**M630 Notes** (Fall 21) **Meihe Chen**

Things marked in gray are definitions not directly from [BC09].

# Contents

# 1 Calculus stuff

*Notation.* My $\mathbb{N}$ includes 0. In consistency with the convention in [BC09], $A^{\sim}$ will be used to indicate the complement of the set $A$.

## 1.1 topology of $\mathbb{R}$

**Definition** (limit suprerior and limit inferior)**.**

$$\overline{\lim_{n\to\infty}}\, x_n = \limsup_{n\to\infty} x_n = \inf_n \sup_{k\geqslant n} x_k$$

$$\underline{\lim_{n\to\infty}}\, y_n = \liminf_{n\to\infty} y_n = \sup_n \inf_{k\geqslant n} y_k$$

In words:
Given $\epsilon > 0$, $\overline{\lim}\, x_n + \epsilon$ is no longer a lower bound for $\{\sup_{k\geqslant n} x_n\}_{n\in\mathbb{Z}_+}$. So there is some $N \in \mathbb{Z}_+$ with

$$\sup_{k\geqslant N} x_k < \overline{\lim}\, x_n + \epsilon,$$

meaning for all $j \geqslant N$,

$$x_j \ \leqslant\ \sup_{k\geqslant N} x_k \ <\ \overline{\lim}\, x_n + \epsilon.$$

OTOH, fix an $N \in \mathbb{Z}_+$, we have

$$\sup_{k\geqslant N} x_k \geqslant \inf_n \sup_{k\geqslant n} x_k > \overline{\lim}\, x_n - \epsilon.$$

So there is some $m \geqslant N$ for which $x_m > \overline{\lim}\, x_n - \epsilon$.

---

Analogously, for the lim inf guy, $\underline{\lim}\, y_n - \epsilon$ is not an upper bound for $\{\inf_{k\geqslant n} x_n\}_{n\in\mathbb{Z}_+}$ and we can find $N$ large such that for all $j \geqslant N$,

$$x_j > \underline{\lim}\, y_n - \epsilon.$$

And if we fix $N \in \mathbb{Z}_+$, we've

$$\inf_{k\geqslant N} y_k \leqslant \sup_n \inf_{k\geqslant n} y_k < \underline{\lim}\, y_n + \epsilon,$$

so there's some $m$ such that $y_m < \overline{\lim}\, x_n + \epsilon$.

**Proposition.** **a.** $\overline{\lim}(-x_n) = -(\underline{\lim}\, x_n)$;

**b.** $\underline{\lim}\, x_n + \underline{\lim}\, y_n \leqslant \underline{\lim}(x_n + y_n) \leqslant \overline{\lim}\, x_n + \underline{\lim}\, y_n \leqslant \overline{\lim}(x_n + y_n) \leqslant \overline{\lim}\, x_n + \overline{\lim}\, y_n$

*Proof.* a) is easy; b) is easy using a. ∎

---

**Proposition 1.2.3** (p. 3, existence of base $n$ expansion)**.** *Suppose $n > 1$ is an integer.*

**a.** *For every $x \in (0,1)$, there is a sequence $\{a_k\}_k \subset [0,n) \cap \mathbb{Z}$ for which*

$$x = \sum_{k=1}^{\infty} \frac{a_k}{n^k}. \tag{1.4}$$

*If $x$ is of the form $m/n^j$ for some $m \in [n^j - 1]$, then there are two such representations. Otherwise 1.4 is unique.*

**b.** *Conversely, given any $\{a_k\} \subset [0,n) \cap \mathbb{Z}$, the sum $\sum_{k=1}^{\infty} a_k/n^k$ converges to some $x \in [0,1]$.*

*Proof.* **a.** Define a sequence $\{a_k\}$ inductively as follows:

$$\begin{aligned}
a_1 &:= \max\{m \in \mathbb{N} \mid m/n \leqslant x\} & r_1 &:= x - a_1/n \\
a_2 &:= \max\{m \in \mathbb{N} \mid m/n^2 + a_1/n \leqslant x\} & r_2 &:= x - a_1/n - a_2/n^2 \\
&\vdots & &\vdots \\
a_k &:= \max\{m \in \mathbb{N} \mid m/n^k + \sum_{j=1}^{k-1} a_j/n^j \leqslant x\} & r_k &:= x - \sum_{j=1}^{k} a_j/n^j \\
&\vdots & &\vdots
\end{aligned}$$

i.e.

$$x = r_k + \sum_{j=1}^{k} \frac{a_j}{n^j}. \tag{1}$$

First we need to check $a_k \in [0 : n-1]$ for each $k$. Rewrite

$$a_k = \max\left\{ m \in \mathbb{N} \,\middle|\, m \leqslant n^k \left( x - \sum_{j=1}^{k-1} \frac{a_j}{n^j} \right) \right\},$$

we have

$$\begin{aligned}
n^k \left( x - \sum_{j=1}^{k-1} \frac{a_j}{n^j} \right) &= n^k \left( x - \frac{a_{k-1}}{n^{k-1}} - \sum_{j=1}^{k-2} \frac{a_j}{n^j} \right) \\
&= n \underbrace{\left[ n^{k-1} \left( x - \sum_{j=1}^{k-2} \frac{a_j}{n^j} \right) - a_{k-1} \right]}_{*}
\end{aligned}$$

But $a_{k-1}$ is really just $\lfloor n^{k-1}\left(x - \sum_{j=1}^{k-2} \frac{a_j}{n^j}\right)\rfloor$, so $(*) < 1$ and we get $a_k < n$.

It's clear that $a_k \geqslant 0$. Now we need to show convergence to $x$.

By our choice of the $a_k$'s, $\sum_{j=1}^{k} a_j < x$ for all $k \in \mathbb{Z}_+$, and so $\sum_{j=1}^{\infty} a_j \leqslant x$.

Suppose $\epsilon > 0$, I want to find some $N$ large such that $\sum_{j=1}^{N} a_j > x - \epsilon$. Then by the monotonicity of the partial sums, we will get $\sum_{j=1}^{\infty} a_j = x$. To see this, I guess the tricky part is realizing we can stick $n$-many $1/n^{k+1}$'s into $1/n^k$. For $k$'s small, our $1/n^k$ segments won't be enough to hit $x$ (and adding one more will go over it), so we need more segments of smaller length to get close to $x$, you know what I mean.

So, let $t \in \mathbb{Z}$ be minimal such that $n^t > 1/\epsilon$. This guarantees the intersection

$$(x - \epsilon, x] \cap \{kn^{-t}\}_{k \in \mathbb{N}}$$

is nonempty. The pain is to show $\sum_{j=1}^{N} a_j/n^j > x - \epsilon$ for appropriate $N$.

Remember how we defined our $a_k$'s:

$$a_k = \max\left\{ m \in \mathbb{N} \;\middle|\; m \leqslant n^k \left( x - \sum_{j=1}^{k-1} \frac{a_j}{n^j} \right) \right\},$$

and remember $a_k < n$ for all $k$. We stare at

$$n^{k+1}\left( x - \sum_{j=1}^{k} a_j/n^j \right) - a_{k+1} < 1$$

and bahaha,

$$\Rightarrow \quad x - \sum_{j=1}^{k} a_j/n^j < \frac{1 + a_{k+1}}{n^{k+1}} < \frac{1+n}{n^{k+1}} < \frac{n^2}{n^{k+1}} = \frac{1}{n^{k-1}}.$$

Setting $N = t + 1$ gives us

$$x - \sum_{j=1}^{N} a_j/n^j < \frac{1}{n^t} < \epsilon$$

and we are done.

For uniqueness.. now I understand why the authors of [BC09] didn't want to put a proof there.

**b.** skip for now.

∎

**Theorem 1.2.4** (Characterization of open sets in $\mathbb{R}$). *Every open set in $\mathbb{R}$ is a countable union of pairwise disjoint open intervals.*

*Proof.* Suppose $U$ is open in $\mathbb{R}$. For each $x \in U$, there's an open connected neighborhood of $x$ contained $U$, allowing us to define $I_x := (a, b)$, where

$$a = \inf\{u \mid (u, x) \subset U\} \quad \text{and} \quad b = \sup\{v \mid (x, v) \subset U\}.$$

To see $I_x \subset U$, suppose $w \in I_x = (a, b)$. By definition of $a, b$, there are some $a \leqslant u < w < v \leqslant b$ such that $(u, x)$ and $(x, v)$ are subsets of $U$. So $w \in (u, v) \subset U$. This show $U \supset \cup_{x \in U} I_x$. Reverse inclusion is clear.

Notice for each $I_x = (a, b)$, necessarily $a, b \notin U$. Indeed, if $a \in U$, then by openness of $U$, there is an interval neighborhood $(a - \delta, a + \delta)$ of $a$ contained in $U$. But then $(a - \delta, x) \subset U$, contradicting the minimality of $a$. Similarly we get $b \notin U$.

Now we show $I_x \cap I_y \neq \varnothing \Leftrightarrow I_x = I_y$. Suppose $I_x = (a, b) \cap I_y = (c, d) \neq \varnothing$. If $a > d$ or $b < c$ then $I_x, I_y$ are disjoint and we are done. So suppose otherwise, i.e. $a \leqslant d$ and $b \geqslant c$. The equalities can't happen at the same time, so let's also assume $a < d$ is strict. We know from last paragraph that $a, b, c, d$ are not in $U$. ... ∎

**Proposition 1.2.5.** *Every open set in $\mathbb{R}^2$ can be writtten as a disjoint union of closed straight line segments.* ∎

**Definition.** Suppose $A \subset X$. A point $x \in X$ is *adherent* to $A$ if every neighborhood of $x$ intersects $A$ at some point. Define $\bar{A}$ to be the set of all adherent points of $A$, called the *closure* of $A$.

We say $x$ is a *limit point* of $A$ if it is adherent to $A \setminus \{x\}$. The set of limit points of $A$ is called the *derived set* of $A$, denoted $A'$.

Z [BC09] defines limit points to be adherent points and accumulation points to be limit points.

## 1.2 annoying sets

**Example 1.2.7** (Cantor 1/3 set)**.** (I copied this from my topology notes.)

Construct $C$ as follows: Set $F_0 = [0,1]$. Delete the middle third $(1/3, 2/3)$ from $F_0$ to get $F_1$. Proceed recursively, given $F_n$, a union of $2^n$ closed intervals. Remove the union $M_n$ of all the open middle thirds of the intervals in $F_n$ to get $F_{n+1} := F_n \setminus M_n$... This gives us a sequence of closed set $F_0 \supset F_1 \supset F_2 \supset \cdots$.

The Cantor set is defined to be

$$C := \bigcap_{n=0}^{\infty} F_n = [0,1] \setminus \bigcup_{n=1}^{\infty} M_n.$$

Note that $C$ contains all the endpoints of the intervals in $F_n$ for all $n$, e.g. at the second step, $0, 1/9, 2/9, 1/3, 2/3, 7/9, 8/9, 1 \in C$. But these are not all the points in $C$!

Given $x \in [0,1]$, write $x$ in ternary (base 3) expansion

$$x = 0.a_1 a_2 a_3 \ldots = \sum_{k=1}^{\infty} a_k / 3^k$$

where $a_j \in \{0,1,2\}$. Then $x \in C$ iff $a_j \in \{0,2\}$ for all $j \in \mathbb{Z}_+$. (If $a_j = 1$, then $x$ lands in the middle third of one of the intervals at step $j$, so it gets removed at the next step.)

*Mini example:* $\frac{2}{9} + \frac{2}{9^2} + \frac{2}{9^3} + \cdots = \frac{2}{9}(1 + \frac{1}{9} + \frac{1}{9^2} + \frac{1}{9^3} + \cdots) = \frac{2}{9}\frac{1}{1-1/9} = 1/4 \Rightarrow 1/4 \in C$.

**Example** (a discrete set in $[0,1]$ whose closure is uncountable)**.** Given any interval $[a,b] \subset \mathbb{R}$, define a linear tansformation $f_{a,b} : [0,1] \to [a,b]$ that sends $t \mapsto a + t(b-a)$. Then the image of $f_{a,b}$ is contained in $[a,b]$, with $f_{a,b}(0) = a$ and $f_{a,b}(1) = b$.

We say the interval $(a,b)$ is *contiguous* to the closed set $F \subset \mathbb{R}$ if $a, b \in F$ and $(a,b) \subset F^{\sim}$. For each contiguous interval $(a,b)$ of $C$, use $f_{a,c}$ to send the set $\{1/n\}_n \subset (0,1)$ into $(a,c)$, where $c = (b-a)/2$ is the halfpoint of the interval. Then we have $\{f_{a,c}(1/n)\}_n$ discrete with $a$ in its closure. Do the same thing to $(c,b)$ using $f_{b,c}$, so $b$ sits in the closure of $\{f_{b,c}(1/n)\}_n$, discrete.

After running $(a,b)$ over all contiguous intervals of $C$, we get a discrete set

$$\bigcup_{(a,b) \text{ contiguous to } C} \{f_{a,c}(1/n)\} \cup \{f_{b,c}(1/n)\}_n$$

whose closure contains $C$, which is uncountable.

**Example 1.2.8** (Perfect symmetric set)**.** Given a sequence $\{\xi_k\}_k \subset (0, 1/2)$, define the *perfect symmetric set* $E \subset [0,1]$ determined by $\{\xi_k\}$ as follows:
Let $E_1^1 = [0, \xi_1]$, $E_1^2 = [1-\xi_1, 1]$ and set $E_1 = E_1^1 \cup E_1^2$.
Let $E_2^1 = [0, \xi_1\xi_2]$, $E_2^2 = [\xi_1 - \xi_1\xi_2, \xi_1]$, $E_2^3 = [1-\xi_1, 1-\xi_1+\xi_1\xi_2]$, $E_2^4 = [1-\xi_1\xi_2, 1]$ and set $E_2 = E_2^1 \cup E_2^2 \cup E_2^3 \cup E_2^4$...

Let me say this: at the $j$th step we take the leftmost and rightmost $\xi_j$ portion of each of the remaining intervals. So after the $j$th removal, the length (a good guess of the "measure" here) of

$E_j^i$ is just $\xi_1\xi_2\ldots\xi_j$. It follows that $E_j = \cup_{i\in[2^j]} E_j^i$ has length

$$m(E_j) = \sum_{i\in[2^j]} m(E_j^i) = 2^j \xi_1\xi_2\cdots\xi_j,$$

and so $E := \cap_{j=0}^{\infty} E_j$ has length

$$m(E) = \lim_{j\to\infty} 2^j \xi_1\xi_2\cdots\xi_j.$$

Since $\{\xi_j\} \subset (0,1/2)$, $\prod_{j=1}^{n+1}(2\xi_j) < \prod_{j=1}^{n}(2\xi_j)$. So in order to get $m(E) > 0$, the $\xi_j$'s better approach $1/2$ quick enough.

But how quickly? [BC09] says taking $\xi_k = 1/2 - 1/2^k$ is enough. This would have a linear rate of convergence with asymptotic error constant $1/2$:

$$\lim_{k\to\infty} \frac{|\xi_{k+1} - 1/2|}{|\xi_k - 1/2|} = \frac{1}{2}.$$

The measure of this particular $E$ is

$$m(E) = \lim_{j\to\infty} 2^j (1/2 - 1/2^2)(1/2 - 1/2^3)\cdots(1/2 - 1/2^j) = \lim_{j\to\infty} \prod_{k=1}^{j-1} 1 - 1/2^k.$$

(the thing in the limit looks like a q-Pochhammer symbol $(1/2;1/2)_{j-2}$.) I don't know.

Anyway, in the case of $m(E) > 0$, for large $n$, the removal process pretty much looks like bisecting the intervals.

*Properties of E.* From the definition we know it is closed. And as its name suggests, $E$ is perfect, i.e., every point in $E$ is a limit point of $E$. It is also nowhere dense (closure has empty interior) in $[0,1]$. This can be realized by looking at the length of $E$ in the $m(E) = 0$ case and the bisection idea in the $m(E) > 0$ case.

## 1.3 annoying functions

The construction of annoying sets allows us to define annoying functions on these sets.

**Example 1.3.1** (Volterra's function)**.** Let $E$ be a perfect symmetric set with $m(E) > 0$. On each contiguous interval, define

$$\varphi(x,a) = (x-a)^2 \sin\left(\frac{1}{x-a}\right).$$

Note that $\varphi(a,a) = 0$. Compute

$$\varphi'(x,a) = 2(x-a)\sin\left(\frac{1}{x-a}\right) - \cos\left(\frac{1}{x-a}\right).$$

For $k$ large enough such that $a + \frac{1}{k\pi} \in (a,b)$, we have

$$\varphi'(a + \frac{1}{k\pi}, a) = -\cos(k\pi) = (-1)^{k+1}.$$

Since $\varphi'$ is continuous, having $\varphi'$ oscillating between $\pm 1$ means that $\varphi'$ achieves 0 infinitely often. Let $a+y$ be the largest 0 of $\varphi'$ left to the half point $(a+b)/2$ of the interval $(a,b)$ and define on $[0,1]$

$$f(x) = \begin{cases} 0 & \text{if } x \in E \\ \varphi(x,a) & \text{if } x \in [a, a+y] \\ -\varphi(x,b) & \text{if } x \in [b-y, b] \\ \varphi(a+y,a) & \text{if } x \in [a+y, b-y] \end{cases}$$

Then we have

i. $f'$ exists on $(0,1)$ and vanishes on $E$;

ii. $f'$ is discontinuous on $E$.

**Example 1.3.7** (Ruler function)**.** Define $r : [0,1] \to [0,1]$ by

$$r(x) = \begin{cases} 0 & \text{if } x \in (0,1) \cap \mathbb{Q}^\sim \\ 1 & \text{if } x = 0 \\ 1/q & \text{if } x = p/q \in (0,1],\ (p,q) = 1 \end{cases}$$

Then $r$ is discontinous on $[0,1] \cap \mathbb{Q}$: for each $x = p/q \in [0,1] \cap \mathbb{Q}$, pick $0 < \epsilon < 1/q$. Then by density of $\mathbb{Q}^\sim$, there is some irrational $y$ sitting in every neighborhood of $x$, and so $|r(x) - r(y)| = |1/q - 0| > \epsilon$.

But $r$ is continuous on $(0,1) \cap \mathbb{Q}^\sim$: fix $x \in (0,1) \cap \mathbb{Q}^\sim$. Given $\epsilon > 0$, there are only finitely many $p, q \in \mathbb{Z}_+$ for which $p/q \in (0,1)$ and $q \leqslant 1/\epsilon$ ($p/q \in [0,1] \Leftrightarrow p \in [q-1]$). Label these rationals $x_1, \ldots, x_r$ and let $\delta$ be small such that $B(x,\delta)$ conatins none of these $x_i$'s. Then for any $y \in B(x,\delta)$, either $y$ is irrational and we are done, or $y$ is rational. In the latter case, by our choice of $\delta$, $y = m/n$ can't be any of the $x_i$'s, so we must have $n > 1/\epsilon$. Hence $|r(x) - r(y)| = |0 - 1/n| = 1/n < \epsilon$.

**Example 1.3.8** (General ruler function)**.** Suppose $\Xi = \{\xi_n\}_n$ is a sequence of nonnegative reals. Assume $\xi_n \to 0$ as $n \to \infty$. Define $r_\Xi : [0,1] \to \mathbb{R}$ by

$$r_\Xi(x) = \begin{cases} 0 & \text{if } x \in (0,1) \cap \mathbb{Q}^\sim \\ 1 & \text{if } x = 0 \\ \xi_q & \text{if } x = p/q \in (0,1],\ (p,q) = 1 \end{cases}$$

As in the ruler case, $r_\Xi$ is conitnuous on the irrationals and discontinuous at $p/q$ when $\xi_q \neq 0$. Proof of discontinuity is the same as the ruler case. For continuity, pick $N$ large such that $\xi_n < \epsilon$ for all $n \geqslant N$. Then there are only finitely many $p,q$'s with $q < N$. Pick $\delta$ small so that $B(x,\delta)$ contains none of these guys, then any rational in the ball must have $q \geqslant N$, consequently its image under $r_\Xi$ is no greater than $\xi_N$ which is less than $\epsilon$.

## 1.4 some theory before more annoying examples.

**Definition** (First category)**.** A set $E$ is of *first category* or *meager* if it can be written as a countable union of nowhere dense sets.

Sets that are not of first category are said to be of second category.

**Definition** (Baire space)**.** A topological space $X$ is a *Baire space* if the union of any countable collection of closed sets with empty interior in $X$ also has empty interior.

Equivalently, the intersection of any countable collection of open dense sets in $X$ is dense.

**Theorem** (Baire category theorem)**.**

*BCT1 Every complete metric space is a Baire space.*

*BCT2 Every locally compact Hausdorff space is a Baire space.*

*Corollary.* $\mathbb{R}$ is a Baire space.

**Example 1.xx** (Hardy function)**.**

**Example 1.xx.**

# 2 Measure theory

I guess this is what I take the class for.

## 2.1 algebra of sets

**Definition 2.1.1** (Rings and algebras of sets)**.**

Given a set $X$, let $\mathcal{P}(X)$ be the power set of $X$. A collection $R \subset \mathcal{P}(X)$ is a *ring* if it is closed under taking finite intersection, union and complements. $R$ is a $\sigma$-ring if it is closed under countable set operations. A ring (resp. $\sigma$-ring) $A$ is an algebra (resp. $\sigma$-algebra) if $X \in A$.

**Definition 2.1.3** (Borel sets)**.** The Borel sets $\mathcal{B} = \mathcal{B}(X)$ in $X \subset \mathbb{R}$ is the smallest $\sigma$-algebra in $\mathcal{P}(X)$ containing the open sets in $X$. It is called the *Borel algebra*.

$F_\sigma$ and $G_\delta$ sets are Borel sets. There are also weird things like $F_{\sigma\delta\sigma}$ sets, which seem impossible to work with.

**Example.**

**Definition.** Suppose $\mathcal{A} \subset \mathcal{P}(X)$. A set function $\mu : \mathcal{A} \to [0, +\infty]$ is *finitely additive* (resp. *$\sigma$-additive*) if for any finite sequence $\{A_1, \ldots, A_n\} \subset \mathcal{A}$ (resp. infinite sequence $\{A_n\}_{n \in \mathbb{Z}_+} \subset \mathcal{A}$) of disjoint sets, we have

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i) \qquad \left(\text{resp.} \quad \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).\ \right)$$

Notice that [BC09] does not require the union to be in the set $\mathcal{A}$! Some authors define $\sigma$-additive set functions to eat countable unions only when the union lands in the set. One can get stuck on exercise 2.14 following the latter approach (which should've been easy using the text's definition). Would not recommend that.

Pedantic comment: since $\bigcup_{i=1}^{n} A_i = A_1 \cup \cdots \cup A_n \cup \varnothing \cup \varnothing \cup \cdots$, every $\sigma$-additive function is finitely additive.

We say $\mu$ is a *measure* if it is $\sigma$-additive on a $\sigma$-algebra.

**Definition.** Suppose $\mu$ is a measure on a ring $\mathcal{R}$. A subset $E \in \mathcal{R}$ has *finite measure* if $\mu(E) < \infty$. $E$ is *$\sigma$-finite* if there is a sequence $\{E_j\}_j \subset \mathcal{R}$ that covers $E$, with the measure of each $E_j$ finite.

If every element of $\mathcal{R}$ is finite (resp. $\sigma$-finite), then we say $\mu$ is finite (resp. $\sigma$-finite) on $\mathcal{R}$.

## 2.2 Lebesgue measure on $\mathbb{R}$

Define for simple intervals $I = (a,b) \subset \mathbb{R}$, $m(I) = b - a$ to be its length.

For a set $A \subset \mathbb{R}$, define the *Lebesgue outer measure* of $A$ to be

$$m^*(A) = \inf_{\substack{\{I_n\}_n \\ \text{countable} \\ \text{cover of} A}} \left\{ \sum_{n=1}^{\infty} m(I_n) \right\}.$$

Easy properties:

**a.** $A \subset B \Rightarrow m^*(A) \leqslant m^*(B)$;

**b.** $m^*(\varnothing) = 0$.

**Proposition 2.2.1.** *Let $I \subset \mathbb{R}$ be an interval. Then $m^*(I) = m(I)$.*

*Proof.* We bound the value from above and below.

Write $I = [a,b]$. For every $\epsilon > 0$, we have $[a,b] \subset (a-\epsilon, b+\epsilon)$. So $m^*(I) \leqslant m((a-\epsilon, b+\epsilon)) = b - a + 2\epsilon$. This gives $m^*(I) \leqslant b - a$.

On the other hand, suppose $\{I_n\}_n$ is an open cover of $I$ made up of open intervals. By compactness of $I$ we can pass to a finite subcover $\{I_{n_j}\}_{j\in[m]}$. Then

$$\sum_{n=1}^{\infty} m^*(I_n) \geqslant \sum_{n=1}^{k} m^*(I_{n_j}).$$

By relabeling, suppose $\{I_{n_j}\} = \{J_j\}$, where the indices $j \in [k]$ are to be determined later. Pick $J_1$ so that $a \in J_1 = (a_1, b_1)$. If $b < b_1$ then

$$\sum_{j=1}^{k} m(J_j) \geqslant m(J_1) = b_1 - a_1 > b - a$$

and we are done. Otherwise $b \geqslant b_1$, then $b_1 \in [a,b]$ and so it's in some $J_2 = (a_2, b_2)$. Again, if $b < b_2$ we are done. Otherwise, pick $J_3$ and conitinue this process. Since $k < \infty$, this has to end at some step $n \leqslant k$. By choice of intervals, we have $a_n < b < b_n$. Hence

$$\sum_{j=1}^{k} m(J_j) \geqslant \sum_{j=1}^{n} m(J_j) = \sum_{j=1}^{n} (b_j - a_j) = b_n - (a_n - b_{n-1}) - \cdots - a_1. \tag{2.2.1}$$

But $a_j < b_{j-1} < b_j$ for each $j$, so each of the things in the parentheses on the RHS of 2.2.1 is greater than 0. Hence 2.2.1$> b_n - a_1 > b - a$, which shows the reverse inequality.

For open intervals $I$ just pick some closed interval $J \supset I$ such that $m(J) > m(I) - \epsilon$. Then

$$m(I) - \epsilon < m(J) = m^*(J) \leqslant m^*(I) \leqslant m^*(\bar{I}) = m(\bar{I}) = m(I).$$

So $m(I) = m^*(I)$. ∎

**Proposition 2.2.2.** ($\sigma$-subadditivity of outer measure)**.** *Suppose $A_n \subset \mathbb{R}$, $n \in \mathbb{Z}_+$, then*

$$m^*(\bigcup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} m^*(A_n)$$

*Proof.* WLOG assume $m^*(A_n) < \infty$ for each $n$. Fix $\epsilon > 0$. Set up a 1-1 correspondence between $A_n$ and $\{I_{n,i}\}_i$, where $\{I_{n,i}\}_i$ is a countable family of open intervals whose union covers $A_n$ and that $\sum_{i=1}^{\infty} m^*(I_{n,i}) < m^*(A_n) + \epsilon/2^n$.

Then

$$\bigcup_{i=1}^{\infty} A_n \subset \bigcup_{n,i} I_{n,i}.$$

But the RHS is a countable union of open intervals, so

$$\begin{aligned} m^*\left(\bigcup_{i=1}^{\infty} A_n\right) &\leqslant m^*\left(\bigcup_{n,i} I_{n,i}\right) \\ &= \sum_{n=1}^{\infty}\left(\sum_{i=1}^{\infty} m^*(I_{n,i})\right) \\ &\leqslant \sum_{n=1}^{\infty}(m^*(A_n) + \epsilon/2^n) \\ &= \epsilon + \sum_{n=1}^{\infty} m^*(A_n) \end{aligned}$$

which shows that $m^*\left(\bigcup_{i=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} m^*(A_n)$. ∎

More properties of $m^*$:

*n+1*) If $S$ is countable, then $m^*(S) = 0$.

*n+2*) card $[0,1] > \aleph_0$.

*n+3*) For any subset $A \subset \mathbb{R}$, any $\epsilon \in \mathbb{R}_+$, there is some open neighborhood $U$ of $B$ satisfying $m^*(U) \leqslant m^*(A) + \epsilon$. [this comes by definition of $m^*$.]

*n+4*) there is some $G_\delta$ set $G$ containing $A$ with $m^*(A) = m^*(G)$.

[Use last item and take intersection over $1/n$]

The following is a very nice characterization of something being measurable.

**Definition** (Carathéodory condition)**.** A set $A \subset \mathbb{R}$ is *Lebesgue measurable* if for every subset $E \subset \mathbb{R}$,

$$m^*(E) = m^*(E \cap A) + m^*(E \cap A^{\sim}).$$

Write $\mathcal{M}(\mathbb{R})$ for the collection of all Lebesgue measurable sets in $\mathbb{R}$.
Some immediate consequences:

° If $m^*(A) = 0$ then $A$ is Lebesgue measurable.

°° If $A \in \mathcal{M}(\mathbb{R})$ then $A^\sim \in \mathcal{M}(\mathbb{R})$. [By Carathéodory.]

°°° If $A, B \in \mathcal{M}(\mathbb{R})$ then $A \cup B \in \mathcal{M}(\mathbb{R})$.
[ Suppose $E \subset \mathbb{R}$. By measurability of $A$ and $B$:

$$\begin{aligned} m^*(E) &= m^*(E \cap A) + m^*(E \cap A^\sim) \\ &= m^*(E \cap A \cap B) + m^*(E \cap A \cap B^\sim) \\ &\qquad + m^*(E \cap A^\sim \cap B) + m^*(E \cap A^\sim \cap B^\sim) \end{aligned}$$

Similarly,

$$\begin{aligned} m^*(E \cap (A \cup B)) &= m^*(E \cap (A \cup B) \cap A \cap B) + m^*(E \cap (A \cup B) \cap A \cap B^\sim) \\ &\qquad + m^*(E \cap (A \cup B) \cap A^\sim \cap B) + m^*(E \cap (A \cup B) \cap A^\sim \cap B^\sim) \\ &= m^*(E \cap A \cap B) + m^*(E \cap A \cap B^\sim) \\ &\qquad + m^*(E \cap A^\sim \cap B) + m^*(\varnothing) \\ &= m^*(E \cap A \cap B) + m^*(E \cap A \cap B^\sim) + m^*(E \cap A^\sim \cap B) \\ &= m^*(E) - m^*(E \cap A^\sim \cap B^\sim) \end{aligned}$$

i.e. $m^*(E) = m^*(E \cap (A \cup B)) + m^*(E \cap (A \cup B)^\sim)$ and we are done.]

°°°° $\mathcal{M}(\mathbb{R})$ is an algebra.
[This comes from that $\mathbb{R}$ is measurable, which follows from parts °° and °°°.]

Now we establish that Lebesgue measurable sets form a $\sigma$-algebra:

**Theorem 2.2.5** (Lebesgue and Borel $\sigma$-algbebras on $\mathbb{R}$)**.**

**a.** $\mathcal{M}(\mathbb{R})$ *is a* $\sigma$*-algebra.*

**b.** $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$.

*Proof.* **a.**

Step 1. (finite additivity) Suppose $\{A_i\}_{i=1}^n \subset \mathcal{M}(\mathbb{R})$ is a disjoint family. We show by induction that $m(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n m(A_i)$. For $n = 1$ we have nothing to prove. Suppose claim holds for $1, \ldots, n-1$. Take $E \subset \mathbb{R}$ and note that

$$E \cap \left( \bigcup_{i=1}^n A_i \right) \cap A_n = E \cap A_n,$$

and

$$E \cap \left( \bigcup_{i=1}^n A_i \right) \cap A_n^\sim = E \cap \left( \bigcup_{i=1}^{n-1} A_i \right)$$

as the $A_i$'s are disjoint. Using properties of outer measure and induction, we obtain

$$m^*\left(E \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) = m^*\left[E \cap \left(\bigcup_{i=1}^{n} A_i\right) \cap A_n\right] + m^*\left[E \cap \left(\bigcup_{i=1}^{n} A_i\right) \cap A_n^{\sim}\right]$$
$$= m^*\left[E \cap A_n\right] + m^*\left[E \cap \left(\bigcup_{i=1}^{n-1} A_i\right)\right]$$
$$= m^*\left(E \cap A_n\right) + \sum_{i=1}^{n-1} m^*\left(E \cap A_i\right).$$

Step 2. ($\sigma$-additivity) Suppose $\{A_i\}_{i=1}^{\infty} \subset \mathcal{M}(\mathbb{R})$ is a disjoint family. Set $A = \bigcup_{i=1}^{\infty} A_i$, $B_n = \bigcup_{i=1}^{n} A_i$. By Step 1 we know $\mathcal{M}$ is an algebra, so $B_n \in \mathcal{M}$. Note that $B_n \subset A \Leftrightarrow B_n^{\sim} \supset A^{\sim}$. Take $E \subset \mathbb{R}$, we have

$$m^*(E) = m^*(E \cap B_n) + m^*(E \cap B_n^{\sim}) \geqslant m^*(E \cap B_n) + m^*(E \cap A^{\sim}).$$

Well,

$$m^*(E \cap B_n) = m^*\left[E \cap \left(\bigcup_{i=1}^{n} A_i\right)\right] = \sum_{i=1}^{n} m^*(E \cap A_i)$$

so

$$m^*(E) \geqslant \sum_{i=1}^{n} m^*(E \cap A_i) + m^*(E \cap A^{\sim}).$$

Send $n \to \infty$ on the RHS and apply subadditivity of outer measure, get

$$m^*(E) \geqslant \sum_{i=1}^{\infty} m^*(E \cap A_i) + m^*(E \cap A^{\sim})$$
$$\geqslant m^*\left[E \cap \left(\bigcup_{i=1}^{\infty} A_i\right)\right] + m^*(E \cap A^{\sim})$$
$$= m^*(E \cap A) + m^*(E \cap A^{\sim})$$

This shows one inequality. The other one follows from subadditivity of $m^*$.

**b.** Since $\mathcal{B}(\mathbb{R})$ is the minimal $\sigma$-algebra containing open rays of the form $(a, \infty)$, it suffices to show the $(a, \infty)$'s are in $\mathcal{M}(\mathbb{R})$.

Take $E \subset \mathbb{R}$ and put $E_1 := E \cap (a, \infty)$, $E_2 := E \cap (\infty, a]$. Again by subadditivity of outer measure, $m^*(E) \leqslant m^*(E_1) + m^*(E_2)$. We need to show the reverse inequality. If $m^*(E) = \infty$ we have nothing to prove. So let's assume $m^*(E) < \infty$.

Suppose $\epsilon > 0$. Use the definition of $m^*$ to produce a family of open intervals $I_j$ whose union covers $E$ and for which

$$\sum_j m(I_j) \geqslant m^*(E) + \epsilon.$$

Set $I_j^1 := I_j \cap (a,\infty)$, $I_j^2 := I_j \cap (\infty, a]$, then

$$m(I_j) = m(I_j^1) + m(I_j^2) = m^*(I_j^1) + m^*(I_j^2).$$

Also, for $i = 1,2$, $E_i \subset \bigcup_j I_j^i$, so

$$m^*(E_i) \leqslant m^*\left(\bigcup_j I_j^i\right) \leqslant \sum_j m^*(I_j^i).$$

Consequently

$$m^*(E_1) + m^*(E_2) \leqslant \sum_j m^*(I_j^1) + \sum_j m^*(I_j^2) = \sum_j m^*(I_j^1) + m^*(I_j^2) = \sum_j m^*(I_j) \leqslant m^*(E) + \epsilon.$$

As $\epsilon$ is arbitrary, we've shown $m^*(E) \geqslant m^*(E_1) + m^*(E_2)$, which finishes the proof.

∎

**Theorem 2.2.6** (sequences of Lebesgue measurable sets)**.** *Suppose $\mathcal{A} = \{A_n\}$ is a family of Lebesgue measurable sets.*

**a.** $m(\bigcup A_n) \leqslant \sum m(A_n)$

**b.** *If $\mathcal{A}$ is disjoint, then $m(\bigcup A_n) = \sum m(A_n)$*

**c.** *If $\mathcal{A}$ is decreasing (i.e. $A_i \supset A_{i+1}$) and $m(A_1) < \infty$, then $m(\bigcap A_n) = \lim m(A_n)$.*

**d.** *If $\mathcal{A}$ is increasing (i.e. $A_i \subset A_{i+1}$), then $m(\bigcup A_n) = \lim m(A_n)$.*

*Proof is one in the same as that for a general measure, see the amazing Theorem 2.4.3.*

**Definition 2.2.7** (Lebesgue measure)**.** The set function $\mu : \mathcal{M}(\mathbb{R}) \to [0, +\infty]$ defined by $m(A) = m^*(A)$ is the *Lebesgue measure* on $\mathbb{R}$.

More examples:

**Example 2.2.9a** (a set of second category of measure zero)**.** Enumerate the rationals in $[0,1]$ by $\{r_n\}_n$. Let $I_{jk} = B(r_j, 1/2^j k)$, $U_k = \cup_{j=1}^{\infty} I_{jk}$ and set

$$D = \bigcap_{k=3}^{\infty} U_k.$$

First note that each $U_k$ is dense open in $[0,1]$ because the rationals are dense. By the Baire category theorem, $[0,1]$ is a Baire space, so $D$ being a countable intersection of dense open sets is dense in $[0,1]$. On the other hand, $U_k$ being dense open in $[0,1]$ also implies that $U_k^\sim$ is nowhere dense. So $\bigcup_{k=3}^{\infty} U_k^\sim$ is of first category, which means $D = \left(\bigcup_{k=3}^{\infty} U_k^\sim\right)^\sim$ is of second category.

Additionally, observe that the $U_k$'s are nested (decreasing in $k$), so by Theorem 2.2.6 c,

$$m(D) = m\Big(\bigcap_{k \geqslant 3} U_k\Big) = \lim_{k\to\infty} m(U_k) \leqslant \lim_{k\to\infty} \sum_{j=1}^{\infty} m(I_{jk}) = \lim_{k\to\infty} \frac{1}{k} = 0.$$

**Example 2.2.9b** (a function continuous strictly on a dense set of measure zero)**.** Define

$$f_k(x) = \begin{cases} 0 & \text{if } x \in U_k \\ 1/2^k & \text{otherwise} \end{cases}$$

and set

$$f = \sum_{k \geqslant 3} f_k.$$

If $x \in D$, then $x \in U_k$ for all $k \geqslant 3$. Consequently $f_k(x) = 0$ for all $k \geqslant 3$ and so $f(x) = 0$. If $y \notin D$, then there is some $r_y$ minimal for which $y \notin U_{r_y+1} \supset U_{r_y+2} \supset \cdots$. This means

$$f(y) = \sum_{k \geqslant 3} f_k(y) = \sum_{k \geqslant r_y+1} 1/2^k = 2^{-r_y} \sum_{k \geqslant 1} 1/2^k = 2^{-r_y}.$$

Given $\epsilon > 0$. Let $N$ be large such that $2^{-N} < \epsilon$. Then for all $y \in U_N$ (note that this implies $r_y \geqslant N$), $f(y) = 2^{-r_y} \leqslant 2^{-N} < \epsilon$. So $U_N$ is a neighborhood of $D$ on which the continuity of $f$ at any $x \in D$ holds for $\epsilon$.

To see discontinuity on $[0,1] \setminus D$, fix $y \notin D$ and put $\epsilon < 2^{-r_y}$. Then since $D$ is dense, for every $\delta > 0$, there is some $x \in D \cap B(y, \delta)$ that gives $|f(y) - f(x)| = 2^{-r_y} > \epsilon$. ∎

Next we show it is impossible to define a translation invariant $\sigma$-additive set function $\mu$ for which $\mu(0,1) = 1$ and for which every subset of $\mathbb{R}$ is measurable. This in particular implies the existence of non-Lebesgue measurable sets.

**Example 2.2.16** (Vitali's nonmeasurable sets)**.** Suppose $x \in \mathbb{R}$. Define an equivalence relation on $\mathbb{R}$ by $x \sim x + q$ where $q \in \mathbb{Q}$. Let $A_x$ be the equivalence class of $x$ (an $\mathbb{R}/\mathbb{Q}$ coset) and put $\mathcal{G} := \mathbb{R}/\sim$. For each class $A \in \mathcal{G}$, pick a representative $r_A$ in $(0, 1/2) \cap A$ and let $V_0$ be the set of all these $r_A$. For each $q \in \mathbb{Q}$, define

$$V_q = V_0 + q = \{r_A + q | A \in \mathcal{G}\} \subset (0, 1/2) + q.$$

Then for distinct rationals $p \neq q$, $V_q$ and $V_p$ are disjoint, otherwise we can find $r_A$, $r_B$ representing classes $A, B$ such that $r_A + p = r_B + q \Leftrightarrow r_A - r_B = q - p \in \mathbb{Q} \Leftrightarrow r_A = r_B \Leftrightarrow p = q$. Corresponding to all the rationals, there are countably infinitely many of these $V_q$'s. By the translation invariance of Lebesgue measure, the $V_q$'s have the same measure (if they were measurable).

For $0 < q < 1/2$, $V_q$ is contained in $(0,1)$. So the sets $V_0, V_{1/2}, V_{1/3}, V_{1/4}, \ldots$ are subsets of $(0,1)$. Consequently

$$\mu\left[V_0 \cup \bigcup_{n=1}^{\infty} V_{1/n}\right] \leqslant \mu(0,1) = 1. \tag{2.2.2}$$

On the other hand,

$$\mu\left[V_0 \cup \bigcup_{n=1}^{\infty} V_{1/n}\right] = \mu(V_0) + \sum_{n=1}^{\infty} \mu(V_{1/n}) = \lim_{n \to \infty} n\mu(V_0)$$

which is either $\infty \Leftrightarrow \mu(V_0) > 0$, or $0 \Leftrightarrow \mu(V_0) = 0$. The first case doesn't happen because of 2.2.2. The second case would mean the union of all the $V_q$'s has measure 0. But $\bigcup_{q \in \mathbb{Q}} V_q \supset \mathbb{R}$, so this doesn't happen either. We've arrived at a contradiciton. ∎

## 2.3 Lebesgue measure on $\mathbb{R}^d$

Similar to the case in $\mathbb{R}$, we now want to define measurable sets in the setting of arbitrary rings.

Suppose $\mathcal{R} \subset \mathcal{P}(\mathbb{R}^d)$ is a ring and $\mu : \mathcal{R} \to [0, +\infty]$ a $\sigma$-additive set function. We want to extend $\mu$ to the $\sigma$-ring $\mathcal{R}_\sigma$ generated by $\mathcal{R}$.

Define the *outer measure* of a subset $A \subset \mathbb{R}^d$ as

$$\mu^*(A) = \inf_{\{I_n\}} \sum_{n=1}^{\infty} \mu(I_n).$$

where the infimum runs over all countable family $\{I_n\} \subset \mathcal{R}$ that covers $A$. Easily seen $\mu^*$ is also finitely additive and satisfies $\mu^*(\varnothing) = 0$ as in the case of $m^*$.

The *hereditary ring*

$$\mathcal{H}(\mathcal{R}) := \left\{ A \subset \mathbb{R}^d \,\middle|\, A \subset \bigcup_{n=1}^{\infty} I_n,\ I_n \in \mathcal{R} \right\}$$

is a $\sigma$-ring. The name comes from that the subsets of elements of $\mathcal{H}(\mathcal{R})$ are also elements of the ring. Note that also $\mathcal{R} \subset \mathcal{H}(\mathcal{R})$, so $\mathcal{R}_\sigma$ by minimality would necessarily be contained in $\mathcal{H}(\mathcal{R})$.

By definition of $\mathcal{H}(\mathcal{R})$, $\mu^*$ is defined there. let's show subadditivity of $\mu^*$ on $\mathcal{H}(\mathcal{R})$:

**Proposition 2.3.2** ($\sigma$-subadditivity of $\mu^*$)**.** *Suppose $\mu$ is a nonnegative $\sigma$-additive set function on a ring $\mathcal{R} \subset \mathcal{P}(\mathbb{R}^d)$ and $\{A_n\}$ a countable family of elements of $\mathcal{H}(\mathcal{R})$. Then*

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n).$$

*Proof.* Pretty much the same as in the case of $m^*$. If any $\mu^*(A_n) = \infty$ then we are done. Otherwise, suppose $\epsilon > 0$. For each $n$, find $\{I_{n,i}\}_i \subset \mathcal{R}$ such that $\bigcup_i I_{n,i} \supset A_n$ and that $\sum_{n=1}^{\infty} \mu(I_n) < \mu^*(A_n) + \epsilon/2^n$. Then

$$\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} I_{n,i}$$

and so by $\sigma$-additivity of $\mu$ on $\mathcal{R}$,

$$\mu^*(\text{LHS}) \leqslant \mu^*(\text{RHS}) = \sum_{n,i} \mu(I_{n,i}) < \sum_{n=1}^{\infty} A_n + \epsilon/2^n = \epsilon + \sum_n A_n.$$

∎

**Proposition 2.3.3.** *Suppose $\mu : \mathcal{R} \to [0, +\infty]$ is a $\sigma$-additive set function. If $A \in \mathcal{R}$ then $\mu^*(A) = \mu(A)$.*

*Proof.* Since $A \subset A$ we get $\mu^*(A) \leqslant \mu(A)$. OTOH, cover $A$ with some family $\{I_n\}_n \subset \mathcal{R}$, then by $\sigma$-additivity of $\mu$ on $\mathcal{R}$, we have

$$\mu(A) \leqslant \sum_{i=1}^{\infty} \mu(I_n).$$

Hence $\mu(A)$ is a lower bound for all such sums $\sum \mu(I_n)$. This gives $\mu(A) \leqslant \mu^*(A)$. ∎

Next we define measurable sets, followed by a sequence of results.

Suppose $\mu : \mathcal{R} \to [0, +\infty]$ is a $\sigma$-additive set function on a ring $\mathcal{R} \subset \mathcal{P}(\mathbb{R}^d)$.

**Definition** (Carathéodory condition)**.** Define a set $A \in \mathcal{H}(\mathcal{R})$ to be *measurable* if for every subset $E \subset \mathcal{H}(R)$,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^\sim).$$

The family of all measurable sets with respect to $\mu$ on $\mathbb{R}$ will be denoted $\mathcal{A}$.

**Proposition 2.3.4.** *$\mathcal{R} \subset \mathcal{A}$.*

*Proof.* Need to show every $A \in \mathcal{R}$ is measurable. Suppose $E \in \mathcal{H}(R)$. By Prop 2.3.2, $E = (E \cap A) \cup (E \cap A^\sim)$ implies $\mu^*(E) \leqslant \mu^*(E \cap A) + \mu^*(E \cap A^\sim)$. To see the reverse inequality, for a given $\epsilon > 0$, find a family $\{I_n\} \subset \mathcal{R}$ that covers $E$, with $\sum_n \mu(I_n) < \mu^*(E) + \epsilon$. Well, $\cup(I_n \cap A) \supset E \cap A$ and $\cup(I_n \cap A^\sim) \supset E \cap A^\sim$. So $\mu(I_n) = \mu(I_n \cap A) + \mu(I_n \cap A^\sim)$ by $\sigma$-additivity of $\mu$ on the two disjoint sets. Then by definition of $\mu^*$,

$$\mu^*(E \cap A) + \mu^*(E \cap A^\sim) \leqslant \sum_n \mu(I_n \cap A) + \sum_n \mu(I_n \cap A^\sim) = \sum_n \mu(I_n) < \mu^*(E) + \epsilon.$$

Take $\epsilon \to 0$ to arrive at $\mu^*(E) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^\sim)$. ∎

**Theorem 2.3.5.** *$\mathcal{A}$ is a ring.*

*Proof.* ∎

**Theorem 2.3.6.** *$\mathcal{A}$ is a $\sigma$-ring.*

*Proof.* ∎

**Theorem 2.3.7.** *$\mu^*$ is $\sigma$-additive on $\mathcal{A}$.*

*Proof.* ∎

**Example 2.3.8** (volumes of parallelepipeds)**.**

**Theorem 2.3.9** (Lesbegue measure on $\mathbb{R}^d$)**.** *There exist a $\sigma$-algebra $\mathcal{M}(\mathbb{R}^d) \supset \mathcal{B}(\mathbb{R}^d)$ and a measure $m^d$ on $\mathcal{M}(\mathbb{R}^d)$ that is (i) translation invariant, and (ii) coincides with volume on cubes in $\mathbb{R}^d$.*

*Proof.* ∎

**Example 2.3.10** (Lebesgue-Stiltjes measure)**.** Suppose $f : \mathbb{R} \to \mathbb{R}$ is an increasing and right-continuous function. For $Q = \prod_{i=1}^d (a_i, b_i] \subset \mathbb{R}^d$, define

$$\mu_f(Q) = \prod_{i=1}^d [f(b_i) - f(a_i)].$$

Then $\mu_f$ is a $\sigma$-additive set function on the ring $\mathcal{Q}$ (Exercise 2.21).

## 2.4 Measure space

*Stressed? I hope it's not because of real analysis – it's such a NATURAL, BEAUTIFUL subject.*

– Wojciech Czaja

Suppose $X$ is a set and $\mathcal{A}$ a $\sigma$-algebra in $\mathcal{P}(X)$. Then we say $(X, \mathcal{A})$ is a *measurable space*. If this gets equipped with a measure $\mu$, then $(X, \mathcal{A}, \mu)$ becomes a *measure space*.

**Theorem 2.4.1** (Carathéodory theorem)**.** *Let $X$ be a set, $\mathcal{R} \subset \mathcal{P}(X)$ an algebra, $\mu : \mathcal{R} \to [0, +\infty]$ a $\sigma$-additive set function. Then $\mu$ extends to a measure on the $\sigma$-algebra of measurable sets $\mathcal{A}$, and to a measure on the $\sigma$-algebra generated by $\mathcal{R}$. If $\mu$ is $\sigma$-finite, then this extension is unique (as a measure to the smallest $\sigma$-algebra containing $\mathcal{A}$).*

*Proof.* Left as an exercise (2.20):

Define for $E \subset P$,

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j) \right\}$$

where the inf runs over all countable families $\{A_j\} \subset \mathcal{R}$ that cover $E$. Define $E \subset P$ to be $\mu^*$-measurable if for all $F \subset X$,

$$\mu^*(F) \geqslant \mu^*(F \cap E) + \mu^*(F \cap E^{\sim}).$$

Claim that the $\mu^*$-measurable sets form a $\sigma$-algebra $\mathcal{A}$ containing $\mathcal{R}$.

⋮

∎

If the $\mu$-measurable sets $\mathcal{A}$ contains the Borel sets, then $\mu$ is a *Borel measure*. Two Borel sets in a measure space $X$ are *congruent* if they are isometric. Remember the Lebesgue measure is translation invariant, so that kind of preserves the congruent classes of Borel sets.

*Remark.* (Fun facts on page 59.) On a countable compact metric space, there is a finitely additive measure for which all congruent Borel sets have the same measure. On a compact metric space, there is a Borel measure for which all open congruent sets have the same measure (Mycielski).

**Example** (measure spaces)**.**

a. The Lebesgue measure space $(\mathbb{R}^d, \mathcal{M}(\mathbb{R}^d), m^d)$.

b. $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m^d)$

c. The *Dirac measure at $x$* on a measurable space $(X, \mathcal{A})$: fix $x \in X$, for $A \in A$,

$$\delta_x(A) = \begin{cases} 0 & \text{if } x \notin A \\ 1 & \text{if } x \in A. \end{cases}$$

Then $(X, \mathcal{A}, \delta_x)$ is a measure space.

d. The counting measure $c$ on a measurable space $(X, \mathcal{A})$:

$$c(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{otherwise} \end{cases}$$

### 2.4.1 Theorems on unions/intersections of measurable sets

The following is presumptively Czaja's favorite theorem throughout the analysis sequence:

→ **Theorem 2.4.3** (measure of unions and intersections)**.** *Suppose $(X, \mathcal{A}, \mu)$ is a measure space.*

**a.** *If $A, B \in \mathcal{A}$ and $A \subset B$, then $\mu(A) \leqslant \mu(B)$.*

**b.** *For each sequence $\{A_n\} \subset \mathcal{A}$,*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu(A_n).$$

**c.** *Suppose $\{A_n\}$ is a descending family of elements in $\mathcal{A}$, i.e. $A_{n+1} \subset A_n$, and that $\mu(A_1) < \infty$, then*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

**d.** *Suppose $\{A_n\}$ is an ascending family of elements in $\mathcal{A}$, i.e. $A_n \subset A_{n+1}$, then*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

*Proof.* **a.** Write $B = A \sqcup (B \setminus A)$ then $\mu(A) \leqslant \mu(A) + \mu(B \setminus A) = \mu(B)$.

**b.** Write $B_n = A_n \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)$. Then the $B_n$'s are disjoint, with $B_n \subset A_n$ and $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$. So

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\coprod_{i=1}^{\infty} B_i\right) = \sum i = 1^{\infty} \mu(B_i) \leqslant \sum i = 1^{\infty} \mu(A_i).$$

**c.** Put $A := \bigcap_{i=1}^{\infty} A_i$, $B_i := A_i \setminus A_{i+1}$. Then $B_i$'s are disjoint, and $\coprod_{i=1}^{\infty} B_i = \bigcup_i A_i \setminus A = A_1 \setminus A$. Write $A_n = A_{n+1} \sqcup B_n$. By monotonicity, $\infty > \mu(A_1) \geqslant \mu(A_2) \geqslant \cdots$. Can subtract and get

$$\mu(A_n) - \mu(A_{n+1}) = \mu(B_n).$$

Write $A_1 = A \sqcup (A_1 \setminus A)$. Then

$$\begin{aligned}
\mu(A_1) &= \mu(A) + \mu(\coprod_{i=1}^{\infty} B_i) \\
&= \mu(A) + \sum_{i=1}^{\infty} \mu(B_i) \\
&= \mu(A) + \sum_{i=1}^{\infty} [\mu(A_i) - \mu(A_{i+1})] \\
&= \mu(A) + \lim_{k\to\infty} \sum_{i=1}^{k} [\mu(A_i) - \mu(A_{i+1})] \\
&= \mu(A) + \mu(A_1) - \lim_{k\to\infty} \mu(A_{k+1}) \\
\Rightarrow \quad \mu(A) &= \lim_{k\to\infty} \mu(A_k).
\end{aligned}$$

**d.** Put $B_1 = A_1$, $B_{i+1} := A_{i+1} \setminus A_i$. Then the $B_i$'s are disjoint, $\coprod_{i=1}^n B_i = \bigcup_{i=1}^n A_i = A_n$ and $\coprod_{i=1}^\infty B_i = \bigcup_{i=1}^\infty A_i$. We have

$$\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{i=1}^n \mu(B_i) = \sum_{i=1}^\infty \mu(B_i) = \mu\left(\coprod_{i=1}^\infty B_i\right) = \mu\left(\bigcup_{i=1}^\infty A_i\right)$$

∎

**Theorem 2.4.4** (Borel-Cantelli lemma I). *Suppose $(X, \mathcal{A}, \mu)$ is a measure space. If $\{A_n\} \subset \mathcal{A}$ is such that $\sum_{j=1}^\infty \mu(A_j) < \infty$, then*

$$\mu\left(\bigcap_{k=1}^\infty \bigcup_{j=k}^\infty A_j\right) = 0.$$

∎

**Definition** (independent sets and probability space). Suppose $(X, \mathcal{A}, \mu)$ is a measure space. The sets $A_1, A_2, \ldots, A_n \subset \mathcal{A}$ are said to be *independent* if

$$\mu(A_1 \cap A_2 \cap \cdots \cap A_n) = \mu(A_1)\mu(A_2)\cdots\mu(A_n).$$

We say $\{A_n\} \subset \mathcal{A}$ is *independent* if every finite subcollection of $\{A_n\}$ is independent.

$X$ is called a *probability space* if $\mu(X) = 1$. In this case, we say that $\mu$ is a *probability measure*.

**Theorem 2.4.5** (Borel-Cantelli lemma II). *Suppose $(X, \mathcal{A}, \mu)$ is a probability space. If $\{A_n\} \subset \mathcal{A}$ is a collection of independent sets such that $\sum_{j=1}^\infty \mu(A_j) = \infty$, then*

$$\mu\left(\bigcap_{k=1}^\infty \bigcup_{j=k}^\infty A_j\right) = 1.$$

∎

**Theorem 2.4.6** (Kolmogorov 0-1 law). *Suppose $(X, \mathcal{A}, \mu)$ is a probability space and $\{A_n\}$ is a sequence of independent sets. Let $\mathcal{A}_m$ be the $\sigma$-algebra generated by $\{A_n\}_{n=m}^\infty$. Then for each $A \in \bigcap_{m=1}^\infty \mathcal{A}_m$, either $\mu(A) = 0$ or $\mu(A) = 1$.* ∎

**Definition 2.4.7** ($\sigma$-/finite, complete measure space). Suppose $(X, \mathcal{A}, \mu)$ is a measure space. We say that

**a.** $\mu$ is *finite* or *bounded* if $\mu(X) < \infty$.

**b.** $\mu$ is *$\sigma$-finite* if there is a countable family $\{A_n\} \subset X$ that covers $X$ and each member of the family has finite measure. $\mu(X) < \infty$

**c.** $\mu$ is *complete* if

$$\big(F \in \mathcal{R},\ E \subset F \text{ and } \mu(F) = 0\big) \quad \Rightarrow \quad E \in \mathcal{R}.$$

**Theorem 2.4.8** (Completeness theorem). *Suppose $(X, \mathcal{A}, \mu)$ is a measure space. There is an complete measure space $(X, \mathcal{A}_0, \mu_0)$ extending $\mu$ in the following way:*

**a.** *$\mathcal{A} \subset \mathcal{A}_0$,*

**b.** *$\mu = \mu_0$ on $\mathcal{A}$,*

**c.** *$A \in \mathcal{A}_0$ iff $A = B \cup E$, where $B \in \mathcal{A}$ and $E \subset D$, for some $D \in \mathcal{A}$ with $\mu(D) = 0$.*

**d.** *if $A \in \mathcal{A}_0$ has measure 0 and contains some set $S$, then $S$ is measurable and has measure zero.*

∎

**Theorem 2.4.9** (induced measure spaces)**.** **a.** *Suppose $\{(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)\}$ is a collection of measure spaces. Define $(X, \mathcal{A}, \mu)$ by: $X = \bigcup X_\alpha$, $\mathcal{A} = \{A \subset X \mid A \cap X_\alpha \in \mathcal{A}_\alpha \,\forall\, \alpha\}$, $\mu(A) = \sum \mu_\alpha(A \cap X_\alpha)$ for $A \in \mathcal{A}$. Then $(X, \mathcal{A}, \mu)$ is a measure space. Moreover, $\mu$ is $\sigma$-finite iff all but countably many $\mu_\alpha$ are zero, and all remaining ones $\sigma$-finite.*

**b.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and $Y \in \mathcal{A}$ a measurable set. Set*

$$\mathcal{A}_Y = \{A \in \mathcal{A} \mid A \subset Y\}$$

*and define $\mu_Y(A) = \mu(A)$ for $A \in \mathcal{A}_Y$. Then $(Y, \mathcal{A}_Y, \mu_Y)$ is a measure space.*

∎

### 2.4.2 Measurable functions

**Proposition / definition 2.4.10** (measurable functions)**.** *Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \to [-\infty, +\infty]$ is a map. TFAE:*

**a.** *$\forall\, r \in \mathbb{R}$, $f^{-1}\,(r, \infty) \in \mathcal{A}$,*

**b.** *$\forall\, r \in \mathbb{R}$, $f^{-1}\,[r, \infty) \in \mathcal{A}$,*

**c.** *$\forall\, r \in \mathbb{R}$, $f^{-1}(-\infty, r) \in \mathcal{A}$,*

**d.** *$\forall\, r \in \mathbb{R}$, $f^{-1}(-\infty, r] \in \mathcal{A}$,*

**e.** *$\forall\, U \subset \mathbb{R}$ open, $f^{-1}(U) \in \mathcal{A}$ and $f^{-1}(\pm\infty) \in \mathcal{A}$.*

If $f$ satisfies any of these equivalent conditions, then we say $f$ is *measurable with respect to* $(X, \mathcal{A}, \mu)$. A complex-valued function is measurable if both its real and imaginary parts are measurable.

*Proof.* ∎

*Remark.* We require measurable functions to be finite $\mu$-a.e., to avoid dealing with infinities.

The definition above can be extended as follows. Let $(X, \mathcal{A})$ and $(Y, \mathcal{E})$ be two measurable spaces. A function $g : X \to Y$ is measurable if $g^{-1}(E) \in \mathcal{A}$ for all $E \in \mathcal{E}$. In other words, the preimage of a measurable set is measurable, just like the way things are (and supposed to be) in topology. Immediately, this means $A \in \mathcal{A}$ iff $\mathbb{1}_A$ is measurable.

If $X$ is a topological space and $(X, \mathcal{A}, \mu)$ is a measure space, we always assume that $\mathcal{A}$ is Borel, i.e. $\mathcal{B}(X) \subset \mathcal{A}$. One can check that the set $\{E \subset Y | f^{-1}(E) \in \mathcal{A}\}$ is a $\sigma$-algebra. If $f$ is continuous, then the $\sigma$-algebra generated by preimages of open sets in $Y$ under $f$ is contained in $\mathcal{B}(X)$. Consequently $f^{-1}$ takes Borel sets in $Y$ to Borel sets in $X$, and $f$ is Borel measurable.

**Proposition.** *The set of $\mathbb{R}$-valued measurable functions on $(X, \mathcal{A}, \mu)$ forms an algebra (ha-ha! In the algebraic sense!) over $\mathbb{R}$.*

**Example 2.4.13** (nonmeasurable subsets of p.s.s.)**.** *There is a perfect symmetric set $E \subset [0,1]$ of positive Lebesgue measure that has a non-Lebesgue measurable subset $N$.*

Suppose for contradiction that all subsets of $E$ are measurable. Take a nonmeasurable subset $S \subset [0,1]$ of $\mathbb{R}$ from example (2.2.16) and let $\{E_n\}$ be a sequence of perfect symmetric nowhere dense sets defined in (1.2.8) such that for each $n$,

$$m(E_n) \in [1 - 1/n, 1].$$

If $S \cap E_n \subset E$ were measurable for each $n$, then

$$\bigcup_{n=1}^{\infty} (S \cap E_n) \in \mathcal{M}(\mathbb{R}).$$

Now, $m(\cup E_n) = 1$, so the set $F := [0,1] \cap (\cup E_n)^{\sim} \subset [0,1]$ has measure 0. We have then $S \cap F \subset F \Rightarrow m(S \cap F) = 0$, which by completeness of Lebesgue measure implies that $S \cap F$ is measurable. But then we see

$$S = (S \cap F) \cup \bigcup_n (S \cap E_n)$$

is measurable too, reaching a contradiction.

**Example 2.4.14** ($(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not complete)**.** For this we'll find a Lebesgue measurable but not Borel measurable set. Let $C$ be the ternary Cantor set. Since $C$ is closed, $C \in \mathcal{B}(\mathbb{R})$, $m(C) = 0$, if we can find a subset $L \subset C$ that's not Borel, we are done.

I get bored so I'll skip this for now.

**Definition 2.4.15** (almost everwhere)**.** Suppose $(X, \mathcal{A}, \mu)$ is a measure space. A statement $S$ is true *almost everywhere* if the complement of the set on which $S$ is true (i.e. the set on which $S$ is false) has measure 0. In this case we write $S$ $\mu$-a.e..

**Proposition 2.4.16.** *Suppose $(X, \mathcal{A}, \mu)$ is a complete measure space. If $f$ is measurable and $f \equiv g$ $\mu$-a.e., then $g$ is measurable.*

*Proof.* ∎

**Example 2.4.17** (noncomplete measure)**.** Let $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(R)$ with Lebesgue measure 0 (see 2.4.14). Let $G$ be a $G_\delta$ set containing $L$ with $m(G) = m(L) = 0$. Set $f = 1$ and

$$g(x) = \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus G \\ 1/2 & \text{if } x \in G \setminus L \\ 0 & \text{if } x \in L \end{cases}$$

Then since $G$ is Borel, $f \equiv g$ m-a.e. in $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$. But $g^{-1}(0) = L \notin \mathcal{B}(\mathbb{R})$. So $g$ is not measurable on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$.

**Example .** Recall if $A$ is Lebesgue measurable then $A = B \cup U$ decomposes as a union of a Borel set $B$ and a set $E$ of Lebesgue measure 0. This is not true for the measure space $(\mathbb{R}, \mathcal{M}(\mathbb{R}), c)$. Take $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ is $c$-measurable, but $c(E) = 0 \Leftrightarrow E = \varnothing$. So $L$ cannot be written as such a union.

**Definition 2.4.18** (outer measure)**.** An outer measure is a nonnegative set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ that satisfies:

i. monotonicity: $A \subset B \Rightarrow \mu^*(A) \leqslant \mu^*(B)$,

ii. $\mu^*(\varnothing) = 0$,

iii. $\sigma$-subadditivity: $\mu^*(\bigcup_{i=1}^{\infty} A_n) \leqslant \sum_{i=1}^{\infty} \mu^*(A_n)$.

**Theorem 2.4.19** (restriction of outer measure)**.** *Suppose $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is an other measure. There is a $\sigma$-algebra $\mathcal{A} \subset \mathcal{P}(X)$ and a nonnegative $\sigma$-additive set funtion $\mu$ on $\mathcal{A}$ that is the restriction of $\mu^*$ to $\mathcal{A}$.*

*Proof.* (Carathéodory's approach) Define $\mathcal{A}$ to be the collection of sets $A$ with the property:

$$\forall\, E \subset X,\ \ \mu^*(\mu) = \mu^*(E \cap A) + \mu^*(E \cap A^{\sim}).$$

Set $\mu = \mu^*|_{\mathcal{A}}$. The rest is left as an exercise. ∎

### 2.4.3 summary of $\mu \leftrightarrow \mu^*$

Let $(X, \mathcal{A}, \mu)$ be a measure space. Define the set function $\mu^* : \mathcal{P} \to [0, \infty)$ by

$$\mu^*(E) := \inf \sum_{j=1}^{\infty} \mu(A_n)$$

where $\{A_n\}$ runs over all countable family of elements of $\mathcal{A}$ that covers $E$. Moreover, the restriction of $\mu^*$ to $\mathcal{A}$ coincides with $\mu$. We say that $\mu^*$ is the outer measure associated with $\mu$.

On the other hand, given an outer measure $\mu^*$, Theorem 2.4.19 gives a way to generate a $\sigma$-algebra $\mathcal{A}$ and a measure $\mu$ on $\mathcal{A}$ that is the restriction of $\mu^*$ to $\mathcal{A}$. The outer measure $\hat{\mu}$ associated with $\mu$ may differ from $\mu^*$ on $\mathcal{P}(X)$. But we do have for all $A \subset X$,

$$\mu^*(A) \leqslant \hat{\mu}(A).$$

## 2.5 approximation of measurable functions

Instead of dividing the range as in Riemann theory, we are going to divide the domain in Lebesgue theory.

**Definition** (simple function)**.** A complex-valued function on a measure space $(X, \mathcal{A}, \mu)$ is *simple* if its image consists of only finitely many points (exluding $\infty$).

If $f$ is simple, there is a natural representation of $f$: let $f(X) = \{a_j\}_{j=1}^{n}$ and $A_j = \{x \in X | f(x) = a_j\}$. Then

$$f = \sum_{j=1}^{n} a_j \mathbb{1}_{A_j}$$

with the $a_j$'s distinct and the $A_j$'s disjoint. This is called the *canonical form* of $f$. Note that $f$ is measurable iff all the $A_j$'s are measurable. The text requires simple functions to be measurable in the definition.

**Theorem 2.5.4** (measurable functions as limits a.e.)**.** *Let $(X, \mathcal{A}, \mu)$ be a complete measure space and $\{f_n\}$ a sequence of $\mathbb{C}$-/$\mathbb{R}^*$-valued measurable funtions, each defined $\mu$-a.e.on $S$. If $f$ is defined $\mu$-a.e.and $f_n \to f$ pointwise $\mu$-a.e., then $f$ is measurable.*

*Proof.* exercise (2.34). ∎

**Theorem 2.5.5** (measurable functions as limits of simple functions)**.** *Let $f : X \to [0, \infty]$ be measurable. There is a sequence $\{s_n\}$ of simple functions on $X$ such that*

**a.** *$s_j \leqslant s_{j+1}$ for all $j \geqslant \mathbb{Z}_+$.*

**b.** *$s_n$ converges to $f$ pointwise as $n \to \infty$.*

*Moreover, if $f$ is bounded on $X$, then $f_j$ convergence to $f$ uniformly.*

*Proof.* For $n \in \mathbb{Z}_+$, set

$$E_{nj} = f^{-1}\Big[\frac{j-1}{2^n}, \frac{j}{2^n}\Big), \quad F_n = [n, \infty);$$

$$r_n = \sum_{j=1}^{n^{2n}} \frac{j-1}{2^n} \mathbb{1}_{E_{nj}} + n\mathbb{1}_{F_n}.$$

Then for each $x \in X$, there is some $n$ large such that $f(x) < n$ ($\Rightarrow x \notin F_n$). So $f(x) \in [\frac{k-1}{2^n}, \frac{k}{2^n})$ for a $k \in [n2^n]$ ($\Rightarrow x \in E_{nk}$). This means

$$r_n(x) = \sum_{j=1}^{n^{2n}} \tfrac{j-1}{2^n} \mathbb{1}_{E_{nj}}(x) + n\mathbb{1}_{F_n}(x) = \tfrac{k-1}{2^n} < x$$

And it's not hard to see for $n$ larger, each of the intervals $E_{nj}$ gets smaller. So $r_n(x) \uparrow x$ as $n \to \infty$. It follows that the functions $s_n := f \circ r_n$ satisfy the requirements. ∎

**Theorem 2.5.7** (Egorov)**.** *Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Let $\{f_n\}$ be a sequence of $\mathbb{C}$ or $[-\infty, +\infty]$-valued measurable functions each defined and finite $\mu$-a.e.on $X$ and $f_n \to f$ $\mu$-a.e.. Then for every $\delta > 0$, there is an $A \in \mathcal{A}$ with $\mu(A) < \delta$ and $f_n \rightrightarrows f$ on $A^\sim$.*

*Proof.* Fix $\epsilon > 0$. Define for $n, k \in \mathbb{Z}_+$,

$$E_n^k = \bigcup_{m \geqslant n} \{x \in X \,:\, |f_m(x) - f(x)| \geqslant 1/k\}.$$

These guys are decreasing in $n$, as we are looking at the tail of the union: $E_n^k \supset E_{n+1}^k$. For a fixed $k$, if $f_n(x) \to f(x)$ pointwise, then there exists some $N$ large such that $|f_n(x) - f(x)| < 1/k$ for $n \geqslant N$, and so $x \notin E_N^k$. Put

$$F_k := \bigcap_{n \in \mathbb{Z}_+} E_n^k$$

consisting of all the points $x$ at which $|f_n(x) - f(x)|$ doesn't stay below $1/k$. By assumption, $f_n \to f$ $\mu$-a.e. implies that $\mu(F_k) = 0$ for every $k$. OTOH, continuity from above gets us (this needs the fact that $\mu(X) < \infty$)

$$\mu(F_k) = \mu\left(\bigcap_{n\in\mathbb{Z}_+} E_n^k\right) = \lim_{n\to\infty} \mu(E_n^k).$$

So for each $k$, we can choose $n_k$ large such that $\mu(E_n^k) < \epsilon 2^{-k}$.

Set $E = \bigcup_{k\in\mathbb{Z}_+} E_{n_k}^k$. This set contains the points $x$ at which we consider the speed of $f_n(x)$ approaching into at least one of these $1/k$-neighborhood of $f(x)$ as too slow. On $F \setminus E$, we have uniform convergence.

Finally,

$$\mu(E) \leqslant \sum_{k=1}^{\infty} E_{n_k}^k < \sum_{k=1}^{\infty} \epsilon/2^k = \epsilon.$$

∎

**Definition 2.5.9** ($L_\mu^\infty(X)$)**.** Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is measurable. Define the essential supremum

$$\|f\|_\infty := \inf\{M \mid \mu\{|f(x)| > M\} = 0\}$$

Now we introduce regularity, which sloppily, means the measure can be approximated from above and from below.

**Definition 2.5.12** (regular Borel measure)**.** Let $X$ be a locally compact Hausdorff space and $(X, \mathcal{A}, \mu)$ a Borel measure space. We say a measurable set $A \in \mathcal{A}$ is *inner regular* if

$$\mu(A) = \sup\{\mu(F) : F \subset A \text{ compact }\};$$

*outer regular* if

$$\mu(A) = \inf\{\mu(G) : G \supset A \text{ open }\};$$

and just *regular* if $A$ is both inner and outer regular.

If all measurable sets are regular, then $\mu$ is said to be a *regular Borel measure*.

**Theorem 2.5.13** (Vitali-Luzin theorem)**.** *Suppose $(X, \mathcal{A}, \mu)$ is a locally compact Hausdorff measure space and $\mu$ is a regular Borel measure. If the support of $f : X \to Y$ measurable is contained in a set $A$ of finite measure, then for every $\epsilon > 0$, there exists a continuous map $g$ with compact support that coincides with $f$ on a compact set $E$ and that $\mu(E^\sim) < \epsilon$.*

The following is a more well-known version.

**Theorem** (Luzin's theorem)**.** $(\mathbb{R}, \mathcal{M}, m)$. *If the support of $f : X \to Y$ is contained in a set $A$ of finite measure, then for every $\epsilon > 0$, there exists a continuous map $g$ with compact support that coincides with $f$ on a compact set $E$ and that $\mu(E^\sim) < \epsilon$.*

*Proof.* Enumerate the rationals in $[0,1]$ and let $\{V_n\}$ be the open intervals with these points as endpoints. Pick compact sets $K_n \subset f^{-1}(V_n)$ and $L_n \subset A \setminus f^{-1}(V_n)$ with $\mu[A \setminus (K_n \cup L_n)] < \epsilon/2^n$. Put $K = \bigcap_{n=1}^{\infty}(K_n \cup L_n)$. Then $\mu(A \setminus K) < \epsilon$.

Suppose $x \in K$, $f(x) \in V_m$, one can show $f$ is continuous when restricted to $K$ by choosing a compact nbhd $N_m$ of $x$ and $f(N_m \cap K) \subset V_m$. ∎

**Theorem** (Urysohn's lemma)**.** *A topological space is normal if and only if any two disjoint closed sets can be separated by a continuous function.*

(In consistence with Munkres's convention, regular spaces are Hausdorff.)

**Theorem** (Urysohn's metrization theorem)**.** *Every second countable regular space is metrizable.*

Punchline: almost everywhere convergence is almost like uniform convergence. Almost everywhere continuity is almost like continuity. For each of these categories, the set of discrpancy points has measure 0, which can kind of be ignored for the purpose of this class (but not the next one :).

# 3 Integration

## 3.1 Why?

Read [BC09] – I thought the book did a good job motivating stuff, which I really appreciate (despite the fact that its untraditional approach possibly induces headaches). That said, having one of the authors lecturing made a noticeable difference.

## 3.2 abstract integration

Around this time, I skipped a couple of classes to go to Patrick Brosnan's office hours (a great tradition of mine). So here I'm not exactly following Czaja's approach. To compensate, I used [Rud15] and [Kna16]. (You mean I actually read texts?!)

Throughout the section, we'll assume $(X, \mathcal{A}, \mu)$ is a measure space.

**Proposition 2.5.1.** *Suppose $f_n : X \to [-\infty, +\infty]$ is measurable for all $n$ and*

$$g = \sup_n f_n, \quad h = \limsup_n f_n$$

*then $g$, $h$ are measurable. (Analogous result holds for* inf *and* lim inf*.)*

*Proof.* Write $g^{-1}(\alpha, \infty] = \bigcup f_n^{-1}(\alpha, \infty]$ and $h^{-1}(\alpha, \infty] = \bigcap_m \bigcup_{n \geqslant m} f_n^{-1}(\alpha, \infty]$. ∎

*Corollary.* **a.** Pointwise limit of a convergent sequence of measurable functions is measurable: If $\{f_n\}$ is convergent, then there is a monotone subsequence converging to the pointwise limit $f$. WLOG assume $\{f_n\}$ is monotone increasing. It follows that $f = \limsup f_n$ is measurable.

**b.** If $f, g$ are measurable, then so are $\max(f, g)$ and $\min(f, g)$. In particular, it is true of the functions

$$f^+ = \max\{f, 0\}, \quad f^- = -\min\{f, 0\}.$$

Note that $f = f^+ - f^-$, $|f| = f^+ + f^-$.

**Proposition.** *If $f = g - h$ with $g \geqslant 0$ and $h \geqslant 0$. Then $f^+ \leqslant g$ and $f^- \leqslant h$.*

**Example 2.5.3** (a.e. limits that's nonmeasurable)**.** Consider $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$. Set $f_n = 0$, $f = \mathbb{1}_L$, where $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ and has Lebesgue measure 0. Then $f_n \equiv f$ m-a.e. (in particular, $f_n \to f$ m-a.e. ) but $f$ is not Borel measuable.

**Definition.** Let $s = \sum_{i=1}^{n} a_j \mathbb{1}_{A_j}$ be a simple $\mathbb{R}$-function, where the $a_j$'s are the distinct values of $s$, $A_j = \{x : s(x) = a_j\} \in \mathcal{A}$, and $\mu(A_j) < \infty$. If $E \in \mathcal{A}$, define

$$\int_E s d\mu = \sum_{j=1}^{a} a_j \mu(A_j \cap E).$$

**Theorem 3.2.1** (measurability of bounded functions)**.** *Suppose $(X, \mathcal{A}, \mu)$ is a complete finite measure space. Let $f : X \to \mathbb{R}$ be a bounded function. Then $f$ is $\mu$-measurable if and only if*

$$\inf\left\{\int_X h\, d\mu : f \leqslant h,\ h \text{ simple}\right\} = \sup\left\{\int_X g\, d\mu : f \geqslant g,\ g \text{ simple}\right\}.$$

*In this case, define the $\mu$-integral of $f$ to be this common value:*

$$\int_X f\, d\mu := \inf\left\{\int_X h\, d\mu : f \leqslant h,\ h \text{ simple}\right\}.$$

**Definition 3.2.3.** Let $f \geqslant 0$ be a $\mu$-measurable function on $(X, \mathcal{A}, \mu)$. Define

$$\int_X f\, d\mu := \sup_{h \leqslant f} \int_X h\, d\mu$$

where $h$ is a bounded $\mu$-measurable function with finite support.

For $X \subset \mathbb{R}^d$, $\mathcal{A} = \mathcal{M}^d(X)$, and $\mu = m^d$, we write

$$\int_X f\, dm^d = \int_X f(x) dx = \int_X f dx = \int_X f.$$

Define

$$\int_X f\, d\mu = \int_X f^+\, d\mu + \int_X f^-\, d\mu$$

whenever the terms on the RHS are both finite.

**Proposition 3.2.4.** *Suppose $f$ is a $\mu$-measurable function that is nonnegative $\mu$-a.e.. Then*

$$\int_X f\, d\mu = 0 \quad \Leftrightarrow \quad f = 0 \ \ \mu\text{-a.e.}$$

*Proof.* (⇒) Let $A_n := \{x | f(x) > 1/n\}$, $A := \{x | f(x) > 0\}$. Then $A = \bigcup_{n=1}^{\infty} A_n$. Let

$$g_n = \tfrac{1}{n}\mathbb{1}_{A_n}.$$

Then $g_n \leqslant f$. So

$$\int_X g_n \, d\mu = \tfrac{1}{n}\int_X \mathbb{1}_{A_n} = \tfrac{1}{n}\mu(A_n).$$

Since $\{A_n\}$ is nested, we have continuity from above:

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

So $A$ has positive measure iff there is some $m$ such that $\mu(A_m) > 0$. But then we would have

$$0 < \mu(A_m) = \int_X g_n \, d\mu \leqslant \int_X f \, d\mu,$$

which is contradictory.

(⇐) Clear. ∎

**Proposition 3.2.x.** *All functions and sets here are assumed to be measurable.*

- **a.** *if* $0 \leqslant f \leqslant g$, *then* $\int_E f d\mu \leqslant \int_E g d\mu$
- **b.** *if* $A \subset B$ *and* $f \geqslant 0$, *then* $\int_A f d\mu \leqslant \int_B f d\mu$
- **c.** *if* $f \geqslant 0$ *and* $c \in [0, \infty)$, *then*

$$\int_E cf \, d\mu = c \int_E f d\mu$$

- **d.** *if* $f \equiv 0$, *then* $\int_E f d\mu = 0$, *even if* $\mu(E) = \infty$
- **e.** *if* $\mu(E) = 0$, *then* $\int_E f d\mu = 0$, *even if* $f \equiv \infty$
- **f.** *if* $f \geqslant 0$, *then* $\int_E f d\mu = \int_X \mathbb{1}_E f d\mu$

### 3.2.1 $\mathcal{L}^1_\mu$ and $L^1_\mu$

**Definition.** Let $\mathcal{L}^1_\mu(X)$ represent thte family of $\mu$-integrable functions on the measure space $(X, \mathcal{A}, \mu)$. Define an equivalence relation $\sim$ on $\mathcal{L}^1_\mu$ by setting

$$f \sim g \quad \Leftrightarrow \quad (f \equiv g \quad \mu\text{-a.e.}).$$

Let $L^1_\mu(X)$ denote the set of equivalence classes under $\sim$.

We can put a norm on $\mathcal{L}^1_\mu(X)$:

$$\|f\|_1 := \int_X |f| \, d\mu.$$

In this part of the course we don't really care about what happens on sets of measure zero. For that, we will be focusing mainly on $L^1_\mu$ and slightly abuse notation by treating the equivalence classes as functions.

### 3.3 Convergence theorems

As usual, $(X, \mathcal{A}, \mu)$ is a measure space.

**Proposition 3.3.x** (linearity). *Suppose $f, g \in L^1_\mu(X)$ and $\alpha, \beta \in \mathbb{R}$. Then*

$$\int_X \alpha f + \beta g \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu.$$

*Proof.* Associate $\{f_n\}$, $\{g_n\}$ monotone simple functions to $f, g$ respectively. Since $f_n$, $g_n$ are simple, we have

$$\int f_n + g_n \, d\mu = \int f_n \, d\mu + \int g_n \, d\mu.$$

So then

$$\alpha \int_X f_n \, d\mu + \beta \int_X g_n \, d\mu = \int_X \alpha f_n + \beta g_n \, d\mu \to \int_X \alpha f + \beta g \, d\mu.$$

But the LHS also goes to $\int_X \alpha f + \beta g \, d\mu$. So done by uniqueness of limit. ∎

*Corollary.* $L^1_\mu(X)$ is a real vector space.

**Proposition.** $f \in L^1_\mu(X) \Leftrightarrow |f| \in L^1_\mu(X)$ *and $f$ is measurable.*

**Example.** Suppose $B \notin \mathcal{A}$, $\mu(X) < 0$. Then

$$f = \begin{cases} 1 & \text{on } B \\ -1 & \text{on } B^{\sim} \end{cases}$$

is not measurable, but $|f|$ is integrable.

**Proposition.** *If $f \in L^1_\mu(X)$, then*

$$\left| \int_X f \, d\mu \right| \leqslant \int_X |f| \, d\mu.$$

*Proof.* Rotate $\left|\int_X f \, d\mu\right|$ onto the positive real axis using a constant $c$ on $\partial B(0,1)$, so that $\left|\int_X f \, d\mu\right| = c \int_X f \, d\mu$. But then

$$\left| \int_X f \, d\mu \right| = c \int_X f \, d\mu = \int_X cf \, d\mu.$$

The LHS is real, so it must be the case that

$$\int_X cf \, d\mu = \int_X \Re(cf) \, d\mu \leqslant \int_X |cf| \, d\mu = \int_X |f| \, d\mu.$$

∎

*Corollary.* In above proposition, equality holds iff there exists $c \in \partial B(0,1)$ such that $cf \geqslant 0$.

**Theorem 3.3.2** (pre-LDCT). *Suppose $(X, \mathcal{A}, \mu)$ is a finite measure space and $\{f_n\}$ is a sequence of measurable functions $X \to [-\infty, +\infty]$ essentially bounded uniformly in $n$:*

$$\sup_n \|f_n\|_\infty = M < \infty.$$

*Suppose $f_n \to f$ pointwise on $X$. Then $f$ is essentially bounded, and $f_n \to f$ in $L^1_\mu$.*

*Proof.* By Prop 2.5.1, $f$ being the pointwise limit of the $f_n$'s is measurable. If $f$ weren't essentially bounded, then on a set $A$ of positive measure, $f$ blows up, which would violate the uniform essentially boundedness of the $f_n$'s. To see the $L^1$ convergence, take $\epsilon > 0$ and use Egorov's theorem to cook up a set $B$ with measure smaller than $\epsilon/4M$ on whose complement $f_n \rightrightarrows f$. Let $N$ be large so that $\max_{B^\sim} |f_n - f| < \frac{\epsilon}{2\mu(X)}$. Then we have for all $n \geqslant N$,

$$\int_X |f_n - f| \leqslant \int_B |f_n| + |f| + \int_{B^\sim} |f_n - f| \leqslant \mu(B)\Big( \|f_n\|_\infty + \|f\|_\infty \Big) + \mu(X) \cdot \frac{\epsilon}{2\mu(X)} = \epsilon.$$

∎

≫ **Theorem 3.3.5** (Fatou's lemma). *Suppose $\{f_n\}$ is a sequence of measurable functions $X \to [-\infty, +\infty]$ bounded below by some $g \in L^1_\mu$. Suppose $f := \underline{\lim}_{n\to\infty} f_n$. Then*

$$\int_X f \, d\mu \leqslant \underline{\lim}_{n\to\infty} \int_X f_n \, d\mu.$$

This more traditional version deviates from the one stated in the text, which assumes $f$ is the pointwise limit of the $f_n$'s. Proof I followed Wikipedia, and does not rely on the LMCT.
*Proof of Fatou's lemma.* First reductions: by subtracting $g$ from the $f_n$'s we can WLOG take $g \equiv 0$. Also if the RHS is $\infty$ there is nothing to prove. So, assume $\underline{\lim}_n \int f_n < \infty$.

Several observations:

- M1. If $f \leqslant g$ on $X$, then $\int_X f \leqslant \int_X g$. [clear from definition]
- M2. If $A \subset B$ and $0 \leqslant f$, then $\int_A f \leqslant \int_B f$. [write $\int_A f = \int_B f \mathbb{1}_A$ and apply M1]
- M3. If $S = \bigcup_n S_n$ where $S_n$ is an increasing family of measurable sets, then $\int_S f = \lim_{n\to\infty} \int_{S_n} f$. [prove this for characteristic functions (use 2.4.3d), then simple functions, and finally any measurable $f$]

Let $g_n := \inf_{k \geqslant n} f_k \leqslant f_n$. Then the $g_n$'s are measurable by 2.5.1, $g_n \leqslant g_{n+1}$ and we have pointwise

$$\lim_{n\to\infty} g_n = \liminf_{n\to\infty} f_n = f.$$

Fix a nonnegative simple function $h \leqslant f$. For $t \in (0,1)$, define

$$A_k^t(h) := \{x : g_k \geqslant t \cdot h\}.$$

As the $g_n$'s are increasing, the $A_k^t(h)$'s are increasing in $k$ as well. Moreover, $\bigcup_k A_k^t(h)$ exhausts $X$: if there were $y \notin \bigcup_k A_k^t(h)$, then

$$f(y) = \lim_{k\to\infty} g_k(y) \leqslant t \cdot h(y) < h(y)$$

would contradict $h \leqslant f$.

Hence M3 implies

$$\int_X h = \lim_{k\to\infty} \int_{A_k^t(h)} h$$

On the other hand, on $A_k^t(h)$, because $g_k \leqslant t \cdot h$, we have

$$t\int_{A_k^t(h)} h \leqslant \int_{A_k^t(h)} g_k \leqslant \int_X g_k.$$

Taking $k \to \infty$ yields

$$t\int_X h \leqslant \lim_k \int_X g_k$$

Taking $t \to 1$ yields

$$\int_X h \leqslant \lim_k \int_X g_k$$

Finally, run $h$ over all simple functions no larger than $f$ and use that $g_k \leqslant f_k$:

$$\int_X f = \sup_h \int h \leqslant \lim_k \int_X g_k = \liminf_k \int_X g_k \leqslant \liminf_k \int_X f_k.$$

And we are done. ∎

**Theorem 3.3.6** (Monotone Convergence Theorem/Levi-Lebesgue). *Suppose $\{f_n\}$ is a sequence of measurable functions $X \to [-\infty, +\infty]$ that's bounded below by some $g \in L^1_\mu(X)$. Let $f$ be the $\mu$-a.e.pointwise limit (which may take $\infty$ as value). If each $f_n \leqslant f$ $\mu$-a.e., then $\int_X f\, d\mu = \lim \int_X f_n\, d\mu$.*

*Proof.* Since $f_n \leqslant f$ $\mu$-a.e., $\int f_n \leqslant \int f$. We get

$$\int f \leqslant \underline{\lim} \int f_n \leqslant \overline{\lim} \int f_n \leqslant \int f.$$

where the first $\leqslant$ comes from Fatou's lemma. ∎

⋙ **Theorem 3.3.7** (Dominated Convergence Theorem). *Suppose $\{f_n\}$ is a sequence of measurable functions $X \to \mathbb{R}$. Assume $f_n \to f$ pointwise $\mu$-a.e., where $f$ is measurable. Also assume there is a $g \in L^1_\mu$ such that $|f_n| \leqslant g$ for all $n$. Then $\int_X f_n\, d\mu \to \int_X f\, d\mu$ and $\|f_n - f\|_1 \to 0$.*

*Proof.* WLOG, suppose the $f_n$'s are nonnegative and $g \geqslant f_n$. Then $g - f_n$ is bounded below by the constant function $0 \in L^1_\mu$.

Because $\int |f|\, d\mu = \int f\, d\mu \leqslant \underline{\lim} \int f_n\, d\mu \leqslant \int g\, d\mu$ and $f \geqslant 0$, Fatou's lemma implies that $f$ is Lebesgue integrable.

Now,

$$\int_X g\, d\mu - \int_X f\, d\mu = \int_X (g - f)\, d\mu \leqslant \underline{\lim}_{n\to\infty} \int_X (g - f_n)\, d\mu = \int_X g\, d\mu - \overline{\lim}_{n\to\infty} \int_X f_n\, d\mu$$

$$\Rightarrow \quad \int_X f\, d\mu \geqslant \overline{\lim}_{n\to\infty} \int_X f_n\, d\mu.$$

OTOH, $g + f_n \geqslant g \in L^1_\mu(X)$. Fatou's lemma again gives

$$\int_X g\, d\mu + \int_X f\, d\mu = \int_X (g+f)\, d\mu \leqslant \varliminf_{n\to\infty} \int_X (g+f_n)\, d\mu = \int_X g\, d\mu + \varliminf_{n\to\infty} \int_X f_n\, d\mu$$

$$\Rightarrow \quad \int_X f\, d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n\, d\mu.$$

Together we get

$$\int_X f\, d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n\, d\mu \leqslant \varlimsup_{n\to\infty} \int_X f_n\, d\mu \leqslant \int_X f\, d\mu$$

and so

$$\int_X f_n\, d\mu \to \int_X f\, d\mu.$$

For the second statement regarding the norm, let $g_n = |f - f_n|$. Then $|g_n| \leqslant |f| + |g| \in L^1_\mu(X)$ and $g_n \to 0$ $\mu$-a.e.. Use $\int_X f_n\, d\mu \to \int_X f\, d\mu$ to get

$$\int_X |f - f_n|\, d\mu \to \int_X 0\, d\mu = 0.$$

∎

**Theorem 3.3.8** (General LDC).

→ **Proposition 3.3.9** (uniform integrability). *Suppose $f \in L^1_\mu(X)$. Then for any $\epsilon$, there is some $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\mu(A) < \delta$, $\int_A |f|\, d\mu < \epsilon$.*

*Proof.* Case 1: $\|f\|_\infty < \infty$. Set $\delta = \epsilon / \max(\|f\|_\infty, 1)$.

Case 2: $\|f\|_\infty = \infty$. Define

$$f_n(x) = \begin{cases} |f|(x) & \text{if} |f(x)| \leqslant n \\ n & \text{if} |f(x)| > n \end{cases}$$

Then $\|f_n\|_\infty \leqslant n$, $|f_n| \leqslant |f| \in L^1_\mu$, and $|f_n| \to |f|$ pointwise. By the LDCT, $\int |f_n| \to \int |f|$, so there's an $N$ such that for all $n \geqslant N$, $\int |f_n| - |f| < \epsilon/2$.

Notice that

$$\frac{\epsilon}{2} > \int |f_N| - |f| = N \cdot \mu(f > N) \quad \Rightarrow \quad \mu(f > N) < \frac{\epsilon}{2N}.$$

Set $\delta < \epsilon/2N$. We then have for $\mu(A) < \delta$,

$$\begin{aligned} \int_A |f|\, d\mu &= \int_A |f| - |f_N|\, d\mu + \int_A |f_N|\, d\mu \\ &\leqslant \underbrace{\int_A |f| - |f_N|\, d\mu}_{<\epsilon/2} + \underbrace{N\mu(A)}_{<\epsilon/2} \quad < \epsilon \end{aligned}$$

∎

**Definition 3.3.10** (absolute continuity). **a.** A measurable function $f$ is *absolutely continuous* with respect to $\mu$ if for every $\epsilon > 0$, there's a $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\mu(A) < \delta$, $\left|\int_A f \, d\mu\right| < \epsilon$. If $f \in L^1_\mu$, then $\int f < \infty$, so automatically $f$ is absolutely continuous.

**b.** A collection of set functions $\{\nu_m\}$ on $\mathcal{A} \subset \mathcal{P}(X)$ is *uniformly absolutely continuous* with respect to $\mu$ if for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\mu(A) < \delta$, $|\nu_m(A)| < \epsilon$ uniformly in $m$.

**b′.** A collection of integrable functions $\{f_\alpha\}$ is *uniformly absolutely continuous* if for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $A \in \mathcal{A}$ with $\mu(A) < \delta$, $\left|\int_A f_\alpha \, d\mu\right| < \epsilon$ for all $\alpha$.

**c.** A set function $\nu$ on $\mathcal{F} \subset \mathcal{P}(X)$ is *Vitali continuous* if for decreasing sequence $\{A_n\} \subset \mathcal{F}$ with $\bigcap A_n = \varnothing$, $\lim_{n\to\infty} \nu(A_n) = 0$.

A family of set functions $\nu_m$ on $\mathcal{F} \subset \mathcal{P}(X)$ is *Vitali equicontinuous* if for decreasing sequence $\{A_n\} \subset \mathcal{F}$ with $\bigcap A_n = \varnothing$, for every $\epsilon > 0$, there is an $N \in \mathbb{Z}_+$ large such that for all $n \geqslant N$, $|\nu_m(A_n)| < \epsilon$ uniformly in $m$.

**Definition** (convergence in measure). A sequence of measurable functions $\{f_n\}$ *converges in measure* to $f$ if for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mu\{x : |f(x) - f_n(x)| > \epsilon\} = 0.$$

For this we sometimes write $f_n \xrightarrow{M} f$.

**Theorem 3.3.11.**

**Theorem 3.3.12.**

**Theorem 3.3.13** (Riesz-Lebesgue). *Suppose $f, f_n$ are measurable functions.*

**a.** *If $f_n \to f$ in measure then there is a subsequence $f_{n_k} \to f$ pointwise a.e.*

**b.** *If the space has finite measure then $f_n \to f$ a.e. implies $f_n \to f$ in measure.*

The following is useful in cooking up counterexamples when dealing with weak convergence.

**Theorem 3.3.14** (Strongly nonconvergent dilations). *Suppose $f$ is a nonconstant bounded 1-periodic Lebesgue measurable function on $\mathbb{R}$. Then the sequence $\{f_n\}$ defined by $f_n := f(n\cdot)$ has no subsequence that converges pointwise a.e. on any bounded interval.*

## 3.4 Riemann vs Lebesgue integrals

**Proposition 3.4.1.** *Suppose $f : [a,b] \to \mathbb{R}$ is bounded and that $R\int_a^b f$ exists. Then $f$ is Lebesgue integrable on $[a,b]$, with $R\int_a^b f = \int_a^b f$.*

*Proof.* From definitions:

$$
\begin{aligned}
R\int_a^b f = \overline{\int_a^b} f = & \inf_{P\text{: partition of } [a,b]} \left\{ R\int_a^b g : g \geqslant f \text{ a simple function subordinated to } P \right\} \\
\geqslant & \inf \left\{ R\int_a^b g : g \geqslant f \text{ a simple function on } [a,b] \right\} \\
\geqslant & \sup \left\{ R\int_a^b h : f \geqslant h \text{ a simple function on } [a,b] \right\} \\
\geqslant & \sup_{P\text{: partition of } [a,b]} \left\{ R\int_a^b h : f \geqslant h \text{ a simple function subordinated to } P \right\} \\
= & \underline{\int_a^b} f = R\int_a^b f
\end{aligned}
$$

∎

**Proposition 3.4.2.** *$f:\mathbb{R}\to\mathbb{R}$ is continuous m-a.e. if and only if for every $V$ open in $\mathbb{R}$, $f^{-1}V = U \cup A$ for some $U$ open and $A \in \mathcal{M}(\mathbb{R})$ with measure zero.*

**Example 3.4.3** (composition of continuous $m$-a.e. functions needs not be continuous)**.** Consider

$$
f(x) = \mathbb{1}_{\{1/n : n\in\mathbb{Z}_+\}}, \quad \text{and} \quad g(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R}\setminus\mathbb{Q} \\ 1 & \text{if } x = 0 \\ 1/|q| & \text{if } x = p/q \text{ in lowest term} \end{cases}
$$

Obviously, $f$ is continuous a.e. and $g$ is so by1.3.7.

However $f \circ g = \mathbb{1}_{\mathbb{Q}}$ is discontinuous everywhere.

The following is important:

**Theorem 3.4.5** (Riemann and continuity m-a.e.)**.** *Suppose $f:[a,b]\to\mathbb{R}$ is bouned. Then $R\int_a^b f$ exists if and only if it is continuous m-a.e.*

**Theorem 3.4.8.** *Suppose $f:[a,b]\to\mathbb{R}$ is such that $\lim_{y\to x\pm} f(y)$ exists for every $x\in(a,b)$. Then $f$ is Riemann integrable.*

### 3.5 Lebesgue-Stieltjes measures and integrals

For an increasing right-continuous function $f:\mathbb{R}\to\mathbb{R}$, define

$$\mu_f(a,b] := f(b) - f(a).$$

One can easily check that $\mu_f$ is a nonnegative, finitely additive set function on the semiring $\mathcal{Q}$ generated by half parallelepipeds $\prod_{i=1}^d (a_i, b_i]$ in $\mathbb{R}^d$.

To see $\mu_f$ is in fact $\sigma$-additive, suppose $A = (a,b] = \bigcup_{j=1}^\infty A_j$ where $\{A_j = (a_j, b_j] : j \in \mathbb{Z}_+\} \subset \mathcal{Q}$ is pairwise disjoint.

One inequality $\sum \mu_f(A_j) \leqslant \mu_f(A)$ is obtained by taking $n \to \infty$ of

$$\sum_{j=1}^{n} \mu_f(A_j) = \mu_f\Big(\bigcup_{j=1}^{n} A_j\Big) \leqslant \mu_f(A).$$

For the other one, suppose $\epsilon > 0$. Produce some $\delta_j$'s such that the right-continuity of $f$ holds in the following manner:

$$\mu_f(a, a+\delta_0] = f(a+\delta_0) - f(a) < \epsilon$$
$$\mu_f(b_j, b_j+\delta_j] = f(b_j+\delta_j) - f(b_j) < \epsilon/2^j$$

Then we see

$$[a+\delta_0, b] \subset (a, b] = \bigcup_{j=1}^{\infty} (a_j, b_j] \subset \bigcup_{j=1}^{\infty} (a_j, b_j+\delta_j).$$

Use the compactness of $[a+\delta_0, b]$ to pass the RHS to a finite subcover (rearranging the indices if necessary)

$$[a+\delta_0, b] \subset \bigcup_{j=1}^{k} (a_j, b_j+\delta_j)$$

Apply finite additivity and nonnegativity of $\mu_f$ again:

$$\begin{aligned}
\mu_f(a,b] \leqslant \mu_f(a, a+\delta_0] + \mu_f(a+\delta_0, b] &\leqslant \mu_f(a, a+\delta_0] + \sum_{j=1}^{k} \mu_f(a_j, b_j+\delta_j] \\
&= \mu_f(a, a+\delta_0] + \sum_{j=1}^{k} \mu_f(a_j, b_j] + \sum_{j=1}^{k} \mu_f(b_j, b_j+\delta_j] \\
&\leqslant f(a+\delta_0) - f(a) + \sum_{j=1}^{\infty} \mu_f(a_j, b_j] + \sum_{j=1}^{k} f(b_j\delta_j) - f(b_j) \\
&< 2\epsilon + \sum_{j=1}^{\infty} \mu_f(a_j, b_j].
\end{aligned}$$

Extending $\mu_f$ to the $\sigma$-algebra $\mathcal{A}$ generated by $\mathcal{Q}$, we get ourselves the *Lebesgue-Stieltjes measure associated with $f$*.

Now, going backwards:

**Proposition 3.5.2.** *Suppose $\mu : \mathcal{Q} \to \mathbb{R}_+$ is $\sigma$-additive. There exists an increasing right-continuous function $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(b) - f(a) = \mu(a, b].$$

*Proof.* Define

$$f(x) = \begin{cases} \mu(0, x] & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\mu(x, 0] & \text{if } x < 0 \end{cases}$$

The only nontrivial thing to check is right-continuity. For this, take a decreasing sequence $\{a_n\} \subset \mathbb{R}$ with $a_n \to b$. Write

$$(b, a_1] = \bigcup_{j=1}^{\infty} (a_{j+1}, a_j]$$

∎

## 3.6 Some applications

### 3.6.1 Fourier stuff

Write $\mathbb{T} = \mathbb{R}/\mathbb{Z} \cong [0,1)$. For $f \in L^1_\mu(\mathbb{T})$, define the *Fourier coefficients* of $f$ to be

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(t) e^{-2\pi i \xi t} dt.$$

(This is just the Fourier transform of $f$) Usually we do this over the integers, i.e. run $\xi$ over $\mathbb{Z}$ and get $c_n = \hat{f}(n)$. The *Fourier series of $f$* is then

$$S(f)(x) = \sum_{n \in \mathbb{Z}} c_n e^{2\pi i x t}.$$

**Theorem 3.6.4** (Riemann-Lebesgue lemma)**.** *For $f \in L^1_m(\mathbb{T})$,*

$$\lim_{|n| \to \infty} \hat{f}(n) = 0.$$

*Proof.* Suppose $(a,b) \subset [0,1)$. Then

$$\lim_{|n| \to \infty} \left| \int_a^b e^{2\pi i n t} dt \right| = \lim_{|n| \to \infty} \left| \frac{e^{2\pi i n b} - e^{2\pi i n a}}{2\pi i n} \right| = 0$$

Consequently the result is true for characteristic functions of intervals. Because integrable functions can be approximated by continuous things, which can be approximated by characteritic functions of intervals, the result holds for arbitrary elements of $L^1_m(\mathbb{T})$. ∎

## 3.7 Fubini and Tonelli

Suppose $(X, \mathcal{A}_1, \mu)$, $(Y, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces.

For $A \in \mathcal{A}_1$ and $B \in \mathcal{A}_2$, we call $A \times B$ a *measurable rectangle*. The collection $\mathcal{R}$ of all measurable rectangles satisfies 1) closed under finite intersections; 2) complement of a rectangle is a finite union rectangles; So $\mathcal{R}$ forms a semialgebra.

Set $\mathcal{A}$ to be the collection of all finite unions of disjoint elements of $\mathcal{R}$. It's easy to see $\mathcal{A}$ is an algebra.

Define $\omega$ on $\mathcal{R}$ by

$$\omega(A \times B) = \mu(A) \cdot \nu(B)$$

and set $0 \cdot \infty = 0$.

**Proposition 3.7.1.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(Y, \mathcal{A}_2, \nu)$ are measure spaces. If $X, Y$ are both finite (resp. $\sigma$-finite, resp. complete $\sigma$-finite) measure spaces, then $(X \times Y, \mathcal{A}_1 \times \mathcal{A}_2, \mu \times \nu)$ is a finite (resp. $\sigma$-finite, resp. complete $\sigma$-finite) measure space.* ∎

**Definition** (sections)**.** For $E \subset X \times Y$, $x \in X$, $y \in Y$ define the *sections* of $E$ at $x, y$ resp. to be

$$E_x := \{y \in Y \,:\, (x,y) \in E\}$$

$$E^y := \{x \in X \,:\, (x,y) \in E\}.$$

For $f : X \times Y \to [-\infty, +\infty]$, define

$$f_x(y) = f^y(x) = (x,y).$$

**Proposition 3.7.2.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(Y, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces. Suppose $E \in \mathcal{A}_1 \times \mathcal{A}_2$. Then for any $x \in X$, $y \in Y$,*

$$E_x \in \mathcal{A}_2 \quad \textit{and} \quad E^y \in \mathcal{A}_1.$$

**Proposition 3.7.3.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces. Let $(X \times Y, \mathcal{A}_1 \times \mathcal{A}_2, \mu \times \nu)$ be the associate $\sigma$-finite product measure space.*

*If $f : X \times Y \to [-\infty, +\infty]$ is a $(\mu \times \nu)$-measurable function, then for all $x \in X$ (resp. $y \in Y$), the section $f_x$ is $\nu$-measurable (resp. $f^y$ is $\mu$-measurable).*

**Theorem 3.7.4.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces. Suppose $E \in \mathcal{A}_1 \times \mathcal{A}_2$ is $(\mu \times \nu)$-measurable. Then the functions $v(E_x) : X \mapsto [-\infty, +\infty]$ and $\mu(E^y) : Y \mapsto [-\infty, +\infty]$ are measurable. Moreover,*

$$\int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) \nu(y) = (\mu \times \nu)(E).$$

**Theorem 3.7.5** (Fubini theorem for integrable functions)**.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces and $h : X \times Y \to [-\infty, +\infty]$ is $(\mu \times \nu)$-integrable.*

- **a.** *For $\mu$-a.e.$x \in X$, $h_x \in L^1_\nu(Y)$ and $\int_Y h_x d\nu \in L^1_\mu(X)$.*
- **b.** *For $\nu$-a.e. $y \in Y$, $h^y \in L^1_\mu(X)$ and $\int_X h^y d\mu \in L^1_\nu(Y)$.*
- **c.** *We have*

$$\begin{aligned} \int\int_{X \times Y} h(x,y) d(\mu \times \nu)(x,y) &= \int_X \left( \int_Y h(x,y) d\nu(y) \right) d\mu(x) \\ &= \int_Y \left( \int_X h(x,y) d\mu(x) \right) d\nu(y). \end{aligned} \tag{3.7.1}$$

**Theorem 3.7.7** (Tonelli theorem for nonnegative functions)**.** *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces and $h : X \times Y \to \mathbb{R}_+$ is $(\mu \times \nu)$-measurable. Then $\int_Y h_x d\nu$ is $\mu$-measurable and $\int_X h^y d\mu$ is $\nu$-measurable. Moreover, equation 3.7.1 holds.*

**Theorem 3.7.8** (Tonelli theorem). *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are $\sigma$-finite measure spaces. Let $h : X \times Y \to [-\infty, +\infty]$ be a $(\mu \times \nu)$-measurable function such that*

$$\int_X \int_Y |h(x,y)| d\nu(y) d\mu(x) < \infty \quad \text{or} \quad \int_Y \int_X h(x,y) d\mu(x) d\nu(y) < \infty.$$

*Then $h \in L^1_{(\mu \times \nu)}(X \times Y)$ and equation 3.7.1 holds.*

**Theorem 3.7.9** (Fubini-Tonelli for complete $\sigma$-finite measure spaces). *Suppose $(X, \mathcal{A}_1, \mu)$, $(X, \mathcal{A}_2, \nu)$ are complete $\sigma$-finite measure spaces. Let $(X \times Y, \mathcal{A}_3, \omega) = (X \times Y, \overline{\mathcal{A}_1 \times \mathcal{A}_2}, \overline{\mu \times \nu})$ be the completion of the product measure space of $X \times Y$. Let $h : X \times Y \to \mathbb{C}$ be $(\mu \times \nu)$-measurable.*

- **a.** *If $h \in L^1_\omega(X \times Y)$, then $h_x \in L^1_\nu(Y)$ for $\mu$-a.e. $x \in X$; $h^y \in L^1_\mu(X)$ for $\nu$-a.e. $y \in Y$. $\int_X h^y d\mu \in L^1_\nu(Y)$ and $\int_Y h_x d\nu \in L^1_m(X)$. Moreover, 3.7.1 holds.*
- **b.** *If $h \geqslant 0$ then $h_x$ is $\nu$-measurable for $\mu$-a.e.$x \in X$ and $h^y$ is $\mu$-measruable or $\nu$-a.e. $y \in Y$. Also, $\int_X h^y d\mu \in L^1_\nu$, $\int_Y h_x d\nu \in L^1_\mu$ and 3.7.1 holds.*

**Theorem 3.7.10** (Lebesuge's first theorem). *Suppose $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ is continuous in $x$ for each fixed $y \in \mathbb{R}$ and continuous in $y$ for each fixed $x$. Then $f$ is Lebesgue measurable.*

**Lemma.** *Suppose $g : [a,b] \to \mathbb{C}$ is bounded and assmue that both*

$$\lim_{n\to\infty} S_{P_n} \quad \text{and} \quad \lim_{n\to\infty} s_{P_n}$$

*exist and are equal for every sequence of partitions $\{P_n\}$ of $[a,b]$ with $\text{card} P_n = n+1$, $|P_n| \to 0$. Then $R\int_a^b g$ exists.*

**Theorem 3.7.12** (Fubini-Tonelli for Lebesgue and Riemann integrable functions). *Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \times [a,b] \to \mathbb{C}$. Suppose the $R\int_a^b f(x,y)dy$ exists $\mu$-a.e.and $f$ is $\mu$-measurable for all $y \in [a,b]$. Furthermore, assume there is an $F \in L^1_\mu(X)$ satisfying for all $(x,y) \in X \times [a,b]$,*

$$|f(x,y)| \leqslant F(x).$$

*Then $\int_X f(x,t) d\mu(x)$ is Riemann integrable and*

$$\int_X \left( R\int_a^b f(x,y)dy \right) d\mu(x) = R\int_a^b \left( \int_X f(x,y) d\mu(x) \right) dy.$$

## 3.8 Selected exercises

### 3.8.1 Convolution

Define the convolution of $f$ and $g$ as

$$f * g(x) = \int f(x-y) g(y) dy.$$

This thing came out of harmonic analysis (I think) and is quite useful in PDE. I don't know a good way to visualize it but it's sort of averaging the two functions. It hurts my brain less when one of the guys is compactly supported. Perhaps a helpful example is when $g = \mathbb{1}_{B(0,r)}$. We have $f * g(x) = \int_{B(x,r)} f(y)dy$. If $f \in L^1_{\text{loc}}$, then by the Lebesgue differentiation theorem (whoops that comes later in 631. But 4.4.5 has a similar flavor), as $r \to 0$, $f * g \to f$ pointwise a.e.

These are more fun when formulated using distribution theory. But let's not get into that...

**Exercise 3.5.** Suppose $f,g$ are Lebesgue integrable functions on $\mathbb{R}$. Then

**a.** For a.e. $x \in \mathbb{R}$, the map $y \mapsto f(x-y)g(y)$ belongs to $L^1$.

**b.** $f * g$ is an element of $L^1$ satisfying the estimate $\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1$.

*Sketch of proof.* First show the function $(x,y) \mapsto f(x-y)g(y)$ is measurable a.e. Then apply Tonelli to obtain $\int \int |f(x-y)g(y)|dydx = \int |g(y)| \int |f(x-y)|dxdy = \|f\|_1 \|g\|_1$. ∎

**Exercise #5, Aug'10 Analysis qual.** Suppose $f,g \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Define

$$h(x) = \int_{\mathbb{R}} f(x+y)g(y)dy.$$

Prove that $h$ is continuous and vanishes at infinity.

*Proof.* ∎

**Exercise 3.6** (Steinhaus)**.** If $X \in \mathcal{M}(\mathbb{R})$ and $m(X) > 0$, then the algebraic difference $X - X = \{x - y | x,y \in X\}$ contains a neighborhood of 0.

*Sketch of proof.* WLOG assume $X$ has finite measure. Let $f := \mathbb{1}_X$ and $g := \mathbb{1}_{(-X)}$. By previous exercise, $f * g$ is continuous. Observe that $0 \in \text{supp}(f * g) \subset X - X$. Hence by continuity one can produce a neighborhood of 0 contained in $\text{supp}(f * g)$ contained in $X - X$. ∎

**Exercise 2.24.** Suppose $E \subset [0,1]$ has strictly positive outer measure. Prove the existence of a nonmeasurable subset of $E$.

*Proof.* As in example 2.2.16, let $\mathcal{G} := \mathbb{R}/\sim$ where $x \sim x + q$, $q \in \mathbb{Q}$. For each $A \in \mathcal{G}$, fix $r_A \in (0, 1/2)$ to be a representative of the class $A$. Let $V_0$ be the set made up of all these $r_A$'s. For $q \in \mathbb{Q}$, define $V_q = V_0 + q$ to be the algebraic sum. Put $E_q := E \cap V_q$. Again, the $V_q$'s are disjoint for distinct $q$'s, thus, so are the $E_q$'s.

Now, if any of the $E_q$'s is nonmeasurable, we are done. So suppose not. Look at

$$V_q - V_q = \{(r+q) - (t+q) | r,t \in V_0\} = \{(r - t | r,t \in V_0\} = V_0 - V_0$$

and realize two things:

- $V_0 - V_0$ is disjoint from $\mathbb{Q}$ because we modded $\mathbb{Q}$ out;
- if $m(V_q) > 0$, then $V_q - V_q = V_0 - V_0$ contains a neighborhood of 0. (previous exercise)

Obviously the second point can't happen or it contradicts the first. So we must have $m(V_q) = 0$ for all $q \in \mathbb{Q}$. But $E = \bigcup E_q$ and that gives

$$0 < m(E) \leqslant \sum_q m(E_q) = 0$$

which is absurd. ∎

# 4 Differentiation

(12/06/21) Hahaha life is falling apart but I'm too busy to be depressed! Spewing out the following as a review for the last impending midterm.
(06/04/22) I'm updating everything because I have no better things to do. I also miss being an undergrad. You know, going to classes and impudently bugging the profs...

This section sets in $(X, \mathcal{M}, m)$.

## 4.1 bounded variation

**Definition** (Total variation)**.** Suppose $f : [a,b] \to \mathbb{R}$. The *total variation* of $f$ on $[a,b]$ is defined to be

$$V(f,[a,b]) = \sup_P \sum_P |f(x_j) - f(x_{j-1})|$$

where the supremum runs over all partitions $P = [a = x_1, x_2, \ldots, x_n = b]$ of $[a,b]$.

We say $f$ is of *bounded variation* on $[a,b]$ if $V(f,[a,b]) < \infty$. Let $BV[a,b]$ denotes the set of all functions of bounded variation on $[a,b]$.

Quickly observe that if $f$ is increasing on $[a,b]$, then the total variation telescopes, so we end up with $V(f,[a,b]) = f(b) - f(a)$.

**Definition 4.1.1** (Variation function)**.** For $f \in BV[a,b]$, define

$$V(f)(x) = \begin{cases} f(0) + V(f,[0,x]) & \text{if } x \geqslant 0 \\ f(0) - V(f,[x,0]) & \text{if } x < 0 \end{cases}.$$

Note that $V(f)(x)$ is increasing.

Define $BV_{\text{loc}}(\mathbb{R})$ to be the space of functions $f$ that is of bounded variation on every interval $[a,b] \subset \mathbb{R}$. Define $BV(\mathbb{R}) \subset BV_{\text{loc}}(\mathbb{R})$ to be the subspace containing $f$ whose variation function on $\mathbb{R}$ satisfies

$$V(f) = V(f,\mathbb{R}) = \sup_{a<b} V(f,[a,b]) < \infty.$$

$BV(\mathbb{R})$ is a real vector space by the following theorem:

**Theorem 4.1.2+4** (Jordan decomposition theorem)**.**

**a.** *The difference $f = g_1 - g_2$ of two increasing functions $g_i$ belongs to $BV_{loc}$. Moreover, on each $[a,b] \subset \mathbb{R}$,*

$$V(f,[a,b]) \leqslant g_1(b) - g_1(a) + g_2(b) - g_2(a).$$

**b.** *If $f \in BV_{loc}$, then*

$$P(x) = \frac{1}{2}[V(f)(x) + f(x)]$$
$$N(x) = \frac{1}{2}[V(f)(x) - f(x)]$$

*are increasing functions, with $f = P - N$, $V(f) = P + N$, $f(0) = P(0) = V(f)(0)$, and $N(0) = 0$.*

**c.** *$f \in BV_{loc}$ iff $f = f_1 - f_2$ for some functions $f_1, f_2$ bounded increasing on $\mathbb{R}$. In this case, $\lim_{x\to\pm\infty} f(x)$ exists.*

**Proposition 4.1.11.** *Suppose $f \in BV[a,b]$. Then*

**a.** *$f(c\pm)$ exists at every $c \in [a,b]$ and $D(f)$ is at most countable.*

**b.** *$f$ is Lebesgue measurable and bounded.*

**Example 4.1.12** (increasing function with discontinuities on $\mathbb{Q}$)**.** See page 172. See also exercise 4.21 or something. See further the Vitali differentiation theorem (exercise 4.19, if I remember correctly).

## 4.2 Decomposition into discrete and continuous parts

**Example 4.2.5** (Hellinger)**.** I went to talk to Patrick and missed the first half of the lecture, which proved impossible for me to follow the rest of the example. I'll try to read the text one day.

**Example** (Cantor function revisit)**.** Increasing, continuous, BV, but not absolutely continuous. Set $\epsilon < 1$. Because $m(C)$ has measure 0, we can produce a cover for $C$ with length as small as we want and still have $\sum |f(x_j) - f(x_j)| = 1$.

Suppose $f : [a,b] \to \mathbb{R}$ is continuous. Set the *Banach indicatrix* $B : \mathbb{R} \to [0,\infty]$ to be

$$B(y) := \operatorname{card} f^{-1}(y).$$

**Theorem 4.2.7** (Banach-Vitali theorem)**.**

**a.** *$B$ is Lebesgue measurable and*

$$\int_c^d B(y)dy = V(f,[a,b])$$

*where $c = \inf_{[a,b]}\{f(\xi)\}$ and $d = \sup_{[a,b]}\{f(\xi)\}$.*

**b.** *$f \in BV[a,b] \Leftrightarrow B \in L^1_m[c,d]$.*

**c.** *If $f \in BV[a,b]$ then $m(\{y : B(y) \geqslant \aleph_0\}) = 0$.*

## 4.3 Differentiation theorems

A family $\mathcal{V}$ of intervals is a *Vitali covering* of $X \subset \mathbb{R}$ if for any $\epsilon > 0$, $x \in X$, there is a member $I \in \mathcal{V}$ containing $x$ and has length $m(I) < \epsilon$.

**Observations.** Let $X \subset \mathbb{R}$, $U \subset \mathbb{R}$ open, $F \subset \mathbb{R}$ closed, $\mathcal{V}$ a Vitali covering of $X$. Then

i. $\{I \in \mathcal{V} : I \subset U\}$ is a Vitali covering of $X \cap U$.

ii. $\{I \in \mathcal{V} : I \subset F\}$ is NOT necessarily a Vitali covering of $X \cap F$. ($\partial$ would be a problem: consider when $F$ is discrete.)

iii. $\{V^\circ : V \in \mathcal{V}\}$ is a Vitali covering of $\limsup \mathcal{V}^\circ =$ points that belong to $\mathcal{V}^\circ$ infinitely often.

**Theorem 4.3.1** (Vitali covering lemma)**.** *Suppopse $X \subset \mathbb{R}$ has finite outer measure and $\mathcal{V}$ is a Vitali covering of X. Then for every $\epsilon > 0$, there is a finite disjoint subcollection $\{I_n\} \subset \mathcal{V}$ with outer measure*

$$m^*\left(X \setminus \bigcup_{n=1}^{N} I_n\right) < \epsilon.$$

*Proof.* Assume $U$ is open and every $I \in \mathcal{V}$ is contained in $U$. (Because outer measure is characterized by that of open intervals on $\mathbb{R}$, we can then approximate $m^*(X)$.) Also assume all elements in $\mathcal{V}$ are closed.

Take any finite disjoint subcollection $\{I_i\}_{i=1}^{n} \subset \mathcal{V}$ with $m(I_i) \geqslant m(I_{i+1})$. If it covers $U$, move on to the next step. Otherwise, put

$$r_n := \sup\{m(I) : I \in \mathcal{V}, I \cap I_i = \varnothing, i \in [n]\}$$

Then $r_n \leqslant m(U) < \infty$. Also $r_n > 0$ because $\mathcal{V}$ covers $U$ and is a Vitali covering.

Let $I_{n+1} \in \mathcal{V}$ be such that

$$m(I_{n+1}) > \tfrac{1}{2} r_n \text{ and } I_{n+1} \cap I_i = \varnothing, i \in [n]$$

and $m(I_{n+1}) \leqslant m(I_n)$. Claim that we can choose appropriate $I_j$'s so that $\{m(I_j)\}$ is decreasing.

I'm not done yet ∎

**Theorem** (VCL for balls)**.** *A more general version for separable metric spaces. This we proved in 631. The eager reader should check this out.*

**Theorem 4.3.2** (Lebesgue differentiation theorem)**.** *Suppose $f \in BV[a,b]$. Then*

- *$f'$ exists m-a.e.*
- *$f' \in L^1_m[a,b]$*
- *If $f$ is increasing then $\int_a^b f' \leqslant f(b) - f(a)$*

**Theorem** (Alternate form of LDT)**.** *Let $f \in L^1_m[a,b]$. The for m-a.e. $x \in [a,b]$,*

$$f(x) = \lim_{r \to 0} \frac{1}{r} \int_x^{x+r} f(y) dm(y).$$

**Theorem** (Lebesgue density theorem)**.** *Let $A \in \mathcal{M}(\mathbb{R})$, $m(A) > 0$. Then for m-a.e. $x \in A$,*

$$\lim_{h \downarrow 0} \frac{m(A \cap (x-h, x+h))}{2h} = 1.$$

## 4.4 FTC I

In this section, unless otherwise specified, $F(x) = \int_a^x f\,dm$. Note that when $f \in L^1_m[a,b]$, $F$ is continuous of BV.

**Proposition 4.4.2.** *Let $f \in L^1_m[a,b]$ is such that $F \equiv 0$ on $[a,b]$. Then $f \equiv 0$ $m$-a.e.*

**Theorem 4.4.3** (FTC I)**.** *Let $f \in L^1_m[a,b]$, $r \in \mathbb{R}$. Define $F : [a,b] \to \mathbb{R}$ by*

$$F(x) := r + \int_a^x f$$

*Then $F' = f$ $m$-a.e.*

The theorem is weaker than the old calculus FTC. Consider the Cantor function $C_C$, continuous of BV, however $C_C' \equiv 0$ $m$-a.e. and $\int_0^1 C_C' = 0 \neq 1 = C_C(1) - C_C(0)$.

The following is not really covered in 630 but after PDE I thought to introduce it early:

**Definition** (Lebesgue points)**.** Given a locally integrable function $f$ on $\mathbb{R}^n$, we say $x$ is a *Lebesgue point of $f$ if*

$$\lim_{r \downarrow 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)| dy = 0. \tag{$\star$}$$

*The set of Lebesgue points of $f$ is called the Lebesgue set of $f$.*

In one dimension, the condition $\star$ translates to

$$\lim_{h \to 0} \frac{1}{h} \int_0^h |f(x+t) - f(x)| dt = 0$$

or more succinctly,

$$\int_0^h |f(x+t) - f(x)| dt = o(h), \quad h \to 0.$$

**Theorem 4.4.5** (Lebesgue's extension of FTC I)**.** *Let $f \in L^1_m[a,b]$. There is a subset $L \subset [a,b]$ of full measure such that for all $x \in L$ and any $r \in \mathbb{R}$,*

$$\lim_{h \to 0} \frac{1}{h} \int_0^h |f(x+t) - r| dt = |f(x) - r|.$$

*Proof.* Fix a countable dense subset $\{r_n\}$ of $\mathbb{R}$. Define $g_n(x) = |f(x) - r_n|$, integrable on $[a,b]$. Let $L_n$ be a set of full measure on which

$$\lim_{h \to 0} \frac{1}{h} \int_x^{x+h} g_n = \lim_{h \to 0} \frac{1}{h} \left( \int_a^{x+h} g_n - \int_a^x g_n \right) = g_n(x)$$

where the second equality is given by the FTC I. Put $L = \bigcap L_n$. Then $m(L^\sim) = m(\bigcup L_n^\sim) \leqslant \sum m(L_n^\sim) = 0$, ie $L$ has full measure.

Suppose $\epsilon > 0$. Using density of $\{r_n\}$, find appropriate $n$ such that $r_n$ lands in $B(r, \epsilon/3)$. Shifting the ball by $f(y)$ yields $\big||f(y) - r| - |f(y) - r_n|\big| < \epsilon/3$, so we have on $L$,

$$\left|\frac{1}{h}\int_x^{x+h}|f(t) - r|dt - |f(x) - r|\right| < \tfrac{2\epsilon}{3} + \left|\frac{1}{h}\int_x^{x+h}\underbrace{|f(t) - r_n|}_{g_n(t)}dt - \underbrace{|f(x) - r_n|}_{g_n(x)}\right|$$

the second term goes to zero by the first equation, $\epsilon$ we can make as small as we like. So done. ∎

**Corollary 4.4.6.** *Let $f \in L^1_m[a,b]$. The Lebesgue set of $f$ has full measure.*

## 4.5 FTC II and absolute continuity

**Definition** (absolute continuity)**.** A function $F : [a,b] \to \mathbb{R}$ is *absolutely continuous* on $[a,b]$ if for every $\epsilon > 0$, there is a $\delta > 0$ such that for all disjoint families $\{(x_j, y_j) : j \in [n]\}$,

$$\sum_{j=1}^{n}(y_j - x_j) < \delta \quad \Rightarrow \quad \sum_{j=1}^{n}\left|F(y_j) - F(x_j)\right| < \epsilon.$$

or equivalently (exercise 4.35 ii)

$$\sum_{j=1}^{n}(y_j - x_j) < \delta \quad \Rightarrow \quad \left|\sum_{j=1}^{n}F(b_j) - F(a_j)\right| < \epsilon.$$

or equivalently (exercise 4.35 iii)

$$\sum_{j=1}^{n}(y_j - x_j) < \delta \quad \Rightarrow \quad \sum_{j=1}^{n}\omega(F, [a_j, b_j]) < \epsilon.$$

(remember the *oscillation* $\omega$ is defined by $\omega(f, I) = \sup_I f - \inf_I f$.)

**Example (exercise 4.33b)** (AC on $(0,1]$ but not on $[0,1]$)**.** Let $f(x) = x\sin(1/x)$ on $[0,1]$. This guy is smooth on $(0,1]$. In particular, its derivative is $L^1$, and so $f$ is AC there.

However $f$ isn't even of $BV[0,1]$: consider the partition of $[0,1]$ given by

$$[0,\ \tfrac{1}{\pi}\tfrac{1}{n+1/2},\ \tfrac{1}{n\pi},\ \tfrac{1}{\pi}\tfrac{1}{(n-1)+1/2},\ \tfrac{1}{(n-1)\pi},\ \cdots,\ \tfrac{1}{\pi}\tfrac{1}{1+1/2},\ \tfrac{1}{\pi},\ 1]$$

$$= [y_0, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, x_{n+1}]$$

Then $|\sin(1/x_i)| = 1$ for $i \in [n]$, $\sin(1/y_i) = 0$ for all $i \in [0:n]$, and $\sin(x_{n+1}) = \sin(1)$. This yields a variation of

$$
\begin{aligned}
&|f(x_{n+1}) - f(y_n)| + \sum_{i=0}^{n-1} |f(y_i) - f(x_{i+1})| + |f(x_{i+1}) - f(y_{i+1})| \\
&= \sin(1) + \sum_{i=0}^{n-1} 2x_{i+1} \\
&= \sin(1) + \tfrac{2}{\pi} \sum_{j=2}^{n+1} \frac{1}{j - 1/2} \\
&\geqslant \sin(1) + \underbrace{\tfrac{2}{\pi} \sum_{j=2}^{n+1} \frac{1}{j}}_{\text{diverges as } n \to \infty}
\end{aligned}
$$

**Proposition 4.5.3.** **a.** *For $f \in L^1_m[a,b]$, let $F(x) = r + \int_a^x f dm$. Then $r = F(a)$ and $F$ is absolutely continuous on $[a,b]$.*

**b.** *If $G : [a,b] \to \mathbb{R}$ is absolutely continuous then $G$ is continuous of bounded variation. In particular, $G'$ exists $m$-a.e. and belongs to $L^1_m[a,b]$.*

**Proposition 4.5.4.** *Suppose $F$ is absolutely continuous on $[a,b]$. If $F' = 0$ $m$-a.e., then $F$ is constant.*

*Proof.* Let $c \in [a,b]$. Take a subset $A$ of full measure and it suffices to consider $F' = 0$ on $A$. Fix $\epsilon > 0$. Let $\delta$ be such that the absolute continuity of $F$ holds for $\epsilon/2$. Pick $x \in (a,c) \cap A$ and find a $y \in (a,c) \cap A$ such that

$$
|F(y) - F(x)| \leqslant \frac{\epsilon}{2} \frac{y - x}{c - a} \quad \text{iff} \quad \left| \frac{F(y) - F(x)}{y - x} \right| \leqslant \frac{\epsilon}{2(c - a)}. \tag{4.5.1}
$$

The intervals $[x,y]$ associated to such $y$'s form a Vitali covering of $A \cap (a,c)$. So by the VCL, there is a finite subcollection of disjoint intervals $\{[x_j, y_j]\}_{[n]}$ with $m\left(A \setminus \bigcup_{[n]} [x_j, y_j]\right) < \delta$. In particular, since $A$ has full measure, $m\left([a,b] \setminus \bigcup_{[n]} [x_j, y_j]\right) < \delta$.

Now, relabel if necessary, assume $x_j < x_{j+1}$ and set set $x_{n+1} = c$, $y_0 = a$. We deduce from above that $\sum_{j=1}^n |x_{j+1} - y_j| < \delta$. Absolute continuity of $F$ then gives

$$
\sum_{j=1}^n |F(x_{j+1}) - F(y_j)| < \epsilon/2.
$$

OTOH, from 4.5.1, we have

$$
\sum_{j=0}^{n+1} |F(y_j) - F(x_j)| \leqslant \frac{\epsilon}{2(c - a)} \underbrace{\sum_{j=0}^{n+1} (y_j - x_j)}_{= c - a} = \epsilon/2.
$$

Combine the two inequalities to conclude $|f(c) - f(a)| < \epsilon$.

∎

⋙ **Theorem 4.5.5** (FTC II). *$F : [a,b] \to \mathbb{R}$ is absolutely continuous on $[a,b]$ if and only if there is an element $f \in L^1_m[a,b]$ such that for all $x \in [a,b]$,*

$$F(x) - F(a) = \int_a^x f.$$

## 4.6 absolutely continuous functions

**Definition** (Luzin N property)**.** A function $f$ is said to satisfy the *Luzin N property* ($N$ stands for null?) if the image of any measure 0 set has measure 0 under $f$.

**Theorem 4.6.2** (Banach-Zaretsky)**.** *Suppose $F \in BV[a,b]$ is continuous. Then F is AC iff it satisfies the Luzin N property.*

**Theorem 4.6.3** (Integration by parts)**.** *Let $f,g \in L^1_m[a,b]$ with*

$$F(x) = r + \int_a^x f \quad \text{and} \quad G(x) = s + \int_a^x g.$$

*Then*

$$\int_a^b fG + \int_a^b gF = F(b)G(b) - F(a)G(a).$$

**Theorem 4.6.4** (measure theoretical characterization of 0-derivative)**.** *Suppose $F : [a,b] \to \mathbb{R}$ has derivative on $A \subset [a,b]$. Then $F' = 0$ m-a.e. iff $m(F(A)) = 0$.*

**Theorem 4.6.7** (Differentiability criteria for AC)**.** *Suppose $F : [a,b] \to \mathbb{R}$ is differentiable. If $F' \in L^1_m[a,b]$, then F is absolutely continuous on $[a,b]$.*

**Example (exercise 4.4 a,b).**

**Theorem 4.6.8.** *Let $f, f' \in L^1_m(\mathbb{R})$ and assume $f'$ exists everywhere. Then $\int f' = 0$.*

*Proof.* By Thm 4.6.7, $f$ is absolutely continuous on any finite interval. Take a sequence of smooth functions $\{F_n\} \subset C^\infty$ such that $\sup_n \|F_n'\|_\infty \leqslant M < \infty$, $0 \leqslant F_n \leqslant 1$, $F_n = 1$ on $[-n,n]$ and $F_n = 0$ off $[-(n+1), n+1]$. Each $F_n$ is absolutely continuous. We have $F_n f' \to f'$ pointwise and $|F_n f'| \leqslant |f'| \in L^1_m(\mathbb{R})$. Apply LDCT to get $\int f' = \lim \int F_n f'$. Since $F_n$, $f$ are both absolutely continuous on $[-(n+1), n+1]$, integrate by parts (Thm 4.6.3) to obtain

$$\int_{-(n+1)}^{n+1} F_n f' = \int_{-(n+1)}^{n+1} F_n' f.$$

Now, note that $F_n' f \to 0$ pointwise and $|F' f| \leqslant M|f| \in L^1_m$. So by LDCT $\int F_n' f \to 0$. Together, we conclude that $\int f' = 0$.

∎

## 4.7 Lipschitz functions

Everybody knows the definition of a Lipschitz function.

**Definition.** A function $f : [a,b] \to \mathbb{R}$ is said to satisfy a Lipschitz condition of order $\alpha > 0$ if there is a constant $C > 0$ such that for all $x, y \in [a,b]$

$$|f(x) - f(y)| \leqslant C|x - y|^{\alpha}.$$

In this case, we say $f$ is a Hölder function of order $\alpha$.

**Fun facts.** (exercise 4.36 on p.213)

*i.* Lipschitz functions are absolutely continuous.

*ii.* The Cantor function $C$ is Hölder of order $\log_3^2$.

*iii.* So $C$ is a continuous function of bounded variation that's not Lipschitz.

*iv.* For $\alpha > 1$, the set of Hölder functions of order $\alpha$ consists only of constants.

## 4.8 Summary

Suppose $f : [a,b] \to \mathbb{R}$ is absolutely continuous, then

- $f$ is of bounded variation;
- $f$ is the difference of two increasing absolutely continuous functions. (Jordan)
- $f$ satisfies the Luzin $N$ property.

**Example (UMD Analysis qual Jan 22 #1)** (composition of two AC functions needs not be AC)**.** Look at $\sqrt{x}$ and $x^2|\sin(x)|$.

## 4.9 Selected exercises

**Exercise 4.12.** Suppose $f : [a,b] \to \mathbb{R}$, $g : [c,d] \to \mathbb{R}$ are continuous of bounded variation and $g([c,d]) \subset [a,b]$. Show that $f \circ g$ iis continuous. Under what condition is $f \circ g$ is of $BV[c,d]$?

**Exercise 4.18a** (Fubini differentiation theorem)**.** Suppose $\{f_n\}$ is a sequence of monotone increasing functions $[a,b] \to \mathbb{R}$ for which

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

exists (converges) everywhere on $[a,b]$. Then $f' = \sum_{n=1}^{\infty} f_n'$ exists a.e.

**Exercise 4.18b.** Let $f \in BV[a,b]$. Then $[V(f)]' = |f'|$ a.e., where $V$ is the variation function.

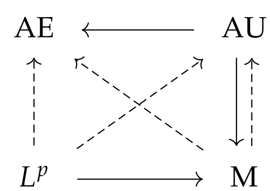
# 5 Modes of convergence

Just so my excellent diagram drawing skill obtained from algebra isn't wasted!

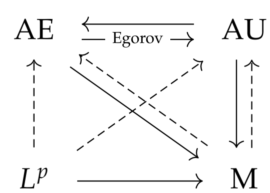
Notation:

- AE = almost everywhere convergence
  (pointwise convergence almost everywhere)
- AU = almost uniform convergence
  (given $\epsilon > 0$, there's $E \in \mathcal{A}$ with $\mu(E) < \epsilon$ such that the convergence on $E^{\sim}$ is uniform)
- $L^p$ = $L^p$ convergence ...
- M = convergence in measure
  (given $\epsilon > 0$, $\lim_{n\to\infty} \mu\{x : |f(x) - f_n(x)| > \epsilon\} = 0$)
- $\longrightarrow$ means "implies"
- $\dashrightarrow$ means "implies existence of a convergent subsequence"

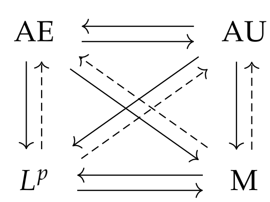
## 5.1 General measure space




## 5.2 Finite measure space




## 5.3 Dominated convergence



## 6 List of commonly used results

The class bugged Czaja for a review sheet before the final. He produced the following shortlist:

- Properties of Lebesgue outer measure
- Completeness of Lebesgue measure (2.2.13)
- Lebesgue measure on $R^d$ (2.3.9)
- Carathéodory theorem (2.4.1)
- Monotonicity and continuity of measures (2.4.3)
- 1st Borel-Cantelli lemma (2.4.4)
- 2nd Borel-Cantelli lemma (2.4.5)
- Kolmogorov law (2.4.6)
- Completeness theorem (2.4.8)
- Characterization of measurable functions (2.4.10)
- Properties of measurable functions (2.4.12, 2.5.1–2.5.5)
- Measures in terms of outer measures (2.4.19)
- Egorov theorem (2.5.7)
- Vitali-Luzin theorem (2.5.13)
- Properties of integral (3.2.6, 3.2.7)
- Special case of LDC (3.3.2)
- Fatou lemma (3.3.5)
- Levi-Lebesgue theorem (3.3.6)
- LDCT (3.3.7)
- General LDCT (3.3.8)
- Uniform integrability (3.3.9)
- Riesz-Lebesgue theorem (3.3.13)
- Riemann integrability (3.4.5)
- Riemann-Stjeltjes integrability (3.5.4)
- Lebesgue-Stjeltjes integrability (3.5.5)

- Fundamental applications of LDCT (3.6.1-3.6.3)
- Rieman-Lebesgue lemma (3.6.4)
- Fubini theorem (3.7.5)
- Tonelli theorem (3.7.7 and 3.7.8)
- Fubini-Tonelli theorem (3.7.9 and 3.7.12)
- Chebyshev inequality (Pr. 3.13)
- Jensen inequality (Pr. 3.43)
- Jordan decomposition theorem (4.1.2 and 4.1.4)
- Equivalence of bounded measures and increasing functions (4.1.9)
- Banach-Vitali theorem (4.2.7)
- Vitali covering lemma (4.3.1)
- Lebesgue differentiation theorem (4.3.2)
- Lebesgue density theorem
- FTC-I (4.4.3)
- FTC-II (4.5.5)
- Banach-Zaretsky theorem (4.6.2)
- Integration by parts (4.6.3)
- Everywhere differentiability and absolute continuity (4.6.7)
- Hölder and Minkowski inequalities (5.5.2)

## References


[BC09] John Benedetto and Wojciech Czaja. *Integration and modern analysis*. Birkhäuser, Boston, 2009.

[Coh80] Donald L. Cohn. *Measure theory*. Birkhäuser, 1980.

[Cza21] Wojciech Czaja. Math 630, Fall 2021.

[Hal74] Paul R. Halmos. *Measure theory*. Springer-Verlag, New York, 1974.

[Ham21] David H. Hamilton. Stat 410, Spring 2021.

[Kna16] Anthony W. Knapp. *Basic real analysis*. 2016.

[Rud15] Walter Rudin. *Real and complex analysis*. McGraw Hill Education, 2015.

[Tam19] Harry Tamvakis. Math 432, Fall 2019.