m \geq N_1.$$

Also, for the same $\varepsilon > 0$, there exists $N_2 \in \mathbb{N}$ such that

$$d(x_{n_k}, x) < \frac{\varepsilon}{2}, \quad \forall\, n_k \geq N_2.$$

Let $N = \max\{N_1, N_2\}$. Then

$$d(x_n, x_m) < \frac{\varepsilon}{2} \quad \text{and} \quad d(x_{n_k}, x) < \frac{\varepsilon}{2} \quad \text{for all } n, m, n_k \geq N.$$

$$\implies d(x_{n_k}, x_m) < \frac{\varepsilon}{2}, \quad \forall\, n_k, m \geq N.$$

Thus,

$$d(x, x_m) \leq d(x, x_{n_k}) + d(x_{n_k}, x_m) < \varepsilon \quad \text{for all } m \geq N.$$

Hence, $x_m \to x$. □

Remark. If $X = (0,1)$ and $d(x,y) = |x-y|$. Then $x_n = \frac{1}{n}$ is a Cauchy sequence, but it has no convergent subsequence.

1.3.7 Closed sets in metric spaces

Definition 1.3.23. A set $F \subset (X,d)$ is said to be closed if F^c is open. that is, for all $x \in F^c = X \setminus F$, $\exists\, \epsilon > 0$ such that $B_\epsilon(x) \subseteq F^c$.
On the other hand, if for each $\epsilon > 0$, $B_\epsilon(x) \cap F \neq \emptyset \implies x \in F$.

Example 1.3.24. The set $A = \{(x,y) : y = \sin\frac{1}{x},\, x \neq 0\}$ is neither open nor closed set in $\mathbb{R}^2$. If $x_n = \frac{1}{n\pi} \neq 0$, $(x_n, y_n) = (\frac{1}{n\pi}, 0) \in A$, but $\lim_{n\to\infty}(x_n, y_n) = (0,0) \notin A$ Since any ball $B_{\frac{1}{n}}(\frac{1}{\pi}, 0) \nsubseteq A \implies A$ is not open in $\mathbb{R}^2$.

Theorem 1.3.25. *Let (X,d) be a metric space and $F \subset X$. Then the following are equivalent (F.A.E):*

1. *F is a closed set (F^c open).*
2. *$\forall \epsilon > 0$, $B_\epsilon(x) \cap F \neq \emptyset \implies x \in F$.*
3. *$\forall$ sequence $(x_n) \in F$ such that $x_n \to x \implies x \in F$.*

Proof. (1) $\implies$ (2): Suppose F is closed. *Claim:* $B_\epsilon(x) \cap F \neq \emptyset$, $\forall \epsilon > 0 \implies x \in F$.
Notice that if $x \notin F \implies x \in F^c$ and F^c is open $\implies \exists \epsilon_0 > 0$ such that

$$B_{\epsilon_0}(x) \subset F^c \implies B_{\epsilon_0}(x) \cap F = \emptyset,$$

which is a contradiction.
(2) $\implies$ *(3):* Let $(x_n) \subset F$ and $x_n \to x$. Then for each $\epsilon > 0$, $x_n \in B_\epsilon(x)$ for all $n \geq n_0$.

$$\implies x_n \in B_\epsilon(x) \cap F \neq \emptyset, \quad \forall \epsilon > 0 \implies x \in F$$

(3) $\implies$ *(1): Claim:* F^c is open. Suppose F^c is not open. Then there exists $x \in F^c$ such that for each $n \in \mathbb{N}$, there will be $x_n \in F$ and $d(x_n, x) < \frac{1}{n}$. By (3), $x \in F$, which is a contradiction. ☐

Example 1.3.26. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then $A = \{x : f(x) = 0\}$ is closed.
Since $x_n \in A$ and $x_n \to x$. So $f(x_n) = 0, \quad \forall n \geq 1 \implies \lim f(x_n) = 0 \implies f(x) = 0$.

1.3.8 Interior points and interior of a set

Let $A \subset X$. Then interior(A) or Int(A) or A° is the largest open set contained in A. That is,

$$A^\circ = \bigcup \{O \subset X : O \text{ open}, O \subseteq A\}$$

$$= \bigcup \{B_\epsilon(x) \subset A : \text{ for } x \in A \text{ and some } \epsilon > 0\} = \text{ union of all open balls contained in } A.$$

1.3.9 Closure and limit points

Let $A \subset (X, d)$. The closure of A or $cl(A)$ or $\overline{A}$ is the smallest closed set containing A. That is,

$$\overline{A} = \bigcap \{F \subset X : F \text{ closed and } A \subset F\}$$

$$= \{x \in X : \exists x_n \in A \text{ with } x_n \to x\}$$

= collection of limits of all convergent sequences in A (limit need not be in the set A).

Example 1.3.27. $A = \left\{(n, \frac{1}{n}) : n \in \mathbb{N}\right\}$. Then closure of A in $(\mathbb{R}, u)$ is $\overline{A} = A$ and $A^\circ = \emptyset$ (Why?).

Example 1.3.28.

1. $A = \{(x, y) : |x| < 1, |y| < 1\}$. Then
$\overline{A} = \{(x, y) : |x| \leq 1, |y| \leq 1\}$.
2. $A = \{(x, y) : y = \sin\left(\frac{1}{x}\right), x \neq 0\}$. Then
$\overline{A} = \{(x, y) : y = \sin\left(\frac{1}{x}\right), x \neq 0\} \cup (\{0\} \times [-1, 1])$.

Example 1.3.29. Let c_{00} = space of all sequences having finitely many non-zero terms.

$$c_{00} = \{x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) : x_i \in \mathbb{R}\}$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i| < \infty.$$

$$\implies c_{00} \subsetneq \ell^{\infty} \quad \text{(proper subspace)}.$$

Let

$$X^n = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots\right) \in c_{00}.$$

Let

$$X = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots\right) \in \ell^{\infty}.$$

then

$$\|X - X^n\|_{\infty} = \sup_{k \geq n} \frac{1}{k+1} = \frac{1}{n+1} \to 0.$$

but $X \notin c_{00}$. Hence c_{00} is not a closed subspace of ℓ^{∞}. In addition c_{00} is not open in ℓ^{∞}. For this, let $\epsilon > 0$ be arbitrary and consider the sequence $y = (\frac{\epsilon}{2}, \frac{\epsilon}{4}, \frac{\epsilon}{8}, \ldots) \in \ell^{\infty}$. Then $\|y\|_{\infty} = \frac{\epsilon}{2} < \epsilon$, so $y \in B_{\epsilon}(0)$, but $y \notin c_{00}$. Therefore, $B_{\epsilon}(0) \nsubseteq c_{00}$ for any $\epsilon > 0$.
For $1 \leq p < \infty$, $c_{00} \subsetneq \ell^p$ and c_{00} is neither closed nor open in ℓ^p. For this, let

$$x_n = \left(\frac{\epsilon^p}{2^{n+1}}\right)^{1/p}, \ 1 \leq p < \infty,$$

and consider $x = (x_1, x_2, \ldots) \in \ell^p$. Then $x \notin c_{00}$ and

$$\|x\|_p^p = \sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} \frac{\epsilon^p}{2^{n+1}} = \frac{\epsilon^p}{2},$$

so $\|x\|_p = \frac{\epsilon}{2^{1/p}} < \epsilon$. Hence $x \in B_{\epsilon}(0)$, and therefore $B_{\epsilon}(0) \nsubseteq c_{00}$ for any $\epsilon > 0$. Consequently, c_{00} is not open in ℓ^p.
To see that c_{00} is not closed in ℓ^p, let $X^n = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) \in c_{00}$. Then $X^n \to x$ in ℓ^p, since

$$\|X^n - x\|_p^p = \sum_{k=n+1}^{\infty} \frac{\epsilon^p}{2^{k+1}} \to 0 \quad \text{as} \quad n \to \infty.$$

But $x \notin c_{00}$.

Proposition 1.3.30. *Let $A \subset (X, d)$. Then $x \in \overline{A}$ if and only if $B_{\epsilon}(x) \cap A \neq \varnothing$, for all $\epsilon > 0$.*

Proof. Let $x \in \overline{A}$. Suppose $\exists\, \epsilon_0 > 0$ such that $B_{\epsilon_0}(x) \cap A = \varnothing$. Then $A \subset (B_{\epsilon_0}(x))^c$, a closed set. By definition of $\overline{A}$, $\overline{A}$ is the smallest closed set containing A. Hence, $\overline{A} \subset (B_{\epsilon_0}(x))^c$. Since $x \in \overline{A}$, but $x \notin (B_{\epsilon_0}(x))^c$, this is a contradiction.

Conversely, suppose $B_{\epsilon}(x) \cap A \neq \varnothing$ for all $\epsilon > 0$. By the previous result, $x \in \overline{A}$ (since $\overline{A}$ is closed). □

Proposition 1.3.31. *$x \in \overline{A}$ if and only if there exists a sequence (x_n) with $x_n \in A$ such that $x_n \to x$.*

Proof. If $x \in \overline{A}$, then for all $n \in \mathbb{N}$, $B_{1/n}(x) \cap A \neq \varnothing$. So, $\exists\, x_n \in B_{1/n}(x) \cap A$. Thus,

$$d(x_n, x) < \frac{1}{n}, \forall n \in \mathbb{N} \implies x_n \to x.$$

Conversely, if there exists $x_n \in A$ with $x_n \to x$. Then for $\epsilon > 0$, $\exists n_0 \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq n_0$, $\implies x_n \in B_\epsilon(x) \cap A \neq \varnothing$ for all $\epsilon > 0$. Thus $x \in \overline{A}$ (by previous result). $\square$

1.4 Complete metric spaces

We have seen that there are Cauchy sequences whose limits need not necessarily belong to the space.
For example, the sequence $\frac{1}{n} \in ((0,1), u)$ under the usual metric, is a Cauchy sequence but the limit $\frac{1}{n} \to 0 \notin (0,1)$.

It is always possible to enlarge the space so that limits of all Cauchy sequences can be accommodated. This process is known as the completion of metric spaces, we shall see later. However, there are many spaces which do accommodate limits of their Cauchy sequences.

Definition 1.4.1. A metric space (X, d) is called complete if every Cauchy sequence in X has its limit in X.

Example 1.4.2. $(\mathbb{R}, u)$ is a complete space.

Let (x_n) be a Cauchy sequence in $\mathbb{R}$. Then it is bounded. And by the Bolzano–Weierstrass theorem, there exists a subsequence $x_{n_k} \to x \in \mathbb{R}$. For any $\epsilon > 0$, there exists a natural number k_0 such that

$$|x_{n_k} - x| < \epsilon \quad \text{for all} \quad k \geq k_0 \tag{1}$$

But the sequence (x_n) is Cauchy, so for all $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $|x_n - x_m| < \epsilon$ for all $n, m \geq n_0$. Let $m \geq n_0$ and $m \geq n_{k_0}$. Then

$$|x_n - x_{n_k}| < \epsilon \quad \text{for any } n \geq n_0 \text{ and } k \geq k_0. \tag{2}$$

From (1) and (2), it follows that:

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < 2\epsilon$$

for $n \geq n_0$ and $n_k \geq n_{k_0}$. Thus, for $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$n \geq n_0 \implies |x_n - x| < \epsilon.$$

Notice that the above discussion can be used to prove the following result.

Proposition 1.4.3. *Let (x_n) be a Cauchy sequence in a metric space (X, d). If (x_n) has a convergent subsequence $x_{n_k} \to x$, then $x_n \to x$. (Proof is similar to the above.)*

Example 1.4.4. $(\mathbb{R}, d)$ with $d(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$ is incomplete.

(*Hint:* $x_n = \tan \frac{\pi}{2}\left(\frac{n}{n+1}\right)$ is Cauchy, but not converging to a point in $\mathbb{R}$.

Example 1.4.5. Every discrete metric space is complete.

Let $X \neq \emptyset$, and $d_0(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$

Suppose $(x_n) \subset X$ is Cauchy. Then for $\epsilon > 0$, $\exists N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq N$.

Now, $d_0(x_n, x_m) = \begin{cases} 0 & \text{if } 0 < \epsilon \leq 1 \\ 0 \quad \text{or} \quad 1 & \text{if } \epsilon > 1. \end{cases}$

But if $d_0(x_n, x_m) = 1$ for only finitely many $n, m > N$ (for some $\epsilon > 1$), then

$$\lim_{n,m\to\infty} d_0(x_n, x_m) = 1 \neq 0 \quad \text{(Why?)}$$

Thus, for all $\epsilon > 0$, $\exists N' \in \mathbb{N}$ such that $d(x_n, x_m) = 0$, for all $n, m \geq N'$.
that is, $(x_n) = (x_1, x_2, \ldots, x'_N, x, x, \ldots) \to x$.
(Thus, every Cauchy sequence in (X, d_0) is eventually constant.)

Example 1.4.6. $(\mathbb{R}^n, \|\cdot\|_p)$ is complete for $1 \leq p \leq \infty$.

Let $1 \leq p < \infty$, and $x^k = (x_1^k, \ldots, x_n^k)$ be a Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_p)$. Then for $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that for all $k, l \geq k_0$,

$$\|x^k - x^l\|_p = \left(\sum_{j=1}^{n} |x_j^k - x_j^l|^p\right)^{1/p} < \epsilon$$

$$\implies |x_j^k - x_j^l| < \epsilon \quad \text{for all } k, l \geq k_0$$

$$\implies (x_j^k) \text{ is a Cauchy sequence in } (\mathbb{R}, u).$$

Hence $x_j^k \to x_j$ for all j. Then for $\epsilon > 0$, there exists $m_j \in \mathbb{N}$ such that $k \geq m_j \implies |x_j^k - x_j| < \epsilon$. Let $m_0 = \max_j\{m_j\}$. Then, for $x = (x_1, \ldots, x_n)$,

$$\|x^k - x\|_p < \epsilon \qquad \text{for } k \geq m_0.$$

Notice that the case $p = \infty$ is similar. We skip its proof here.

Example 1.4.7. Let $1 \leq p \leq \infty$. Then $(\ell^p, \|\cdot\|_p)$ is complete.

Let $1 \leq p < \infty$, and let $x^k = (x_1^k, x_2^k, \ldots)$ be a Cauchy sequence in $(\ell^p, \|\cdot\|_p)$. Then for $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $\forall k, l \geq n_0 \implies \|x^k - x^l\|_p < \epsilon$

$$\implies \sum_{j=1}^{n} |x_j^k - x_j^l|^p < \epsilon^p \tag{1}$$

For each fixed n, this reduces to $(\mathbb{R}^n, \|\cdot\|_p)$, which we know is complete. Hence $x_j^k \to x_j$; $j = 1, 2, \ldots, n$. Thus, letting $k \to \infty$ in (1), it follows that

$$\sum_{j=1}^{n} |x_j^l - x_j|^p < \epsilon^p, \quad \forall l \geq n_0 \tag{2}$$

But the left-hand side of (2) is an increasing sequence and bounded above, hence, letting $n \to \infty$, we get

$$\sum_{j=1}^{\infty} |x_j^l - x_j|^p < \epsilon^p$$

$$\|x^l - x\|_p \leq \epsilon, \quad \forall l \geq n_0$$

where $x = (x_1, x_2, \ldots, x_n, \ldots)$. Notice that

$$\|x\|_p \leq \|x - x^{n_0}\|_p + \|x^{n_0}\|_p < \epsilon + \|x^{n_0}\|_p < \infty \implies x \in \ell^p.$$

Proposition 1.4.8. *Every closed subset of a complete metric space is complete.*

Proof. Let F be a closed subset of a complete metric space (X, d). Then $(x_n) \subset F$ is a Cauchy sequence, it follows that (x_n) is a Cauchy sequence in X. Hence $x_n \to x \in X$. But F is closed, it implies that $x \in F$.
In fact, if (X, d) is complete, then F is closed if and only if F is complete. (*Hint:* it follows easily.) □

Example 1.4.9. Show that $(c_0, \|\cdot\|_\infty)$ is a proper closed subspace of $(\ell^\infty, \|\cdot\|_\infty)$.

We know that $c_0 \subsetneq \ell^\infty$. Now, let $x^k = (x_1^k, \ldots, x_j^k, \ldots)$ be a sequence in c_0 such that $x^k \to x = (x_1, \ldots, x_j, \ldots)$. That is, for every $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that $\forall k > k_0 \implies \|x^k - x\|_\infty < \epsilon$ which implies

$$|x_j^k - x_j| < \epsilon \quad \text{for each } j \geq 1\text{and} \quad \forall k > k_0. \tag{1}$$

Since $x_j^k \in c_0 \implies \lim_{j\to\infty} x_j^k = 0$ for each k. For $\epsilon > 0$, there exists $j_0 \in \mathbb{N}$ such that

$$|x_j^k| < \epsilon \quad \forall j \geq j_0 \quad \text{and} \quad k \geq k_0. \tag{2}$$

It follows from (1) and (2) that

$$|x_j| < |x_j^{k_0} - x_j| + |x_j^{k_0}| < 2\epsilon \quad \forall j > J_0,$$

i.e., $|x_j| < 2\epsilon$ for all $j > J_0$, which means $\lim_{j\to\infty} x_j = 0$. Hence c_0 is a closed subspace of ℓ^∞. Thus, c_0 is complete in its own right.

Example 1.4.10. The space $(C[a,b], \|\cdot\|_\infty)$ is a complete normed linear space.

Let (f_n) be a Cauchy sequence in $(C[a,b], \|\cdot\|_\infty)$. Then for $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $\forall n, m \geq n_0 \implies \|f_n - f_m\|_\infty < \epsilon$ which implies

$$|f_n(t) - f_{n_0}(t)| < \epsilon \quad \forall n \geq n_0, \forall t \in [a,b]. \tag{1}$$

So $(f_n(t))$ is a Cauchy sequence in $(\mathbb{R}, u)$ for each fixed $t \in [a,b]$. Hence $f_n(t) \to f(t)$.
Letting $n \to \infty$ in (1), we get $|f(t) - f_{n_0}(t)| \leq \epsilon \quad \forall t \in [a,b]$. (Notice that n_0 is free of choice of t). Since f_{n_0} is continuous, for each fixed t and $\epsilon > 0$, there exists $\delta > 0$ such that $|s-t| < \delta$ implies $|f_{n_0}(s) - f_{n_0}(t)| < \epsilon$. Hence,

$$\begin{aligned} |f(s) - f(t)| &< |f(s) - f_{n_0}(s)| + |f_{n_0}(s) - f_{n_0}(t)| + |f_{n_0}(t) - f(t)| \\ &< 3\epsilon \end{aligned}$$

So f is continuous on $[a,b]$.
However, the space $(C[a,b], \|\cdot\|_1)$ is not complete. For this, we consider the following: Consider

$$f_n(t) = \begin{cases} nt & 0 \leq t < \frac{1}{n} \\ 1 & \frac{1}{n} \leq t \leq 1 \end{cases}$$

It is easy to see that for $\frac{1}{m} < \frac{1}{n}$,

$$\|f_n - f_m\|_1 = \left(\int_0^{1/m} + \int_{1/m}^{1/n} + \int_{1/n}^{1}\right) |f_n(t) - f_m(t)| dt$$

$$= \int_0^{1/m} (mt - nt)dt + \int_{1/m}^{1/n} (1 - nt)dt + \int_{1/n}^{1} (1-1)dt$$
$$= \frac{1}{2}\left(\frac{1}{m} - \frac{1}{n}\right) \to 0 \text{ as } n < m \to \infty$$

Thus (f_n) is a Cauchy sequence in $(C[0,1], \|\cdot\|_1)$. But the pointwise limit:

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 1 & 0 < t \leq 1 \\ 0 & t = 0 \end{cases}$$

(*Hint:* $f_n(0) = 0$ and $f_n(1) = 1$ for all n, so $f(0) = 0$ and $f(1) = 1$. For $0 < t_0 < 1$, we can find large n such that $0 < \frac{1}{n} < t_0 < 1$. Hence $f_n(t_0) = 1$ for large n. Thus $f(t_0) = 1$.) However, f is not continuous, hence $(C[0,1], \|\cdot\|_1)$ is not complete.

1.4.1 Dense subsets and separability

A set $A \subset (X, d)$ is said to be *dense* in X if $\bar{A} = X$. (that is, $\forall x \in X$, $\exists x_n \in A$ such that $x_n \to x$, or $\forall x \in X$, $B_\epsilon(x) \cap A \neq \varnothing$, $\forall \epsilon > 0$.)

Example 1.4.11. $\overline{\mathbb{Q}} = \mathbb{R}$ with usual metric $u(x, y) = |x - y|$.
Let $x \in \mathbb{R}$, $x = [x] + \alpha$, $0 < \alpha < 1$. But $\alpha = 0.x_1x_2\ldots$ with $x_i \in \{0, 1, 2, \ldots, 9\}$. $\implies x = x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} + \cdots \infty$. Let $x_n = x_0 + \frac{x_1}{10} + \cdots + \frac{x_n}{10^n}$. Then $x_n \in \mathbb{Q}$, and

$$|x - x_n| = \frac{x_{n+1}}{10^{n+1}} + \cdots \to 0.$$

Thus $x_n \in \mathbb{Q}$ and $x_n \to x \in \mathbb{R}$.

Example 1.4.12. If $1 \leq p < \infty$, then $\overline{c_{00}} = \ell^p$.

Let $x \in \ell^p$, $x = (x_1, x_2, \ldots, x_n, \ldots)$. Write $X^n = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$. Then $X^n \in c_{00}$, $\forall n \geq 1$. Now,

$$\|x - X^n\|_p = \left(\sum_{k=n}^{\infty} |x_{k+1}|^p\right)^{\frac{1}{p}} \to 0 \text{ as } n \to \infty$$

Thus, $X^n \to x$.

Example 1.4.13. $\overline{c_{00}} = c_0$. Let $x \in c_0$. Then $x = (x_1, x_2, \ldots, x_n, \ldots)$ and $\lim_{n\to\infty} x_n = 0$.

For $\epsilon > 0$, $\exists N \in \mathbb{N}$ such that $|x_n| < \frac{\epsilon}{2}$, $\forall n \geq N \cdots (1)$.
Write $X^n = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$, $n \geq N$. Then $X^n \in c_{00}$ and

$$\|x - X^n\|_\infty = \sup_{n \geq N} |x_{n+1}| \leq \frac{\epsilon}{2}, \quad \forall n \geq N \quad \text{(by (1))}$$

Thus, $X^n \to x$.

Remark: $\overline{c_{00}} = c_0 \subsetneq \ell^\infty$. That is, c_{00} is not dense in ℓ^∞.

1.4.2 Continuous maps between metric spaces

A function $f : (X, d) \to (\mathbb{R}, u)$ is said to be continuous at $x_0 \in X$ if for all $\epsilon > 0$, there exist $\delta > 0$ such that $d(x_0, y) < \delta \implies |f(x_0) - f(y)| < \varepsilon$.

$$\implies f(B_\delta(x_0)) \subseteq (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$$

Theorem 1.4.14. *Let $f : (X, d) \to (\mathbb{R}, u)$ or $(\mathbb{R}$, usual metric$)$. Then the following are equivalent:*

(i) f is continuous on X (with $\varepsilon - \delta$ definition).

(ii) For any sequence $x_n \in X$ such that $x_n \to x \implies f(x_n) \to f(x)$.

(iii) $f^{-1}(O)$ is open in (X, d), for every open set $O \subseteq \mathbb{R}$.

(iv) $f^{-1}(F)$ is closed in (X, d), for every closed set $F \subseteq \mathbb{R}$.

(Proof is similar as to $f : \mathbb{R} \to \mathbb{R}$ when $d \to u$, $u(x, y) \to d(x, y)$.)

Example 1.4.15. For $x, y, z \in (X, d)$, we get

$$|d(x, y) - d(x, z)| \le d(y, z) \quad \text{(by triangle inequality)}$$

Thus, for $f(y) = d(x_0, y)$

$$|f(y) - f(x)| < d(y, z) \to 0 \quad \text{as} \quad y \to z$$

Hence, f is continuous on (X, d) to $(\mathbb{R}, u)$.

1.4.3 Uniform continuity

Definition 1.4.16. A function $f : A(\subset (X, d)) \to \mathbb{R}$ is said to be *uniformly continuous* on A if for each $\epsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in A$,

$$d(x, y) < \delta \implies |f(x) - f(y)| < \epsilon$$

Notice that δ is free of choice of locations of points $x, y \in A$; it only depends on their separation.

Example 1.4.17. For $x_0 \in X$, let $f(x) = d(x, x_0)$. Then f is uniformly continuous on X. (Hint: $d(x, x_0) \le d(x, y) + d(y, x_0) \implies f(x) - f(y) < d(x, y)$.) Similarly, by replacing x with y, it follows.

Example 1.4.18. For $x \in X$, $A \subset X$, define $d(x, A) = \inf\{d(x, a) : a \in A\}$, which is called the *distance of A from x*, and is uniformly continuous as a function of x. (Hint: $d(x, a) \le d(x, y) + d(y, a)$.) Thus, $d(x, A) \le d(x, y) + d(y, A)$ and so,

$$|f(x) - f(y)| \le d(x, y) \quad (\because x \leftrightarrow y)$$

Example 1.4.19. The function $f : (0, 1) \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$ is continuous on $(0, 1)$, but not uniformly continuous.

Let $x_0 \in (0, 1)$. Then for $\epsilon > 0$, there exists $n \in \mathbb{N}$ such that $(x_0 - \frac{\epsilon}{n},\, x_0 + \frac{\epsilon}{n}) \subset (0, 1)$. Suppose $\left|\frac{1}{x_0} - \frac{1}{y}\right| < \epsilon$ for $y \in (x_0 - \frac{\epsilon}{n},\, x_0 + \frac{\epsilon}{n}) =: I_{x_0}$. Then $|x_0 - y| < \epsilon x_0 y$. Let $\delta = \min_{y \in I_{x_0}}\{\epsilon x_0 y\} =$

$\epsilon x_0(x_0 - \epsilon/n) > 0$. If $|x_0 - y| < \delta$. Then

$$\left|\frac{1}{x_0} - \frac{1}{y}\right| = \frac{|x_0 - y|}{x_0 y} < \frac{\delta}{x_0 y} \leq \frac{\epsilon x_0(x_0 - \epsilon/n)}{x_0 y} < \epsilon.$$

Hence, f is continuous at each $x_0 \in (0, 1)$.
f is not uniformly continuous: Let $\epsilon = \frac{1}{2}$, $x = \frac{1}{n}$, $y = \frac{1}{n+1}$, $n \in \mathbb{N}$. Then for any $\delta > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$|x - y| = \left|\frac{1}{n} - \frac{1}{n+1}\right| < \delta$$

but

$$|f(x) - f(y)| = 1 \not< \frac{1}{2}.$$

Hence, f is not uniformly continuous on $(0, 1)$. From the above argument, we can prove the following result.

Theorem 1.4.20. *Let $f : A(\subset (X, d)) \to \mathbb{R}$. Then f is uniformly continuous on A if and only if for every pair of sequences $x_n, y_n \in A$ with $d(x_n, y_n) \to 0$, implies $|f(x_n) - f(y_n)| \to 0$.*

Proof. Suppose f is uniformly continuous on A. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$d(x, y) < \delta \implies |f(x) - f(y)| < \epsilon. \tag{1}$$

Let $x_n, y_n \in A$ such that $d(x_n, y_n) \to 0$. Then for $\delta > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$d(x_n, y_n) < \delta \implies |f(x_n) - f(y_n)| < \epsilon. \quad \text{(from (1))},$$

That is, if $d(x_n, y_n) \to 0$, then $|f(x_n) - f(y_n)| \to 0$. Conversely, suppose that f is not uniformly continuous. Then there exists $\epsilon_0 > 0$ such that for every $\delta > 0$ there exist $x, y \in A$ with $d(x, y) < \delta$ but $|f(x) - f(y)| \geq \epsilon_0$. Now, let $\delta = \frac{1}{n}$ for $n \in \mathbb{N}$. Then there exist $x_n, y_n \in A$ such that

$$d(x_n, y_n) < \frac{1}{n}\ , \forall n \in \mathbb{N}, \quad \text{but} \quad |f(x_n) - f(y_n)| \geq \epsilon_0.$$

That is, $d(x_n, y_n) \to 0$ but $\underline{\lim}|f(x_n) - f(y_n)| \geq \epsilon_0$, is a contradiction. Hence, f is uniformly continuous. □

Exercise 1.4.21. Show that a uniformly continuous function on a metric space (X, d) sends Cauchy sequences to Cauchy sequences. (*Hint:* If $f : (X, d) \to \mathbb{R}$ is uniformly continuous, so for $d(x_n, x_m) \to 0 \implies |f(x_n) - f(x_m)| \to 0$.)

Theorem 1.4.22. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is uniformly continuous.*

Proof. On contrary, suppose f is not uniformly continuous on $[a, b]$. Then there exists $\epsilon_0 > 0$ such that for every $\delta > 0$, there exist $x, y \in [a, b]$ with $|x - y| < \delta$ but $|f(x) - f(y)| \geq \epsilon_0$. For $\delta = \frac{1}{n}$, there exist $x_n, y_n \in [a, b]$ such that $|x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| \geq \epsilon_0$. By the

Bolzano–Weierstrass theorem, x_n, y_n have convergent subsequences, say $x_{n_k} \to x$ and $y_{n_k} \to y$. Now,

$$|x - y| = \lim_{k\to\infty} |x_{n_k} - y_{n_k}| \le \lim_{k\to\infty} \frac{1}{n_k} = 0,$$

so $x = y$. Since f is continuous, $f(x_{n_k}) - f(y_{n_k}) \to f(x) - f(y) = 0$, but $|f(x_{n_k}) - f(y_{n_k})| \ge \epsilon_0$, contradiction. □

Example 1.4.23. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous such that $\lim_{|x|\to\infty} f(x) = 0$. Then f is uniformly continuous.

Proof. For $\epsilon > 0$, there exists $[-a, a]$ such that $|f(x)| < \epsilon/2$ if $x \in [-a, a]^c$. Hence, if $x, y \in [-a, a]^c$, then

$$|f(x) - f(y)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \qquad (1)$$

Since f is uniformly continuous on $[-a, a]$. For $\epsilon > 0$, there exists $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \epsilon \qquad (2)$$

Since (1) holds true for x, y with $|x - y| < \delta$. It follows that for $\epsilon > 0$, we get $\delta > 0$ such that $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$ (for any $x, y \in \mathbb{R}$). Hence, f is uniformly continuous on $\mathbb{R}$. □

Notice that if $f \in C_0(\mathbb{R})$, that is f is continuous and $\lim_{|x|\to\infty} f(x) = 0$ and hence f is uniformly continuous. But if f is continuous and bounded, then f need not be uniformly continuous on $\mathbb{R}$.

Example 1.4.24. $f(x) = \sin x^2$, which is continuous and bounded but not uniformly continuous on $\mathbb{R}$. (Hint: Take $x^2 = n\pi$ and $y^2 = n\pi + \frac{1}{2}\pi$.)

Example 1.4.25. Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded continuous function. If f is monotone, then f is uniformly continuous on $\mathbb{R}$. Since f is bounded, let $\inf f(x) = L, \quad \sup f(x) = M$. For $\epsilon > 0$, there exist $x_0, y_0 \in \mathbb{R}$ such that $f(x_0) < L + \epsilon$ and $f(y_0) > M - \epsilon$.
If f is monotone increasing, then for $x, y \in [x_0, y_0]^c$ and $x, y \ge y_0$

$$f(y) - f(x) \le M - f(y_0) < M - (M - \varepsilon) = \varepsilon.$$

Similarly, if $x, y \le x_0$ then

$$f(y) - f(x) \le L + \varepsilon - f(x_0) < L + \varepsilon - L = \varepsilon.$$

Thus, for $x, y \in [x_0, y_0]^c$, we get $|f(x) - f(y)| < \varepsilon \qquad (1)$.
Since f is continuous on $[x_0, y_0]$, f is uniformly continuous on $[x_0, y_0]$. For any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$x, y \in [x_0, y_0],\ |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon \qquad (2)$$

Notice that (1) also holds for $x, y \in [x_0, y_0]^c$ with $|x-y| < \delta$. Thus, we get single $\delta > 0$ such that

$$|x-y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

Exercise 1.4.26. If $f : \mathbb{R} \to \mathbb{R}$ is a bounded continuous function then for f monotone, it follows that

$$\lim_{x \to -\infty} f(x) = \text{finite}, \qquad \lim_{x \to +\infty} f(x) = \text{finite}.$$

(*Hint:* For any sequence $x_n \to \infty$, $f(x_n)$ is bounded and $\lim_{n\to\infty} f(x_n) = \sup_n f(x_n)$, for f is increasing.)

Example 1.4.27. Let $f : (a, b] \to \mathbb{R}$ and $f : (b, c) \to \mathbb{R}$ be uniformly continuous. Then $f : (a, c) \to \mathbb{R}$ is uniformly continuous.

Proof. Since f is uniformly continuous on $(a, b]$ and (b, c), for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $x, y \in (a, b]$ or $x, y \in (b, c)$ with $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Now, let $x, y \in (a, c)$, with $|x - y| < \delta$. Then $|x - b| < \delta$ and $|y - b| < \delta$. Hence,

$$|f(x) - f(y)| < |f(x) - f(b)| + |f(b) - f(y)| < 2\varepsilon.$$

Thus, f is uniformly continuous on (a, c). □

We see that a uniformly continuous function can be extended uniformly to the closure of the set.

Theorem 1.4.28. *Let $f : A(\subset \mathbb{R}) \to \mathbb{R}$ be uniformly continuous on A. Then f can be extended uniformly to $\overline{A}$, and this extension is unique.*

Proof. Let $x \in \overline{A}$. Then there exists $x_n \in A$ such that $x_n \to x$. Now, $f(x_n)$ is a bounded sequence in $\mathbb{R}$. Hence, by Bolzano-Weierstrass theorem, $f(x_n)$ has a convergent subsequence. Without loss of generality we can assume that $f(x_n)$ is convergent. Let $\tilde{f}(x) = \lim f(x_n)$ ($\because \lim f(x_n)$ exists). Notice that $\tilde{f}$ is well defined, because f is uniformly continuous on A. If $x_n, y_n \to x$, then $x_n - y_n \to 0 \implies f(x_n) - f(y_n) \to 0$ i.e. $\lim f(x_n) = \lim f(y_n)$ ($\because \lim f(x_n)$ and $\lim f(y_n)$) both exist). Hence $\tilde{f} : \overline{A} \to \mathbb{R}$ is well defined. Suppose $x, y \in \overline{A}$ and they are close enough to each other. Then there exist $x_n, y_n \in A$ such that $x_n \to x$ and $y_n \to y$. Hence,

$$\begin{aligned} \tilde{f}(x) - \tilde{f}(y) &= \tilde{f}(x) - f(x_n) + f(x_n) - f(y_n) + f(y_n) - \tilde{f}(y) \\ \implies |\tilde{f}(x) - \tilde{f}(y)| &\le |\tilde{f}(x) - f(x_n)| + |f(x_n) - f(y_n)| + |f(y_n) - \tilde{f}(y)| \end{aligned}$$

Notice that $|\tilde{f}(x) - f(x_n)| < \varepsilon$ and $|\tilde{f}(y) - f(y_n)| < \varepsilon$ for $n \ge n_0$ (say). Let $|x - y| < \delta$ (small enough). Then there exists $n' \in \mathbb{N}$ such that $|x_n - y_n| < \delta$ for $n \ge n'$. Since f is uniformly continuous on A, it follows that $|f(x_n) - f(y_n)| < \epsilon$ for $n \ge n'$. Thus for sufficiently large

$n \geq \max(n_0, n')$.

$$|\tilde{f}(x) - f(y)| \leq 3\epsilon, \quad \text{where } |x-y| < \delta.$$

Hence, $\tilde{f}$ is uniformly continuous on $\overline{A}$.
This extension of f is unique: If there exists $\tilde{g} : \overline{A} \to \mathbb{R}$ which is uniformly continuous and $\tilde{g} = f$ on A, then for $x \in \overline{A}$, there is a sequence $x_n \in A$ such that $x_n \to x$. Hence,

$$\tilde{f}(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g(x_n) = \tilde{g}(x)$$

($\because$ g is uniformly continuous extension). □

Next, we shall see that uniformly continuous function grows slower than a straight line.

Theorem 1.4.29. *Let $f : \mathbb{R} \to \mathbb{R}$ be uniformly continuous,then there exist constants $A, B \geq 0$ such that $|f(x)| \leq A|x| + B$ for all $x \in \mathbb{R}$.*

Proof. For any $\epsilon > 0$, there exists $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < 1$. We divide the proof into two parts: one is near "0" and other is away from "0". Let $a > 0$. Then $|f(x)| \leq A < \infty$ for $x \in [-a, a]$. Now, consider $f : [a, \infty) \to \mathbb{R}$.Then for $x \in [a, \infty)$, we can find $n \in \mathbb{N}$ such that $x \in [a + n\delta,\, a + (n + 1)\delta]$. Then,

$$f(x) - f(a) = f(x) - f(a + n\delta) + f(a + n\delta) - f(a)$$

$$= f(x) - f(a + n\delta) + \sum_{j=1}^{n} [f(a + j\delta) - f(a + (j + 1)\delta)]$$

$$\Rightarrow |f(x)| < 1 + n + |f(a)|$$

$$\Rightarrow \left|\frac{f(x)}{x}\right| < \frac{(n+1) + |f(a)|}{a + n\delta} < \frac{(n+1) + |f(a)|}{n\delta} < \left(1 + \frac{1}{n}\right)\frac{1}{\delta} + \frac{|f(a)|}{n\delta} \leq B < \infty.$$

Notice that B is independent of n, hence B is independent of x. That is, $|f(x)| \leq B|x|$ if $x > a$. Hence, we can summarize that $|f(x)| \leq B|x| + A \qquad$ for all $x \in \mathbb{R}$. □

Example 1.4.30. Notice that $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$, as it cannot satisfies the conclusion of the above theorem.

Example 1.4.31. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and its derivative is bounded. Then f is uniformly continuous on $\mathbb{R}$. For any $x, y \in \mathbb{R}$, by the Mean Value Theorem,

$$|f(x) - f(y)| = |f'(t)(x - y)| \leq M|x - y|$$

where t is between x and y, and M is an upper bound for $|f'(t)|$. However, $f(x) = \sqrt{x}$ for $x \in (0, \infty)$ is uniformly continuous, but its derivative is $f'(x) = \frac{1}{2\sqrt{x}}$, is not bounded.

Example 1.4.32. Let $f : (X, d) \to \mathbb{R}$ be uniformly continuous, then f sends Cauchy sequence in X to Cauchy sequence in $\mathbb{R}$.

Let (x_n) be a Cauchy sequence in (X, d). Since f is uniformly continuous, for $\varepsilon > 0$, there exists $\delta > 0$ such that $d(x, y) < \delta \implies |f(x) - f(y)| < \varepsilon$. For $\delta > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \delta$ for all $n, m \geq N$, $\implies |f(x_n) - f(x_m)| < \varepsilon, \quad \forall n, m \geq N$. Therefore, $(f(x_n))$ is a Cauchy sequence in $\mathbb{R}$.

1.4.4 Compactness in metric spaces

Definition 1.4.33. Let (X, d) be a metric space. A subset $K \subset X$ is called *compact* if every open cover of K admits a finite subcover.

Theorem 1.4.34 (Sequential compactness)**.** *If (X, d) is a metric space and $K \subset X$, then K is compact if and only if every sequence in K has a convergent subsequence with limit in K.*

Remark 1.4.35. In $\mathbb{R}^n$ equipped with the Euclidean metric, the Heine–Borel theorem asserts that a set is compact if and only if it is closed and bounded.

Theorem 1.4.36. *Every compact metric space is complete. Moreover, if $f : X \to Y$ is continuous, then $f(K)$ is compact whenever K is compact.*

Proof. If (x_n) is a Cauchy sequence in a compact metric space, then (x_n) has a convergent subsequence $(x_{n_k}) \to x$. The Cauchy property forces $x_n \to x$, proving completeness. The continuous image statement follows from the open-cover definition. □

1.4.5 The contraction mapping principle

Fixed point searching is an idea to solve equation of the form $\varphi(x) = x$. This helps solving a range of problems, including approximation theory, differential equations etc. Fixed points can be obtained via iterations, i.e. if the function "shrinks nicely", then we get fixed points via iteration. That is, if x_0 is a point in the space X, then $x_0 \to \varphi^1(x_0) \to \varphi^2(x_0) \to \cdots$ where φ^n denotes n-times composition of φ. If the sequence $(\varphi^n(x_0))$ is convergent and φ is continuous, then $\varphi^n(x_0) \to x$ and thus $\varphi(x) = \varphi(\lim_{n\to\infty} \varphi^n(x_0)) = x$. However, if the space is complete, we only need to verify $\varphi^n(x_0)$ to be a Cauchy sequence. Nicely shrinking function, we mean here with contraction mapping.

Definition 1.4.37. A function $\varphi : (X, d) \to (X, d)$ is called contraction if there exists $0 < \alpha < 1$ such that

$$d(\varphi(x), \varphi(y)) \leq \alpha\, d(x, y), \quad \forall x, y \in X.$$

Theorem 1.4.38. *Let (X, d) be a complete metric space. If $\varphi : (X, d) \to (X, d)$ is a contraction, then φ has a unique fixed point.*

Proof. Let $0 < \alpha < 1$ be such that

$$d(\varphi(x), \varphi(y)) \leq \alpha\, d(x, y), \quad \forall x, y \in X.$$

For a point $x_0 \in X$, let

$$\varphi^0(x_0) = x_0, \quad \varphi^1(x_0) = \varphi(x_0) \quad etc.$$

Then

$$d(\varphi^{n+1}(x_0), \varphi^n(x_0)) \leq \alpha\, d(\varphi^n(x_0), \varphi^{n-1}(x_0)) \leq \alpha^n d(\varphi(x_0), x_0).$$

We show that $\varphi^n(x_0)$ is a Cauchy sequence. Let $m > n$. Then

$$d(\varphi^n(x_0), \varphi^m(x_0)) \leq (\alpha^n + \cdots + \alpha^{m-1})\, d(\varphi(x_0), x_0)$$

$$\leq \frac{\alpha^n}{1-\alpha} d(\varphi(x_0), x_0) \qquad (\because 0 < \alpha < 1)$$

$$\to 0 \quad \text{as } n \to \infty.$$

Since (X, d) is complete, $\varphi^n(x_0) \to x \in X$ (say).

$$\implies \varphi(x) = \varphi\left(\lim_{n\to\infty} \varphi^n(x_0)\right) = \lim_{n\to\infty} \varphi^{n+1}(x_0)$$

$$\implies \varphi(x) = x.$$

If $\exists y \in X$ such that $\varphi(y) = y$, then

$$d(x, y) = d(\varphi(x), \varphi(y)) \leq \alpha d(x, y)$$

$$\iff x = y \quad (\because 0 < \alpha < 1)$$

This establishes that φ has unique fixed point. □

Remark: If $\Omega \subset \mathbb{R}^n$ is open, then any contraction mapping $f : \Omega \to \Omega$ can have at most one fixed point.

Notice that completeness property of the space is a sufficient condition for existence of fixed point. For example,

$$\varphi : (0, \infty) \to (0, \infty)$$

$$\varphi(x) = \frac{1}{2}(x + \frac{a}{x}), \qquad a > 0$$

satisfies $\varphi(\sqrt{a}) = \sqrt{a}$.

Notice that φ above is not a contraction mapping, since

$$|\varphi(x) - \varphi(y)| = \frac{1}{2}\, |1 - \frac{a}{xy}|\, |x - y|$$

because the function $|1 - \frac{a}{xy}|$ is not bounded near zero.

Example 1.4.39. $\varphi : (0, 2\pi) \to (0, 2\pi),\ \varphi(x) = \sin\frac{x}{2}$.

$$|\varphi(x) - \varphi(y)| \leq \frac{1}{2}|x - y| \quad \text{(By Mean Value Theorem)}$$

Thus, φ is a contraction mapping, but φ has no fixed point in $(0, 2\pi)$.

Exercise 1.4.40. If (X, d) is a complete metric space and $f : X \to X$ is such that f^k is a contraction, then show that f has a unique fixed point. (*Hint:* do for $k = 2$, use the fact that f^k cannot have two fixed points. If $f^2(x_0) = x_0$ and $y_0 = f(x_0)$(say), implies that $f(y_0) = y_0 \implies y_0 = x_0$).

Exercise 1.4.41. Let $T : C[0, 1] \to C[0, 1]$ be defined by

$$T(f)(x) = \int_0^x f(t)dt.$$

Show that T^2 is a contraction but T is not a contraction.

Notice that the above fact in these example is also clear from the fact that in the convergence of $\varphi^n(x_0)$, we can ignore finitely many steps.
Now, we shall try to understand the existence and uniqueness of the initial value problem:

$$\begin{cases} y' = f(x, y) \\ y(0) = y_0 \end{cases} \tag{*}$$

with the help of fixed point theorem.
Suppose f is a continuous function in some rectangle containing the interval $(0, y_0)$ in its interior, and f is Lipschitz in the second variable, i.e.,

$$|f(x, y_1) - f(x, y_2)| \leq K|y_1 - y_2|,$$

where K is a fixed constant. Then the equation (*) has a unique solution in some neighborhood of $x = 0$. Notice that solving $(*)$ is equivalent to solve

$$\int_0^x y'(t)dt = \int_0^x f(t, y(t))dt$$

i.e.,

$$y(x) = y_0 + \int_0^x f(t, y(t))dt \tag{**}$$

That is, we want $y(t)$ such that (**) holds. In other words, we want to get fixed point for the map $\varphi \mapsto F(\varphi)$, where

$$F(\varphi)(x) = y_0 + \int_0^x f(t, \varphi(t))dt,$$

with $\varphi \in C[-\delta, \delta]$ for some $\delta > 0$, which we get very soon. Now,

$$|F(\varphi)(x) - F(\psi)(x)| \leq \int_0^x |f(t, \varphi(t)) - f(t, \psi(t))| dt,$$

$$\leq K \int_0^x |\varphi(t) - \psi(t)| dt$$

$$\leq K \cdot 2\delta \cdot \|\varphi - \psi\|_\infty.$$

Thus, $F : C[-\delta, \delta] \to C[-\delta, \delta]$ is a contraction as long as $2K\delta < 1$, i.e. if $\delta < \frac{1}{2K}$. Hence F has a unique fixed point in $C[-\frac{1}{2K}, \frac{1}{2K}]$. That is, (*) has a unique solution in $|x| < \frac{1}{2K}$.

Example 1.4.42. Consider $y' = 2x(1 + y), \qquad y(0) = 0$. Then

$$\varphi(x) = \int_0^x 2t(1 + \varphi(t)) dt.$$

With the initial guess $\varphi^0 \equiv 0$, we get

$$\varphi^1(x) = \int_0^x 2t(1 + 0)\, dt = x^2,$$

$$\varphi^2(x) = \int_0^x 2t(1 + t^2)\, dt = x^2 + \frac{x^4}{2},$$

$$\varphi^3(x) = x^2 + \frac{x^4}{2} + \frac{x^6}{6}.$$

Thus, by induction,

$$\varphi^n(x) = \sum_{k=1}^{n} \frac{x^{2k}}{k!} \longrightarrow e^{x^2} - 1, \tag{*}$$

and $\varphi(x) = e^{x^2} - 1$ is a solution, which is same as method of separation of variables. Notice that the series (*) converges uniformly on every interval $[-a, a]$, or on any interval $[a, b]$. On the other hand, $\varphi'(x) = 2x(1 + \varphi(x))$ has unique solution in neighborhood of any point x_0, i.e., $[x_0 - \delta, x_0 + \delta]$ with $\delta < \frac{1}{4}$. (*Hint:* Lipschitz constant = 2.)

1.5 Uniform convergence

1.5.1 Uniform convergence of sequences of functions

Notice that in the previous exercises, we have seen that $(C([0, 1]), \|\cdot\|_\infty)$ is complete. That is, if $\|f_n - f_m\|_\infty \to 0$, then there exists $f \in C([0, 1])$ such that $\|f_n - f\|_\infty \to 0$. But then,

$$|f_n(t) - f(t)| < \|f_n - f\|_\infty \to 0, \quad \forall t \in [0, 1],$$

i.e., $f_n(t) \to f(t)$ for each $t \in [0,1]$. We say that $f_n \to f$ uniformly if

$$\sup_t |f_n(t) - f(t)| \to 0.$$

But there are sequence of functions which converge pointwise but not uniformly.

Example 1.5.1. Let $f_n(t) = t^n,\ \ t \in [0,1]$. Then,

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 0 & 0 \le t < 1 \\ 1 & t = 1 \end{cases}$$

So,

$$\sup_t |f_n(t) - f(t)| = 1 \not\to 0.$$

Example 1.5.2. Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by

$$f_n(t) = e^{-nt^2}, \quad n \in \mathbb{N}$$

Then,

$$f(t) = \lim_{n\to\infty} f_n(t) = \begin{cases} 1 & t = 0 \\ 0 & |t| > 0 \end{cases}$$

Notice that for $t = 0$, $|f_n(0) - f(0)| = |1 - 1| = 0 < \epsilon, \quad \forall n \in \mathbb{N}$ If $|t_0| > 0$, $t_0^2 > 0$. Then for $|f_n(t_0) - 0| < \epsilon$, we get

$$e^{-nt_0^2} < \epsilon \implies n > \frac{\log \frac{1}{\epsilon}}{t_0^2}$$

Let $n_0 = \left\lceil \frac{\log \frac{1}{\epsilon}}{t_0^2} \right\rceil + 1$. Then, $|f_n(t_0) - f(t_0)| < \epsilon$ for $n \ge n_0$

Notice that $n_0 = n_0(\epsilon, t_0)$ and n_0 is large for $|t_0|$ close to zero. Thus, n_0 cannot be free from t_0. Therefore, $f_n \to f$ pointwise but not uniformly. Also,

$$\|f_n - f\|_\infty = \sup_{t\in\mathbb{R}} e^{-nt^2} = 1 \not\to 0$$

If $f_n(t) = e^{-nt}$ for $t \in [1, \infty)$, then

$$\sup_t |f_n(t) - 0| = e^{-n} \to 0 \implies e^{-nt} \xrightarrow[{[1,\infty)}]{\text{unif.}} 0$$

Exercise 1.5.3. Let $f_n, f : A(\subseteq \mathbb{R}) \to \mathbb{R}$ be such that $f_n \to f$ uniformly on A. Then for $|f_n(t)| \le M_n$ (i.e. f_n's are bounded), that implies f is bounded.
(*Hint:* $|f(t)| \le |f_{n_0}(t) - f(t)| + |f_{n_0}(t)| < \epsilon + M_{n_0} < \infty \quad \forall t \in A$)

We shall see later that uniform convergent sequences is a good carrier for many underline properties.

Theorem 1.5.4. *Let $f, f_n : A(\subset \mathbb{R}) \to \mathbb{R}$ be such that $f_n \to f$ uniformly. Then f is continuous if f_n's are continuous (i.e. the uniform limit of a sequence of continuous functions is continuous).*

Proof. For $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $\sup_{t \in A} |f_{n_0}(t) - f(t)| < \epsilon$ Thus,

$$|f_{n_0}(t) - f(t)| < \epsilon, \quad \forall t \in A$$

Since f_{n_0} is continuous on A, for fixed t and for $\epsilon > 0$, there exists $\delta > 0$ such that if $|t - s| < \delta \implies |f_{n_0}(t) - f_{n_0}(s)| < \epsilon$. Thus,

$$|f(s) - f(t)| < |f(s) - f_{n_0}(s)| + |f_{n_0}(s) - f_{n_0}(t)| + |f_{n_0}(t) - f(t)| < 3\epsilon$$

□

Theorem 1.5.5. *Let $\mathcal{R}[a, b]$ denote the space of all Riemann integrable functions on $[a, b]$. Let $f_n, f \in \mathcal{R}[a, b]$ and $f_n \to f$ uniformly. Then,*

$$\int_a^b f_n \to \int_a^b f$$

that is,

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b \lim_{n\to\infty} f_n$$

Proof.

$$\left| \int_a^b (f_n - f) \right| \le \int_a^b |f_n - f| \le \|f_n - f\|_\infty (b - a) \to 0$$

□

Corollary 1.5.6. *If $f_n \in \mathcal{R}[a, b]$ such that $S_n = f_1 + f_2 + \ldots + f_n$ converges uniformly to S, then*

$$\int_a^b \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_a^b f_n$$

(Obvious from the previous result).

Theorem 1.5.7. *Let $f_n \in C^1[a, b]$ be such that $f_n' \to g$ uniformly. If there exists $x_0 \in [a, b]$ such that $f_n(x_0)$ converges, then there exists $f \in C^1[a, b]$ such that $f_n \to f$ uniformly and $f' = g$.*

Proof. Since $f_n' \to g$ uniformly and f_n is continuous, g will be continuous. Define

$$f : [a, b] \to \mathbb{R} \quad \text{by} \quad f(x_0) = \lim_{n\to\infty} f_n(x_0)$$

and

$$f(x) = \begin{cases} f(x_0) + \int_{x_0}^{x} g(t)\, dt, & \text{if } x > x_0 \\ f(x_0) - \int_{x}^{x_0} g(t)\, dt, & \text{if } x < x_0 \end{cases}$$

Then $f'(x) = g(x)$ for every $x \in [a, b]$. Hence, $f \in C^1[a, b]$. Now,

$$\begin{aligned} f_n(x) - f_m(x) &= f_n(x) - f_m(x) - (f_n(x_0) - f_m(x_0)) + (f_n(x_0) - f_m(x_0)) \\ &= (x - x_0)(f_n'(t) - f_m'(t)) + (f_n(x_0) - f_m(x_0)) \end{aligned}$$

Therefore,

$$\|f_n - f_m\|_\infty \leq (b - a)\|f_n' - f_m'\|_\infty + |f_n(x_0) - f_m(x_0)| \to 0,$$

as $n, m \to \infty$. Hence, (f_n) is a Cauchy sequence in $(C[a, b], || \cdot ||_\infty)$. Therefore, f_n converges uniformly. Again, since $f_n' \to g = f'$ uniformly, it follows that

$$\int_{x_0}^{x} f_n'(t)\, dt \to \int_{x_0}^{x} f'(t)\, dt.$$

$$\lim_{n\to\infty} [f_n(x) - f_n(x_0)] = f(x) - f(x_0)$$

$$\lim_{n\to\infty} f_n(x) = f(x) \quad (\because \lim_{n\to\infty} f_n(x_0) = f(x_0))$$

□

Remark 1.5.8. Convergence of $(f_n(x_0))$ is necessary in the above result. Consider

$$f_n(t) = \sqrt{t + n}, \qquad t \in [0, 1]$$

Then f_n does not converge at any point of $[0, 1]$, but

$$f_n'(t) = \frac{1}{2\sqrt{t + n}} \xrightarrow{\text{unif.}} 0$$

Since

$$\sup_{t\in[0,1]} |f_n'(t) - 0| = \sup_{t\in[0,1]} \frac{1}{2\sqrt{t + n}} = \frac{1}{2\sqrt{n}} \to 0.$$

Exercise 1.5.9. Let $f_n : \mathbb{R} \to \mathbb{R}$. Check for uniform convergence of f_n to some f:

1. $f_n(t) = \frac{\sin(nt)}{\sqrt{n}}$.
2. $f_n(t) = n^2 t(1 - t^2)^n$.
3. $f_n(t) = te^{-nt}$.

Also, verify for term-by-term integration and differentiation for each of the above.

Theorem 1.5.10. *Let $E \subseteq \mathbb{R}$, and $f_n \to f$ uniformly on E. For a limit point x of E. Suppose*

$$\lim_{t\to x} f_n(t) = A_n \quad \textit{(finite)} \tag{*}$$

Then (A_n) is convergent and

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n.$$

That is,

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

Proof. Since $f_n \to f$ uniformly on E. For each $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$|f_n(t) - f_m(t)| < \epsilon, \quad \forall n, m \geq n_0,\ \forall t \in E \tag{*}$$

By (*), it implies that $|A_n - A_m| < \epsilon, \quad \forall n, m \geq n_0$. So (A_n) is Cauchy, hence convergent $\implies$ $A_n \to A$ (Say). Now,

$$\begin{aligned}|f(t) - A| &= |f(t) - f_n(t) + f_n(t) - A_n + A_n - A|\\ &\leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|\\ &< \epsilon + \epsilon + \epsilon\end{aligned}$$

for $t \in (x-\delta, x+\delta) \setminus x$ and $n \geq n_0$ (free of t)

$$\lim_{t\to x} f(t) = A = \lim_{n\to\infty} A_n$$

$$\text{Thus,} \quad \lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

□

Theorem 1.5.11. *Let $f_n : [a,b] \to \mathbb{R}$ be such that (f_n') converges uniformly. If there exists $x_0 \in [a,b]$ such that $(f_n(x_0))$ is convergent, then (f_n) is uniformly convergent, and*

$$\lim_{n\to\infty} f_n'(x) = \left(\lim_{n\to\infty} f_n(x)\right)'$$

(i.e. limit and derivative commute).

Proof. The first part of the proof is as earlier. By the Mean Value Theorem, it follows that

$$|f_n(x) - f_m(x)| \leq (b-a)\|f_n' - f_m'\| + |f_n(x_0) - f_m(x_0)|$$

Since f_n' converges uniformly and $f_n(x_0)$ is convergent, it follows that $f_n \to f$ (say) uniformly.
Claim: $\lim_{n\to\infty} f_n'(x) = f'(x)$.
Notice that f_n' need not be continuous, hence Fundamental Theorem of Calculus cannot be applied. Therefore, we need to exploit the differentiability of f. For $x \in [a,b]$, define

$$\varphi_n(t) = \frac{f_n(x) - f_n(t)}{x - t}, \qquad t \in [a,b] \setminus \{x\}$$

Then

$$\lim \varphi_n(t) = \frac{f(x) - f(t)}{x - t} =: \varphi(t)$$

Notice that $\lim_{t\to x} \varphi_n(t) = f_n'(x)$ (finite). Also,

$$|\varphi_n(t) - \varphi_m(t)| = |f_n'(x) - f_m'(x)| < \epsilon \quad \text{(by MVT)}$$

for $n, m \geq n_0$ and for all $t \in [a, b] \setminus \{x\}$. Thus, $\varphi_n \to \varphi$ uniformly on $[a, b] \setminus \{x\}$. Apply previous theorem with $E = [a, b]$. Then,

$$\lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty} \lim_{t\to x} \varphi_n(t) = \lim_{t\to x} \lim_{n\to\infty} \varphi_n(t) = \lim_{t\to x} \varphi(t) = f'(x).$$

Thus,

$$\lim_{n\to\infty} f_n'(x) = \Big(\lim_{n\to\infty} f_n(x)\Big)'$$

□

1.5.2 Term-by-term differentiation

Let $S_n = f_1 + f_2 + \cdots + f_n$, where each $f_i : [a, b] \to \mathbb{R}$ such that $S_n' \xrightarrow{\text{unif}} S$ and $S_n(x_0) \to L$. Then, $\lim(S_n') = (\lim S_n)'$. That is,

$$f_1' + f_2' + \cdots + f_n' + \cdots = (f_1 + f_2 + \cdots + f_n + \cdots)'.$$

This raises a very fundamental question: When does

$$\left(\int_a^x f(t)\,dt\right)' = \int_a^x f'(t)dt \tag{**}$$

hold? Notice that if f' is continuous then for

$$F(x) = \int_a^x f'(t)\,dt,$$

by the Fundamental Theorem of Calculus, $F'(x) = f'(x)$.

$$(F - f)' = 0$$

By the Mean Value Theorem, $F - f$ is constant. So $F(x) = f(x) - f(a)$ ($\because F(a) = 0$). However, if f' is not continuous, i.e. $f' \in \mathcal{R}[a, b]$, then $(**)$ need not be true.

Consider the sequence $f_n : A \subset \mathbb{R} \to \mathbb{R}$. We say f_n converges to $f : A \subset \mathbb{R} \to \mathbb{R}$ pointwise if for any $t_0 \in A$, and $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$|f_n(t_0) - f(t_0)| < \varepsilon, \quad \forall n \geq N$$

Notice that $N = N(\varepsilon, t_0)$.

Example 1.5.12. $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(t) = e^{-nt^2},\ n \in \mathbb{N}$. Then

$$f(t) = \begin{cases} 1 & t = 0 \\ 0 & |t| > 0 \end{cases}$$

$$|f_n(0) - f(0)| = |1 - 1| = 0 < \varepsilon, \quad \forall n \geq 1$$

Now, if $|t_0| > 0$, $t_0^2 > 0$. Then for

$$|f_n(t_0) - 0| < \varepsilon \implies e^{-nt_0^2} < \varepsilon$$

$$\implies n > \frac{\log \frac{1}{\varepsilon}}{t_0^2}$$

Let $N_0 = \left\lceil \frac{\log \frac{1}{\varepsilon}}{t_0^2} \right\rceil + 1$. Then $N_0 = N(\varepsilon, t_0)$ and N_0 is larger when $|t_0|$ is close to 0. Thus, N_0 *cannot* be free of t_0.

However, if it happens that N_0 is free of choice of $t_0 \in A$. Then, we say, f_n converges to f *uniformly.*
Note: $f_n \to f$ uniformly if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$|f_n(t) - f(t)| < \varepsilon, \quad \forall n \geq N, \forall t \in A.$$

Then

$$\sup_{t \in A} |f_n(t) - f(t)| \leq \varepsilon, \quad \forall n \geq N$$

or

$$\|f_n - f\|_\infty \leq \varepsilon, \quad \forall n \geq N,$$

So,

$$\|f_n - f\|_\infty \to 0 \text{ as } n \to \infty$$

If $f_n(t) = e^{-nt^2}$, $t \in \mathbb{R}$, $n \in \mathbb{N}$, $\sup_{t \in \mathbb{R}} |f_n(t) - f(t)| = 1 \not\to 0$. Hence $f_n \to f$ pointwise but not uniformly.

Example 1.5.13. If $f_n, f : A \subset \mathbb{R} \to \mathbb{R}$ be such that $f_n \to f$ uniformly. Then for $|f_n(t)| \leq M_n$ implies f is bounded.

$$|f(t)| \leq |f(t) - f_N(t)| + |f_N(t)| < 1 + M_N$$

Example 1.5.14. If $f_n \to f$ uniformly and f_n are continuous/uniformly continuous, then f is continuous/uniformly continuous.

Theorem 1.5.15. *Let $f_n, f \in \mathcal{R}[a,b]$ be such that $f_n \to f$ uniformly on $[a,b]$. Then*

$$\int_a^b f_n \to \int_a^b f \quad \left(\lim \int_a^b f_n = \int_a^b \lim f_n\right)$$

Proof.

$$\left|\int_a^b (f_n - f)\right| \leq \int_a^b |f_n - f| \leq \|f_n - f\|_\infty (b-a)$$

Since $f_n \to f$ uniformly $\implies \|f_n - f\|_\infty < \varepsilon$, for any $\varepsilon > 0$, for all $n \geq N$.
Therefore,

$$\left|\int_a^b f_n - \int_a^b f\right| < \varepsilon(b-a), \quad \forall n \geq N$$

Thus,

$$\int_a^b f_n \to \int_a^b f$$

□

Corollary 1.5.16. *If $f_n \in \mathcal{R}[a,b]$ and $S_n = f_1 + \cdots + f_n \to S$ uniformly, then*

$$\int_a^b \sum f_n = \sum \int_a^b f_n$$

(This follows immediately from the previous result.)

Theorem 1.5.17. *Let $f_n \in C^1[a,b]$ be such that $f_n' \to g$ uniformly. If there exists $x_0 \in [a,b]$ such that $f_n(x_0)$ converges, then there exists $f \in C^1[a,b]$ such that $f_n \to f$ uniformly and $f' = g$.*

Remark 1.5.18. Convergence of $(f_n(x_0))$ is necessary in the above result. Consider

$$f_n(t) = \sqrt{t+n}, \qquad t \in [0,1]$$

Then f_n does not converge at any point of $[0,1]$, but

$$f_n'(t) = \frac{1}{2\sqrt{t+n}} \xrightarrow{\text{unif.}} 0$$

Since

$$\sup_{t\in[0,1]} |f_n'(t) - 0| = \sup_{t\in[0,1]} \frac{1}{2\sqrt{t+n}} = \frac{1}{2\sqrt{n}} \to 0.$$

Exercise 1.5.19. Let $f_n : \mathbb{R} \to \mathbb{R}$. Check for uniform convergence of f_n to some f:

1. $f_n(t) = \frac{\sin(nt)}{\sqrt{n}}$.
2. $f_n(t) = n^2 t(1-t^2)^n$.
3. $f_n(t) = te^{-nt}$.

Also, verify for term-by-term integration and differentiation for each of the above.

Chapter 2

Function of Several Variables

This chapter extends one-variable calculus to functions on $\mathbb{R}^n$. After fixing notation and basic limit/continuity concepts, we study partial and directional derivatives and the precise notion of differentiability via linear approximation. The chain rule is developed in a form suitable for compositions and coordinate changes. We then establish Taylor's theorem as a higher-order approximation scheme, and conclude with two central structural results: the inverse mapping theorem and the implicit function theorem, which explain when nonlinear maps are locally invertible and when level sets can be described as graphs.

2.1 Syllabus map

This chapter develops multivariable calculus from a rigorous analytic viewpoint. We proceed from limits and continuity to differentiability, and then to the inverse and implicit function theorems.

2.2 Limits and continuity

2.2.1 Notation and basic definitions in Euclidean space

For $n \in \{1, 2, \ldots\} = \mathbb{N}$; $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ copies}}$. Let $\mathbf{x} \in \mathbb{R}^n$, then $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Let $\mathbf{0} \in \mathbb{R}^n$, represent as $\mathbf{0} = (0, 0, \ldots, 0)$. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (x_1 + y_1,\, x_2 + y_2,\, \ldots,\, x_n + y_n) \\ \lambda \mathbf{x} &= (\lambda x_1,\, \lambda x_2,\, \ldots,\, \lambda x_n) \end{aligned}$$

Define the standard inner product $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Then inner product $\langle \cdot, \cdot \rangle$ satisfies,

(i) $\langle x, x\rangle = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0$.

(ii) $\langle x, x\rangle = 0$ if and only if $x = 0$.

(iii) For $\alpha, \beta \in \mathbb{R}$ and $x, y, z \in \mathbb{R}^n$:

$$\langle x,\, \alpha y + \beta z\rangle = \alpha\langle x, y\rangle + \beta\langle x, z\rangle$$
$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$$

Therefore, $\langle \cdot, \cdot \rangle$ is a bilinear map and is called the *inner product*.
Let $x \in \mathbb{R}^n$. Define the norm:

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

For $x, y \in \mathbb{R}^n$, then

$$|\langle x, y\rangle| \leq \|x\|\,\|y\| \quad \textit{(Cauchy-Schwarz inequality)}.$$

If $x \neq 0$, $y \neq 0$, then $\|x\| \neq 0$, $\|y\| \neq 0$.

$$\left|\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle\right| \leq 1$$

But $\left\|\frac{x}{\|x\|}\right\| = 1$ and $\left\|\frac{y}{\|y\|}\right\| = 1$. We need to prove the inequality when $\|x\| = 1$, $\|y\| = 1$. For any $t \in \mathbb{R}$, $\langle x - ty,\, x - ty\rangle = \|x - ty\|^2 \geq 0$.
Let $P(t) = \langle x - ty,\, x - ty\rangle$. Then

$$\begin{aligned} P(t) &= \langle x, x\rangle - 2t\langle x, y\rangle + t^2\langle y, y\rangle \\ &= 1 - 2t\langle x, y\rangle + t^2 \cdot 1 \qquad (\text{since } \|x\| = \|y\| = 1) \\ &= t^2 - 2t\langle x, y\rangle + 1 \geq 0 \end{aligned}$$

Take $t_0 = \langle x, y\rangle$, then $P(t_0) = t_0^2 - 2t_0^2 + 1 = 1 - t_0^2 \geq 0 \implies t_0^2 \leq 1 \implies |t_0| \leq 1$ that is $|\langle x, y\rangle| \leq 1$.
Notice $|\langle x, y\rangle| = 1$ if and only if $x = \alpha y$ or $y = \alpha x$ for some $\alpha \in \mathbb{R}$. Suppose $y = \alpha x$, then

$$|\langle x, \alpha x\rangle| = |\alpha||\langle x, x\rangle| = |\alpha| \cdot 1 \cdot 1$$

$$= \|\alpha x\|.\|x\| = \|y\|.\|x\|$$

$$\implies |\langle x, \alpha x\rangle| = 1$$

Suppose $|\langle x, y\rangle| = 1 \cdots (1)$. *Claim:* $y = \alpha x$, for some α.
Let $p(t) = t^2 - 2t\langle x, y\rangle + 1$. If we take $t_0 = \langle x, y\rangle$, then

$$p(t_0) = \langle x, y\rangle^2 - 2\langle x, y\rangle^2 + 1 = 0 \quad \text{(by (1))}$$

But $p(t_0) = \|x - t_0 y\|^2 = 0$ if and only if $x = t_0 y$. Thus, x and y are linearly dependent.

Theorem 2.2.1. *: $|\langle x, y\rangle| \leq \|x\|\,\|y\|, \forall x, y \in \mathbb{R}^n$ and $|\langle x, y\rangle| = \|x\|\,\|y\|$ if and only if there exist $\alpha \in \mathbb{R}$ such that $x = \alpha y$. that is, x and y are linearly dependent. (Explain linear dependent sets and so forth).*
For $x, y \in \mathbb{R}^n$:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| \quad (since |\langle x, y\rangle| < \|x\|\|y\|) \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Therefore, $\|x + y\| \leq \|x\| + \|y\|$ (Triangle Inequality).

Bolzano-Weierstrass Theorem: *Every bounded sequence $(a_n) \subset \mathbb{R}$ has a convergent subsequence.*

Bolzano-Weierstrass Theorem for $\mathbb{R}^2$:
Let $\{X_n\} = \{(x_n, y_n)\}$. $\|X_n\| = \sqrt{x_n^2 + y_n^2} \leq M, \quad \forall n \geq 1$.

$$|x_n| \leq \sqrt{x_n^2 + y_n^2} \leq M$$

$$|y_n| \leq \sqrt{x_n^2 + y_n^2} \leq M$$

By Bolzano-Weierstrass theorem $x_{n_k} \to x$ and $\{(x_{n_k}, y_{n_k})\}$ is bounded. So y_{n_k} is bounded. So by Bolzano-Weierstrass theorem $y_{n_{k,l}} \to y$. Hence, $(x_{n_{k,l}}, y_{n_{k,l}}) \to (x, y)$.
$\underline{\mathbb{R}^n}$: Let $X_k = (x_1^k, x_2^k, \ldots, x_n^k)$. If $\{X_k\}$ is a bounded sequence in $\mathbb{R}^n$, then there exists a subsequence $\{X_{k_l}\}$ such that $X_{k_l} \to X \in \mathbb{R}^n$.

2.2.2 Limits in Euclidean space

Suppose $f : (a, b)(\subset \mathbb{R}) \to \mathbb{R}$. If $\lim_{h\to 0} f(x + h)$ and $\lim_{h\to 0} f(x - h)$ both exist and are equal, then we say the limit at x exists.

Suppose $f : D(\subseteq \mathbb{R}^2) \to \mathbb{R}$. $\lim_{(x,y)\to(0,0)} f(x, y) =$ finite and equal along all paths joining (x, y) and $(0, 0)$. Let $x = r\cos\theta$, $y = r\sin\theta$, so $(x, y) \to (0, 0) \iff x^2 + y^2 \to 0$ that is $r^2 \to 0$ or $r \to 0$ $(since r > 0)$. $\lim_{r\to 0} f(r\cos\theta, r\sin\theta) =$ finite, we say limit at $(0, 0)$ exists.

Let $D = (a_1, b_1) \times \cdots \times (a_n, b_n)$, and $f : D(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$, $f(X) = f(x_1, x_2, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$. Then f is said to be continuous at $X \in D$, if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $Y \in D$ with $\|X - Y\| < \delta$ implies $\|f(X) - f(Y)\| < \varepsilon$

$$\implies \left(\sum_{i=1}^{m} |f_i(X) - f_i(Y)|^2 \right)^{1/2} < \varepsilon$$

$$\implies |f_i(X) - f_i(Y)| < \varepsilon, \ \forall i = 1, 2, \ldots, m$$

Thus, f continuous at X implies *each component* f_i is continuous at X.

Conversely, if each f_i for $i = 1, 2, \ldots, m$ is continuous, then for $\varepsilon > 0$, there exists $\delta > 0$ such that $\|X - Y\| < \delta \implies |f_i(X) - f_i(Y)| < \frac{\varepsilon}{\sqrt{m}} \implies \|f(X) - f(Y)\| < \varepsilon$. Thus, it is enough to consider $f : \mathbb{R}^2 \to \mathbb{R}$ for questions result regarding $f : \mathbb{R}^n \to \mathbb{R}^m$.

2.2.3 Continuity in Euclidean space

Definition 2.2.2. Let $D(\subseteq \mathbb{R}^2)$ and $f : D \to \mathbb{R}$. Then f is said to be continuous at $X_0 = (x_0, y_0) \in D$, if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $X = (x, y) \in D$, $\|X - X_0\| < \delta \implies |f(X) - f(X_0)| < \varepsilon$ that is, $\lim_{X \to X_0} f(X) = f(X_0)$

Negation of Continuity: $\exists \varepsilon_0 > 0$ such that $\forall \delta > 0$, $\exists X \in D$ such that $\|X - X_0\| < \delta$ but $|f(X) - f(X_0)| \geq \varepsilon_0$.

Proposition 2.2.3. *If $f : D(\subset \mathbb{R}^2) \to \mathbb{R}$ is continuous at X_0 if and only if for every sequence $X_n \to X_0$, implies $f(X_n) \to f(X_0)$.*

Proof. Let $X_0 = (x_0, y_0)$, $X_n = (x_n, y_n)$. Suppose f is continuous at X_0. Then for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|X - X_0\| < \delta \implies |f(X) - f(X_0)| < \varepsilon. \tag{1}$$

Let $X_n \to X_0$. Then for $\delta > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$n \geq n_0 \implies \|X_n - X_0\| < \delta \implies |f(X_n) - f(X_0)| < \varepsilon \qquad \text{(by (1))} \tag{2}$$

Thus, $X_n \to X_0 \Longrightarrow f(X_n) \to f(X_0)$.

Conversely, suppose (2) holds, but f is not continuous at X_0, then $\exists \, \varepsilon_0 > 0$ such that $\forall \, \delta > 0$, there exists $X \in D$ such that $\|X - X_0\| < \delta$ but $|f(X) - f(X_0)| \geq \varepsilon_0$. Take $\delta = \frac{1}{n} > 0$, then there exists $X_n \in D$ such that $\|X_n - X_0\| < \frac{1}{n}$ but $|f(X_n) - f(X_0)| \geq \varepsilon_0$. So $X_n \to X_0$, but $f(X_n) \not\to f(X_0)$. □

Example 2.2.4. Define

$$f(x,y) = \begin{cases} 1 & \text{if } xy \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist. But if $xy \neq 0$ is replaced by $xy = 1$, it exists.

Exercise 2.2.5. Let $f : \mathbb{R}^2 \to \mathbb{R}$, check the continuity of f at $(0,0)$.

1. $f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } x^2+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$

2. $f(x,y) = \dfrac{\sin^2(x-y)}{\sqrt{x^2+y^2}}, \qquad f(0,0) = 0.$

3. $f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y} & \text{if } x^2+y \neq 0 \\ 0 & \text{otherwise} \end{cases}$

4. $f(x,y) = \begin{cases} \dfrac{x^2y}{x^4+y^2} & \text{if } x^4+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$

5. $f(x,y) = \begin{cases} \dfrac{\sin xy}{xy} & \text{if } xy \neq 0 \\ 0 & \text{otherwise} \end{cases}$

Using the epsilon-delta definition: Let $f(x,y) = \dfrac{xy}{x^2+y^2}$, $f(0,0) = 0$. For $x = y$, $f(x,x) = \frac{1}{2}$. Thus, $|f(x,x) - f(0,0)| = \frac{1}{2}$. Take $\varepsilon = \frac{1}{4}$, then there does not exist any $\delta > 0$ such that $\sqrt{x^2+y^2} < \delta \implies |f(x,y) - f(0,0)| < \frac{1}{4}$.

Composition of Two Continuous Functions:

Let $f : D(\subset \mathbb{R}^2) \to \mathbb{R}$ and $g : I(\subset \mathbb{R}) \to \mathbb{R}$ be continuous, where $f(x) \in I$ for each x. Then $g \circ f$ is continuous.

Proof. Since f is continuous at $x \in D$, for $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|x - y\| < \delta \implies |f(x) - f(y)| < \varepsilon. \tag{1}$$

Similarly, g is continuous at $f(x)$, so for $\eta > 0$, there exists $\mu > 0$ such that

$$|t - f(x)| < \mu \implies |g(t) - g(f(x))| < \eta.$$

Given $\varepsilon > 0$, choose $\eta = \varepsilon$. Then from (1), $\|x - y\| < \delta \implies |g(f(x)) - g(f(y))| < \eta$. Thus, $g \circ f$ is continuous at x.

Alternatively, let $x_n \to x$, then $f(x_n) \to f(x)$ and hence $g(f(x_n)) \to g(f(x))$.

□

Example 2.2.6.

$$f(x,y) = \begin{cases} \dfrac{\sin xy}{xy} & \text{if } xy \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

$$f(x,y) = p \circ g(x,y), \quad \text{where } p(t) = \begin{cases} \dfrac{\sin t}{t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

2.3 Differentiation in $\mathbb{R}^n$

2.3.1 Partial derivatives

Let $D = (a,b) \times (c,d)$ (or in general open set in $\mathbb{R}^2$). Let $f : D \to \mathbb{R}$. Let $x_0 = (x_0, y_0)$,

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

If this exists, we say f has partial derivative parallel to the x-axis at (x_0, y_0), and we denote it by $\frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0)$. In other words, for $\epsilon > 0$, there exists $\delta > 0$ such that

$$|h| < \delta \implies \left| \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} - f_x(x_0, y_0) \right| < \epsilon$$

$$f(x_0 + h, y_0) - f(x_0, y_0) = h f_x(x_0, y_0) + h\eta(h)$$

where $\eta(h) \to 0$ as $h \to 0$ (let $h\eta(h) = \gamma(h)$). $f(x_0+h, y_0) - f(x_0, y_0) = hf_x(x_0, y_0) + \gamma(h)$ where $\gamma(h) \to 0$ as $h \to 0$.

Similarly, $f(x_0, y_0 + k) - f(x_0, y_0) = kf_y(x_0, y_0) + \gamma(k)$ where $\gamma(k) \to 0$ as $k \to 0$.

Note: From the accompanying graph, one sees that the existence of the partial derivative in the direction parallel to the x-axis depends only on the values of f along an appropriate line segment through (x_0, y_0); it does not require f to be defined on an open disk around (x_0, y_0).

Example 2.3.1. $f(x,y) = \dfrac{xy}{x^2 + y^2}$, $f(0,0) = 0$. Then $f_x(0,0) = 0 = f_y(0,0)$ but f is not continuous at $(0,0)$.

2.3.2 Directional derivatives

Directional derivative is the rate of change of a function parallel to a given direction.

Let $x_0 \in D$ (rectangle or open set) and $f : D(\subset \mathbb{R}^2) \to \mathbb{R}$. Let $\mathbf{v} = (v_1, v_2)$, $|\mathbf{v}| = \sqrt{v_1^2 + v_2^2} = 1$. Then the directional derivative of f at x_0 along $\mathbf{v}$ is defined by

$$D_{\mathbf{v}} f(x_0) = \lim_{t \to 0} \frac{f(x_0 + t\mathbf{v}) - f(x_0)}{t}$$

Note: The existence of the directional derivative of f at x_0 in the direction $\mathbf{v}$ depends only on the values of f along a line segment through x_0 parallel to $\mathbf{v}$; it does not require f to be defined on an open neighborhood of x_0.

Example 2.3.2.

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{if } x^4+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$D_{\mathbf{v}} f(0,0) = \lim_{t\to 0} \frac{t^2 v_1^2 v_2}{t^4 v_1^4 + t^2 v_2^2} = \begin{cases} 0 & v_2 = 0 \\ \frac{v_1^2}{v_1^2} & v_2 \neq 0 \end{cases}$$

But f is not continuous at $(0,0)$, for $y = mx$ and so forth.

Example 2.3.3. Let $D = (a,b) \times (c,d)$ (or open convex set in $\mathbb{R}^2$), that is, $(x, y \in D \implies \lambda x + (1-\lambda)y \in D$, $\forall \lambda \in [0,1])$. Suppose $f : D(\subset \mathbb{R}^2) \to \mathbb{R}$ such that $f_x(x,y) = 0 = f_y(x,y)$, $\forall x, y \in D$. Then f is constant.

Since D is convex, $(a,s) \times \{y\} \subset D$. Thus,

$$\int_a^s f_x(x,y)dx = 0$$

$$f(s,y) = f(a,y)$$

Let $g(y) = f(a,y)$. Then $0 = \frac{\partial}{\partial y} f(s,y) = g'(y) \implies \int_c^t g'(y)dy = 0 \implies g(y) = g(c)$. Thus, $f(s,y) = f(a,y) = g(y) = g(c)$ for all $(s,y) \in D \implies f$ is constant on D.
Remark: A similar proof will work for D open and convex.

2.3.3 Differentiability

Let D be an open set in $\mathbb{R}^2$. Let $H = (h,k)$, $X_0 = (x_0, y_0)$. Then f is said to be differentiable at $X_0 \in D$ if there exists $L \in \mathbb{R}^2$ such that

$$\epsilon_L(H) = \frac{f(X_0 + H) - f(X_0) - L \cdot H}{\|H\|} \to 0 \quad \text{as} \quad \|H\| \to 0. \tag{*}$$

Notice that, since we need limit in $(*)$ exists in a δ-neighborhood of X_0, it means f is differentiable along all directions including parallel to x-axis and y-axis.
The vector L is unique. Suppose not, then there exist $M \in \mathbb{R}^2$ such that (*) holds. Thus,

$$\frac{(L-M)\cdot H}{\|H\|} = \epsilon_L(H) - \epsilon_M(H) \to 0 \quad \text{as} \quad \|H\| \to 0.$$

Set $H = tV$, $V \neq 0$ in $\mathbb{R}^2$. Then,

$$\lim_{t\to 0} \frac{|t|\,|(L-M)\cdot V|}{|t|\|V\|} = 0 \implies |(L-M)\cdot V| = 0, \quad \forall V \in \mathbb{R}^2$$

Consider $V = L - M$, then $\|L - M\| = 0 \implies L = M$. Hence, the derivative of f at X_0 is unique and we write $L = f'(X_0)$. Since $\epsilon(H) = \frac{f(X_0+H)-f(X_0)-H\cdot f'(X_0)}{\|H\|} \to 0$ as $\|H\| \to 0$. Set $H = tV$, $\|V\| = 1$.

$$\epsilon(tV) = \frac{f(X_0+tV) - f(X_0) - tV\cdot f'(X_0)}{|t|} \to 0$$

as $t \to 0$. Thus, $V \cdot f'(X_0) = D_V f(X_0)$. Put $V = (1,0)$, then $D_V f(X_0) = f_x(X_0)$. Similarly, $V = (0,1)$, $D_V f(x_0) = f_y(X_0)$.

Example 2.3.4. Let D be an open set in $\mathbb{R}^2$ and $f : D(\subset \mathbb{R}^2) \to \mathbb{R}$ be such that f_x and f_y both are bounded on D. Then f is continuous.

Proof.

$$\begin{aligned}
&f(x_0+h, y_0+k) - f(x_0,y_0)\\
&= f(x_0+h, y_0+k) - f(x_0, y_0+k) + f(x_0, y_0+k) - f(x_0,y_0)\\
&= hf_x(x_0+\theta_1 h, y_0+k) + kf_y(x_0, y_0+\theta_2 k) \quad \text{(By Mean Value Theorem of one variable).}
\end{aligned}$$

where $\theta_1, \theta_2 \in (0,1)$.

Hence, $|f(x_0+h, y_0+k) - f(x_0,y_0)| \leq |h|M_1 + |k|M_2 \leq \sqrt{h^2+k^2}\sqrt{M_1^2+M_2^2}$ where $|f_x(x,y)| \leq M_1$, $|f_y(x,y)| \leq M_2$ for all $(x,y) \in D$. Thus, $|f(x_0+h, y_0+k) - f(x_0,y_0)| \to 0$ as $\sqrt{h^2+k^2} \to 0$. Therefore, f is continuous at (x_0,y_0).

□

Exercise 2.3.5. Let $\nabla f = (f_x, f_y)$, as along as $f_x(X_0)$ and $f_y(X_0)$ just exist, then f need not be differentiable at X_0.

Note: If f is differentiable,

$$D_V f(X_0) = f'(X_0) = (f_x(X_0), f_y(X_0)) = \nabla f(X_0)$$

Example 2.3.6.

$$f(x,y) = \begin{cases} \frac{y}{|y|}\sqrt{x^2+y^2} & \text{if } y \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then f is continuous at $(0,0)$ and $D_v f(0,0) = \frac{v_2}{|v_2|} = 2$ or 0 if $v_2 = 0$. But f is not differentiable at $(0,0)$.

$$\epsilon(h,k) = \frac{\frac{k}{|k|}\sqrt{h^2+k^2} - k}{\sqrt{h^2+k^2}}$$

$$= \frac{k}{|k|} - \frac{k}{\sqrt{h^2+k^2}}$$

For $h = mk, m, k > 0$,

$$\epsilon(mk,k) = 1 - \frac{1}{\sqrt{1+m^2}} \not\to 0 \quad \text{as} \quad k \to 0$$

Exercise 2.3.7. Prove that

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{x^2+y^2} & \text{if } x^2+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is differentiable at $(0,0)$ and $f'(0,0) = (0,0)$. But none of f_x and f_y is continuous at $(0,0)$.

Theorem 2.3.8. *Let D be an open set in $\mathbb{R}^2$. Suppose f_x and f_y are continuous in a neighbourhood of $(x_0,y_0) \in D$. Then f is differentiable at (x_0,y_0).*

Proof. Since $(x_0,y_0) \in D$ and D is open, $\exists\, \delta > 0$ such that $B_\delta(x_0,y_0) \subset D$. Let $(x_0+h, y_0+k) \in B_\delta(x_0,y_0)$. Then consider

$$\epsilon(h,k) = \frac{f(x_0+h,y_0+k) - f(x_0,y_0) - hf_x(x_0,y_0) - kf_y(x_0,y_0)}{\sqrt{h^2+k^2}}$$

Since f_x and f_y exist in $B_\delta(x_0,y_0)$ (say), one can apply the Mean Value Theorem coordinate-wise. Thus,

$$\epsilon(h,k) = \frac{hf_x(x_0+\theta_1 h, y_0+k) + kf_y(x_0, y_0+\theta_2 k) - hf_x(x_0,y_0) - kf_y(x_0,y_0)}{\sqrt{h^2+k^2}}$$

where $0 < \theta_1, \theta_2 < 1$.

$$|\epsilon(h,k)| < \sqrt{h^2+k^2}\left((f_x(x_0+\theta_1 h, y_0+k) - f_x(x_0,y_0))^2 + (f_y(x_0,y_0+\theta_2 k) - f_y(x_0,y_0))^2\right)$$

Since f_x and f_y are continuous in $B_\delta(x_0,y_0)$, $|\epsilon(h,k)| \to 0$ as $\sqrt{h^2+k^2} \to 0$. Thus f is differentiable at (x_0,y_0). $\square$

Geometric Interpretation of Derivative:

For function from $\mathbb{R}^n \to \mathbb{R}$. Let $y = f(x_0) + f'(x_0)(x-x_0)$ For $n=1$, $y = f(x_0)+f'(x_0)(x-x_0)$ (line passing through $(x_0, f(x_0))$). For $n=2$, $z = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$ (a plane passing through $(x_0, y_0, f(x_0,y_0))$).

2.3.4 Chain rule

$$I \subset \mathbb{R} \xrightarrow{g} J \xrightarrow{f} \mathbb{R}.$$

Let $F = f \circ g$. If f and g are both differentiable, then $f \circ g$ is differentiable.

Proof. Since f is differentiable at $y = g(x)$,

$$f(y+k) - f(y) - f'(y)k = k\eta(k) \tag{1}$$

when $\eta(k) \to 0$ as $k \to 0$. Since g is differentiable at x, g is continuous. Set $k = g(x+h) - g(x)$, then $h \to 0 \implies k \to 0$. Since g is differentiable,

$$k = g(x+h) - g(x) = hg'(x) + h\mu(h),$$

where $\mu(h) \to 0$ as $h \to 0$.
Consider

$$\begin{aligned}\epsilon(h) &= \frac{f \circ g(x+h) - f \circ g(x) - f'(g(x))g'(x)h}{h} \\ &= \frac{f(y+k) - f(y) - f'(y)(k - h\mu(h))}{h}\end{aligned}$$

Since $\frac{1}{h} = \frac{g'(x)+\mu(h)}{k}$,

$$\epsilon(h) = \eta(k)(g'(x) + \mu(h)) + f'(y)\mu(h)$$

Since $h \to 0 \implies k \to 0 \implies \eta(k) \to 0$. So $\epsilon(h) \to 0$. Thus $f \circ g$ is differentiable and

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

□

Chain Rule for $\mathbb{R}^2 \to \mathbb{R}$: If f and g both are differentiable, then $f \circ g$ is differentiable and

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

Proof.

$$\eta(k) = \frac{f(y+k) - f(y) - f'(y)k}{\|k\|} \to 0$$

where $\|k\| \to 0$. Since g is continuous, set $K = g(x+h) - g(x)$, then $\|k\| \to 0$ as $|h| \to 0$. Since g is differentiable at x,

$$k = g(x+h) - g(x) = hg'(x) + |h|\mu(h)$$

that is,

$$\|k\| \leq |h|\|g'(x)\| + |h|\|\mu(h)\|$$

Now,

$$\epsilon(h) = \frac{f \circ g(x+h) - f \circ g(x) - f'(g(x))g'(x)h}{|h|}$$

$$|\epsilon(h)| \leq |\eta(k)|(\|g'(x)\| + \|\mu(h)\|) + \|f'(y)\|\|\mu(h)\|$$

$$\to 0 \text{ as } h \to 0, \text{ because } h \to 0 \implies k \to 0$$

Thus,

$$(f \circ g)'(x) = \underbrace{f'(g(x))}_{1\times 2}\underbrace{g'(x)}_{2\times 1}$$

□

Mean Value Theorem for Convex Domain:

Let $\mathcal{D}$ be an open and convex set in $\mathbb{R}^2$. Suppose $f : \mathcal{D} \to \mathbb{R}$ is differentiable. Then for any $x, y \in \mathcal{D}$, there exists $c \in \mathcal{D}$ such that $f(x) - f(y) = (x - y) \cdot f'(c)$ where $c \in (x, y) = \{\lambda x + (1 - \lambda)y : 0 < \lambda < 1\}$.

Proof. Consider

$$\varphi(t) = f((1 - t)x + ty)$$

By the chain rule, φ is differentiable on $(0, 1)$ and

$$\varphi'(t) = f'((1 - t)x + ty) \cdot (y - x)$$

By the Mean Value Theorem for one variable,

$$\varphi(1) - \varphi(0) = \varphi'(\lambda)(1 - 0)$$

that is,

$$f(y) - f(x) = f'((1 - \lambda)x + \lambda y))(y - x)$$

□

Function from $\mathbb{R}^n$ to $\mathbb{R}^m$: Let $\mathcal{D}$ be an open set in $\mathbb{R}^n$ and $f : \mathcal{D}(\subset \mathbb{R}^n) \to \mathbb{R}^m$ be differentiable. Then

$$f'(x_0) = \left(\frac{\partial f_i(x_0)}{\partial x_j}\right)_{m\times n}$$

Proof. We know that $f : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x_0 if there exists a $A_{m\times n}$ matrix such that

$$\epsilon(h) = \frac{f(x_0 + h) - f(x_0) - Ah}{\|h\|} \to 0 \quad \text{as } \|h\| \to 0, \tag{1}$$

Let $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$ be the free standard basis for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. If $f = (f_1, \ldots, f_m)$, then $f_i(x) = f(x) \cdot u_i$. In (1) substitute $h = h_j e_j$, $\|h\| = |h_j|$,

$$\epsilon(h_j e_j) = \frac{f(x_0 + h_j e_j) - f(x_0) - h_j f'(x_0)e_j}{|h_j|} \to 0 \quad \text{as} \quad h_j \to 0$$

$$\iff \lim_{h_j \to 0} \frac{f(x_0 + h_j e_j) - f(x_0)}{h_j} = f'(x_0) e_j$$

$$\implies \left(\frac{\partial f_i(x_0)}{\partial x_j} \right) \quad \text{exists} \quad \text{and}$$

$$f'(x_0) = \left(\frac{\partial f_i(x_0)}{\partial x_j} \right)_{m \times n}$$

$$= \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \cdots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \cdots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix}_{m \times n}$$

□

Write $\mathcal{J}_f(x_0) = \left(\frac{\partial f_i(x_0)}{\partial x_j} \right)_{m \times n}$. Then $\mathcal{J}_f$ is called the *Jacobian matrix* of f.
Note: Existence of $\frac{\partial f_i(x_0)}{\partial x_j}$ does not imply that $f'(x_0)$ exists.

Example 2.3.9. $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y) = \begin{cases} \left(\frac{x^2 y}{x^2 + y^2}, \frac{x y^2}{x^2 + y^2} \right) & \text{if } x^2 + y^2 \neq 0 \\ (0, 0) & \text{otherwise} \end{cases}$$

Then $f = (g, h)$.

$$\mathcal{J}_f(0, 0) = \begin{pmatrix} g_x & g_y \\ h_x & h_y \end{pmatrix} (0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

But f is *not* differentiable at $(0, 0)$.

$$\|\epsilon(h, k)\| = \frac{\left\| \left(\frac{h^2 k}{h^2 + k^2}, \frac{h k^2}{h^2 + k^2} \right) - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \right\|}{\sqrt{h^2 + k^2}}$$

$$\|\epsilon(h, k)\| = \frac{|hk|}{h^2 + k^2} \not\to 0 \text{ as } \sqrt{h^2 + k^2} \to 0$$

Therefore, f is not differentiable at $(0, 0)$.

Example 2.3.10. Let $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x, y) = (\mathrm{e}^x \cos y, \mathrm{e}^x \sin y)$$

$\det(J_f(x, y)) = \mathrm{e}^{2x} \neq 0 \implies J_f(x, y)$ is non-singular matrix $\forall (x, y) \in \mathbb{R}^2$, but f is not one-to-one on $\mathbb{R}^2$, since $f(x, 2\pi + y) = f(x, y)$.

Norm of a matrix (or linear map):

Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear. Then $A = (R_1, R_2, \ldots, R_m)^T$, where R_i's are rows of A. Let $x \in \mathbb{R}^n$. Then

$$Ax = (R_1 x, R_2 x, \ldots, R_m x) \in \mathbb{R}^m$$

and

$$\|Ax\| = \sqrt{\sum |R_i x|^2} \leq \left(\sqrt{\sum \|R_i\|^2}\right) \|x\|$$

If $x \neq 0$,

$$\frac{\|Ax\|}{\|x\|} \leq \sqrt{\sum \|R_i\|^2}$$

Therefore, $\{\frac{\|Ax\|}{\|x\|} : x \neq 0\}$ is bounded in $\mathbb{R}$. Hence, it has a supremum. Let

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} < \infty.$$

Then

(i) $\|Ax\| \leq \|A\| \|x\|, \quad \forall x \in \mathbb{R}^n$.

(ii) $\|A\| = \sup_{\|x\|=1} \|Ax\|$.

Example 2.3.11. Let $A : \mathbb{R}^2 \to \mathbb{R}$, $A(x, y) = 4x + 3y$. Then

$$\|A\| = \sup_{x^2+y^2=1} |4x + 3y| = \sup_{-1 \leq x \leq 1} |4x + 3\sqrt{1 - x^2}|$$

Example 2.3.12. Let $A : \mathbb{R}^2 \to \mathbb{R}^2$, $A(x, y) = (3x, 4y)$. Then

$$\|A\| = \sup_{x^2+y^2=1} \|(3x, 4y)\| = \sup_{x^2+y^2=1} \sqrt{9x^2 + 16y^2} = \sup_{0 \leq x \leq 1} \sqrt{9x^2 + 16(1 - x^2)}$$

Chain rule for functions from $\mathbb{R}^n \to \mathbb{R}^m$:

Let D be an open set in $\mathbb{R}^n$ and $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable and $g : f(D) \to \mathbb{R}^l$ be differentiable. Then $g \circ f : D \to \mathbb{R}^l$ is differentiable and

$$(g \circ f)'(x) = g'(f(x)) f'(x)$$

(where $g'(f(x))$ is an $l \times m$ matrix and $f'(x)$ is an $m \times n$ matrix).

Proof.

$$\eta(k) = \frac{g(y + k) - g(y) - g'(y)k}{\|k\|} \to 0 \quad \text{as} \quad \|k\| \to 0$$

Since $y = f(x)$ and f is continuous at x, set $k = f(x + h) - f(x)$. Then $\|h\| \to 0 \implies \|k\| \to 0$.

Also,

$$\|k\| = \|f(x+h) - f(x)\| = \|f'(x)h + \|h\|\epsilon(h)\|$$

(since f is differentiable at x)

$$\le \|f'(x)\|\|h\| + \|h\|\|\epsilon(h)\|$$

that is,

$$\frac{1}{\|h\|} \le \frac{1}{\|k\|}\{\|f'(x)\| + \|\epsilon(h)\|\}$$

Now,

$$\begin{aligned}\mu(h) &= \frac{g\circ f(x+h) - g\circ f(x) - g'(f(x))f'(x)h}{\|h\|} \\ &= \frac{g(y+k) - g(y) - g'(y)(k - \|h\|\epsilon(h))}{\|h\|} \\ &= \frac{\|k\|\eta(h) - \|h\|g'(y)\epsilon(h)}{\|h\|}\end{aligned}$$

$$\|\mu(h)\| \le \|\eta(h)\|\{\|f'(x)\| + \|\epsilon(h)\|\} + \|g'(y)\|\|\epsilon(h)\| \to 0$$

as $\|h\| \to 0$. Hence, $(g\circ f)'(x)$ exists and $(g\circ f)'(x) = g'(f(x))\, f'(x)$. □

Example 2.3.13. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and $F : \mathbb{R}^n \to \mathbb{R}$ be defined by $F(x) = f(\|x\|^2)$. Then F is differentiable and $F'(x) = 2f'(\|x\|^2)x$.
Let $g(x) = \|x\|^2 = x_1^2 + \cdots + x_n^2$. $g'(x) = (2x_1, 2x_2, \ldots, 2x_n)$ Thus,

$$F(x) = (f\circ g)(x)$$

By the chain rule, since F is differentiable and

$$F'(x) = f'(g(x))\, g'(x)$$

that is,

$$F'(x) = 2f'(\|x\|^2)x$$

Exercise 2.3.14. Let $F(x) = f(\|x\|^{2k})$. Prove that $F'(x) = 2k\|x\|^{2k-2}f'(\|x\|^{2k})x$.

Euler's Formula. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable and $f(rx) = r^\alpha f(x)$, $\forall r > 0$ and some $\alpha \in \mathbb{R}$. Then $f'(x)x = \alpha f(x)$.

Proof. Since $f(rx) = r^\alpha f(x)$, $\forall r > 0$, differentiate both sides with respect to r.

$$f'(rx)\frac{d}{dr}(rx) = \alpha r^{\alpha-1} f(x)$$

$$f'(rx)x = \alpha r^{\alpha-1} f(x)$$

Putting $r = 1$,

$$\implies f'(x)x = \alpha f(x)$$

For $n = 2$,

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = \alpha f(x, y)$$

□

Example 2.3.15. If $\alpha > 0$, f is continuous at 0. If $\alpha > 1$, f is differentiable at 0.

Proof. (i) If $\alpha > 0$, $f(0 + h) - f(0) = f(h)$. Take $h = \|h\|v$, with $\|v\| = 1$

$$\|f(0 + h) - f(0)\| = \|h\|^\alpha \|f(v)\| \to 0 \quad \text{as } \|h\| \to 0$$

(ii) If $\alpha > 1$,

$$\frac{\partial f}{\partial x_j}(0) = \lim_{h\to 0} \frac{f(0 + h_j e_j) - f(0)}{h_j}$$

$$= \lim_{h\to 0} \frac{|h_j|^\alpha f(e_j)}{h_j} \to 0 \text{ as } h_j \to 0 \quad (since \alpha > 1)$$

$\implies J_f(0) = 0$ ($m \times n$ matrix)

$$\epsilon(h) = \frac{f(0 + h) - f(0) - J_f(0)h}{\|h\|} = \frac{\|h\|^\alpha f(v)}{\|h\|}, \|v\| = 1, h = \|h\|v$$

$$\to 0 \text{ as } \|h\| \to 0$$

□

Mixed Derivatives: Let $D \subset \mathbb{R}^n$ (or $\mathbb{R}^2$) be an open set.

$$f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$$

$$f_{xy} = \frac{\partial}{\partial y}(f_x) = \frac{\partial^2 f}{\partial y \partial x}$$

Example 2.3.16.

$$f(x, y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, & x^2 + y^2 \neq 0 \\ 0, & x^2 + y^2 = 0 \end{cases}$$

$$f_{yx}(0, 0) = \frac{\partial f_y}{\partial x}(0, 0) = \lim_{h\to 0} \frac{f_y(h, 0) - f_y(0, 0)}{h}$$

But

$$f_y(h, 0) = \lim_{k\to 0} \frac{f(h, k) - f(h, 0)}{k} = h$$

So,

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{h-0}{h} = 1$$

Similarly, $f_x(0,0) = -1 \neq f_{yx}(0,0)$.

Notations: $C^1(\mathcal{D})$ — set of all continuously differentiable functions on $\mathcal{D}$ whose derivative is continuous (that is, f_x and f_y both are continuous).

$C^2(\mathcal{D})$ – set of all functions on $\mathcal{D}$ whose partial derivatives up to second order are continuous. (that is, f_x, f_y, f_{xy},f_{yx}, f_{xx}, f_{yy} are continuous.)

Theorem 2.3.17. *If $\mathcal{D}$ is open and $f \in C^2(\mathcal{D})$, then $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$.*

Proof. Since $\mathcal{D}$ is open and $(x_0, y_0) \in \mathcal{D}$, there exists an open ball $B_\delta(x_0, y_0) \subset \mathcal{D}$ or one can draw a rectangle. Let

$$F(x, y) = f(x, y) - f(x_0, y) + f(x_0, y_0) - f(x, y_0) \tag{1}$$

Again, let $A(x, y) = f(x, y) - f(x_0, y)$. From (1), we get $F(x, y) = A(x, y) - A(x, y_0)$ By the mean value theorem,

$$F(x,y) = \frac{\partial A}{\partial y}(x,\eta)(y-y_0) = (\frac{\partial f}{\partial y}(x,\eta) - \frac{\partial f}{\partial y}(x_0,\eta))(y-y_0) \quad \text{where} \quad \eta = y_0 + (y-y_0)\theta_1, 0 < \theta_1 < 1$$

$$F(x,y) = \frac{\partial^2 f}{\partial x \partial y}(\xi, \eta)(x - x_0)(y - y_0)$$

where $\xi = x_0 + (x - x_0)\theta_2$, $0 < \theta_2 < 1$.

$$\frac{F(x,y)}{(x-x_0)(y-y_0)} = \frac{\partial^2 f}{\partial x \partial y}(\xi, \eta)$$

Since $(x, y) \to (x_0, y_0) \implies (\xi, \eta) \to (x_0, y_0)$ and $\frac{\partial^2 f}{\partial x \partial y}$ is continuous at (x_0, y_0),

$$\lim_{(x,y)\to(x_0,y_0)} \frac{F(x,y)}{(x-x_0)(y-y_0)} = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \tag{2}$$

Similarly, let $B(x, y) = f(x, y) - f(x, y_0)$. Then $F(x, y) = B(x, y) - B(x_0, y)$. It is straightforward to verify that

$$\lim_{(x,y)\to(x_0,y_0)} \frac{F(x,y)}{(x-x_0)(y-y_0)} = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \tag{3}$$

Thus from (2) and (3),

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

Note that if $f \in C^2(\mathcal{D}), \mathcal{D} \subset \mathbb{R}^n$, then

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}, \quad \forall j, k = 1, 2, \dots, n$$

□

2.3.5 Taylor's theorem

Theorem 2.3.18 (Taylor's Theorem)**.** *Let $\mathcal{D}$ be an open set in $\mathbb{R}^2$ and $f \in C^2(\mathcal{D})$. Then there exist $\lambda \in (0, 1)$ such that*

$$f(X + H) = f(X) + f'(X)H + H^t f''(C)H,$$

where $C = X + \lambda H$ and $\|H\| < \delta$.

Proof. Let $g(t) = f(X + tH)$, so $g(t) = f \circ \varphi(t)$ where $\varphi(t) = X + tH$.

$$g'(t) = f'(\varphi(t))\varphi'(t) = f'(\varphi(t))H$$

$$= hf_x(\varphi(t)) + kf_y(\varphi(t))$$

$$g''(t) = h(f_x)'(\varphi(t))\varphi'(t) + k(f_y)'(\varphi(t))\varphi'(t)$$

$$= h(f_{xx}(\varphi(t))f_{xy}(\varphi(t)))H + k(f_{yx}(\varphi(t))f_{yy}(\varphi(t)))H$$

$$= H^t \begin{pmatrix} f_{xx}(\varphi(t)) & f_{xy}(\varphi(t)) \\ f_{yx}(\varphi(t)) & f_{yy}(\varphi(t)) \end{pmatrix} H \quad \text{where} \quad H^t = (h\ k) \quad \text{(row vector)}$$

Since $g(0) = f(X)$, $g(1) = f(X + H)$, the Mean Value Theorem for one variable gives:

$$g(1) = g(0) + g'(0) \cdot 1 + \frac{1}{2}g''(\lambda).1^2$$

So,

$$f(X + H) = f(X) + f'(X)H + \frac{1}{2}H^T f''(C)H$$

where $C = X + \lambda H$ and $\|H\| < \delta$. □

Theorem 2.3.19. *Let $f : [a, b] \to \mathbb{R}^n$ be differentiable on (a, b) and continuous on $[a, b]$. Then there exists $\lambda \in (a, b)$ such that $\|f(b) - f(a)\| \leq \|f'(\lambda)\|(b - a)$*

Proof. Let $g(t) = (f(b) - f(a)) \cdot f(a + (b - a)t)$. Then $g'(t) = (f(b) - f(a)) \cdot f'(a + (b - a)t)(b - a)$ (by chain rule). Since $g : [a, b] \to \mathbb{R}$ is differentiable, by the Mean Value Theorem, there exists $\lambda \in (a, b)$ such that

$$g(b) - g(a) = g'(\lambda)(b - a)$$

$$\begin{aligned}\|f(b)-f(a)\|^2 &= (f(b)-f(a))\cdot f'(\lambda)(b-a)\\ &\le \|f(b)-f(a)\|\cdot\|f'(\lambda)\|(b-a)\end{aligned}$$

Thus,

$$\|f(b)-f(a)\| \le \|f'(\lambda)\|(b-a)$$

□

Theorem 2.3.20. *Let $\mathcal{D}$ be open in $\mathbb{R}^n$ and $f:\mathcal{D}\subset\mathbb{R}^n\to\mathbb{R}^m$ be differentiable at $X\in\mathcal{D}$. Then there exist $\lambda\in(0,1)$ such that $\|f(X+H)-f(X)\|\le\|f'(C)\|\,\|H\|$, where $C=X+\lambda H,\quad \|H\|<\varepsilon$ (for some $\lambda>0$).*
Note: Equality need not hold. For $g:(-1,1)\to\mathbb{R}^2$,

$$g(t)=(t^3,1-t^2)$$

Suppose

$$g(1)-g(-1)=g'(\lambda)(1-(-1))$$
$$(2,0)=2(3\lambda^2,-2\lambda)\implies \lambda=0,\pm\frac{1}{\sqrt{3}}$$

But $x=t^3,\ y=1-t^2,\quad x^2=(1-y)^3$, has no tangent parallel to x-axis.

Proof. Let $g(t)=f(X+tH)$. Then $g:[0,1]\to\mathbb{R}^m$ is differentiable. By previous Mean Value Theorem, $\exists\ \lambda\in(0,1)$ such that

$$\|g(1)-g(0)\|\le\|g'(\lambda)\|(1-0)$$

$$\|f(X+H)-f(X)\|\le\|g'(\lambda)\|\le\|f'(c)\|\|H\|,\qquad C=X+\lambda H$$

where $g'(\lambda)=f'(X+\lambda H)H$. □

Notations:

(i) $L_n(\mathbb{R})$ = space of all linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$.

(ii) $GL_n(\mathbb{R})=\{A\in L_n(\mathbb{R}):AA^{-1}=I\}$= set of all invertible matrices.

Proposition 2.3.21. *Let $A\in GL_n(\mathbb{R})$ and $B\in L_n(\mathbb{R})$ be such that $\|B-A\|<\frac{1}{\|A^{-1}\|}$. Then*
(i) $B\in GL_n(\mathbb{R})$ (that is, $GL_n(\mathbb{R})$ is open in $L_n(\mathbb{R})$).
(ii) $A\mapsto A^{-1}$ is continuous on $GL_n(\mathbb{R})$.

Proof. Let $\alpha=\frac{1}{\|A^{-1}\|}$, $\beta=\|B-A\|$. Then $\beta<\alpha$. For $x\in\mathbb{R}^n$, write

$$\alpha\|x\|=\alpha\|A^{-1}Ax\|\le\alpha\|A^{-1}\|\|Ax\|$$

that is,

$$\alpha\|x\| \le \|Ax\| = \|(A-B)x + Bx\| \le \|A-B\|\|x\| + \|Bx\|$$

$$\implies (\alpha-\beta)\|x\| \le \|Bx\| \tag{1}$$

(i) If $Bx = 0$, then $(\alpha-\beta)\|x\| = 0 \implies x = 0$. Since B is a one-to-one linear map from $\mathbb{R}^n \to \mathbb{R}^n$, so B is onto.
(ii) Put $x = B^{-1}y$ in (1), then

$$\frac{\|B^{-1}y\|}{\|y\|} \le \frac{1}{\alpha-\beta}, \quad y \neq 0.$$

$$\sup_{y\neq 0} \frac{\|B^{-1}y\|}{\|y\|} \le \frac{1}{\alpha-\beta} \implies \|B^{-1}\| < \frac{1}{\alpha-\beta}$$

Now,

$$\|B^{-1} - A^{-1}\| = \|B^{-1}(A-B)A^{-1}\| \le \|A-B\|\frac{1}{2(\alpha-\beta)} \to 0 \quad \text{as} \quad A \to B$$

Hence, the map $A \mapsto A^{-1}$ is continuous. □

Note: $A \mapsto A^{-1}$ is one-to-one map, because $A^{-1} = B^{-1} \implies A = B$.

Example 2.3.22. Let $f : \mathbb{R} \to \mathbb{R}$ be one-to-one and onto, and f is continuously differentiable at $x_0 \in \mathbb{R}$ such that $f'(x_0) \neq 0$. Then f^{-1} is differentiable at $y_0 = f(x_0)$ and

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)}.$$

Proof.

$$\epsilon(k) = \frac{f^{-1}(y_0+k) - f^{-1}(y_0) - \dfrac{k}{f'(x_0)}}{|k|}$$

Let $h = f^{-1}(y_0+k) - f^{-1}(y_0)$, $\quad y_0 + k = f(x_0+h)$ and $k = f(x_0+h) - f(x_0) \implies h \cdot f'(x_0+\theta h)$ for some θ.

Since $f'(x_0) \neq 0$, $\exists\, \delta > 0$ such that $f'(x) \neq 0$ for all $x \in [x_0-\delta, x_0+\delta]$. So $|f'(x)| > m > 0$ for all $x \in [x_0-\delta, x_0+\delta]$. Choose h small such that $x_0+\theta h \in [x_0-\delta, x_0+\delta]$. $|k| > |h|m$. Thus $k \to 0 \implies h \to 0$.

$$|\epsilon(k)| = \frac{\left|h - \frac{f(x_0+h)-f(x_0)}{f'(x_0)}\right|}{|f(x_0+h)-f(x_0)|} = \frac{|f'(x_0) - f'(x_0+\theta h)|}{|f'(x_0+\theta h)||f'(x_0)|} \to \frac{0}{|f'(x_0)|} = 0 \quad (since f' \text{ is continuous at } x_0)$$

□

Note: If f^{-1} is differentiable, then $f^{-1} \circ f(x) = x$ and $(f^{-1})'(f(x_0))f'(x_0) = 1$.

2.4 Inverse and implicit function theorems

2.4.1 Inverse function theorem

Theorem 2.4.1 (Inverse Function Theorem)**.** *Let Ω be an open set in $\mathbb{R}^n$. Suppose $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map such that $\det f'(x_0) \neq 0$. Then*

(i) $\exists$ open sets U and $V \subset \mathbb{R}^n$ such that $f : U \to V$ $(= f(U))$ is bijective.

(ii) f^{-1} is a C^1 map on V, and

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Proof. Let $A = f'(x_0)$. For $y \in \mathbb{R}^n$, define $\varphi : \Omega \to \mathbb{R}^n$ by

$$\varphi(x) = x + A^{-1}(y - f(x)) \tag{1}$$

Then $\varphi(x) = x$ if and only if $y = f(x)$ (that is, x is the fixed point of φ if and only if $y = f(x)$). Since f' is continuous at x_0, for $\epsilon = \frac{1}{2\|A^{-1}\|} > 0$, there exists $\delta > 0$ such that

$$\|x - x_0\| < \delta \implies \|f'(x) - f'(x_0)\| < \frac{1}{2\|A^{-1}\|}.$$

Let $U = B_\delta(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \delta\}$ and $V = f(U)$.

(i) *Claim:* f is one-to-one on U.

Now, $\varphi'(x) = I + A^{-1}f'(x) = A^{-1}(A - f'(x))$. Thus, $\|\varphi'(x)\| \leq \|A^{-1}\|\|A - f'(x)\| < \frac{1}{2}$.

If $x_1, x_2 \in U$, by the Mean Value Theorem for φ,

$$\|\varphi(x_1) - \varphi(x_2)\| \leq \|\varphi'(x_1 + \lambda(x_2 - x_1))\|\|x_1 - x_2\| < \frac{1}{2}\|x_1 - x_2\|.$$

So, φ is a contraction on U. Hence, φ can have only one fixed point. Hence, $y = f(x)$ for at most one $x \in U$. Therefore, f is one-to-one on U.

(ii) *Claim:* V is open.

Let $y^* \in V$. Then $y^* = f(x^*)$ for some $x^* \in U$. Then $\exists r > 0$ such that $B_r(x^*) = \{x \in U : \|x - x^*\| < r\} \subset U$.

Now, it is enough to prove that, whenever

$$\|y - y^*\| < \frac{r}{2\|A^{-1}\|} \implies y \in V \tag{2}$$

Suppose $\|y - y^*\| < \frac{r}{2\|A^{-1}\|}$. Then

$$\begin{aligned}\|\varphi(x^*) - x^*\| &= \|A^{-1}(y - y^*)\| \\ &\leq \|A^{-1}\|\|y - y^*\| < \frac{r}{2}.\end{aligned}$$

If $x \in \overline{B}_r(x^*) = \{x \in \Omega : \|x - x^*\| \leq r\}$, then

$$\begin{aligned}\|\varphi(x) - x^*\| &\leq \|\varphi(x) - \varphi(x^*)\| + \|\varphi(x^*) - x^*\| \\ &< \frac{1}{2}\|x - x^*\| + \frac{r}{2} < r.\end{aligned}$$

So $x \in \overline{B}_r(x^*) \implies \varphi(x) \in \overline{B}_r(x^*)$. $\varphi : \overline{B}_r(x^*) \to \overline{B}_r(x^*)$ is a contraction mapping. Then φ has a fixed point $x \in \overline{B}_r(x^*)$ such that $\varphi(x) = x$ if and only if $y = f(x)$. Now $y = f(x) \subset f(\overline{B}_r(x^*)) \subset f(U) = V$. Thus, V is open and hence $f : U \to V$ is one-to-one and onto (with $V = f(U)$ open).

(iii) *Claim:* $f^{-1} : V \to U$ is differentiable at $f(x_0)$.

Let $y \in V$, then $y + k \in V$ (since V is open) for small $\|k\|$.

Let $h = f^{-1}(y + k) - f^{-1}(y)$. Then $k = f(x + h) - f(x)$ (since $f^{-1}(y) = x$). Now,

$$\varphi(x + h) - \varphi(x) = h + A^{-1}(f(x) - f(x + h)) = h - A^{-1}k.$$

$$\begin{aligned}\implies \|h - A^{-1}k\| &\leq \frac{1}{2}\|h\| \\ \implies \|h\| &\leq \|h - A^{-1}k\| + \|A^{-1}k\| \\ &\leq \frac{1}{2}\|h\| + \|Ah\|\end{aligned}$$

that is,

$$\begin{aligned}\frac{1}{2}\|h\| &\leq \|A^{-1}k\| \\ &\leq \|A^{-1}\|\|k\|\end{aligned} \tag{3}$$

Now,

$$\begin{aligned}\eta(k) &= \frac{f^{-1}(y_0 + k) - f^{-1}(y_0) - (f^{-1}(x_0))^{-1}k}{\|k\|} \\ &= \frac{(f'(x_0))^{-1}(f'(x_0)h - (f(x_0 + h) - f(x_0)))}{\|k\|} \\ \|\eta(k)\| &\leq \frac{\|(f'(x_0))^{-1}\|\|f(x_0 + h) - f(x_0) - f'(x_0)h\|}{\frac{\|h\|}{2\|A^{-1}\|}} \\ &\to 0 \quad \text{as} \quad h \to 0 \quad (sincek \to 0 \implies h \to 0)\end{aligned}$$

$$\implies (f^{-1})'(f(x_0)) = (f'(x_0))^{-1}$$

(iv) f^{-1} is continuously differentiable that is, $(f^{-1})'$ is continuous. Need to prove $(f^{-1})'(y_0) = (f'(x_0))^{-1}$. Since $A \mapsto A^{-1}$ is continuous on $GL_n(\mathbb{R})$, $(f^{-1})'$ is continuous.

□

Example 2.4.2. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f(x, y) = (x - e^{-y},\ y - e^x)$. Then

$$f'(0,0) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \det f'(0,0) = 2 \neq 0$$

Hence f is one-to-one in a neighborhood of $(0,0)$ and

$$(f^{-1})'(f(0,0)) = (f'(0,0))^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^{-1}$$

2.4.2 Implicit function theorem

Consider $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x, y) = x^2 + y^2 - 1$. Then $f'(x, y) = (2x, 2y)$.

$$\left.\frac{\partial f}{\partial x}\right|_{(1,0)} = 2, \qquad \left.\frac{\partial f}{\partial y}\right|_{(1,0)} = 0$$

Then one can draw a ball centered at $(1, 0)$ such that of radius $r < 1$ such that $f(\varphi(y), y) = 0$, that is,

$$x = \varphi(y), \quad |y| < r < 1, \quad \varphi(y) = \sqrt{1 - y^2}$$

However, one cannot draw a ball of any radius around $(1, 0)$ such that $f(x, \psi(x)) = 0$, that is, $y = \psi(x)$ for $|x| < r$, even r very small. Because, for any $r > 0$, one cannot write $\psi(x) = \sqrt{1 - x^2}$ as $x > 1$ will be included in any ball around $(1, 0)$.

However, at any point on the circle, other than $(\pm 1, 0)$ and $(0, \pm 1)$. One can solve x and y simultaneously in a small neighborhood of the point.

Now, consider a linear map

$$A : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$

Then $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$, $(h, k) = (h, 0) + (0, k)$.

$$A(h, k) = A(h, 0) + A(0, k) = A_x h + A_y k \quad \text{(say)}$$

Lemma 2.4.3. *If A_x is invertible ($A_x \in L_n(\mathbb{R})$), then for each $k \in \mathbb{R}^m$, there exist a unique $h \in \mathbb{R}^n$ such that $h = -A_x^{-1} A_y k$*

Proof. $A(h,k) = 0 \iff A_x h + A_y k = 0$. Since A_x is invertible, $h = -A_x^{-1} A_y k$. Now, let $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set and $f : \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be differentiable.

$$f = (f_1, \dots, f_n)$$

$$f_i : \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$$

$$f_i'(x,y) = \left(\frac{\partial f_i}{\partial x_1}(x,y) \cdots \frac{\partial f_i}{\partial x_n}(x,y) \cdots \frac{\partial f_i}{\partial y_1}(x,y) \cdots \frac{\partial f_i}{\partial y_m}(x,y) \right)$$

$$\begin{aligned} f' &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} & \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{bmatrix}_{n \times (n+m)} \\ &= \left[\left(\frac{\partial f_i}{\partial x_j} \right)_x \left(\frac{\partial f_i}{\partial y_k} \right)_y \right]_{n \times (n+m)} \\ &= (A_x \quad A_y) \end{aligned}$$

Then $A_x : \mathbb{R}^n \to \mathbb{R}^n$ is linear and $A_y : \mathbb{R}^m \to \mathbb{R}^n$ is linear, where $A_x = \left(\frac{\partial f_i}{\partial x_j} \right)_x$, $A_y = \left(\frac{\partial f_i}{\partial y_k} \right)_y$. ☐

Theorem 2.4.4. ***Implicit Function Theorem:*** *Let Ω be an open subset in $\mathbb{R}^n \times \mathbb{R}^m$. If $f : \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be a C^1 map, with $f(x_0, y_0) = 0$ and $\det [f'(x_0,y_0)]_x \neq 0$ for some $(x_0, y_0) \in \Omega$. Then*

(i) There exist open sets $U \subset \mathbb{R}^n \times \mathbb{R}^m$ and $W \subset \mathbb{R}^m$ such that for all $y \in W$ there exist a unique $x \in \mathbb{R}^n$ with $(x,y) \in U$ and $f(x,y) = 0$.

(ii) If $x = g(y)$, then $g : W \subset \mathbb{R}^n \to \mathbb{R}^n$ is C^1 map, $g(y_0) = x_0$, $f(g(y), y) = 0$ for all $y \in W$ and $g'(y_0) = -A_x^{-1} A_y$, where $A_x = f_x'$, $A_y = f_y'$.

that is, f will vanish on a curve $x = g(y)$.

Proof. (i) Let $F : \Omega \to \mathbb{R}^n \times \mathbb{R}^m$ by $F(x,y) = (f(x,y), y)$. Then F is a C^1-map, and

$$F'(x_0, y_0) = \begin{bmatrix} \{f'(x_0,y_0)\}_x & \{f'(x_0,y_0)\}_y \\ 0 & I \end{bmatrix}$$

$\det F'(x_0, y_0) \neq 0$. Therefore, by the Inverse Mapping Theorem, there exist open sets $U \subset \mathbb{R}^n \times \mathbb{R}^m$ and $V \subset \mathbb{R}^n \times \mathbb{R}^m$ such that $F : U \to V$ is a one-one onto C^1-map.

Let $W = \{y \in \mathbb{R}^m : (0,y) \in V\}$. Then W is open, because V is open. Since F is onto, for $y \in W$,

$$(0,y) = F(x,y) \implies (x,y) \in U.$$

$$\implies f(x,y) = 0, \quad \forall y \in W.$$

Suppose, for this y, there exist $(x', y) \in U$ such that $f(x', y) = 0$. Then

$$F(x', y) = (f(x', y), y) = (f(x, y), y) = F(x, y).$$

Since F is one-to-one on $U \implies x' = x$.

(ii) Define $x = g(y)$ for $y \in W$. Then

$$(g(y), y) \in U \quad \text{and} \quad f(g(y), y) = 0 \tag{*}$$

$$\implies F(g(y), y) = (0, y) \quad \forall y \in W.$$

that is, $F^{-1}(0, y) = (g(y), y)$.

By the Inverse Mapping Theorem, F^{-1} is a C^1-map, hence g is a C^1-map.

To compute $g'(y_0)$, consider $f(g(y), y) = 0, \quad y \in W$. Differentiating with respect to y and using the chain rule, we get

$$f'(g(y_0), y_0) \begin{pmatrix} g'(y_0) \\ I \end{pmatrix} = 0$$

$$f'(x_0, y_0) \begin{pmatrix} g'(y_0) \\ I \end{pmatrix} = 0$$

Let $A := f'(x_0, y_0)$, then

$$\begin{pmatrix} A_x & A_y \end{pmatrix} \begin{pmatrix} g'(y_0) \\ I \end{pmatrix} = 0$$

$$\implies A_x g'(y_0) + A_y = 0 \implies g'(y_0) = -A_x^{-1} A_y$$

□

Example 2.4.5. Prove that $x^2 + ye^x - \sin(xy) = 0$ can be solved for y in a neighborhood of $(0, 0)$, but cannot be solved for x in any neighborhood of $(0, 0)$.

$$F(x, y) = x^2 + ye^x - \sin(xy) \tag{1}$$

(i) $F(0, 0) = 0$, $\frac{\partial F}{\partial y}|_{(0,0)} = 1 \neq 0$. By the implicit function theorem, there exists a ball around $(0, 0)$ and an interval for x such that $F(x, g(x)) = 0$ or $y = g(x)$ for $|x| < r$.

(ii) $\frac{\partial F}{\partial x}|_{(0,0)} = 0$. Hence, the implicit function theorem *cannot* be applied.

On the contrary, suppose $x = \phi(y)$, then $0 = \phi(0)$ and

$$(\phi(y))^2 + ye^{\phi(y)} - \sin(\phi(y)y) = 0$$

for $|y| < r$ for some $r > 0$. Then

$$2\phi(0)\phi'(0) + 1 \cdot e^{\phi(0)} + 0 \cdot e^{\phi(0)}\phi'(0) - \cos(\phi(0)0)\left(\phi'(0)0 + \phi(0) \cdot 1\right) = 0$$

$$\implies 1 = 0 \quad \text{(contradiction)}$$

Example 2.4.6. Let $f : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y, z) = (xe^y + ye^z,\ xe^z + ze^y)$$

Then f is a C^1-map.

$$f'(x, y, z) = \begin{pmatrix} e^y & xe^y + e^z & ye^z \\ e^z & ze^y & xe^z + e^y \end{pmatrix}$$

$$f(-1, 1, 1) = (0, 0)$$

Let $f = (f_1, f_2)$. Then

$$\begin{pmatrix} \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{pmatrix}(-1, 1, 1) = \begin{pmatrix} 0 & e \\ e & 0 \end{pmatrix}$$

By the implicit function theorem, there exists an open ball U in $\mathbb{R}^3$ and open ball V in $\mathbb{R}^2$, such that

$$(y, z) = (\phi(x), \psi(x)), \quad |x| < r \quad \text{for some } r > 0.$$

Exercise 2.4.7. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^1-map such that $f(0, o) = 0, f_x(0, 0) = 1$. Let $F(x, y) = (f(x, y), y)$. Prove that F is injective in some neighborhood of $(0, 0)$. Does F remain injective in any neighborhood of $(0, 0)$?

Remark: Condition in implicit function theorem or inverse mapping theorem on derivatives are sufficient.

Example 2.4.8. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^3$.

$$f(0, 0) = 0,$$
$$\frac{\partial f}{\partial y}(0, 0) = 0,$$

but $y = x^{2/3}$ is a solution of $f(x, y) = 0$ near $(0, 0)$.

Example 2.4.9. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x^3, y^3)$. Then $\det f'(0, 0) = 0$ but f is one-to-one, onto.

Chapter 3

Lebesgue Measure and Integral

This chapter presents the measure-theoretic foundation of modern integration. Motivated by limitations of the Riemann integral, we build Lebesgue outer measure, deduce its main properties, and define Lebesgue measurable sets, including instructive examples such as the Cantor set and the existence of non-measurable sets. We introduce measurable and simple functions, define the Lebesgue integral, and develop the main convergence principles—the monotone convergence theorem, Fatou's lemma, the dominated convergence theorem, and the bounded convergence theorem—together with a useful estimate (Chebyshev's inequality) for controlling the size of level sets.

3.1 Syllabus map

We introduce the measure-theoretic approach to integration: we build Lebesgue measure from outer measure, define measurable functions, construct the Lebesgue integral, prove convergence theorems, and introduce the L^p spaces.

3.2 From Riemann to Lebesgue

3.2.1 Limitations of the Riemann integral

Let $f : [a, b] \to \mathbb{R}$ and f is bounded on $[a, b]$. Then $f \in \mathcal{R}[a, b]$ (that is, f is Riemann integrable) if and only if f is almost continuous. However, there are functions which are neither almost continuous nor bounded and so forth.

(I) $\quad f : [0, 1] \to \mathbb{R}, \quad f(x) = \begin{cases} 1 & x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [0, 1] \\ 0 & x \in \mathbb{Q} \cap [0, 1] \end{cases}$

Then $\inf U(P, f) = 1$ and $\sup L(P, f) = 0$. $\implies \quad f \notin \mathcal{R}[0, 1]$.

(II) $\displaystyle\int_0^1 \frac{1}{\sqrt{t}}\,dt,\qquad f(t)=\frac{1}{\sqrt{t}}$ is not bounded near "0". However,

$$\int_{\frac{1}{n}}^1 \frac{1}{\sqrt{t}}\,dt = 2\left(1-\frac{1}{\sqrt{n}}\right) \le 2.$$

Question is should we write

$$\int_0^1 \frac{1}{\sqrt{t}}\;dt = \sup_n \int_{\frac{1}{n}}^1 \frac{1}{\sqrt{t}}\;dt = 2\ ?$$

(III) $\displaystyle\int_0^\infty \frac{1}{1+t^2}\,dt,\qquad \int_0^n \frac{1}{1+t^2}\,dt = \tan^{-1} n \le \frac{\pi}{2}.$
Does it suitable to rewrite

$$\int_0^\infty \frac{1}{1+t^2}\,dt = \sup_n \int_0^n \frac{1}{1+t^2}\,dt = \frac{\pi}{2}\ ?$$

3.3 Measure and measurability

3.3.1 Sigma-algebras and measures

Definition 3.3.1. Let X be a nonempty set. A collection $\mathcal{A} \subset \mathcal{P}(X)$ is called a *σ-algebra* on X if.

(i) $X \in \mathcal{A}$;

(ii) if $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$;

(iii) if $(E_n)_{n\ge1} \subset \mathcal{A}$, then $\bigcup_{n\ge1} E_n \in \mathcal{A}$.

The pair $(X, \mathcal{A})$ is called a *measurable space*. Elements of $\mathcal{A}$ are called *measurable sets*.

Definition 3.3.2. A function $\mu : \mathcal{A} \to [0,\infty]$ is a *measure* if $\mu(\emptyset) = 0$ and μ is countably additive. The triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.

3.3.2 Lebesgue outer measure

For open (closed) interval $I = (a, b)$ assign the length $\ell(I) = b - a$. For $I = (a,\infty)$ or $(-\infty, b)$, we assign $\ell(I) = \infty$. Now, the question is to assign an appropriate length to an arbitrary subset of $\mathbb{R}$. If $O \subset \mathbb{R}$ is open, then $O = \bigcup_n I_n$, $I_n = (a_n, b_n)$ and $I_n \cap I_m = \emptyset$ if $n \ne m$. In this case, one can consider $\ell(O) = \sum_{n=1}^\infty \ell(I_n)$. However, if $A \subseteq \mathbb{R}$, $A \subseteq O \subset \mathbb{R}$. Hence, $A \subset \bigcup_{n=1}^\infty I_n$. Thus, we have an over-estimate for length of A. that is,

$$\ell(A) \le \sum \ell(I_n),\ \text{such that } A \subset \bigcup_{n=1}^\infty I_n.$$

Therefore, we assign a number to A via

$$m^*(A) := \inf \left\{ \sum \ell(I_n) : A \subset \bigcup_n I_n \right\}$$

where $m^*(A)$ denotes the outer measure of A.
Notice that we do not require disjointness in the cover $\{I_n : n \in \mathbb{N}\}$ of A. Moreover, I_n could be any type of interval, for example, (a_n, b_n) or $[a_n, b_n)$ or $[a_n, b_n]$ or $(a_n, b_n]$.
Since $\phi \subset (0, \epsilon), \forall \epsilon > 0$. Then $m^*(\phi) \leq \epsilon, \forall \epsilon > 0$. Hence $m^*(\phi) = 0$.
For $a \in \mathbb{R}$,

$$\begin{aligned} \{a\} &\subset (a - \frac{\epsilon}{2}, a + \frac{\epsilon}{2}) \\ \implies m^*(\{a\}) &\leq \epsilon, \quad \forall \epsilon > 0 \\ \implies m^*(\{a\}) &= 0. \end{aligned}$$

3.3.3 Basic properties of outer measure

(i) If $A \subset B$, then $m^*(A) \leq m^*(B)$.
Let $B \subset \bigcup_n I_n$, then $A \subset \bigcup_n I_n$. By definition $m^*(A) \leq \sum \ell(I_n)$; $B \subset \bigcup_n I_n$.
$\implies m^*(A) \leq \inf \{\sum \ell(I_n) : \bigcup_n I_n \supset B\} \implies m^*(A) \leq m^*(B)$.

(ii) If $\{A_n\}_{n=1}^{\infty}$ is a sequence of subsets in $\mathbb{R}$, then

$$m^* \left(\bigcup_n A_n \right) \leq \sum m^*(A_n)$$

By definition of infimum, for $\epsilon > 0$, $\exists$ a cover $\{I_{n,k}\}_{k=1}^{\infty}$of A_n such that
$\sum_{k=1}^{\infty} \ell(I_{n,k}) < m^*(A_n) + \frac{\epsilon}{2^n}$ (if $m^*(A_n) < \infty$).
Thus, $\{I_{n,k} : k = 1, 2, \ldots, n = 1, 2, \ldots\}$ is a cover of $\bigcup_{n=1}^{\infty} A_n$.

Therefore,

$$m^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \ell(I_{n,k}) \leq \sum_{n=1}^{\infty} \left(m^*(A_n) + \frac{\epsilon}{2^n} \right) \leq \sum_{n=1}^{\infty} m^*(A_n) + \epsilon, \quad \forall \epsilon > 0.$$

Thus,

$$m^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} m^*(A_n)$$

Example 3.3.3. If $A \subset \mathbb{R}$ is countable, then $A = \{a_1, a_2, \ldots\} = \bigcup_{i=1}^{\infty} \{a_i\}$

$$m^*(A) \leq \sum m^*(\{a_i\}) = 0 \implies m^*(A) = 0.$$

Thus $m^*(\mathbb{Q}) = 0$. Alternatively, one can think,

$$\mathbb{Q} \subset \bigcup_{n\in\mathbb{Z}} \left(r_n - \frac{\epsilon}{2^{|n|+1}},\ r_n + \frac{\epsilon}{2^{|n|+1}}\right)$$

$$\implies m^*(\mathbb{Q}) \leq \sum \ell\left(r_n - \frac{\epsilon}{2^{|n|+1}},\ r_n + \frac{\epsilon}{2^{|n|+1}}\right) = \frac{\epsilon}{2},\ \forall\epsilon > 0.$$

Proposition 3.3.4. *If I is any interval with end points a and b. Then $m^*(I) = b - a$.*

Proof. We prove the result for each type of interval. Suppose $I = [a, b]$ and $m^*(I) = b - a$. Then for $I = (a, b)$, one can deduce that

$$\left[a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}\right] \subset (a, b)$$

$$\text{therefore } m^*\left(\left[a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}\right]\right) \leq m^*\{(a, b)\}$$

that is,

$$b - a \leq m^*\{(a, b)\}$$

Now, (a, b) is a cover of itself, so

$$m^*\{(a, b)\} \leq \ell\{(a, b)\} = b - a$$

Other covering can be done in similar way. Now, consider the case of proving $m^*([a, b]) = b - a$.

$$[a, b] \subset \left(a - \frac{1}{n}, b + \frac{1}{n}\right), \quad \forall n \geq 1$$

$$m^*([a, b]) \leq b - a + \frac{2}{n} \to b - a$$

On the other hand, suppose $[a, b] \subset \bigcup_{n=1}^{\infty} I_n$. Then $[a, b] \subset \bigcup_{n=1}^{k} I_n$ (Exercise)
(*Hint:* use Bolzano–Weierstrass theorem.)

$$\implies (a, b) \subset \bigcup_{n=1}^{k} I_n$$

By induction,

$$b - a \leq \sum_{n=1}^{k} \ell(I_n).$$

(if $[a, b] \subset \bigcup_{n=1}^{k} I_n \sqcup I_{k+1}$. Then $(a, b) \subset \bigcup_{n=1}^{k} I_n$ or $(a, b) \subset I_{k+1}$. Thus

$$b - a \leq \sum_{n=1}^{k+1} \ell(I_n).)$$

$$\implies b - a \le \sum_{n=1}^{\infty} \ell(I_n) \quad \text{for} \quad \{I_n\}_{n=1}^{\infty} \quad \text{that cover} \quad [a,b].$$

Hence,

$$b - a \le m^*([a,b]) \le b - a.$$

□

Example 3.3.5. Let $A \subset \mathbb{R}$ and $x \in \mathbb{R}$. Then for $A + x = \{a + x : a \in A\}$, we have

$$m^*(A + x) = m^*(A).$$

Let $A \subset \bigcup_n I_n$. Then $A + x \subset \bigcup_n (I_n + x)$ that is, $\{I_n + x\}_{n=1}^{\infty}$ is a covering of $A + x$. Hence

$$m^*(A + x) \le \sum \ell(I_n + x) = \sum \ell(I_n)$$

for all cover $\{I_n\}$ of A. Therefore, $m^*(A+x) \le m^*(A)$. By replacing $x \to -x$, $m^*(A-x) \le m^*(A)$. Replacing A by $A+x$, $m^*(A) \le m^*(A+x)$. Thus, $m^*(A+x) = m^*(A)$, that is, m^* is translation invariant.

Proposition 3.3.6. *Let $A \subset \mathbb{R}$ and $\epsilon > 0$. Then $\exists$ an open set $O \supset A$ such that $m^*(O) < m^*(A) + \epsilon$ that is, $m^*(A) = \inf\{m^*(O) : O \supset A, O \text{ open}\}$*

Proof. By definition, for $\epsilon > 0$, $\exists \{I_n\}$ that cover A such that

$$\sum \ell(I_n) < m^*(A) + \epsilon \quad (\text{if} \quad m^*(A) < \infty.)$$

But $m^*\left(\bigcup I_n\right) \le \sum \ell(I_n) < m^*(A) + \epsilon$. Let $O = \bigcup I_n$. Then $m^*(O) < m^*(A) + \epsilon$. □

Theorem 3.3.7. *If $A \subset \mathbb{R}$, then $\exists$ a G_δ-set $G \subset \mathbb{R}$ such that $m^*(A) = m^*(G)$.*

Proof. By the previous result for $\epsilon = \frac{1}{n}$, $\exists$ an open set $O_n \supset A$ such that

$$m^*(O_n) < m^*(A) + \frac{1}{n}$$

Let $G = \bigcap O_n$ (a G_δ-set in $\mathbb{R}$). Then $A \subset G \subset O_n$. Thus

$$m^*(A) \le m^*(G) \le m^*(O_n) < m^*(A) + \frac{1}{n}$$

So $m^*(A) \le m^*(G) \le m^*(A) + \frac{1}{n}$, $\forall n \ge 1 \implies m^*(A) = m^*(G)$ □

Example 3.3.8. Let $E = \bigcup E_n$, $E_n \subset \mathbb{R}$. Then $m^*(E) = 0$ if and only if $m^*(E_n) = 0$ for all $n \in \mathbb{N}$.

Solution: $m^*(E) \le \sum m^*(E_n)$ If each of $m^*(E_n) = 0 \implies m^*(E) = 0$.

Conversely, suppose $m^*(E) = 0$ and $m^*(E_{n_0}) > 0$ for some $n_0 \in \mathbb{N}$. Then for $\epsilon = \frac{1}{2}m^*(E_{n_0}) > 0$, $\exists$ a cover $\{I_k\}$ of E such that

$$\sum \ell(I_k) < m^*(E) + \frac{1}{2}m^*(E_{n_0})$$

But $E_{n_0} \subset E \subset \bigcup I_k \implies m^*(E_{n_0}) < \sum \ell(I_k)$, that is,

$$m^*(E_{n_0}) < \frac{1}{2}m^*(E_{n_0})$$

which is a contradiction.

Example 3.3.9. Let $O = \bigcup I_n$, I_n open intervals. Then $m^*(O) = \sum \ell(I_n)$.
For $\epsilon > 0$, $\exists$ a cover $\{J_k\}$ of O such that

$$\sum \ell(I_k) < m^*(O) + \epsilon \tag{1}$$

Now, $\bigcup I_n = O \subset \bigcup J_k$. Since I_n's are disjoint, each $I_n \subset J_{k,n}$,

$$\ell(I_n) \le \ell(J_{k,n})$$

$$\begin{aligned}
&\implies \sum_{n=1}^{\infty} \ell(I_n) < \sum_{n=1}^{\infty} \ell(J_{k,n}) < \sum_{n=1}^{\infty} \ell(J_k) < m^*(O) + \varepsilon \\
&\implies \sum_{n=1}^{\infty} \ell(I_n) < m^*(O) + \varepsilon, \quad \forall\, \varepsilon > 0 \\
&\implies \sum_{n=1}^{\infty} \ell(I_n) \le m^*(O) \le \sum_{n=1}^{\infty} \ell(I_n) \\
\text{So,} \qquad & m^*\left(\bigcup_{n=1}^{\infty} I_n\right) = \sum_{n=1}^{\infty} \ell(I_n) = \sum_{n=1}^{\infty} m^*(I_n).
\end{aligned}$$

Corollary 3.3.10. *If $\{O_i\}_{i=1}^{\infty}$ is a family of disjoint open sets in $\mathbb{R}$, then*

$$m^*\left(\bigcup_{i=1}^{\infty} O_i\right) = \sum_{i=1}^{\infty} m^*(O_i).$$

$$m^*\left(\bigcup_{i=1}^{\infty} O_i\right) = m^*\left(\bigcup_{i=1}^{\infty}\bigcup_{n=1}^{\infty} I_{i,n}\right) = \sum_{i=1}^{\infty}\sum_{n=1}^{\infty} \ell(I_{i,n}).$$

$$\textit{So, } m^*\left(\bigcup_{i=1}^{\infty} O_i\right) = \sum_{i=1}^{\infty} m^*(O_i).$$

Question 3.3.11. What are *all* those sets for which m^* is *countably additive*, that is,

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m^*(E_n)?$$

Example 3.3.12. Suppose G is an open and bounded set in $\mathbb{R}$. Then for $\forall \varepsilon > 0$, there exists a compact set $K \subset G$ such that $m^*(K) > m^*(G) - \varepsilon$.
Since G is bounded, $G \subset [\alpha, \beta] \implies m^*(G) \leq \beta - \alpha < \infty$. Further, G is open, therefore

$$G = \bigcup I_n \implies m^*(G) = \sum \ell(I_n) < \infty$$

So for $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} \ell(I_n) < \frac{\varepsilon}{2} \tag{1}$$

Let

$$K = \bigcup_{n=1}^{N} \left[a_n + \frac{\varepsilon}{4N},\ b_n - \frac{\varepsilon}{4N}\right], \quad I_n = (q_n, b_n)$$

Then

$$\begin{aligned} m^*(K) &= \sum_{n=1}^{N} m^* \left[a_n + \frac{\varepsilon}{4N},\ b_n - \frac{\varepsilon}{4N}\right] \\ &= \sum_{n=1}^{N} \left(\ell(I_n) - \frac{\varepsilon}{2N}\right) = \sum_{n=1}^{N} \ell(I_n) - \frac{\varepsilon}{2} \end{aligned}$$

Therefore,

$$\begin{aligned} m^*(K) &= \sum_{n=1}^{N} \ell(I_n) + \frac{\varepsilon}{2} - \varepsilon \\ &> \sum_{n=1}^{N} \ell(I_n) + \sum_{n=N+1}^{\infty} \ell(I_n) - \varepsilon \\ &= m^*(G) - \varepsilon \end{aligned}$$

Proposition 3.3.13. *If* $[a, b] \cap [c, d] = \emptyset$ *then*

$$m^*([a, b] \cup [c, d]) = m^*([a, b]) + m^*([c, d]).$$

Proof. Since $[a, b] \cap [c, d] = \emptyset$. Then $[a, b]$ and $[c, d]$ will be separated by some distance $\varepsilon > 0$. (Why?)

Suppose $[a, b] \cup [c, d] \subset \bigcup I_n$. Then

$$[a, b] \subset \bigcup (I_n \cap (a - \varepsilon, b + \varepsilon)) = \bigcup I_n' \quad \text{(Say)}$$
$$[c, d] \subset \bigcup (I_n \cap (c - \varepsilon, d + \varepsilon)) = \bigcup I_n'' \quad \text{(Say)}$$

Then $I_n' \cap I_l'' = \emptyset$, for all $n, m \geq 1$.

$$\implies m^*([a, b]) + m^*([c, d]) \leq \sum \ell(I_n') + \sum \ell(I_n'')$$

$$= \sum \ell(I_n \sqcup I_n'') = \sum \ell\{I_n \cap ((a - \varepsilon, b + \varepsilon) \cup (c - \varepsilon, d + \varepsilon))\}$$

$$m^*([a, b]) + m^*([c, d]) < \sum_{n=1}^{\infty} \ell(I_n)$$

$$m^*([a, b]) + m^*([c, d]) \leq m^*([a, b] \cup [c, d])$$

Since m^* is countably subadditive, other inequality holds. □

Observation: If G is an open and bounded subset of $\mathbb{R}$, then for each $\varepsilon > 0$, there is an open set O and a compact set K such that $K \subset G \subset O$ and $m^*(O) - m^*(K) < \varepsilon$.
In general, we fail to write

$$m^*(B \setminus A) = m^*(B) - m^*(A)$$

for $A \subseteq B$ (we shall see example later).

3.3.4 Lebesgue measurable sets

A set $E \subset \mathbb{R}$ is said to be Lebesgue measurable, if $\forall \varepsilon > 0$, there exists open set O and closed set F such that

$$F \subset E \subset O \text{ and } m^*(O \setminus F) < \varepsilon$$

Note: $m^*(O \setminus E) \leq m^*(O \setminus F) < \varepsilon$ and $m^*(F \setminus E) \leq m^*(O \setminus E) < \varepsilon$.
Thus, one can interpretate that Lebesgue measurable sets are approximately open and closed.

Proposition 3.3.14. *Let $\mathcal{M}$ denote the class of all Lebesgue measurable subsets of $\mathbb{R}$. Then*

(i) If $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$. $O^c \subset E^c \subset F^c$ and $m^*(F^c \setminus O^c) < \varepsilon$.

(ii) If $m^*(E) = 0$. Then $E \in \mathcal{M}$.

For $\varepsilon > 0$, there exist $O \supset E$ such that $m^*(O) < 0 + \varepsilon$. Let F be any closed set in E. Then $m^*(F) \leq m^*(E) = 0$.

therefore $m^*(O \setminus F) \leq m^*(O) < \varepsilon$. Thus, $E \in \mathcal{M}$.

(iii) If $\{E_n\}_{n=1}^{\infty} \subset \mathcal{M}$, then $E = \bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$.

Write $E_n' = E_n \setminus \bigcup_{i=1}^{n-1} E_i$, *then* $\bigcup E_n' = \bigcup E_n$, *where* E_n' *are pairwise disjoint sets (that is,* $E_n' \cap E_m' = \emptyset$ *for* $n \neq m$*). Thus, without loss of generality, one can assume* $E = \bigcup_{n=1}^{\infty} E_n$, $E_n \cap E_m = \emptyset$ *if* $n \neq m$.

Suppose $m^*(E) < \infty$, *then* $m^*(E_n) \leq m^*(E) < \infty$.

For $\varepsilon > 0$, $\exists\, F_n \subset E_n \subset O_n$ *such that* $m^*(O_n \setminus F_n) < \frac{\varepsilon}{2^n}$. *Now,*

$$\begin{aligned}\sum_{n=1}^{k} m^*(O_n) &\leq \sum_{n=1}^{k} m^*(O_n \setminus F_n) + \sum_{n=1}^{k} m^*(F_n) \\ &< \sum_{n=1}^{k} \frac{\varepsilon}{2^n} + m^*\left(\bigcup_{n=1}^{k} F_n\right) \qquad [\text{ since } F_n \text{ is closed and bounded}] \\ &< \varepsilon + m^*(E) < \infty, \qquad \forall k \geq 1.\end{aligned}$$

that is, $\sum_{n=1}^{\infty} m^*(O_n) < \infty$. *For* $\varepsilon > 0$, $\exists\, n_0 \in \mathbb{N}$ *such that* $\sum_{n=n_0+1}^{\infty} m^*(O_n) < \varepsilon,$.

Let $O = \bigcup_{n=1}^{\infty} O_n$ *and* $F = \bigcup_{n=1}^{n_0} F_n$. *Then,*

$$\begin{aligned}m^*(O \setminus F) &= m^*\left(\left(\bigcup_{n=1}^{n_0} O_n\right) \bigcup \left(\bigcup_{n=n_0+1}^{\infty} O_n\right) \setminus \left(\bigcup_{n=1}^{n_0} F_n\right)\right) \\ &\leq m^*\left(\bigcup_{n=1}^{n_0} (O_n \setminus F_n)\right) + m^*\left(\bigcup_{n=n_0+1}^{\infty} O_n\right) \quad (\text{since} A \cup B \setminus C = (A \setminus C) \cup (B \setminus C)) \\ &\leq \sum_{n=1}^{n_0} m^*(O_n \setminus F_n) + \sum_{n=n_0+1}^{\infty} m^*(O_n) \quad (F_n \subset E_n \subset O_n) \\ &< \sum_{n=1}^{n_0} \frac{\varepsilon}{2^n} + \varepsilon \\ &< 2\varepsilon.\end{aligned}$$

that is, $F \subset E \subset O$ *and* $\forall \varepsilon > 0$, $m^*(O \setminus F) < 2\varepsilon \implies E \in \mathcal{M}$.

If $m^*(E) = \infty$, *write* $E = \bigcup_{k \in \mathbb{Z}} E \cap [k, k+1) = \bigcup_{k \in \mathbb{Z}} A_k$ *and can be done in similar way.*

(iv) If $E_1, E_2 \in \mathcal{M}$, *then* $E_1 \cup E_2 = E_1 \sqcup (E_2 \setminus E_1)$. *But for* $\varepsilon > 0$, $\exists\, O_i \supset E_i \supset F_i$ *such that* $m^*(O_i \setminus F_i) < \frac{\varepsilon}{2}$; $i = 1, 2$.

For $O = O_1 \cup O_2$, $F = F_1 \cup F_2$,

$$O \setminus F = \bigcup_{i=1}^{2} (O_i \setminus F_i) \implies m^*(O \setminus F) < \varepsilon.$$

$$(E_1 \cap E_2)^c = E_1^c \cup E_2^c \in \mathcal{M}, \ \textit{since } E \in \mathcal{M}$$

$$\implies m^*(O \setminus F) < \varepsilon, \quad O^c \subseteq E^c \subseteq F^c$$

$$m^*(F^c \setminus O^c) = m^*(F^c \cap O) < \varepsilon$$

Thus, $\mathcal{M}$ is closed under countable union/intersection and complement.

Note: such family of sets is called a σ-algebra.

Definition 3.3.15. If $\mathcal{J} \subset \mathcal{P}(\mathbb{R})$ such that

(i) $A \in \mathcal{J} \implies A^c \in \mathcal{J}$.

(ii) $A_i \in \mathcal{J} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{J}$, then $\mathcal{J}$ is called a σ-algebra of sets.

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(a,b)\colon a,b \in \mathbb{R},\, a < b < \infty\}) \quad \textit{(Borel } \sigma\textit{-algebra)}$$

is the σ-algebra generated by countable union and complement of sets of type (a,b) and $a,b < \infty$.

Proposition 3.3.16. *Let $a,b \in \mathbb{R}$ and $a < b$, $b - a < \infty$. Then $I = (a,b) \in \mathcal{M}$.*

Proof. For $\varepsilon > 0$, $[a+\varepsilon, b-\varepsilon] \subset (a,b)$ and

$$m^*\{(a,b) \setminus [a+\varepsilon, b-\varepsilon]\} = m^*\{(a, a+\varepsilon) \sqcup (b-\varepsilon, b)\} \text{ (for small } \varepsilon > 0)$$

$$\leq m^*\{(a, a+\varepsilon)\} + m^*\{(b-\varepsilon, b)\}$$

$$= 2\varepsilon$$

Since I is open, it follows that $(a,b) \in \mathcal{M}$. Now $[a,b) = \{a\} \cup (a,b)$ and $m^*(\{a\}) = 0 \implies \{a\} \in \mathcal{M}$ and $(a,b) \in \mathcal{M}$. $\implies [a,b)$ and $[a,b] \in \mathcal{M}$
Thus, any open set $O = \bigcup_n I_n \in \mathcal{M}$. Since $\mathcal{M}$ is closed under complement, any closed set $F \in \mathcal{M}$. □

Example 3.3.17. If $A, B \subset \mathbb{R}$ such that $m^*(A) = 0$. Then $m^*(A \cup B) = m^*(B)$.

$$\text{since } m^*(A \cup B) \leq m^*(A) + m^*(B) = m^*(B) \leq m^*(A \cup B)$$

.

Proposition 3.3.18. *Let $x \in \mathbb{R}$ and $E \in \mathcal{M}$. Then $x + E \in \mathcal{M}$.*

Proof. For $\varepsilon > 0$, there exist $F \subset E \subset O$, O open, F closed such that $m^*(O \setminus F) < \varepsilon$.
But $F + x$ is closed and $O + x = \bigcup (I_n + x)$ is open with $F + x \subset E + x \subset O + x$.
Now, $m^*\left(O + x \setminus (F + x)\right) = m^*(O \setminus F) < \varepsilon$. □

Example 3.3.19. Verify that.

(i) $(F + x)^c = F^c + x$.

(ii) $(O+x)\cap(F+x)^c = O\cap F^c + x$.

(*Hint:* $z\notin F+x \implies z-x\notin F \implies z-x\in F^c \implies z\in F^c$ and so forth).

Theorem 3.3.20. *If* $E=\bigcup_{n=1}^{\infty}E_n$*,* $E_n\in\mathcal{M}$*. Then*

$$m^*\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}m^*(E_n)$$

(i) Suppose E is bounded, then $m^*(E)<\infty \implies m^*(E_n)<\infty$.
For $\varepsilon>0$, $\exists F_n\subset E_n\subset O_n$ such that $m^*(O_n\setminus F_n)<\frac{\varepsilon}{2^n}$.
Now,

$$\sum_{n=1}^{k}m^*(E_n)\le\sum_{n=1}^{k}(m^*(F_n)+m^*(O_n\setminus F_n))$$

$$<\sum_{n=1}^{k}m^*(F_n)+\sum_{n=1}^{k}\frac{\varepsilon}{2^n}<\sum_{n=1}^{k}m^*(F_n)+\varepsilon$$

$$(since E_n=(E_n\setminus F_n)\cup F_n\subseteq(O_n\setminus F_n)\cup F_n)$$

Since F_n's are compact (closed and bounded).

$$\sum_{n=1}^{k}m^*(E_n)<\sum_{n=1}^{k}m^*(F_n)+\varepsilon=\left(\bigcup_{n=1}^{k}F_n\right)+\varepsilon$$

$$\text{that is}\quad\sum_{n=1}^{k}m^*(E_n)<m^*(E)+\varepsilon,\quad\forall k\ge 1.$$

$$\implies\sum_{n=1}^{k}m^*(E_n)\le m^*(E)\le\sum_{n=1}^{k}m^*(E_n).$$

Now, suppose E is not bounded. Then, as

$$\mathbb{R}=\bigcup_{k=1}^{\infty}(k,k+1],$$

let

$$A_k=E\cap(k,k+1],\qquad E_{n,k}=E_n\cap(k,k+1].$$

Then

$$E=\bigcup_{k\in\mathbb{Z}}A_k,\qquad E_n=\bigcup_{k\in\mathbb{Z}}E_{n,k}.$$

Now,

$$\sum_{n=1}^{\infty}m^*(E_n)\le\sum_{n=1}^{\infty}\sum_{k\in\mathbb{Z}}m^*(E_{n,k})\tag{1}$$

Since $A_k = \bigcup_{n=1}^{\infty} E_{n,k}$*,* A_k *is bounded.*

$$m^*(A_k) = \sum_{n=1}^{\infty} m^*(E_{n,k}) \tag{2}$$

$$\textit{therefore } \sum_{n=1}^{\infty} m^*(E_n) \leq \sum_{k=-\infty}^{\infty} m^*(A_k) \tag{3}$$

Now,

$$\sum_{k=-l}^{l} m^*(A_k) = m^*\left(\bigcup_{k=-l}^{l} A_k\right) \leq m^*(E), \quad \forall l \geq 1$$

If $m^*(E) = \infty$*, okay, identity holds trivially. As*

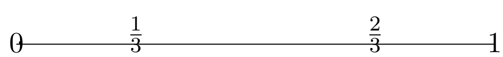
$$m^*(E) \leq \sum_{n=1}^{\infty} m^*(E_n), \quad \textit{let } m^*(E) < \infty,$$

$$\implies \sum_{k=-\infty}^{\infty} m^*(A_k) \leq m^*(E) \tag{4}$$

$$\implies \sum_{n=1}^{\infty} m^*(E_n) \leq m^*(E) \leq \sum_{n=1}^{\infty} m^*(E_n).$$

3.3.5 The Cantor set

The Cantor set is an *uncountable set* in $[0,1]$ having zero length with many peculiar properties, answering some of the difficult questions related to topology of real line.
Let $C_0 = [0,1]$.



Delete middle one-third open interval $J_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$ from C_0. Then

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$$

Delete one-third open interval from each section of C_1, and let

$$J_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$$

Then,

$$C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$$

Thus,

- $C_0 = [0, 1]$, one closed interval of length 1.
- $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, two closed disjoint intervals each of length $\frac{1}{3}$.
- $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, four closed disjoint intervals each of length $\frac{1}{9}$.

By induction, one can construct C_n with 2^n disjoint closed intervals each of length 3^{-n}.

(i) C_n is a decreasing sequence of closed and bounded sets, thus $C_n \in \mathcal{M}$.

(ii) Let $C = \bigcap_{n=1}^{\infty} C_n$, then C contains all the end-points of the intervals.

(iii) $C = [0, 1] \setminus \left\{ \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \dots \right\}$.

(iv) Since $C \subset C_n$, $\forall n \geq 0$,

$$m^*(C) \leq m^*(C_n) = 2^n \cdot \frac{1}{3^n} \to 0.$$

Thus, $m^*(C) = 0$.

(v) C is nowhere dense in $[0, 1]$, that is $(\overline{C})^o = (C^o) = \varnothing$.

If not so, then $C^o \neq \varnothing$ and $x \in C^o$. But C^o is open, there exist $(y, z) \subset C^o \subset C$, $y < z$. Thus, $m^*\{(y, z)\} \leq m^*(C) = 0$, contradiction.

(vi) <u>Cantor set is uncountable:</u>

Consider the endpoint $\frac{1}{3} \in C$. One can write

$$\frac{1}{3} = \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \dots \infty = (0.222\dots)_3$$

end point $x = \frac{2}{3} = (0.2)_3$. Similarly, we shall prove that each endpoint can be written as

$$x = \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots \infty, \quad a_i \in \{0, 2\}.$$

For this, consider the set

$$F = \left\{ x \in [0, 1] : x = \sum_{i=1}^{\infty} \frac{a_i}{3^i}, \quad a_i \in \{0, 1, 2\} \right\} \setminus \{\text{end points}\}$$

For $x \in F$, we have

$$x = \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots$$

Notice that $a_1 = 1$ if and only if $x \in \left(\frac{1}{3}, \frac{2}{3}\right)$ if and only if $x \notin C$.

$a_1 \neq 1$, $a_2 = 1$ if and only if $x \in \left(\frac{1}{9}, \frac{2}{9}\right) \sqcup \left(\frac{7}{9}, \frac{8}{9}\right)$ if and only if $x \notin C$.

Thus, if $a_i = 1$ for some i, if and only if $x \notin C$.

$\implies C = \{x \in [0,1] : x = \sum_{i=1}^{\infty} \frac{a_i}{3^i},\ a_i \in \{0,2\}\}$.

Define $f : C \to [0,1]$ by

$$f(x) = f\left(\sum_{i=1}^{\infty} \frac{a_i}{3^i}\right) = \sum_{i=1}^{\infty} \frac{a_i}{2} 2^{-i}$$

Then $\frac{a_i}{2} \in \{0,1\}$. Thus $f(x) \in [0,1]$.

<u>f is not one-one:</u>

$$f\left(\frac{1}{3}\right) = f\left((0.022\ldots)_3\right) = (0.011\ldots)_2 = (0.1)_2 = \frac{1}{2}$$

and

$$f\left(\frac{2}{3}\right) = f\left((0.2)_3\right) = (0.1)_2 = \frac{1}{2}$$

$$\implies f\left(\frac{1}{3}\right) = f\left(\frac{2}{3}\right)$$

Exercise 3.3.21. Prove that $f(x) = f(y)$ if and only if x, y are end points of one of the deleted open interval.

<u>f is an onto map:</u> Here $f : C \to [0,1]$ and let $y \in [0,1]$ such that

$$f(x) = y = \sum_{i=1}^{\infty} a_i \frac{1}{2^i}$$

Let

$$x = \sum \frac{2a_i}{3^i}$$

then $f(x) = y$ holds. Thus, f is onto. Therefore, C is an uncountable set, having outer measure zero.

3.3.6 Nonmeasurable sets

For $x, y \in \mathbb{R}$, define $x \sim y$ if and only if $x - y \in \mathbb{Q}$. Then $\sim$ is an equivalence relation on $\mathbb{R}$. Hence, it partitions $\mathbb{R}$ into disjoint equivalence classes.
Let $x + \mathbb{Q} = \{x + r : r \in \mathbb{Q}\}$. Then $x + \mathbb{Q}$ is an equivalence class under $\sim$.

(i) $(x + \mathbb{Q}) \cap [0,1] \neq \varnothing$ (easy).

(ii) Let E be a subset of $[0,1]$ that contains exactly one member from each $x+\mathbb{Q}$, $x \in \mathbb{R}$.

Let $\mathbb{Q} \cap [-1,1] = \{r_1, r_2, \dots\}$ and write $E_i = E + r_i, \qquad i = 1, 2, \dots$

(iii) $E_i \cap E_j = \varnothing$, if $i \neq j$.

If $z \in E_i \cap E_j$, then $z = x + r_i = y + r_j$

$$\implies x - y = r_j - r_i \in \mathbb{Q}$$

So $x \sim y$, contradiction to the definition of E, as E contains exactly one member from each $x + \mathbb{Q}$.

(iv) $[0,1] \subset \bigcup_{i=1}^{\infty} E_i \subset [-1,2]$.

Let $x \in [0,1]$. Then $x + \mathbb{Q}$ must contains a point of E. That is, there exists unique $y \in (x+\mathbb{Q}) \cap E$, $y - x \in \mathbb{Q} \cap [-1,1]$. Thus, $y - x = r_{i_0} \implies x = y - r_{i_0} \in E_{i_0}$.

The set E is not Lebesgue measurable. On the contrary, if $E \in \mathcal{M}$, then

$$1 \leq m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq 3$$

$$1 \leq \sum_{i=1}^{\infty} m^*(E) \leq 3$$

which is not possible, because either $m^*(E) > 0$. If $m^*(E) = 0$, then $m^*(E_i) = 0$. But $[0,1] \subseteq \bigcup E_i \implies 1 \leq \sum m^*(E_i) = 0$, which is a contradiction.

Remark 3.3.22. (i) m^* is not countably additive.

Let $A = \bigcup_{i=1}^{\infty} E_i$. Then $1 \leq m^*(A) \leq 3$. But $\sum_{i=1}^{\infty} m^*(E_i) = \infty$. Thus,

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq 3 < \infty = \sum_{i=1}^{\infty} m^*(E_i)$$

(ii) Whether m^* is finitely additive?

Suppose $m^*\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} m^*(A_i)$ for any $A_1, \dots, A_n \in \mathcal{P}(\mathbb{R})$ = power set of $\mathbb{R}$. (in other words, let m^* be finitely additive).

Now,

$$m^*(E) = \sum_{i=1}^{n} m^*(E_i)$$

$$= m^*\left(\bigcup_{i=1}^{n} E_i\right) \leq m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq 3$$

So,

$$m^*(E) < \frac{3}{n}, \quad \forall n \in \mathbb{N}$$

$\implies m^*(E) = 0$, contradiction.

Therefore, m^* cannot be finitely additive.

(iii) Suppose $A \subset E$ and $A \in \mathcal{M}$, then $m^*(A) = 0$.

For this, let $A_i = A + r_i,\ r_i \in \mathbb{Q} \cap [-1, 1]$.

Then,

$$\bigcup_{i=1}^{n} A_i \subset \bigcup_{i=1}^{\infty} E_i \subset [-1, 2]$$

Since A is Lebesgue measurable, each of $A_i \in \mathcal{M}$. Thus,

$$m^*\left(\bigcup_{i=1}^{n} A_i\right) \leq 3$$

$$\sum_{i=1}^{n} m^*(A_i) \leq 3$$

So,

$$m^*(A) \leq 3 \implies m^*(A) \leq \frac{3}{n},\ \forall n \geq 1$$
$$\implies m^*(A) = 0$$

We know that $m^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$. Restrict m^* to $\mathcal{M}$. Then for $E \in \mathcal{M}$, we write $m^*(E) = m(E)$. that is, $m^*|_{\mathcal{M}} = m \quad$ (say).

Theorem 3.3.23. *Let $(E_n) \subset \mathcal{M}$ be an increasing sequence of sets. Then*

$$\lim_{n\to\infty} m(E_n) = m\left(\bigcup_{n=1}^{\infty} E_n\right) \tag{*}$$

Proof. Let $E = \bigcup_n E_n$. If $m(E) = \infty$, then some of $m(E_{n_0}) = \emptyset$. Hence (*) holds. Therefore, suppose $m(E_n) < \infty, \quad \forall n \geq 1$. Since $m(E_n)$ is an increasing sequence.

$$\lim_{n\to\infty} m(E_n) = \sup_n m(E_n) \leq m(E).$$

Now,

$$\bigcup_{n=1}^{\infty} E_n = E_1 \bigcup_{n=1}^{\infty} (E_{n+1} \setminus E_n)$$

Thus,

$$m(E) = m(E_1) + \sum_{n=1}^{\infty} m(E_{n+1} \setminus E_n)$$

$$= m(E_1) + \lim_{k\to\infty} \sum_{n=1}^{k} (m(E_{n+1}) - m(E_n))$$
$$= \lim_{k\to\infty} m(E_{k+1})$$

□

Theorem 3.3.24. *Let $(E_n) \subset \mathcal{M}$ be a decreasing sequence of sets such that $m(E_1) < \infty$. Then*

$$\lim_{n\to\infty} m(E_n) = m\left(\bigcap_{n=1}^{\infty} E_n\right)$$

Proof. Since $m(E_n) \geq m(E_{n+1}) \geq m\left(\bigcap_{n=1}^{\infty} E_n\right)$,

$$\lim_{n\to\infty} m(E_n) = \inf_n m(E_n) \geq m\left(\bigcap_{n=1}^{\infty} E_n\right)$$

$$E_1 \setminus \bigcap_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} (E_n \setminus E_{n+1}) \qquad \text{(Exercise)}$$
$$m\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n \setminus E_{n+1})$$
$$m(E_1) - m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{k\to\infty} \sum_{n=1}^{k} (m(E_n) - m(E_{n+1}))$$
$$= m(E_1) - \lim_{k\to\infty} m(E_{k+1})$$
$$\implies m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{k\to\infty} m(E_{k+1})$$

Alternative: $E_1 \setminus E_n$ is a increasing in n.

$$\lim_{n\to\infty} m(E_1 \setminus E_n) = m\left(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\right)$$

$$m(E_1) - \lim_{n\to\infty} m(E_n) = m\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right)$$
$$= m(E_1) - m\left(\bigcap_{n=1}^{\infty} E_n\right)$$

So,

$$\lim_{n\to\infty} m(E_n) = m\left(\bigcap_{n=1}^{\infty} E_n\right)$$

□

Exercise 3.3.25. $E \in \mathcal{M}$ if and only if $E \cap (a,b) \in \mathcal{M}$, for all $a, b \in \mathbb{R}$.
If $E \in \mathcal{M}$, it follows immediately that $E \cap (a,b) \in \mathcal{M}$, for any $a, b \in \mathbb{R}$, because $(a,b) \in \mathcal{M}$.
Suppose $E \cap (a,b) \in \mathcal{M}$, for all $a, b \in \mathbb{R}$.
Then $E \cap (k, k+1] = E \cap (k, k+1) \cup (E \cap \{k+1\}) \in \mathcal{M}$ (since $m^*(E \cap \{k+1\}) = 0$). But $E = \bigcup_{k \in \mathbb{Z}} (E \cap (k, k+1]) \in \mathcal{M}$.

Theorem 3.3.26. *$E \in \mathcal{M}$ if and only if for all $A \subset \mathbb{R}$, we have*

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E) \tag{1}$$

But, proving (*), it is enough to prove

$$m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E)$$

Proof. If $m^*(A) = \infty$, then (1) is true.
Suppose $m^*(A) < \infty$, and $E \in \mathcal{M}$. Then there exist G_δ set $G \supseteq A$ such that $m^*(A) = m^*(G)$. (since $G = \bigcap_{n=1}^{\infty} O_n$)

therefore $m^*(A \cap E) + m^*(A \setminus E) \leq m^*(G \cap E) + m^*(G \setminus E) = m^*((G \cap E) \sqcup (G \setminus E)) = m^*(G) = m^*(A)$.

Now, let (1) holds. *Claim:* $E \in \mathcal{M}$.
First consider $m^*(E) < \infty$. Then there exist G_δ set G such that $E \subseteq G$ and $m^*(G) = m^*(E) < \infty$.
Since (1) is true for all $A \subset \mathbb{R}$,

$$m^*(G) = m^*(G \cap E) + m^*(G \setminus E)$$

$$\text{that is} \quad m^*(G) = m^*(G) + m^*(G \setminus E)$$

So, $m^*(G \setminus E) = 0 \implies G \setminus E \in \mathcal{M}$.
But $G \setminus (G \setminus E) = E \implies E \in \mathcal{M}$. If $m^*(E) = \infty$, then, write $E = \bigcup_{n \in \mathbb{Z}} (E \cap (n, n+1]) = \bigcup_{n \in \mathbb{Z}} E_n$.
We claim that $E \in \mathcal{M}$. For this, we all need to prove that if E_1, E_2 satisfy (1), then $E_1 \cap E_2$ satisfies (1). From the bounded case $(n, n+1] \in \mathcal{M} \iff (n, n+1]$ satisfies (1)). Thus,

$$m^*(A) = m^*(A \cap E_n) + m^*(A \setminus E_n)$$

Since $E_n = E \cap (n, n+1]$. Hence, by the bounded case $E_n \in \mathcal{M}$. Since $E = \bigcup E_n \implies E \in \mathcal{M}$.
Now,

$$m^*(A) = m^*(E_1 \cap A) + m^*(A \setminus E_1) \tag{2}$$

$$m^*(A) = m^*(E_2 \cap A) + m^*(A \setminus E_2) \tag{3}$$

Replace A in (3) by $A \cap E_1$ and $A \setminus E_1$ and use them in (2). Then R.H.S. of (1)

$$= m^*(E_1 \cap E_2 \cap A) + m^*(A \cap E_1 \setminus E_2) + m^*(E_2 \cap (A \setminus E_1)) + m^*(A \setminus E_2 \setminus E_1)$$

$$\geq m^*((E_1 \cap E_2 \cap A) \cup (A \cap E_1 \setminus E_2) \cup (E_2 \cap (A \setminus E_1)) \cup (A \setminus E_2 \setminus E_1))$$

$$\geq m^*(A) \quad \text{(using (1))}$$

Thus, $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$ will satisfy (1), as (1) is closed under complement. (1) is called *Carathéodory's criterion of measurability.* □

3.4 Measurable functions and integration

3.4.1 Measurable functions

Let $\mathcal{J}_u$ = collection of all open subsets of $\mathbb{R}$ with respect to the usual metric u on $\mathbb{R}$.

$$\{\mathcal{O} \subset \mathbb{R} : \mathcal{O} = \bigcup_{n=1}^{\infty} I_n, \quad I_n = (a_n, b_n)\}$$

and $\mathcal{M}$ = class of all Lebesgue measurable subsets of $\mathbb{R}$.

$\mathcal{J}_{d_0}$ = collection of all open sets of $\mathbb{R}$ with respect to d_0 — the discrete metric on $\mathbb{R} = \mathcal{P}(\mathbb{R})$.
$\implies \mathcal{J}_u \subsetneq \mathcal{M} \subsetneq \mathcal{J}_{d_0} = \mathcal{P}(\mathbb{R})$.
Since $\mathcal{J}_u$ is *not* closed under countable intersections (and complements) of open sets,
$\implies \mathcal{J}_u \subsetneq \mathcal{M}$ and $\mathcal{M} \subsetneq \mathcal{J}_{d_0}$, because every subset need *not* be Lebesgue measurable.
Consider $f : (\mathbb{R}, \mathcal{J}_u) \to (\mathbb{R}, \mathcal{J}_u)$ continuous. Then $f^{-1}(\mathcal{O}) \in \mathcal{J}_u,\ \forall\ \mathcal{O} \in \mathcal{J}_u$ (from range side).
Now, if $f : (\mathbb{R}, \mathcal{M}) \to (\mathbb{R}, \mathcal{J}_u)$, what happen to $f^{-1}(\mathcal{O})$? If f is continuous on $(\mathbb{R}, \mathcal{J}_u)$, then $f^{-1}(\mathcal{O})$ is open and hence $f^{-1}(\mathcal{O}) \in \mathcal{M}$.
In addition, consider $f(x) = \dfrac{1}{x},\ x \in \mathbb{R} \setminus \{0\}$, then f cannot be made continuous at 0 but $f(x) = \infty$ if and only if $x = 0$. (important!)
If we want to take $f(x) = \dfrac{1}{x}$ into consideration, we here to extend the range $(-\infty, \infty)$ to $[-\infty, \infty]$.
Let $\mathbb{R} = (-\infty, \infty)$ and $\overline{\mathbb{R}} = [-\infty, \infty]$. Therefore, the sets $[-\infty, a)$ and $(b, \infty]$ for $a, b \in \mathbb{R}$ should be added to $\mathcal{J}_u$. That is,

$$\overline{\mathcal{J}}_u = \mathcal{J}_u \cup \{[-\infty, a) \cup (b, \infty] : a, b \in \mathbb{R}\}$$

Definition 3.4.1. Let $f : (\mathbb{R}, \mathcal{M}) \to (\overline{\mathbb{R}}, \overline{\mathcal{J}}_u)$ is said to be Lebesgue measurable if $f^{-1}(\mathcal{O}) \in \mathcal{M}$, for all $\mathcal{O} \in \overline{\mathcal{J}}_u$.
Since $\mathcal{O} \in \overline{\mathcal{J}}_u$ can be expressed as the countable union/intersection of sets of the form $[-\infty, a)$ and $(b, \infty]$ and $\mathcal{M}$ is closed under countable union/intersection, it is enough to consider $\mathcal{O} = (b, \infty]$ or $[-\infty, a)$.

Thus, $f : (\mathbb{R}, \mathcal{M}) \to (\overline{\mathbb{R}}, \overline{\mathcal{J}}_u)$ or $\overline{\mathbb{R}}$ is Lebesgue measurable if $f^{-1}\{(\alpha, \infty]\} \in \mathcal{M}$, $\forall \alpha \in \mathbb{R}$.

Proposition 3.4.2. *If* $f : (\mathbb{R}, \mathcal{M}) \to \overline{\mathbb{R}} = [-\infty, \infty]$. *Then the following are equivalent:*

1. $f^{-1}\{(\alpha, \infty]\} \in \mathcal{M}$, *for all* $\alpha \in \mathbb{R}$.
2. $f^{-1}\{[\alpha, \infty]\} \in \mathcal{M}$, *for all* $\alpha \in \mathbb{R}$.
3. $f^{-1}\{[-\infty, \alpha)\} \in \mathcal{M}$, *for all* $\alpha \in \mathbb{R}$.
4. $f^{-1}\{[-\infty, \alpha]\} \in \mathcal{M}$, *for all* $\alpha \in \mathbb{R}$.
5. $f^{-1}\{\pm\infty\} \in \mathcal{M}$ *and* $f^{-1}\{(a, b)\} \in \mathcal{M}$, *for all* $a, b \in \mathbb{R}$.

Proof. $(i) \implies (ii)$:

$$[\alpha, \infty] = \bigcap_{n=1}^{\infty} (\alpha - \tfrac{1}{n}, \infty] \ni x,$$

let $x \notin [\alpha, \infty] \implies \alpha > x > \alpha - \frac{1}{n}, \quad \forall n \geq 1 \implies \alpha = x = \alpha$, which is a contradiction.
Since $\mathcal{M}$ is closed under complement, so $(ii) \implies (iii)$.
Now, $(iii) \implies (iv)$, because

$$[-\infty, \alpha] = \bigcap_{n=1}^{\infty} [-\infty, \alpha + \tfrac{1}{n})$$

$(iv) \implies (i)$ as $\mathcal{M}$ is closed under complements (since $\mathcal{M}^c = \mathcal{M}$).
Thus, (i) to (iv) are equivalent. Hence,

$$f^{-1}(\infty) = \bigcup f^{-1}\{(n, \infty]\} \in \mathcal{M} \quad \text{(by (i))}$$

$$f^{-1}(-\infty) = \bigcup f^{-1}\{[-\infty, -n)\} \in \mathcal{M} \quad \text{(by (iii))}$$

$$(a, b) = (a, \infty] \cap [-\infty, b)$$

$$\implies f^{-1}\{(a, b)\} \in \mathcal{M}, \quad \forall\, a, b \in \mathbb{R}$$

□

Example 3.4.3. Let $E \in \mathcal{M}$, define

$$f(x) = \chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

$$f^{-1}(\{(\alpha, \infty]\} = \begin{cases} E & \alpha = 0 \\ E & 1 > \alpha > 0 \\ \varnothing & \alpha \geq 1 \\ \mathbb{R} & \alpha < 0 \end{cases}$$

Example 3.4.4. $f : \mathbb{R} \to \overline{\mathbb{R}},\ f(x) = k$ is Lebesgue measurable.

$$f^{-1}\{(\alpha, \infty]\} = \begin{cases} \varnothing & \text{if } \alpha \geq k \\ \mathbb{R} & \text{if } \alpha < k \end{cases}$$

f k-finite. If $k = \infty$, $f(x) = \infty, \quad \forall x \in \mathbb{R}$. Then $f^{-1}\{(\alpha, \infty]\} = \mathbb{R}$.
Notice that for $\alpha \in \mathbb{R}$, $\exists r_j \in \mathbb{Q}$ such that r_j increases to α.
$f(x) \geq \alpha \implies f(x) \geq \alpha > r_j, \quad \forall j$.
So $\{x : f(x) > \alpha\} = \bigcap_{j=1}^{\infty}\{x : f(x) > r_j\}$. Thus, f is Lebesgue measurable if and only if $f^{-1}\{(r_j, \infty]\} \in \mathcal{M}$, for all $r_j \in \mathbb{Q}$.

Example 3.4.5. If $f, g : \mathbb{R} \to \overline{\mathbb{R}}$ be Lebesgue measurable such that $f(x) + g(x) \neq \infty - \infty$, for any $x \in \mathbb{R}$. Then $f + g$ is Lebesgue measurable.
Thus, we need to show the following sets to be Lebesgue measurable.

$$A = \{x \in \mathbb{R} : f(x) + g(x) = \pm\infty\}$$

$$B = \{x \in \mathbb{R} : \infty > f(x) + g(x) > \alpha\}, \quad \forall \alpha \in \mathbb{R}$$

$$A = \{x \in \mathbb{R} : f(x) = \pm\infty + g(x)\} \text{ if } g(x) \text{ are finite (or otherwise)}$$

For $x \in B$, $\infty > f(x) + g(x) > \alpha$, $\exists r_x$ such that $f(x) > r_x > \alpha - g(x)$

$$x \in \bigcup_{r \in \mathbb{Q}} (\{x : f(x) > r\} \bigcap \{x : g(x) > \alpha - r\})$$

$$\implies B = \bigcup_{r \in \mathbb{Q}} (\{x \in \mathbb{R} : f(x) > r\} \bigcap \{x \in \mathbb{R} : g(x) > \alpha - r\}) \implies B \in \mathcal{M}$$

Exercise 3.4.6. $\{x : f^2(x) > \alpha\} = \{x : f(x) > \sqrt{\alpha}\} \bigcup \{x : -f(x) > \sqrt{\alpha}\} \in \mathcal{M}$.

Exercise 3.4.7. $4fg = (f + g)^2 - (f - g)^2 \implies$ if f, g are Lebesgue measurable, then f^2, fg are Lebesgue measurable.

Definition 3.4.8. A property P is called "holding almost everywhere" if the places (or points) where it false have Lebesgue measure zero, that is, P is true almost everywhere.

$$m^*(\{x \in \mathbb{R} : P \text{ is false}\}) = 0$$

If $f = g$ almost everywhere on $\mathbb{R}$, then

$$m^*(\{x \in \mathbb{R} : f(x) \neq g(x)\}) = 0$$

Example 3.4.9. If $f : \mathbb{R} \to \overline{\mathbb{R}}$ and $f(x) = 0$ for almost everywhere $x \in \mathbb{R}$, then f is Lebesgue measurable.

Let $E = \{x \in \mathbb{R} : f(x) \neq 0\}$, then $m^*(E) = 0 \implies E, E^c \in \mathcal{M}$, and so forth.

Proposition 3.4.10. *If f, g are Lebesgue measurable, then*

$$\max\{f,g\} = \frac{f+g+|f-g|}{2}$$

$$\min\{f,g\} = \frac{f+g-|f-g|}{2}$$

$\sup f_n$*,* $\inf f_n$*,* $\limsup f_n$*,* $\liminf f_n$*,* $lim f_n$ *are all Lebesgue measurable.*

Proposition 3.4.11. *If $f : \mathbb{R} \to \overline{\mathbb{R}}$ be such that $f(x) \neq 0$, $\forall x \in \mathbb{R}$, then $\frac{1}{f}$ is measurable.*

Proof.

$$\left\{x : \frac{1}{f(x)} > \alpha\right\} = \{x : f(x) > \frac{1}{\alpha}, \alpha < 0\} \bigcup \{x : f(x) < \frac{1}{\alpha}, \alpha > 0\} \in \mathcal{M}$$

□

3.4.2 Simple functions

Let $E_i \in \mathcal{M}$ and $\alpha_i \in \overline{\mathbb{R}}$. Then $\varphi = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$ is called a simple function.

Example 3.4.12. $\varphi = 1 \cdot \chi_{[0,1]} + 2 \cdot \chi_{[2,3]}$

Theorem 3.4.13. *Let $f : \mathbb{R} \to [0, \infty]$ be a measurable function. Then there exist a sequence (φ_n) of simple functions such that:*

(i) $\varphi_n \uparrow$ *and* $\varphi_n \leq f$*.*

(ii) $\varphi_n \to f$ *pointwise.*

(iii) $\varphi_n \to f$ *uniformly on any set A where f is bounded.*

Proof. We first divide the image of f in $[0, 2^n]$ into 2^{2n} disjoint parts. $k = 0, 1, 2, \ldots, 2^{2n} - 1$.

$$f^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right) = E_{n,k} \quad \text{and} \quad f^{-1}\left([2^n, \infty]\right) = F_n$$

Then *(i)* $\varphi_n \geq 0$, *(ii)* $E_{n,k}$'s are disjoint measurable sets, *(iii)* $\varphi_n \uparrow$ on $[0, \infty]$.
Claim: $\varphi_n(x) \leq \varphi_{n+1}(x)$.
If $x \in E_{n,k} = \left\{x : \frac{2k}{2^{n+1}} \leq f(x) < \frac{2k+2}{2^{n+1}}\right\} = E_{n+1,2k} \bigcup E_{n+1,2k+1}$.
For $x \in E_{n+1,2k}$, $\varphi_n(x) = \frac{k}{2^n} = \frac{2k}{2^{n+1}} = \varphi_{n+1}(x)$.
For $x \in E_{n+1,2k+1}$, $\varphi_n(x) = \frac{2k+1}{2^{n+1}} = \varphi_{n+1}(x)$.
If $x \in F_n$, then $x \in (F_n \setminus F_{n+1}) \bigcup F_{n+1}$.
For $x \in F_{n+1}$, $\varphi_n(x) = 2^n < 2^{n+1} = \varphi_{n+1}(x)$.

For $x \in F_n \setminus F_{n+1}$, we have

$$2^n = \frac{2^{2n+1}}{2^{n+1}} \leq f(x) < 2^{n+1} = \frac{2^{2n+2}}{2^{n+1}}$$

that is, $x \in E_{n+1,2^{2n+1}} \bigcup \ldots \bigcup E_{n+1,2^{2n+2}-1}$. Then, $\varphi_{n+1}(x) \in \left\{\frac{2^{2n+1}}{2^{n+1}}, \ldots, \frac{2^{2n+2}-1}{2^{n+1}}\right\}$. Thus,

$$\varphi_n(x) = 2^n = \frac{2^{2n+1}}{2^{n+1}} \leq \varphi_{n+1}(x).$$

That is, $\varphi_n \uparrow$ and $\varphi_n \leq f$.
(iv) $\varphi_n \to f$ pointwise.

Let $f(x) < \infty$. Then

$$\{x : f(x) < \infty\} = \bigcup_{m=1}^{\infty} \{x : f(x) < 2^n\}$$

Therefore, $f(x) < 2^n$, for some n, and hence $x \in E_{n,k} \implies \varphi_n(x) = \frac{k}{2^n}$

$$\text{therefore } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \implies 0 \leq f(x) - \varphi_n(x) < \frac{1}{2^n}, \quad n \geq 1 \implies \varphi_n \to f \quad \text{pointwise.} \tag{*}$$

If $f(x) = \infty$, for some x. Then $\{x : f(x) = \infty\} = \bigcap_{n=1}^{\infty}\{x : f(x) \geq 2^n\}$.
So, $\varphi_n(x) = 2^n \to \infty = f(x)$.
(v) $\varphi_n \to f$ uniformly on a set where f is bounded.
Let $E = \{x : f(x) \leq M\}$. Then, $\exists n_0$ such that $f(x) < 2^n$, $\forall n \geq n_0$,
Hence, from (*), $0 \leq f(x) - \varphi_n(x) < \frac{1}{2^n}, \quad \forall n \geq n_0$
Notice that n_0 is free (or unique on E) of $x \in E$. Thus,

$$0 \leq \sup(f(x) - \varphi_n(x)) \leq \frac{1}{2^n} \to 0.$$

Hence, $\varphi_n \to f$ uniformly on E. □

Corollary 3.4.14. *If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable. Then there exists a sequence of simple functions such that $|\varphi_n| \uparrow |f|$ pointwise.*

Proof. $f = f^+ - f^-$. Then there exist $\varphi_n^+ \uparrow f^+$ and $\varphi_n^- \uparrow f^-$. That is,

$$\varphi_n = \varphi_n^+ - \varphi_n^- \to f^+ - f^- = f$$

$$|\varphi_n| = \varphi_n^+ + \varphi_n^- \leq f^+ + f^- \quad \text{and} \quad |\varphi_n| \uparrow |f|$$

In this case,

$$|f - \varphi_n| = |f^+ - \varphi_n^+ + f^- - \varphi_n^-| \to 0$$

and $\varphi_n \to f$ uniformly on $E = \{x : |f(x)| < M\}$. $\square$

Note that,

$$f^+ = \max\{f, 0\}, \quad f^- = -\min\{f, 0\}.$$

Adoptions: $0 \cdot \infty = 0, \quad \infty \cdot 0 = 0$.
Example: $0 \cdot m(\mathbb{R}) = 0$, $\infty \cdot m(\mathbb{Q}) = 0$.
Avoidation: $\infty - \infty$.

3.4.3 The Lebesgue integral

Let $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ such that

$$\varphi = \sum_{j=1}^{m} \alpha_j \chi_{E_j}, \quad \alpha_j \in [0, \infty],$$

and $E_j \in \mathcal{M}$ and $m(E_j) \leq \infty$. Then we write

$$\int_{\mathbb{R}} \varphi \, dm = \sum_{j=1}^{m} \alpha_j m(E_j).$$

Remark 1. $\int_{\mathbb{R}} \varphi \, dm = 0$ if and only if $\varphi = 0$.
Now, if $E \in \mathcal{M}$, then $\varphi|_E = \sum_{j=1}^{m} \alpha_j \chi_{E_j \cap E}$, hence

$$\int_{E} \varphi \, dm = \sum_{j=1}^{m} \alpha_j m(E_j \cap E)$$

Notice that $(\mathbb{R}, \mathcal{M}, m)$ is called Lebesgue measure space. If $E \in \mathcal{M}$, then for

$$\mathcal{M}_E = \{F \cap E : F \in \mathcal{M}\}, \quad (E, \mathcal{M}_E, m)$$

is also a Lebesgue measure space on E.
Remark 2. Since $E_1 \cup E_2 = (E_1 \setminus E_2) \cup (E_1 \cap E_2) \cup (E_2 \setminus E_1)$, in the definition of φ, one can assume $\{E_j : j = 1, 2, \ldots, n\}$ is a disjoint family, that is, $E_j \cap E_i = \varnothing$ if $i \neq j$.

Now, let $f : \mathbb{R} \to [0, \infty]$ be measurable, then there exists a sequence of simple functions $\varphi_n \uparrow f$ pointwise. Hence $\int_{\mathbb{R}} \varphi_n \, dm \uparrow$ sequence in $\overline{\mathbb{R}}$.
Thus, we define

$$\int_{\mathbb{R}} f \, dm := \sup_{n \geq 1} \int_{\mathbb{R}} \varphi_n \, dm$$

or

$$\int_{\mathbb{R}} f \, dm = \sup \left\{ \int_{\mathbb{R}} \varphi \, dm : \varphi \leq f \right\}$$

If $f : \mathbb{R} \to \overline{\mathbb{R}}$ measurable, then $f = f^+ - f^-$. We write

$$\int_{\mathbb{R}} f\,dm = \int_{\mathbb{R}} f^+\,dm - \int_{\mathbb{R}} f^-\,dm,$$

if at least either of $\int_{\mathbb{R}} f^+\,dm$ or $\int_{\mathbb{R}} f^-\,dm$ is finite.
Let

$$L^+(\mathbb{R}, \mathcal{M}, m) = \{f : \mathbb{R} \to [0,\infty] : f \quad \text{measurable}\}$$

Proposition 3.4.15. *For φ, ψ simple functions in $L^+(\mathbb{R}, \mathcal{M}, m)$ and $c \in \mathbb{R} = [0,\infty]$,*

(i) $\int_{\mathbb{R}} c\varphi = c\int_{\mathbb{R}} \varphi$.

(ii) $\int_{\mathbb{R}}(\varphi + \psi) = \int_{\mathbb{R}} \varphi + \int_{\mathbb{R}} \psi$.

(iii) If $\varphi \leq \psi$, then $\int_R \varphi\,dm \leq \int_R \psi\,dm$.

Proof. (i) is trivial.

(ii) Let $\varphi = \sum_{j=1}^m \alpha_j \chi_{E_j}, \quad \psi = \sum_{k=1}^m \beta_k \chi_{F_k}$.
Notice that by assigning 0 on $\left(\bigcup_{j=1}^n E_j\right)^c$, one can assume that $\mathbb{R} = \bigcup_{j=1}^n E_j$, $\mathbb{R} = \bigcup_{k=1}^m F_k$.
Then $E_j = \bigcup_{k=1}^m (E_j \cap F_k), \quad F_k = \bigcup_{j=1}^n (E_j \cup F_k)$.
Now,

$$\begin{aligned}\int_{\mathbb{R}} \varphi\,dm + \int_{\mathbb{R}} \psi\,dm &= \sum_{j=1}^n\sum_{k=1}^m \alpha_j m(E_j \cap F_k) + \sum_{k=1}^m\sum_{j=1}^n \beta_k m(E_j \cap F_k)\\ &= \sum_{k=1}^m\sum_{j=1}^n (\alpha_j + \beta_k) m(E_j \cap F_k) \qquad (1)\end{aligned}$$

$$\begin{aligned}\int_{\mathbb{R}} (\varphi+\psi)dm &= \int_{\mathbb{R}} \sum_{k=1}^m\sum_{j=1}^n (\alpha_j + \beta_k)\chi_{E_j \cap F_k}\,dm\\ &= \int_{\mathbb{R}} \varphi\,dm + \int_{\mathbb{R}} \psi\,dm \quad \text{(by (1))}\end{aligned}$$

(iii) If $\varphi \leq \psi$, then $\alpha_j \leq \beta_k$, when $E_j \cap F_k \neq \varnothing$.

$$\int_{\mathbb{R}} \varphi\,dm = \sum_{j=1}^n\sum_{k=1}^m \alpha_j m(E_j \cap F_k) \leq \sum_{j=1}^n\sum_{k=1}^m \beta_k m(E_j \cap F_k) = \int_{\mathbb{R}} \psi\,dm.$$

□

Proposition 3.4.16. *If $f, g \in L^+(\mathbb{R}, \mathcal{M}, m)$, then for $f \leq g$, $\int_{\mathbb{R}} f\,dm \leq \int_{\mathbb{R}} g\,dm$.*

For this, let $\varphi \leq f$, φ simple, then $\varphi \leq g$

$$\implies \int_{\mathbb{R}} f\,dm = \sup_{\varphi \leq f} \int_{\mathbb{R}} \varphi\,dm \leq \sup_{\varphi \leq g} \int_{\mathbb{R}} \varphi\,dm = \int_{\mathbb{R}} g\,dm.$$

Proposition 3.4.17. *If $f, g \in L^+(\mathbb{R}, \mathcal{M}, m)$, then*

$$\int_{\mathbb{R}} (f+g)\,dm = \int_{\mathbb{R}} f\,dm + \int_{\mathbb{R}} g\,dm.$$

(We prove it later!)

3.5 Convergence theorems and L^p spaces

3.5.1 Monotone convergence theorem

Theorem 3.5.1 (Monotone Convergence Theorem)**.** *Let $f_n, f \in L^+(\mathbb{R}, \mathcal{M}, m)$ be such that $f_n \uparrow f$ pointwise. Then*

$$\int_{\mathbb{R}} f\,dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n\,dm.$$

Proof: Since $f_n \leq f_{n+1} \leq f$, the limit of $\int_{\mathbb{R}} f_n$ will be bounded above by $\int_{\mathbb{R}} f$. Hence,

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \leq \int_{\mathbb{R}} f.$$

In order to show the other inequality, it is enough to show that for each $\epsilon > 0$,

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \geq (1-\epsilon) \int_{\mathbb{R}} f,$$

or for $\varphi \leq f$,

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \geq (1-\epsilon) \int_{\mathbb{R}} \varphi.$$

Let $E_n = \{x \in \mathbb{R} : f_n(x) \geq (1-\epsilon)\varphi(x)\}$. Since $f_n \uparrow f$, $E_n \subseteq E_{n+1}$. Moreover, $\mathbb{R} = \bigcup_{n=1}^{\infty} E_n$. For, let $x \in \mathbb{R}$, then $f_n(x) \uparrow f(x)$, and so for some n, $f_n(x) \geq (1-\epsilon)\varphi(x)$. If not, $f_n(x) < (1-\epsilon)\varphi(x)$ for all n, so $f(x) \leq (1-\epsilon)\varphi(x)$, $\varphi \leq f \Rightarrow$ Contradiction. Let $\nu(E_n) = \int_{E_n} \varphi$. Then ν becomes a measure on $(\mathbb{R}, \mathcal{M})$ and $E_n \uparrow \mathbb{R}$. Hence,

$$\lim_{n\to\infty} \nu(E_n) = \nu(\mathbb{R}).$$

Thus,

$$(1-\epsilon) \int_{\mathbb{R}} \varphi = \lim_{n\to\infty} \int_{E_n} (1-\epsilon)\varphi \leq \lim_{n\to\infty} \int_{E_n} f_n \leq \lim_{n\to\infty} \int_{\mathbb{R}} f_n.$$

Remark 3.5.2. $f_n \uparrow f$ is necessary in monotone convergence theorem.

Example 3.5.3. $f_n = \frac{1}{n}\chi_{[0,n]} \to 0$.

$$\int_{\mathbb{R}} f_n \, dm = 1 \neq 0 = \int_{\mathbb{R}} \lim f_n \, dm.$$

Example 3.5.4. Verify MCT for $f_n : \mathbb{R} \to [0, \infty]$, given by.

(i) $f_n = \chi_{(n,n+1)}$.

(ii) $f_n = n\chi_{(0,\frac{1}{n})}$.

Remark 3.5.5. Integration is a linear map on $L^+(\mathbb{R}, \mathcal{M}, m)$, that is, $f \mapsto \int_{\mathbb{R}} f \, dm$ is linear. Let $f, g \in L^+(\mathbb{R}, \mathcal{M}, m)$. Then there exists $\varphi_n \uparrow f$ and $\psi_n \uparrow g$. By MCT,

$$\begin{aligned}\int_{\mathbb{R}} (f+g) \, dm &= \lim_{n\to\infty} \int_{\mathbb{R}} (\varphi_n + \psi_n) \, dm \\ &= \lim_{n\to\infty} \left(\int_{\mathbb{R}} \varphi_n \, dm + \int_{\mathbb{R}} \psi_n \, dm \right) \\ &= \int_{\mathbb{R}} f \, dm + \int_{\mathbb{R}} g \, dm.\end{aligned}$$

Example 3.5.6. For $E \in \mathcal{M}$, and $f \in L^+(\mathbb{R}, \mathcal{M}, m)$, if $\int_E f \, dm = 0$, then $f = 0$, provided $m(E) > 0$.

$$\int_E f \, dm = \sup_{\varphi \leq f} \int_E \varphi = 0 \implies \int_E \varphi = 0 \implies \varphi = 0.$$

Corollary to MCT: Let $f_n, f \in L^+(\mathbb{R}, \mathcal{M}, m)$ be such that $f_n \uparrow f$ pointwise almost everywhere on $\mathbb{R}$. Then

$$\int_{\mathbb{R}} f \, dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n \, dm.$$

Proof. Let $f_n \uparrow f$ pointwise on A, then $m^*(A^c) = 0$. So, $A, A^c \in \mathcal{M}$. That is, $\chi_E f_n \to \chi_E f$. By MCT,

$$\int \chi_A f = \lim_{n\to\infty} \int \chi_A f_n$$

$$\implies \int_A f \, dm = \lim_{n\to\infty} \int_A f_n \, dm.$$

Now,

$$\int_{\mathbb{R}} f \, dm = \int_A f \, dm + \int_{A^c} f \, dm = \lim_{n\to\infty} \int_A f_n \, dm + \int_{A^c} f_n \, dm$$

Thus,

$$\int_{\mathbb{R}} f \, dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n \, dm$$

□

Theorem 3.5.7. *Let $f \in L^+(\mathbb{R}, \mathcal{M}, m)$. Then*

$$\int_{\mathbb{R}} f\, dm = 0 \iff f = 0 \text{ almost everywhere on } \mathbb{R}.$$

Proof. For $f = \varphi = \sum_{j=1}^n \alpha_j \chi_{E_j}$,

$$\int_{\mathbb{R}} \varphi\, dm = 0 \iff \text{either } \alpha_j = 0 \text{ or } m(E_j) = 0, \quad \forall j = 1, 2, \ldots, n$$

$$\text{that is} \quad \int_{\mathbb{R}} \varphi\, dm = 0 \iff \varphi = 0 \quad \text{almost everywhere}$$

Now, if $f = 0$ almost everywhere,

$$\int_{\mathbb{R}} f\, dm = \sup_{\varphi \leq f} \int_{\mathbb{R}} \varphi\, dm = 0 \qquad \text{(by previous case)}$$

Suppose $\int_{\mathbb{R}} f\, dm = 0$. Then consider

$$E = \{x \in \mathbb{R} : f(x) > 0\} = \bigcup_{n=1}^{\infty} \left\{x \in \mathbb{R} : f(x) > \frac{1}{n}\right\} = \bigcup_{n=1}^{\infty} E_n \quad \text{(say)}.$$

Now, $m(E_n) = n \int_{E_n} \frac{1}{n}\, dm \leq n \int_{E_n} f\, dm \leq n \int_{\mathbb{R}} f\, dm = 0$.
Thus, $m(E) = 0 \implies f = 0$ almost everywhere □

3.5.2 Fatou's lemma

Lemma 3.5.8 (Fatou's Lemma)**.** *Let $f_n \in L^+(\mathbb{R}, \mathcal{M}, m)$. Then*

$$\int_{\mathbb{R}} \underline{\lim} f_n\, dm \leq \underline{\lim} \int_{\mathbb{R}} f_n\, dm$$

Proof. Let $g_k = \inf_{n \geq k} f_n$. Then $g_k \leq f_j$, for all $j \geq k$. Thus, $\int_{\mathbb{R}} g_k \leq \inf_{j \geq k} \int_{\mathbb{R}} f_j$.
Now,

$$g_k \uparrow \sup_{k \geq 1} \left(\inf_{n \geq k} f_n \right)$$

By the Monotone Convergence Theorem (MCT),

$$\int_{\mathbb{R}} \underline{\lim} f_n\, dm = \int_{\mathbb{R}} \lim_{k \to \infty} g_k\, dm = \lim_{k \to \infty} \int_{\mathbb{R}} g_k\, dm \leq \lim_{k \to \infty} \inf_{j \geq k} \int_{\mathbb{R}} f_j\, dm$$

□

<u>*Remark 1:*</u> Strict inequality *can* hold.
For $f_n = \frac{1}{n} \chi_{[0,n]} \to 0$ uniformly, then $\int_{\mathbb{R}} \underline{\lim} f_n\, dm = 0 < 1 = \underline{\lim} \int_{\mathbb{R}} f_n\, dm$.

<u>*Remark 2:*</u> Fatou's Lemma need not hold beyond non-negative functions.

Example 3.5.9. let $f_n = -\frac{1}{n}\chi_{[n,2n]}, \quad \forall n \geq 1$.
Now, $\inf_{n\geq k} f_n(x) = \inf_{n\geq k}\left\{-\frac{1}{n}\right\} = -\frac{1}{k}$.

$$\sup_{k\geq 1}\left(\inf_{n\geq k} f_n(x)\right) = 0 \quad \text{that is} \quad \underline{\lim} f_n(x) = 0$$

$$\int_{\mathbb{R}} \underline{\lim} f_n = 0 > -1 = \underline{\lim}\int_{\mathbb{R}} f_n.$$

Let $f : (\mathbb{R}, \mathcal{M}, m) \to \overline{\mathbb{R}} = [-\infty, \infty]$ be measurable. Then $f = f^+ - f^-$ and f^+, f^- are $\mathcal{L}$-measurable.

Definition 3.5.10. If $\int_{\mathbb{R}} f^+ \, dm < \infty$ and $\int_{\mathbb{R}} f^- \, dm < \infty$ both hold, then we say f is *integrable*, and

$$\int_{\mathbb{R}} f \, dm = \int_{\mathbb{R}} f^+ \, dm - \int_{\mathbb{R}} f^- \, dm$$

Since $|f| = f^+ + f^-$. It follows that $\int_{\mathbb{R}} f \, dm$ is finite if and only if $\int_{\mathbb{R}} |f| \, dm$ is finite.
Let

$$L^1(\mathbb{R}, \mathcal{M}, m) = \left\{f : \mathbb{R} \to \overline{\mathbb{R}} : f \quad \text{measurable} \quad \text{and} \quad \int_{\mathbb{R}} |f| < \infty\right\}$$

We also use the symbols $L^1(\mathbb{R})$ or $L^1(\mathbb{R}, m)$ or $L^1(\mathbb{R}, \mathcal{M}, m)$.
Notice that L^1 is a linear space over $\mathbb{R}$.
Since

$$\int_{\mathbb{R}} |f| = 0 \iff |f| = 0 \text{almost everywhere} \iff f = 0 \text{almost everywhere}$$

If we adopt $f = 0$ if and only if $f = 0$ almost everywhere Then $L^1(\mathbb{R}, \mathcal{M}, m)$ is a normed linear space with $\|f\|_1 = \int_{\mathbb{R}} |f| \, dm$.

Proposition 3.5.11. *If $f \in L^1(\mathbb{R}, \mathcal{M}, m)$, then*

$$\left|\int_{\mathbb{R}} f \, dm\right| \leq \int_{\mathbb{R}} |f| \, dm$$

Proof.

$$\left|\int_{\mathbb{R}} f \, dm\right| = \left|\int_{\mathbb{R}} f^+ \, dm - \int_{\mathbb{R}} f^- \, dm\right| \leq \left|\int_{\mathbb{R}} f^+ \, dm\right| + \left|\int_{\mathbb{R}} f^- \, dm\right| = \int_{\mathbb{R}} f^+ \, dm + \int_{\mathbb{R}} f^- \, dm = \int_{\mathbb{R}} |f| \, dm$$

□

3.5.3 Chebyshev's inequality

Let $f \in L^1(\mathbb{R}, \mathcal{M}, m)$. Then $m\left(\{x \in \mathbb{R} : |f(x)| \geq \alpha\}\right) \leq \frac{1}{\alpha}\|f\|_1$.

Proof.

$$\text{the left-hand side} = \frac{1}{\alpha}\int_{\{x:|f(x)|\geq\alpha\}} \alpha \, dm \leq \frac{1}{\alpha}\int_{\{x:|f(x)|\geq\alpha\}} |f(x)| \, dm \leq \frac{1}{\alpha}\int_{\mathbb{R}} |f(x)| \, dm = \frac{1}{\alpha}\|f\|_1$$

. □

Corollary 3.5.12. *If $f \in L^1(\mathbb{R}, \mathcal{M}, m)$, then $m\{x \in \mathbb{R} : |f(x)| = \infty\} = 0$ that is, an L^1-function is almost finite.*

Proof. $m\{x : |f(x)| = \infty\} = m\{\bigcap\{x : |f(x)| \geq n\}\}$. But $m\{x : |f(x)| \geq n\} \leq \frac{1}{n}\|f\|_1$.
So, $m\{x : |f(x)| = \infty\} \leq m\{x : |f(x)| \geq n\} \leq \frac{1}{n}\|f\|_1 \to 0$ as $n \to \infty$. □

3.5.4 Dominated convergence theorem

Theorem 3.5.13 (Dominated Convergence Theorem)**.** *Let $f_n : (\mathbb{R}, \mathcal{M}, m) \to \overline{\mathbb{R}}$ be a sequence of measurable functions such that*

(i) $f_n(x) \to f(x)$ *pointwise, for all* $x \in \mathbb{R}$.

(ii) $|f_n| \leq g \in L^1(\mathbb{R}, \mathcal{M}, m)$.

Then

$$\int_{\mathbb{R}} f\, dm = \lim_{n\to\infty} \int_{\mathbb{R}} f_n\, dm$$

Proof. Since $f_n \to f$ pointwise and $|f_n| \leq g \in L^1(\mathbb{R}, \mathcal{M}, m)$,

$$\implies |f_n| \to |f| \implies |f| \leq g \in L^1 \implies f \in L^1$$

Now,

$$0 \leq g + f_n \to g + f \quad \text{pointwise}$$

$$0 \leq g - f_n \to g - f \quad \text{pointwise}$$

By Fatou's Lemma,

$$\int_{\mathbb{R}} (g + f)\, dm = \int_{\mathbb{R}} \lim_{n\to\infty} (g + f_n)\, dm \leq \underline{\lim} \int_{\mathbb{R}} (g + f_n)\, dm$$

$$\implies \int_{\mathbb{R}} f\, dm \leq \underline{\lim} \int_{\mathbb{R}} f_n\, dm \qquad (since \int_{\mathbb{R}} g < \infty)$$

Similarly,

$$\int_{\mathbb{R}} (g - f)\, dm = \int_{\mathbb{R}} \lim_{n\to\infty} (g - f_n)\, dm \leq \underline{\lim} \int_{\mathbb{R}} (g - f_n)\, dm$$

$$-\int_{\mathbb{R}} f\, dm \leq -\overline{\lim} \int_{\mathbb{R}} f_n\, dm$$

$$\text{that is} \quad \int_{\mathbb{R}} f\, dm \geq \overline{\lim} \int_{\mathbb{R}} f_n\, dm$$

So,

$$\overline{\lim} \int_{\mathbb{R}} f_n\, dm \leq \int_{\mathbb{R}} f\, dm \leq \underline{\lim} \int_{\mathbb{R}} f_n\, dm$$

$$\implies \lim \int_{\mathbb{R}} f_n = \int_{\mathbb{R}} f.$$

□

Exercise 3.5.14. Verify the Dominated Convergence Theorem for $f_n : (\mathbb{R}, \mathcal{M}, m) \to \mathbb{R}$, where

(i) $f_n = n\chi_{[0,\frac{1}{n}]}$.

(ii) $f_n = \frac{1}{n}\chi_{[n,n+1]}$.

(iii) $f_n = \chi_{[n,n+1]}$.

(*Hint:* $f_n \to 0$, $\int_{\mathbb{R}} f_n = 1$).

3.5.5 Bounded convergence theorem

Theorem 3.5.15 (Bounded Convergence Theorem)**.** *Let $E \in \mathcal{M}$ and $0 < \mu(E) < \infty$. If $f_n, f : (E, \mathcal{M}_E, m) \to \overline{\mathbb{R}}$ be such that*

(i) $|f_n(x)| \leq M$, $\forall n \in \mathbb{N}, \forall x \in E$.

(ii) $f_n \to f$ pointwise.

Then

$$\int_E f = \lim_{n\to\infty} \int_E f_n$$

Proof.

$$\int_E |f_n| \leq \int_E M = Mm(E) < \infty$$

So, f_n are dominated by M. And by Dominated Convergence Theorem,

$$\int_E f = \lim_{n\to\infty} \int_E f_n$$

□

Theorem 3.5.16. *If f is bounded. Then $f \in \mathcal{R}[a,b]$ if and only if f is continuous on $[a,b]$ almost everywhere, that is, there exists $g : [a,b] \to \mathbb{R}$ continuous such that $f = g$ almost everywhere.*

Theorem 3.5.17. *Every Riemann integrable function is Lebesgue integrable, that is, $\mathcal{R}[a,b] \subset \mathcal{L}^1[a,b]$.*
If $f \in \mathcal{R}[a,b]$, then $f = g$ almost everywhere, where g is continuous on $[a,b]$. Therefore, g is measurable and hence f is measurable.
If $f \in \mathcal{R}[a,b]$, then

$$\inf_P U(P,f) = \overline{\int_a^b} f(x)\,dx$$

$$\sup_P L(P,f) = \underline{\int_a^b} f(x)\,dx$$

both exist and are equal to $\int_a^b f(x)dx$. *But for Lebesgue integration, we only want*

$$\sup_P L(P,f) = \int f\,dm$$

Hence $f \in \mathcal{R}[a,b] \implies f \in \mathcal{L}^1[a,b]$.
(Note that this is just an intuition and not a proof.)

Theorem 3.5.18. *Let* $f \in \mathcal{R}[a,b]$. *Then* $f \in \mathcal{L}^1[a,b]$ *and*

$$\int_{[a,b]} f\,dm = \int_a^b f(x)\,dx$$

Proof. Let $I = [a,b]$ and $f \in \mathcal{R}(I)$, then there exists an increasing sequence of partitions P_n of I such that

$$\lim U(P_n,f) = \lim L(P_n,f) = \int_a^b f(x)dx.$$

For a partition P of $[a,b]$, denote

$$\varphi_P = \sum_{j=1}^{k} M_j \chi_{(t_{j-1},t_j]}, \qquad M_j = \sup_{[t_{j-1},t_j]} f(x)$$

and

$$\psi_P = \sum_{j=1}^{k} m_j \chi_{(t_{j-1},t_j]}, \qquad m_j = \inf_{[t_{j-1},t_j]} f(x),$$

where $P = \{a = t_0 < t_1 < \cdots < t_{j-1} < t_j < \cdots t_k = b\}$.
Then $\varphi_P \downarrow$ sequence and $\psi_P \uparrow$ sequence. Since $f \in \mathcal{R}(I)$, $\exists M, m > 0$ such that $m \le f(x) \le M$ but then

$$m \le \psi_{P_n}(x) \le f(x) \le \varphi_{P_n}(x) \le M \tag{1}$$

For each fixed $x \in I$, $\varphi_{P_n}(x) \downarrow$ sequence bounded below by m and $\psi_{P_n}(x) \uparrow$ sequence bounded above by M. Let

$$\lim_{n\to\infty} \varphi_{P_n}(x) = \varphi(x), \qquad \lim_{n\to\infty} \psi_{P_n}(x) = \psi(x)$$

Then

$$m \le \psi(x) \le f(x) \le \varphi(x) \le M \tag{2}$$

Then ψ and φ being limit of simple functions are measurable.
By Bounded Convergence Theorem,

$$\int_I \varphi\,dm = \lim_{n\to\infty} \int_I \varphi_n\,dm = \lim_{n\to\infty} U(P_n,f) = \int_a^b f(x)dx$$

Similarly,

$$\int_I \psi \, dm = \lim_{n\to\infty} \int_I \psi_n \, dm = \lim_{n\to\infty} L(P_n, f) = \int_a^b f(x)dx$$

Therefore

$$\int_I (\varphi - \psi) \, dm = 0 \iff \varphi - \psi = 0 \quad \text{almost everywhere} \quad (\textit{since } \varphi - \psi \geq 0)$$

From $\psi(x) \leq f(x) \leq \varphi(x)$ almost everywhere. So $f(x) = \psi(x)$ almost everywhere $\implies$ f is measurable. Thus,

$$\int_I f \, dm = \int_I \psi \, dm = \int_a^b f(x)dx$$

Note: $\mathcal{R}[a,b] \subsetneq \mathcal{L}^1[a,b]$. Since $f = \chi_{(\mathbb{R}\setminus\mathbb{Q})\cap[0,1]}$, $\int_{[0,1]} f \, dm = 1$ but $L(P,f) = 0$ and $U(P,f) = 1$, $\forall P$. □

3.5.6 L^p spaces

Definition 3.5.19. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $1 \leq p < \infty$. The space $L^p(X,\mu)$ consists of measurable functions f for which $\int_X |f|^p \, d\mu < \infty$, modulo equality μ-almost everywhere. The norm is

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$

For $p = \infty$, we set $\|f\|_\infty := \inf\{M > 0 : \ |f| \leq M \ \ \mu\text{-almost everywhere}\}$.

Theorem 3.5.20. *For each $1 \leq p \leq \infty$, the space $L^p(X,\mu)$ is complete.*