MATH 5215 - MISSOURI S&T
INTRODUCTION TO REAL ANALYSIS

JASON MURPHY

Contents

References.

- *Principles of Mathematical Analysis,* Rudin
- *Measure and Integral*, Wheeden and Zygmund

Course outline: We will cover as much of the following as possible:

- Rudin Chapter 7
- Wheeden–Zygmund Chapters 2–5 (omitting selected sections)
- Wheeden–Zygmund Chapter 8
- Wheeden–Zygmund Chapters 6–7 (time permitting)

These lecture notes contain the material from Rudin Chapter 7 and Wheeden–Zygmund Chapters 2–8, in the order presented in class.

Prerequisites. The prerequisite for this class is Math 4209, Advanced Calculus I. The catalog description for that course is as follows:

> Completeness of the set of real numbers, sequences and series of real numbers, limits, continuity and differentiability, uniform convergence, Taylor series, Heine-Borel theorem, Riemann integral, fundamental theorem of calculus, Cauchy-Riemann integral.

Familiarity with these topics will be assumed.

1. Sequences and series of functions

Reference: Rudin Chapter 7

1.1. **Pointwise convergence.**

Suppose $\{f_n\}_{n=1}^\infty$ is a sequence of real-valued functions defined on some subset $E \subset \mathbb{R}$. That is, for each n, we have

$$f_n : E \to \mathbb{R}.$$

Suppose that for each $x \in E$, the sequence $\{f_n(x)\}_{n=1}^\infty \subset \mathbb{R}$ converges. We can then define

$$f : E \to \mathbb{R} \quad \text{via} \quad f(x) := \lim_{n\to\infty} f_n(x) \quad \text{for each} \quad x \in E.$$

In this case, we say $\{f_n\}$ **converges (pointwise) on** E and that the function f is the **limit** of the sequence $\{f_n\}$. We may write $f_n \to f$ pointwise.

Remark 1.1. We focus on the case of real-valued functions on $E \subset \mathbb{R}$; however, one can also consider arbitrary metric spaces E and complex-valued functions.

Similarly, suppose the infinite sum

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

converges for each $x \in E$. Then we can define the function

$$f : E \to \mathbb{R} \quad \text{via} \quad f(x) := \sum_{n=1}^\infty f_n(x) \quad \text{for each} \quad x \in E.$$

In this case, we call f the **sum** of the series $\sum f_n$.

Question. Which properties of $\{f_n\}$ are 'inherited' by the limit functions introduced above?

For example, suppose $\{f_n\}$ is a sequence of *continuous* functions on E that converges pointwise to f. Is the limit f continuous on E? This is equivalent to asking if

$$\lim_{y\to x} f(y) = f(x) \quad \text{for all} \quad x \in E.$$

Recalling that $f(x) = \lim_{n\to\infty} f_n(x)$ and that each $\{f_n\}$ is continuous, this is equivalent to asking whether

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

for each $x \in E$. In particular, we are led to the question of the **interchange of limit operations**.

Let us work through several examples to see that in general, we cannot freely exchange the order of limits.

Example 1.1. Let

$$s_{m,n} = \frac{m}{m+n}, \quad m,n \in \mathbb{N}.$$

Then

$$\lim_{n\to\infty} \lim_{m\to\infty} s_{m,n} = \lim_{n\to\infty} 1 = 1,$$

while

$$\lim_{m\to\infty} \lim_{n\to\infty} s_{m,n} = \lim_{m\to\infty} 0 = 0.$$

Example 1.2. Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by

$$f_n(x) = \frac{x^2}{(1+x^2)^n}.$$

Each f_n is continuous. Now define

$$f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}.$$

Since $f_n(0) = 0$, we have $f(0) = 0$.

For $x \neq 0$, this is a geometric series that sums to $1 + x^2$ (cf. $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$).

Thus

$$f(x) = \begin{cases} 0 & x = 0 \\ 1 + x^2 & x \neq 0. \end{cases}$$

We conclude that a convergent series of continuous functions may be discontinuous.

Example 1.3. Define $f_m : \mathbb{R} \to \mathbb{R}$ by

$$f_m(x) = \lim_{n\to\infty} [\cos(m!\pi x)]^{2n}$$

for $m \in \mathbb{N}$. Note that

$$f_m(x) = \begin{cases} 1 & \text{if} \quad m!x \quad \text{is an integer} \\ 0 & \text{otherwise.} \end{cases}$$

Note that f_m is continuous except at countably many points.

Now define the limit function

$$f(x) = \lim_{m\to\infty} f_m(x).$$

We claim that

$$f(x) = \lim_{m\to\infty} \lim_{n\to\infty} [\cos(m!\pi x)]^{2n} = \begin{cases} 1 & \text{if} \quad x \quad \text{is rational,} \\ 0 & \text{if} \quad x \quad \text{is irrational.} \end{cases}$$

Indeed, if x is irrational then $m!x$ is never an integer, so that $f(x) = 0$. On the other hand, if $x = p/q \in \mathbb{Q}$ then $m!x$ is an integer whenever $m \geq q$, so that $f(x) = 1$.

The limit function is everywhere discontinuous and not Riemann integrable.

Example 1.4. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by
$$f_n(x) = \tfrac{\sin(nx)}{\sqrt{n}}$$
for $n \in \mathbb{N}$. Each f_n is differentiable on $\mathbb{R}$, with
$$f_n'(x) = \sqrt{n}\cos(nx).$$
The limit function f satisfies
$$f(x) = \lim_{n\to\infty} f_n(x) = 0 \quad \text{for every} \quad x \in \mathbb{R}.$$
In particular, f is also differentiable on $\mathbb{R}$, with $f' \equiv 0$ on $\mathbb{R}$.

In particular we deduce
$$\lim_{n\to\infty} \tfrac{d}{dx} f_n \neq \tfrac{d}{dx} \lim_{n\to\infty} f_n.$$
For example, $f_n'(0) = \sqrt{n} \to \infty$ as $n \to \infty$

Example 1.5. Let $f_n : [0,1] \to \mathbb{R}$ be defined by
$$f_n(x) = nx(1-x^2)^n.$$
The limit function $f : [0,1] \to \mathbb{R}$ satisfies
$$f(x) = \lim_{n\to\infty} f_n(x) = 0 \quad \text{for all} \quad x \in [0,1].$$
In particular,
$$\int_0^1 f(x)\,dx = 0.$$
On the other hand, a simple substitution (e.g. $u = 1 - x^2$) reveals
$$\int_0^1 f_n(x)\,dx = \tfrac{n}{2(n+1)} \to \tfrac{1}{2} \quad \text{as} \quad n \to \infty.$$
Thus
$$\int_0^1 \lim_{n\to\infty} f(x)\,dx \neq \lim_{n\to\infty} \int_0^1 f_n(x)\,dx.$$
In fact, considering the example $f_n(x) = n^2 x(1-x^2)^n$ shows that we may even have
$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = \infty \quad \text{while} \quad \int_0^1 \lim_{n\to\infty} f_n(x)\,dx = 0.$$

The takeaway of these examples is that one cannot always freely interchange limit operations.

At least, we have seen that *pointwise* convergence is too weak to allow us to make such interchanges.

1.2. **Uniform convergence.** We first revisit the definition of pointwise convergence: a sequence of functions $f_n : E \to \mathbb{R}$ converges pointwise to $f : E \to \mathbb{R}$ if

$$\begin{aligned}&\text{for all} \quad x \in E \quad \text{and for all} \quad \varepsilon > 0 \quad \text{there exists} \quad N = N(x, \varepsilon)\\ &\text{such that} \quad n \geq N \implies |f_n(x) - f(x)| < \varepsilon.\end{aligned}$$

We now introduce a stronger notion of convergence, namely *uniform* convergence.

Definition 1.2. Let $\{f_n\}$ be a sequence of functions $f_n : E \to \mathbb{R}$. We say f_n converges uniformly to $f : E \to \mathbb{R}$ if

$$\begin{aligned}&\text{for all} \quad \varepsilon > 0 \quad \text{there exists} \quad N = N(\varepsilon) \quad \text{such that for all} \quad x \in E,\\ &n \geq N \implies |f_n(x) - f(x)| < \varepsilon.\end{aligned}$$

We write $f_n \to f$ uniformly on E.

This convergence is uniform in the sense that a single choice of $N = N(\varepsilon)$ works *uniformly* over all choices of $x \in E$.

Uniform convergence is stronger than pointwise convergence (that is, uniform convergence implies pointwise convergence).

Example 1.6. Let $f_n : (0, 1) \to \mathbb{R}$ be given by $f_n(x) = x^n$.

Then $f_n \to 0$ pointwise on $(0, 1)$ but not uniformly.

However, $f_n \to 0$ uniformly on any interval of the form $(0, \delta)$ with $\delta < 1$.

Definition 1.3. Let $\{f_n\}$ be a sequence of functions $f_n : E \to \mathbb{R}$. A series of functions $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on E if the sequence of partial sums

$$s_n : E \to \mathbb{R} \quad \text{defined by} \quad s_n(x) = \sum_{i=1}^{n} f_i(x)$$

converges uniformly on E.

A sequence of functions that is 'uniformly Cauchy' converges uniformly.

Theorem 1.4 (Cauchy criterion for uniform convergence)**.** *A sequence of functions $f_n : E \to \mathbb{R}$ converges uniformly on E if and only if the following holds:*

$$\text{for every} \quad \varepsilon > 0 \quad \text{there exists} \quad N = N(\varepsilon) \quad \text{such that for all} \quad x \in E,$$
$$m, n \geq N \implies |f_n(x) - f_m(x)| < \varepsilon. \tag{1.1}$$

Proof. $\implies$: Suppose $\{f_n\}$ converges uniformly to f. Then for any $\varepsilon > 0$ there exists $N = N(\varepsilon)$ so that

$$|f_n(x) - f(x)| < \tfrac{1}{2}\varepsilon \quad \text{for any} \quad n \geq N, \quad x \in E.$$

Then for $n, m \geq N$ we have

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon$$

for any $x \in E$. This implies the uniform Cauchy condition.

$\Longleftarrow$: Suppose the Cauchy condition holds. In particular, for each $x \in E$, the sequence $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$.

Consequently, the sequence f_n converges *pointwise* to a function $f : E \to \mathbb{R}$.

Now let $\varepsilon > 0$ and choose N as in (1.1). Fix $n \geq N$ and $x \in E$. Then for any m, we may write

$$|f_n(x) - f(x)| \leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)|.$$

Taking the limsup as $m \to \infty$ and using (1.1) and pointwise convergence now yields

$$|f_n(x) - f(x)| < \varepsilon + 0.$$

This completes the proof. □

The following result follows from the definition of uniform convergence:

Theorem 1.5. *Suppose $f_n \to f$ pointwise on a set E. Define*

$$M_n := \sup_{x \in E} |f_n(x) - f(x)|.$$

Then $f_n \to f$ uniformly on E if and only if $\lim_{n\to\infty} M_n = 0$.

The following test for uniform convergence is due to Weierstrass.

Theorem 1.6. *Suppose $f_n : E \to \mathbb{R}$ is a sequence of functions satisfying*

$$\sup_{x \in E} |f_n(x)| \leq M_n$$

for some $\{M_n\} \subset \mathbb{R}$.

$$\text{If} \quad \sum_n M_n \quad \text{converges, then} \quad \sum_n f_n \quad \text{converges uniformly.}$$

Proof. Suppose $\sum_n M_n$ converges and let $\varepsilon > 0$. Then for $n \geq m$ sufficiently large, we have

$$\left|\sum_{i=m}^{n} f_i(x)\right| \leq \sum_{i=m}^{n} M_i < \varepsilon$$

for any $x \in E$. Using Theorem 1.5, this implies that $\sum f_n$ converges uniformly. □

Uniform limits inherit continuity. This will be a consequence of the following theorem.

Theorem 1.7. *Suppose $f_n \to f$ uniformly on an open set E. Suppose $x \in E$ and*

$$\lim_{y\to x} f_n(y) = A_n.$$

Then $\{A_n\}$ converges, with

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

That is,

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

Proof. Let $\varepsilon > 0$. By uniform convergence, there exists $N = N(\varepsilon)$ so that

$$n, m \geq N \quad \text{and} \quad y \in E \implies |f_n(y) - f_m(y)| < \varepsilon.$$

Taking the limsup as $y \to x$ yields

$$|A_n - A_m| < \varepsilon.$$

Thus $\{A_n\}$ is Cauchy, and hence convergent. Denote $A = \lim_{n\to\infty} A_n$.

Next, for any n and $y \in E$, we have

$$|f(y) - A| \leq |f(y) - f_n(y)| + |f_n(y) - A_n| + |A_n - A|.$$

Given $\varepsilon > 0$, we may choose n large enough that

$$|f(y) - f_n(y)| < \tfrac{1}{3}\varepsilon \quad \text{for all} \quad y \in E.$$

Choosing n possibly larger, we may also guarantee

$$|A_n - A| < \tfrac{1}{3}\varepsilon.$$

Finally, for this (fixed) n, we choose a neighborhood $U \ni x$ so that

$$|f_n(y) - A_n| < \tfrac{1}{3} \quad \text{for} \quad y \in U.$$

Continuing from above, we have

$$|f_n(y) - A| < \varepsilon \quad \text{for} \quad y \in U,$$

which completes the proof. □

This implies the following:

Theorem 1.8. *If $\{f_n\}$ is a sequence of continuous functions on E and $f_n \to f$ uniformly on E, then f is continuous on E.*

Proof. Let $x \in E$ be a limit point of E. Then by the previous theorem and continuity of the $\{f_n\}$, we have

$$\lim_{y\to x} f(y) = \lim_{y\to x}\lim_{n\to\infty} f_n(y) = \lim_{n\to\infty}\lim_{y\to x} f_n(y) = \lim_{n\to\infty} f_n(x) = f(x).$$

This implies f is continuous at each $x \in E$. □

Remark 1.9. The limit function may be continuous, even if the convergence is not uniform. See Example 1.5.

There is a case when the converse is true:

Theorem 1.10. *Let $K \subset \mathbb{R}$ be compact. Suppose*

- *$\{f_n\}$ are continuous functions on K,*
- *$f_n \to f$ pointwise on K, with f continuous,*
- *$f_n(x) \geq f_{n+1}(x)$ for $x \in K$ and $n \geq 1$.*

Then $f_n \to f$ uniformly.

Proof. The functions $g_n = f_n - f$ are continuous, $g_n \to 0$ pointwise, and $g_n \geq g_{n+1}$.

Let $\varepsilon > 0$ and define

$$K_n = \{x \in K : g_n(x) \geq \varepsilon\}.$$

As g_n is continuous, we have that K_n is closed and hence compact.

As $g_n \geq g_{n+1}$, we have $K_n \supset K_{n+1}$.

Now consider any $x \in K$. Since $g_n(x) \to 0$, we have $x \notin K_n$ for n large enough.

As x was arbitrary, we conclude that $\cap_{n=1}^{\infty} K_n = \emptyset$.

As $K_n \supset K_{n+1}$, this implies $K_N = \emptyset$ for some N (and hence for all $n \geq N$).

This implies $0 \leq g_n(x) < \varepsilon$ for all $x \in K$ and $n \geq N$.

This implies $g_n \to 0$ uniformly, which completes the proof. □

Compactness is necessary. Indeed, $f_n(x) = \frac{1}{nx+1}$ converges to zero monotonically for $x \in (0,1)$, but not uniformly.

We next introduce the space $C(X)$.

Definition 1.11. Let $X \subset \mathbb{R}$. We let $C(X)$ denote the set of all real-valued, continuous, bounded functions on X.

For $f \in C(X)$, we define the **supremum norm** by

$$\|f\| = \sup_{x\in X} |f(x)|.$$

Note that $\|f\| < \infty$ for all $f \in C(X)$.

The quantity $\|\cdot\|$ satisfies the definitions of a norm, namely:

- $\|f\| = 0$ implies $f \equiv 0$,
- $\|f + g\| \leq \|f\| + \|g\|$,

- $\|cf\| = |c|\|f\|$ for $c \in \mathbb{R}$.

Furthermore, $(f,g) \mapsto \|f-g\|$ defines a **metric** on $C(X)$.

Remark 1.12.

(i) This definition makes sense for an arbitrary metric space X (and complex-valued functions).

(ii) If X is compact, then the boundedness assumption is redundant.

(iii) Theorem 1.5 may be restated as follows: $f_n \to f$ uniformly on X if and only if $f_n \to f$ in the metric of $C(X)$.

We close this section with the following result:

Theorem 1.13. *The space $C(X)$ is a complete metric space.*

Proof. Let $\{f_n\}$ be a Cauchy sequence in $C(X)$. Then for any $\varepsilon > 0$, there exists N such that $\|f_n - f_m\| < \varepsilon$ for all $n, m > N$.

Then by Theorem 1.4, f_n converges uniformly to some $f : X \to \mathbb{R}$.

Moreover, by Theorem 1.8, f is continuous.

Finally, since each f_n is bounded and there exists n such that

$$|f_n(x) - f(x)| < 1 \quad \text{for all} \quad x \in X,$$

we deduce f is bounded. Thus $f \in C(X)$ and $\|f_n - f\| \to 0$ as $n \to \infty$. □

1.3. **Uniform convergence and integration/differentiation.** We recall the definition of Riemann integration, including upper and lower sums (with respect to a given partition), and upper and lower integrals (denoted by $\overline{\int}$ and $\underline{\int}$).

Theorem 1.14. *Suppose f_n are Riemann integrable functions on an interval $[a,b]$ and $f_n \to f$ uniformly on $[a,b]$. Then f is Riemann integrable and*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Proof. Define

$$\varepsilon_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|.$$

In particular,

$$f_n - \varepsilon_n \le f \le f_n + \varepsilon_n,$$

so that

$$\int_a^b [f_n(x) - \varepsilon_n]\,dx \le \underline{\int} f(x)\,dx \le \overline{\int} f(x)\,dx \le \int_a^b [f_n(x) + \varepsilon_n]\,dx.$$

In particular,

$$0 \le \overline{\int} f(x)\,dx - \underline{\int} f(x)\,dx \le 2\varepsilon_n[b-a].$$

Uniform convergence implies $\varepsilon_n \to 0$ as $n \to \infty$, and hence the upper and lower integrals of f are equal.

Therefore f is Riemann integrable, and

$$\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| \leq \varepsilon_n[b-a] \to 0 \quad \text{as} \quad n \to \infty.$$

This completes the proof. □

Corollary 1.15. *Suppose f_n are Riemann integrable on $[a,b]$ and the series*

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

converges uniformly on $[a,b]$. Then

$$\int_a^b f(x)\,dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,dx.$$

That is, the series may be integrated term-by-term.

We turn to the question of differentiation.

Theorem 1.16. *Let f_n be differentiable functions on an interval $[a,b]$. Suppose $f_n(x_0)$ converges for some $x_0 \in [a,b]$. Suppose further that f_n' converges uniformly on $[a,b]$. Then f_n converges to a function f on $[a,b]$, and $f_n' \to f'$.*

Proof. Let $\varepsilon > 0$ and choose N so that

$$m,n \geq N \implies |f_n(x_0) - f_m(x_0)| < \tfrac{1}{2}\varepsilon \quad \text{and} \quad |f_n'(t) - f_m'(t)| < \tfrac{1}{2(b-a)}\varepsilon$$

for all $t \in [a,b]$.

By the mean value theorem (applied to $f_n - f_m$),

$$|f_n(x) - f_m(x) - [f_n(t) - f_m(t)]| < \tfrac{\varepsilon}{2(b-a)}|x-t| < \tfrac{1}{2}\varepsilon \tag{1.2}$$

for any $x,t \in [a,b]$ and $n,m \geq N$.

Thus, by the triangle inequality,

$$\begin{aligned} &|f_n(x) - f_m(x)| \\ &\quad \leq |f_n(x) - f_m(x) - [f_n(x_0) - f_m(x_0)]| + |f_n(x_0) - f_m(x_0)| < \varepsilon. \end{aligned}$$

for any $x \in [a,b]$ and $n,m \geq N$.

Therefore $f_n \to f$ converges uniformly on $[a,b]$ for some function f.

We now show $f_n' \to f'$. Fix $x \in [a,b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}$$

for $t \in [a,b]\backslash\{x\}$. We have

$$\lim_{t \to x} \phi_n(t) = f_n'(x).$$

By (1.2),
$$|\phi_n(t)-\phi_m(t)| < \tfrac{1}{2(b-a)}\varepsilon \quad \text{for} \quad n,m\geq N,$$
which shows that $\{\phi_n\}$ converges uniformly for any $t\neq x$.

Since $f_n\to f$, we see that the (uniform) limit of $\phi_n(t)$ must be
$$\frac{f(t)-f(x)}{t-x}.$$

We now apply Theorem 1.7 to $\{\phi_n\}$ to deduce
$$f'(x)=\lim_{t\to x}\tfrac{f(t)-f(x)}{t-x}=\lim_{t\to x}\lim_{n\to\infty}\phi_n(t)=\lim_{n\to\infty}\lim_{t\to x}\phi_n(t)=\lim_{n\to\infty}f_n'(x),$$
as desired. □

If one assumes the f_n' are continuous, there is a much simpler proof using the fundamental theorem of calculus. [See homework.]

We close this section with the following interesting construction.

Proposition 1.17. *There exists a real-valued continuous function that is nowhere differentiable.*

Proof. Let $\phi(x)=|x|$ for $x\in[-1,1]$. Extend ϕ to $x\in\mathbb{R}$ by imposing
$$\phi(x+2)=\phi(x).$$

For all $s,t\in\mathbb{R}$, we have $|\phi(s)-\phi(t)|\leq|s-t|$, which shows that ϕ is continuous.

Let
$$f(x)=\sum_{n=0}^{\infty}(\tfrac{3}{4})^n\phi(4^n x).$$
Using $0\leq\phi\leq 1$, Theorem 1.6 implies that the series converges uniformly on $\mathbb{R}$, and hence f is continuous on $\mathbb{R}$.

Let $x\in\mathbb{R}$ and for $m\in\mathbb{N}$ define
$$\delta_m=\pm\tfrac{1}{2}\cdot 4^{-m},$$
where the sign is chosen to that
$$(4^m x, 4^m(x+\delta_m))\cap\mathbb{Z}=\emptyset. \tag{1.3}$$
(That this is possible follows from the fact that $4^m|\delta_m|=\frac{1}{2}$).

Next define
$$\gamma_n=\frac{\phi(4^n(x+\delta_m))-\phi(4^n x)}{\delta_m}.$$
When $n>m$, $4^n\delta_m$ is an even integer, and hence $\gamma_n=0$.

On the other hand, when $0\leq n\leq m$, we have $|\gamma_n|\leq 4^n$.

Finally, note that (1.3) implies
$$|\gamma_m|=\left|\frac{\phi(4^m x\pm\frac{1}{2})-\phi(4^m x)}{\pm\frac{1}{2}4^{-m}}\right|=4^m.$$

Thus

$$\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{m}(\tfrac{3}{4})^n\gamma_n\right| \geq 3^m - \sum_{n=0}^{m-1}3^n = \tfrac{1}{2}(3^m+1).$$

Noting that $\delta_m \to 0$ but $3^m \to \infty$, we deduce that f is not differentiable at x. □

1.4. **Equicontinuous families of functions.**

Definition 1.18. Let $\{f_n\}$ be a sequence of functions on $E \subset \mathbb{R}$.

We call $\{f_n\}$ **pointwise bounded** if $\{f_n(x)\}$ is a bounded sequence for each $x \in E$, that is, if there exists $\phi : E \to \mathbb{R}$ so that

$$|f_n(x)| \leq \phi(x) \quad \text{for all} \quad x \in E \quad \text{and} \quad n \geq 1.$$

We call $\{f_n\}$ **uniformly bounded** if there exists M so that

$$|f_n(x)| \leq M \quad \text{for all} \quad x \in E \quad \text{and} \quad n \geq 1.$$

Theorem 1.19. *If $\{f_n\}$ is a pointwise bounded sequence on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ that converges on E.*

Proof. Write $E = \{x_j\}_{j=1}^{\infty}$.

As $\{f_n(x_1)\}$ is bounded, there exists a subsequence denoted $\{f_{1,k}\}$ so that $f_{1,k}(x_1)$ converges.

Similarly, the sequence $\{f_{1,k}(x_2)\}$ is bounded, and hence there exists a further subsequence denoted $\{f_{2,k}\}$ so that $\{f_{2,k}(x_j)\}$ converges for $j = 1, 2$.

Proceeding in this way yields subsequences $\{f_{n,k}\}$ such that $\{f_{n,k}(x_j)\}$ converges for each $j = 1, 2, \ldots, n$.

Now consider the subsequence $\{f_{k,k}\}$. This sequence satisfies that $\{f_{k,k}(x_j)\}$ converges for each j. □

Definition 1.20. A family $\mathcal{F}$ of functions f defined on a set $E \subset \mathbb{R}$ is **equicontinuous** on E if

$$\text{for all} \quad \varepsilon > 0 \quad \text{there exists} \quad \delta > 0 \quad \text{such that}$$
$$|x-y| < \delta \implies |f(x)-f(y)| < \varepsilon \quad \text{for all} \quad f \in \mathcal{F}.$$

Remark 1.21. Every element of an equicontinuous family is uniformly continuous.

Theorem 1.22. *If $K \subset \mathbb{R}$ is compact, $\{f_n\} \subset C(K)$, and $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.*

Proof. Let $\varepsilon > 0$. By uniform convergence, there exists N so that

$$n \geq N \implies \|f_n - f_N\| < \tfrac{1}{3}\varepsilon.$$

As continuous functions on compact sets are uniformly continuous, there exists $\delta > 0$ so that

$$|x-y|<\delta \implies |f_i(x)-f_i(y)|<\tfrac{1}{3}\varepsilon \quad \text{for all} \quad 1\le i\le N.$$

This gives equicontinuity for $\{f_i\}_{i=1}^N$, while if $n>N$ and $|x-y|<\delta$, then

$$|f_n(x)-f_n(y)|\le |f_n(x)-f_N(x)|+|f_N(x)-f_N(y)|+|f_N(y)-f_n(y)|<\varepsilon.$$

The result follows. □

The following result is known as the **Arzelá–Ascoli theorem**.

Theorem 1.23. *Let $K\subset\mathbb{R}$ be compact and $\{f_n\}\subset C(K)$. If $\{f_n\}$ is pointwise bounded and equicontinuous on K, then:*

- *$\{f_n\}$ is uniformly bounded on K,*
- *$\{f_n\}$ has a uniformly convergent subsequence.*

Proof. Let $\varepsilon>0$ and choose $\delta>0$ so that

$$|x-y|<\delta \implies |f_n(x)-f_n(y)|<\tfrac{1}{3}\varepsilon \quad \text{for all} \quad n.$$

By compactness of K, there exist $\{p_i\}_{i=1}^r\subset K$ so that

$$K\subset \cup_{i=1}^r (p_i-\delta,p_i+\delta).$$

As $\{f_n\}$ is pointwise bounded, for each i there exist M_i so that $|f_n(p_i)|<M_i$ for all n.

Writing $M=\max\{M_i\}$, we deduce

$$|f(x)|<M+\varepsilon \quad \text{for all} \quad x\in K,$$

giving uniform boundedness.

Next, let E be a countable dense subset of K. Then by Theorem 1.19, $\{f_n\}$ has a subsequence (which we also denote by f_n) such that $f_n(x)$ converges for every $x\in E$.

We will show (the subsequence) f_n converges uniformly on K.

Let $\varepsilon>0$ and pick $\delta>0$ as above. For $x\in E$, let

$$V(x,\delta)=\{y\in K: |x-y|<\delta\}.$$

As E is dense in K and K is compact, there exist $\{x_i\}_{i=1}^m\subset E$ so that

$$K\subset \cup_{i=1}^m V(x_i,\delta).$$

As $\{f_n(x)\}$ converges for $x\in E$, there exists N so that

$$|f_i(x_k)-f_j(x_k)|<\tfrac{1}{3}\varepsilon \quad \text{for} \quad i,j\ge N \quad \text{and} \quad 1\le k\le m.$$

Now let $x\in K$. Then $x\in V(x_k,\delta)$ for some k, so that for $i,j\ge N$, we have

$$|f_i(x)-f_j(x)|\le |f_i(x)-f_i(x_k)|+|f_i(x_k)-f_j(x_k)|+|f_j(x_k)-f_j(x)|<\varepsilon,$$

which completes the proof. □

1.5. **The Stone–Weierstrass theorem.** For this section we will consider complex-valued functions.

We begin with the following approximation theorem.

Theorem 1.24 (Weierstrass theorem)**.** *Let $f : [a, b] \to \mathbb{C}$ be continuous. Then there exists a sequence of polynomials so that $P_n \to f$ uniformly on $[a, b]$.*

Remark 1.25. This result holds for real-valued functions (with real polynomials) as well.

Proof. Without loss of generality, take $[a, b] = [0, 1]$. We may also assume $f(0) = f(1) = 0$, for then we may apply the result to

$$g(x) = f(x) - f(0) - x[f(1) - f(0)].$$

We set $f \equiv 0$ for $x \notin [0, 1]$, making f uniformly continuous on $\mathbb{R}$.

For $n \geq 1$, define

$$Q_n(x) = c_n(1 - x^2)^n, \quad \text{where} \quad c_n = \frac{1}{\int_{-1}^{1}(1 - x^2)^n\,dx},$$

so that

$$\int_{-1}^{1} Q_n(x)\,dx \equiv 1.$$

Note that

$$\begin{aligned}\int_{-1}^{1}(1 - x^2)^n\,dx &\geq 2\int_0^{\frac{1}{\sqrt{n}}}(1 - x^2)^n\,dx \\ &\geq 2\int_0^{\frac{1}{\sqrt{n}}}(1 - nx^2)\,dx = \tfrac{4}{3\sqrt{n}},\end{aligned}$$

which implies $c_n < \sqrt{n}$. Here we used $(1 - x^2)^n \geq 1 - nx^2$ on $(0, 1)$.

We deduce that for $\delta > 0$ and $|\delta| < |x| \leq 1$,

$$Q_n(x) \leq \sqrt{n}(1 - \delta^2)^n,$$

so that $Q_n \to 0$ uniformly for $\delta \leq |x| \leq 1$.

Now define

$$P_n(x) = \int_{-1}^{1} f(x + t)Q_n(t)\,dt, \quad x \in [0, 1].$$

In particular, since $f = 0$ outside $[0, 1]$,

$$P_n(x) = \int_{-x}^{1-x} f(x + t)Q_n(t)\,dt = \int_0^1 f(t)Q_n(t - x)\,dt.$$

which shows that P_n is a polynomial in x. (Furthermore, $P_n \in \mathbb{R}$ if $f \in \mathbb{R}$.)

We now claim that $P_n \to f$ uniformly. To this end, we let $\varepsilon > 0$ and choose $\delta > 0$ so that

$$|x - y| < \delta \implies |f(x) - f(y)| < \tfrac{1}{2}\varepsilon.$$

Let $M = \sup|f|$. Using $Q_n \geq 0$ and $\int Q_n = 1$, we have for $x \in [0,1]$:

$$\begin{aligned}
|P_n(x) - f(x)| &= \left|\int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)\,dt\right| \\
&\leq \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)\,dx \\
&\leq 2M\int_{-1}^{-\delta} Q_n(t)\,dt + 2M\int_{\delta}^{1} Q_n(t)\,dt + \tfrac{1}{2}\varepsilon\int_{-\delta}^{\delta} Q_n(t)\,dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \tfrac{\varepsilon}{2},
\end{aligned}$$

so that

$$|P_n(x) - f(x)| < \varepsilon \quad \text{for all} \quad x \in [-1,1] \quad \text{and} \quad n \quad \text{large enough.}$$

This completes the proof. □

Corollary 1.26. *For any $a > 0$, there exists a sequence of real polynomials P_n so that $P_n(0) = 0$ and $P_n(x) \to |x|$ uniformly on $[-a,a]$.*

Proof. Let P_n^* be the polynomials given by Theorem 1.24, and set $P_n(x) = P_n^*(x) - P_n^*(0)$. □

This approximation theorem can be generalized.

Definition 1.27. A family A of complex functions on a set E is an **algebra** if for all $f, g \in A$ and $c \in \mathbb{C}$,

- $f + g \in A$,
- $fg \in A$,
- $cf \in A$.

We can also consider algebras of real-valued functions (in which we only consider $c \in \mathbb{R}$).

If A is closed under uniform convergence, then we call A **uniformly closed**.

The **uniform closure** of A is the set of all uniform limits of sequences in A.

The Weierstrass theorem states that the set of continuous functions on $[a,b]$ is the uniform closure of the algebra of polynomials on $[a,b]$.

The following is left as an exercise:

Theorem 1.28. *Let B be the uniform closure of an algebra A of bounded functions. Then B is a uniformly closed algebra.*

Definition 1.29. A family of functions A defined on a set E is said to **separate points** if for every $x_1 \neq x_2 \in E$ there exists $f \in A$ so that $f(x_1) \neq f(x_2)$.

If for each $x \in E$ there exists $g \in A$ so that $g(x) \neq 0$, we say A **vanishes at no point of E**.

For example, the algebra of polynomials has these properties on $\mathbb{R}$. However, the algebra of even polynomials on $[-1,1]$ does not separate points (since $f(x) = f(-x)$ for every f in this algebra).

The following is also left as an exercise:

Theorem 1.30. *Suppose A is an algebra of functions on E that separates points and vanishes at no point of E. For any $x_1 \neq x_2 \in E$ and $c_1, c_2 \in \mathbb{C}$, there exists $f \in A$ so that*

$$f(x_1) = c_1 \quad \textit{and} \quad f(x_2) = c_2.$$

If A is real, then this holds for $c_1, c_2 \in \mathbb{R}$.

We can now state the generalization of Weierstrass's theorem. It gives conditions for an algebra of functions on a compact set K to be dense in $C(K)$.

Theorem 1.31 (Stone–Weierstrass, real version)**.** *Let A be an algebra of real-valued continuous functions on a compact set K. If A separates points on K and vanishes at no point of K, then the uniform closure B of A consists of all real continuous functions on K.*

Proof. The proof proceeds in four steps.

1. If $f \in B$ then $|f| \in B$.

Let $a = \sup_{x\in K} |f(x)|$ and $\varepsilon > 0$. By the corollary above, there exist $\{c_i\}_{i=1}^n$ so that

$$\left| \sum_{i=1}^n c_i y^i - |y| \right| < \varepsilon \quad \text{for} \quad y \in [-a,a].$$

As B is an algebra, the function

$$g = \sum_{i=1}^n c_i f^i$$

belongs to B. Thus

$$\big| |g(x) - |f(x)| \big| < \varepsilon \quad \text{for} \quad x \in K.$$

This implies that we may find $g_n \in B$ so that $g_n \to |f|$ uniformly. As B is uniformly closed, this implies that $|f| \in B$.

2. If $f \in B$ and $g \in B$, then $\max\{f,g\}$ and $\min\{f,g\}$ belong to B.

This follows from Step 1 and the fact that

$$\max\{f,g\} = \tfrac{1}{2}(f+g) + \tfrac{1}{2}|f-g|, \quad \min\{f,g\} = \tfrac{1}{2}(f+g) - \tfrac{1}{2}|f-g|.$$

By iterating this, we can extend Step 2 to any finite collection of functions in B.

3. For $f \in C(K)$, $x \in K$, and $\varepsilon > 0$, there exists $g_x \in B$ so that

$$g_x(x) = f(x) \quad \text{and} \quad g_x(t) > f(t) - \varepsilon \quad \text{for} \quad t \in K.$$

As $A \subset B$ and A satisfies the hypotheses of the preceding theorem, so does B. Thus for $y \in K$ we may find $h_y \in B$ so that

$$h_y(x) = f(x) \quad \text{and} \quad h_y(y) = f(y).$$

By continuity of h_y, there exists open $U_y \ni y$ so that

$$h_y(t) > f(t) - \varepsilon \quad \text{for} \quad t \in U_y.$$

As K is compact, there exists $\{y_1, \ldots, y_n\}$ so that

$$K \subset \cup_{j=1}^n U_{y_j}.$$

Now the function $g_x = \max\{h_{y_j}\} \in B$ has the desired properties.

4. For $f \in C(K)$ and $\varepsilon > 0$, there exists $h \in B$ so that $\|h - f\| < \varepsilon$.

This implies that we may find $h_n \in B$ so that $h_n \to f$ uniformly. As B is uniformly closed, this implies the theorem.

Let $\varepsilon > 0$ and for each $x \in K$ define $g_x \in B$ as in Step 3. By continuity, there exist open sets $U_x \ni x$ so that

$$g_x(t) < f(t) + \varepsilon \quad \text{for} \quad t \in U_x.$$

By compactness of K, there exists $\{x_i\}_{i=1}^m$ so that

$$K \subset \cup_{i=1}^m U_{x_i}.$$

Now set $h = \min\{g_{x_i}\} \in B$. Then by Step 3, we have $h(t) > f(t) - \varepsilon$ on K, while by construction $h(t) < f(t) + \varepsilon$ on K. This implies the result. □

The analogue of Theorem 1.31 for complex-valued functions requires an additional assumption, namely that the algebra is **self-adjoint**. This means that the algebra is closed under complex conjugation.

We leave the complex version of Theorem 1.31 as an exercise. It can be deduced from Theorem 1.31.

Theorem 1.32 (Stone–Weierstrass, complex version)**.** *Let A be a self-adjoint algebra of complex-valued continuous functions on a compact set K. If A separates points on K and vanishes at no point of K, then the uniform closure B of A consists of all complex continuous functions on K.*

2. Functions of bounded variation

Reference: Wheeden–Zygmund Chapter 2

2.1. Functions of bounded variation.

Definition 2.1. Let $f : [a,b] \to \mathbb{R}$, and let

$$\Gamma = \{x_0, \ldots, x_m\}$$

be a partition of $[a,b]$. Define

$$S_\Gamma = S_\Gamma[f; a, b] = \sum_{i=1}^{m} |f(x_i) - f(x_{i-1})|.$$

The **variation of f over** $[a,b]$ is defined by

$$V = V[f; a, b] = \sup_\Gamma S_\Gamma.$$

As $0 \le S_\Gamma < \infty$, we have $V \in [0, \infty]$. If $V < \infty$, we say f is of **bounded variation**. We may write $f \in BV([a,b])$ and $V = \|f\|_{BV}$. Otherwise, we say f is of **unbounded variation**.

If we simply write S_Γ, V, etc., then we assume that we are working with some real-valued function f defined on an interval $[a,b]$.

Example 2.1. If f is monotone on $[a,b]$, then $S_\Gamma \equiv |f(b) - f(a)|$ and hence $V = |f(b) - f(a)|$.

Example 2.2. If we can write $[a,b] = \cup_{i=1}^{k} [a_i, a_{i+1}]$ with f monotone on each subinterval, then

$$V = \sum_{i=1}^{k} |f(a_{i+1}) - f(a_i)|$$

(see below).

Example 2.3. Let $f(x) = 0$ when $x \neq 0$ and $f(0) = 1$. Let $[a,b]$ be any interval with $0 \in (a,b)$. Then $S_\Gamma \in \{0, 2\}$, depending on whether or not $0 \in \Gamma$. Thus $V[a,b] = 2$.

If $\Gamma = \{x_0, \ldots, x_m\}$ is a partition of $[a,b]$, then we define the **norm** of Γ to be

$$|\Gamma| = \max_i [x_i - x_{i-1}].$$

If f is continuous on $[a,b]$ and $|\Gamma_j| \to 0$, then we will see that

$$V = \lim_{j \to \infty} S_{\Gamma_j}.$$

The previous example shows that this may fail if there is even a single discontinuity.

Example 2.4. Let f be the Dirichlet function: $f(x) = 1$ for $x \in \mathbb{Q}$ and $f(x) = 0$ for $x \in \mathbb{R}\backslash\mathbb{Q}$. Then $V[a,b] = \infty$ for any interval.

Example 2.5. Continuity does not imply bounded variation:

Let $\{a_j\}$ and $\{d_j\}$ be decreasing sequences in $(0,1]$ with $a_1 = 1$, $a_j, d_j \to 0$, and $\sum d_j = \infty$.

Construct f as follows. On each $[a_{j+1}, a_j]$, the graph of f consists of the sides of the isosceles triangle with base $[a_{j+1}, a_j]$ and height d_j.

Then $f(a_j) = 0$ and $f(m_j) = d_j$, where m_j is the midpoint of a_{j+1} and a_j.

Setting $f(0) = 0$, we have that f is continuous on $[0,1]$.

Let Γ_k be the partition defined by $0, \{a_j\}_{j=1}^{k+1}$, and $\{m_j\}_{j=1}^{k}$. Then $S_{\Gamma_k} = 2\sum_{j=1}^{k} d_j$, whence $V[f;0,1] = \infty$.

Example 2.6. A function $f : [a,b] \to \mathbb{R}$ is **Lipschitz** if there exists $C > 0$ such that

$$|f(x) - f(y)| \leq C|x - y|, \quad x, y \in [a,b].$$

Lipschitz implies bounded variation, with $V[f;a,b] \leq C(b-a)$.

If f has a continuous derivative on $[a,b]$, it is Lipschitz by the mean value theorem (C can be taken to be the maximum of f').

The following theorem is left as an exercise:

Theorem 2.2.

- *If f is of bounded variation on $[a,b]$, then f is bounded on $[a,b]$.*
- *The family of bounded variation functions on $[a,b]$ is an algebra.*
- *If f and g are of bounded variation and there exists $\varepsilon > 0$ so that $|g| > \varepsilon$, then f/g is of bounded variation.*

Definition 2.3. Let Γ be a partition. Another partition $\bar{\Gamma}$ is a **refinement** of Γ if $\Gamma \subset \bar{\Gamma}$.

Note that if $\bar{\Gamma}$ is a refinement of Γ, then (by the triangle inequality) $S_\Gamma \leq S_{\bar{\Gamma}}$.

Theorem 2.4.

- *If $[a',b'] \subset [a,b]$, then $V[a',b'] \leq V[a,b]$.*

- *Variation is additive on adjacent intervals:* $V[a,b] = V[a,c] + V[c,b]$ *whenever* $a < b < c$.

Proof. If Γ' is any partition of $[a',b']$, then $\Gamma = \Gamma' \cup \{a,b\}$ is a partition of $[a,b]$ and

$$S_{\Gamma'}[a',b'] \le S_\Gamma[a,b] \le V[a,b].$$

This implies $V[a',b'] \le V[a,b]$.

Write $I = [a,b]$, $I_1 = [a,c]$, and $I_2 = [b,c]$. Let $V = V[a,b]$, $V_j = V[I_j]$.

If Γ_1, Γ_2 are partitions of I_1, I_2, then $\Gamma = \Gamma_1 \cup \Gamma_2$ is a partition of I, with

$$S_\Gamma[I] = S_{\Gamma_1}[I_1] + S_{\Gamma_2}[I_2] \le V.$$

Taking the supremum over Γ_1 and Γ_2 yields $V_1 + V_2 \le V$.

On the other hand, suppose Γ is a partition of I. Let $\bar\Gamma = \Gamma \cup \{c\}$. Then

$$S_\Gamma[I] \le S_{\bar\Gamma}[I].$$

Note $\bar\Gamma$ splits into partitions Γ_1 of I_1 and Γ_2 of I_2 (e.g. take $\Gamma_1 = \bar\Gamma \cap I_1$). Thus

$$S_\Gamma[I] \le S_{\bar\Gamma}[I] = S_{\Gamma_1}[I_1] + S_{\Gamma_2}[I_2] \le V_1 + V_2.$$

Taking the supremum over all partitions Γ yields $V \le V_1 + V_2$. Thus $V = V_1 + V_2$. □

Given $x \in \mathbb{R}$, let

$$x^+ = \begin{cases} x & x > 0 \\ 0 & x \le 0 \end{cases} \quad \text{and} \quad x^- = \begin{cases} 0 & x > 0 \\ -x & x \le 0. \end{cases}$$

These are called the **positive** and **negative** parts of x. They satisfy

$$x^+, x^- \ge 0, \quad |x| = x^+ + x^-, \quad x = x^+ - x^-.$$

For a function f and a partition $\Gamma = \{x_i\}_{i=0}^m$ of $[a,b]$, let

$$P_\Gamma = P_\Gamma[f;a,b] = \sum_{i=1}^m [f(x_i) - f(x_{i-1})]^+,$$

$$N_\Gamma = N_\Gamma[f;a,b] = \sum_{i=1}^m [f(x_i) - f(x_{i-1})]^-.$$

Thus $P_\Gamma, N_\Gamma \ge 0$, with

$$S_\Gamma = P_\Gamma + N_\Gamma, \quad P_\Gamma - N_\Gamma = f(b) - f(a).$$

The **positive variation** and **negative variation** are defined by

$$P = P[f;a,b] = \sup_\Gamma P_\Gamma, \quad N = N[f;a,b] = \sup_\Gamma N_\Gamma.$$

Then $P, N \in [0,\infty]$.

Theorem 2.5. *If any one of P, N, or V are finite, then all three are finite, with*

$$P + N = V \quad \text{and} \quad P - N = f(b) - f(a).$$

Equivalently,

$$P = \tfrac{1}{2}[V + f(b) - f(a)], \quad N = \tfrac{1}{2}[V - (f(b) - f(a))].$$

Proof. As $P_\Gamma + N_\Gamma = S_\Gamma$ for any partition Γ, we have

$$P_\Gamma + N_\Gamma \leq V.$$

Because $P_\Gamma, N_\Gamma \geq 0$, this implies $P \leq V$ and $N \leq V$. Thus, finiteness of V implies finiteness of P, N.

Using $P_\Gamma + N_\Gamma = S_\Gamma$ again, we see that $S_\Gamma \leq P+N$ and hence $V \leq P+N$.

On the other hand, since $P_\Gamma - N_\Gamma = f(b) - f(a)$, we see that finiteness of P or N implies finiteness of the other, and hence finiteness of V. This completes the first part of the theorem.

Now assume $P_{\Gamma_k} \to P$. Then $N_{\Gamma_k} \to N$ (since $P_\Gamma - N_\Gamma$ is constant for any partition). Sending $k \to \infty$, we deduce

$$P - N = f(b) - f(a), \quad P + N \leq V.$$

Recalling $V \leq P + N$, the theorem follows. □

Corollary 2.6 (Jordan's theorem)**.** *A function is of bounded variation on $[a, b]$ if and only if it can be written as the difference of two bounded increasing functions on $[a, b]$.*

Proof. $\Longleftarrow$ Bounded monotone functions are of bounded variation, and differences of bounded variation functions are of bounded variation.

$\Longrightarrow$ Suppose f is of bounded variation on $[a, b]$. Then f is of bounded variation on every $[a, x]$ for $x \in [a, b]$.

Let $P(x)$ and $N(x)$ denote the positive and negative variations of f on $[a, x]$.

Noting that P, N also increase on increasing intervals (like V), we have that P, N are bounded and increasing on $[a, b]$. By the previous theorem,

$$f(x) = [P(x) + f(a)] - N(x) \quad \text{for} \quad x \in [a, b].$$

The corollary follows. □

We can rephrase the corollary by saying that f is the sum of a bounded increasing function and a bounded decreasing function.

We turn to a continuity property of bounded variation functions. We say that a discontinuity is of the **first kind** if it is a jump or removable discontinuity.

Theorem 2.7. *Every function of bounded variation has at most a countable number of discontinuities, all of which are of the first kind.*

Proof. Let f be of bounded variation on $[a,b]$. Using Jordan's theorem, we may assume f is bounded and increasing on $[a,b]$. Then the only discontinuities of f are of the first kind; in fact, they are all jump discontinuities. However, each jump continuity defines a distinct interval, which contains a rational number; thus there can be at most countably many. □

Theorem 2.8. *If f is continuous on $[a,b]$, then*

$$V = \lim_{|\Gamma|\to 0} S_\Gamma.$$

That is, for $M < V$, there exists $\delta > 0$ so that $|\Gamma| < \delta \implies S_\Gamma > M$.

Proof. Let $M < V$ and let $\mu > 0$ so that $M + \mu < V$. Choose $\bar{\Gamma} = \{\bar{x}_j\}_{j=0}^k$ so that

$$S_{\bar{\Gamma}} > M + \mu.$$

By uniform continuity of f on $[a,b]$, choose $\eta > 0$ so that

$$|x - y| < \eta \implies |f(x) - f(y)| < \tfrac{\mu}{2(k+1)}.$$

Now take a partition $\Gamma = \{x_i\}_{i=0}^m$ satisfying

$$|\Gamma| < \eta \quad \text{and} \quad |\Gamma| < \min\{\bar{x}_j - \bar{x}_{j-1}\}.$$

We will show that $S_\Gamma > M$, which will complete the proof.

We have

$$S_\Gamma = \sum_{i=1}^m |f(x_i) - f(x_{i-1})| = \Sigma_1 + \Sigma_2,$$

where Σ_2 is the sum over i such that $(x_{i-1}, x_i) \cap \bar{\Gamma} \neq \emptyset$.

By construction, (x_{i-1}, x_i) can contain at most the point $\bar{x}_j$ from $\bar{\Gamma}$. Thus Σ_2 has at most $k+1$ summands.

Now, we may write

$$S_{\Gamma \cup \bar{\Gamma}} = \Sigma_1 + \Sigma_3,$$

where Σ_3 is obtained from Σ_2 by replacing each term by

$$|f(x_i) - f(\bar{x}_j)| + |f(\bar{x}_j) - f(x_{i-1})|.$$

By uniform continuity, each of these is less than $\frac{\mu}{2(k+1)}$, and thus

$$\Sigma_3 < \mu.$$

Therefore

$$S_\Gamma = \Sigma_1 + \Sigma_2 \geq \Sigma_1 = S_{\Gamma \cup \bar{\Gamma}} - \Sigma_3 > S_{\Gamma \cup \bar{\Gamma}} - \mu \geq S_{\bar{\Gamma}} - \mu > M,$$

as desired. □

Corollary 2.9. *If f has a continuous derivative f' on $[a,b]$, then*

$$V = \int_a^b |f'(x)|\, dx, \quad P = \int_a^b \{f'(x)\}^+\, dx, \quad N = \int_a^b \{f'(x)\}^-\, dx.$$

Proof. Using the mean-value theorem,

$$S_\Gamma = \sum_{i=1}^m |f'(\xi_i)|(x_i - x_{i-1})$$

for some $\xi_i \in (x_{i-1}, x_i)$. Thus, by the definition of the Riemann integral,

$$V = \lim_{|\Gamma|\to 0} S_\Gamma = \lim_{|\Gamma|\to 0} \sum_{i=1}^m |f'(\xi_i)|(x_i - x_{i-1}) = \int_a^b |f'(x)|\,dx.$$

Moreover, using $\frac{1}{2}(|y| + y) = y^+$,

$$P = \tfrac{1}{2}[V + f(b) - f(a)] = \tfrac{1}{2}\left[\int_a^b |f'(x)|\,dx + \int_a^b f'(x)\,dx\right] = \int_a^b [f'(x)]^+\,dx.$$

A similar argument yields the formula for N. □

The notion of bounded variation makes sense in the setting of open intervals, infinite intervals, half-open intervals, complex-valued functions, etc.

2.2. **Rectifiable curves.** A **curve** C in the plane is two parametric equations

$$x = \phi(t), \quad y = \psi(t), \quad t \in [a, b].$$

The **graph** of C is

$$\{(x, y) : x = \phi(t), \quad y = \psi(t), \quad t \in [a, b]\}.$$

For a partition $\Gamma = \{t_i\}_{i=0}^m$ of $[a, b]$, we define

$$\ell(\Gamma) = \sum_{i=1}^m \sqrt{[\phi(t_i) - \phi(t_{i-1})]^2 + [\psi(t_i) - \psi(t_{i-1})]^2}.$$

The **length** of C is defined by

$$L = L(C) = \sup_\Gamma \ell(\Gamma).$$

We call C **rectifiable** if $L < \infty$.

Theorem 2.10. *A curve C is rectifiable if and only if ϕ and ψ are of bounded variation. Moreover,*

$$V(\phi),\ V(\psi) \le L \le V(\phi) + V(\psi).$$

Proof. We will use

$$|x|, |y| \le \sqrt{x^2 + y^2} \le |x| + |y| \quad \text{for} \quad x, y \in \mathbb{R}.$$

As

$$\ell(\Gamma) = \sum \sqrt{[\phi(t_i) - \phi(t_{i-1})]^2 + [\psi(t_i) - \psi(t_{i-1})]^2} \le L,$$

we have

$$\sum |\phi(t_i) - \phi(t_{i-1})| \le L \quad \text{and} \quad \sum |\psi(t_i) - \psi(t_{i-1})| \le L.$$

This implies

$$V(\phi), V(\psi) \le L.$$

Conversely,

$$\ell(\Gamma) \le \sum |\phi(t_i) - \phi(t_{i-1})| + \sum |\psi(t_i) - \psi(t_{i-1})| \le V(\phi) + V(\psi),$$

and hence $L \le V(\psi) + V(\psi)$. □

If ϕ is a bounded function that is **not** of bounded variation, the the curve $x = y = \phi(t)$ is not rectificable. However, the graph lies in a finite segment of the line $y = x$.

Thus the length of the graph of a curve is not necessarily equal to the length of the curve.

If C is given by $y = f(x)$, then the theorem reduces to the statement that C is rectifiable if and only if f is of bounded variation.

These ideas generalize to curves in $\mathbb{R}^n$ as well.

2.3. The Riemann–Stieltjes Integral.

Definition 2.11. Let $f, \phi : [a, b] \to \mathbb{R}$. Let $\Gamma = \{x_i\}_{i=0}^m$ be a partition of $[a, b]$ and let $\{\xi_i\}_{i=1}^m$ satisfy

$$x_{i-1} \le \xi_i \le x_i \quad \text{for each} \quad i.$$

The quantity

$$R_\Gamma := \sum_{i=1}^m f(\xi_i)[\phi(x_i) - \phi(x_{i-1})]$$

is called a **Riemann–Stieltjes sum** for Γ.

If

$$I = \lim_{|\Gamma| \to 0} R_\Gamma \tag{2.1}$$

exists and is finite, then I is called the **Riemann–Stieltjes integral of f with respect to ϕ on** $[a, b]$, denoted

$$I = \int_a^b f(x)\, d\phi(x) = \int_a^b f\, d\phi.$$

The condition (2.1) means that for any $\varepsilon > 0$, there exists $\delta > 0$ so that

$$|\Gamma| < \delta \implies |I - R_\Gamma| < \varepsilon$$

(for any choice of ξ_i). Equivalently, the integral exists if and only if for any $\varepsilon > 0$, there exists $\delta > 0$

$$|\Gamma|, |\Gamma'| < \delta \implies |R_\Gamma - R_{\Gamma'}| < \varepsilon.$$

Here are some properties of the integral:

- If $\phi(x) = x$, then the Riemann–Stieltjes integral is simply the Riemann integral.

- If f is continuous on $[a,b]$ and ϕ is continuously differentiable on $[a,b]$, then
$$\int_a^b f\,d\phi = \int_a^b f\phi'\,dx.$$
Indeed, the essential fact is the mean value theorem:
$$\sum f(\xi_i)[\phi(x_i)-\phi(x_{i-1})] = \sum f(\xi_i)\phi'(\eta_i)(x_i - x_{i-1})$$
- Suppose ϕ is a **step function**, that is, there exists partition $\{\alpha_i\}_{i=0}^m$ of $[a,b]$ such that ϕ is constant on each (α_{i-1},α_i). Define the left and right limits at α_i by
$$\begin{aligned} \phi_{\alpha_i+} &= \lim_{x\to\alpha_i+} \phi(x) \quad \text{for} \quad i=0,\dots,m-1, \\ \phi_{\alpha_i-} &= \lim_{x\to\alpha_i-} \phi(x) \quad \text{for} \quad i=1,\dots,m. \end{aligned}$$
Define the jumps of ϕ by
$$d_i = \begin{cases} \phi(\alpha_i+)-\phi(\alpha_i-), & i=1,\dots,m-1 \\ \phi(\alpha_0+)-\phi(\alpha_0), & i=0 \\ \phi(\alpha_m)-\phi(\alpha_m-), & i=m. \end{cases}$$
For $f\in C([a,b])$, one can check that
$$\int_a^b f\,d\phi = \sum_{i=0}^m f(\alpha_i)d_i.$$
- The most important cases occur when ϕ is monotone (or of bounded variation).
- If $\int_a^b f\,d\phi$ exists, then f and ϕ have no common points of discontinuity.

Proof. Suppose f,ϕ are both discontinuous at $\bar{x}\in(a,b)$.

Suppose the discontinuity of ϕ is not removable.

Then there exists $\varepsilon_0>0$ so that for any $\delta>0$ there exist $\bar{x}_1,\bar{x}_2$ with

$\bar{x}-\frac12\delta<\bar{x}_1<\bar{x}<\bar{x}_2<\bar{x}+\frac12\delta$ and $|\phi(\bar{x}_2)-\phi(\bar{x}_1)|>\varepsilon_0$.

Given $\delta>0$, take a partition $\Gamma=\{x_i\}$ of $[a,b]$ so that $|\Gamma|<\delta$, with $x_{i_0-1}=\bar{x}_1$ and $x_{i_0}=\bar{x}_2$ for some i_0.

Let $\xi_i\in[x_{i-1},x_i]$ for $i\neq i_0$ and $\xi_{i_0}\neq\xi_{i_0}'\in[x_{i_0-1},x_{i_0}]$

Let R_Γ be the Riemann–Stieltjes sum using ξ_i in $[x_{i-1},x_i]$ and $\xi_{i_0}\in[x_{i_0-1},x_{i_0}]$, and define $R_{\Gamma'}$ similarly but using $\xi_{i_0}'\in[x_{i_0-1},x_{i_0}]$. Then
$$|R_\Gamma - R_{\Gamma'}| > \varepsilon_0|f(\xi_{i_0})-f(\xi_{i_0}')|.$$
As f is discontinuous at $\bar{x}$, we can choose ξ_{i_0},ξ_{i_0}' so that
$$|f(\xi_{i_0})-f(\xi_{i_0}')|>\mu$$

for some μ (independent of ε). It follows that

$$R_\Gamma - R_{\Gamma'} \not\to 0 \quad \text{as} \quad |\Gamma|, |\Gamma'| \to 0.$$

Similar arguments treat the case of a removable discontinuity at $\bar{x}$, or with $\bar{x} \in \{a, b\}$. □

The following theorem follows from the definition of the integral and is left as an exercise.

Theorem 2.12 (Linearity).

(i) *If $\int_a^b f\,d\phi$ exists, then for any $c \in \mathbb{R}$*

$$\int_a^b cf\,d\phi = \int_a^b f\,d(c\phi) = c\int_a^b f\,d\phi.$$

(In particular, the first two integrals exist.)

(ii) *If $\int_a^b f_1\,d\phi$ and $\int_a^b f_2\,d\phi$ exist then*

$$\int_a^b (f_1 + f_2)\,d\phi = \int_a^b f_1\,d\phi + \int_a^b f_2\,d\phi.$$

(In particular, the integral exists.)

(iii) *If $\int_a^b f\,d\phi_1$ and $\int_a^b f\,d\phi_2$ exist, then*

$$\int_a^b f\,d(\phi_1 + \phi_2) = \int_a^b f\,d\phi_1 + \int_a^b f\,d\phi_2.$$

(In particular, the integral exists.)

We also have the following:

Theorem 2.13 (Additivity). *If $\int_a^b f\,d\phi$ exists and $c \in (a, b)$, then*

$$\int_a^b f\,d\phi = \int_a^c f\,d\phi + \int_c^b f\,d\phi.$$

(In particular, the latter two integrals exist.)

Proof. Denote a sum corresponding to a partition of $[a, b]$ by $R_\Gamma[a, b]$, and similarly with other intervals.

Let $\varepsilon > 0$. Choose $\delta > 0$ so that for any partitions Γ_1' and Γ_2' of $[a, b]$ with $|\Gamma_1'|, |\Gamma_2'| < \delta$, we have

$$|R_{\Gamma_1'}[a, b] - R_{\Gamma_2'}[a, b]| < \varepsilon. \tag{2.2}$$

Now let Γ_1, Γ_2 be partitions of $[a, c]$ and let Γ' be a partition of $[c, b]$. Let

$$\Gamma_1' = \Gamma_1 \cup \Gamma', \quad \Gamma_2' = \Gamma_2 \cup \Gamma'.$$

Then

$$\begin{aligned} R_{\Gamma_1'}[a, b] &= R_{\Gamma_1}[a, c] + R_{\Gamma'}[c, b], \\ R_{\Gamma_2'}[a, b] &= R_{\Gamma_2}[a, c] + R_{\Gamma'}[c, b]. \end{aligned} \tag{2.3}$$

Now assume $|\Gamma_1|, |\Gamma_2| < \delta$ and choose Γ' with $|\Gamma'| < \delta$. Then $|\Gamma_1'|, |\Gamma_2'| < \delta$ and (2.2) implies

$$|R_{\Gamma_1}[a,c] - R_{\Gamma_2}[a,c]| < \varepsilon.$$

This gives existence of $\int_a^c f\,d\phi$. Existence of $\int_c^b f\,d\phi$ follows similarly. Moreover, (2.3) implies

$$\int_a^b f\,d\phi = \int_a^c f\,d\phi + \int_c^b f\,d\phi.$$

□

We turn to an **integration by parts** formula.

Theorem 2.14. *If $\int_a^b f\,d\phi$ exists, then so does $\int_a^b \phi\,df$, and*

$$\int_a^b f\,d\phi = [f(b)\phi(b) - f(a)\phi(a)] - \int_a^b \phi\,df.$$

Proof. Let $\Gamma = \{x_i\}_{i=1}^m$ be a partition of $[a,b]$ and $\xi_i \in [x_{i-1}, x_i]$. Then

$$\begin{aligned} R_\Gamma &= \sum_{i=1}^m f(\xi_i)[\phi(x_i) - \phi(x_{i-1})] \\ &= \sum_{i=1}^m f(\xi_i)\phi(x_i) - \sum_{i=1}^m f(\xi_i)\phi(x_{i-1}) \\ &= \sum_{i=1}^m f(\xi_i)\phi(x_i) - \sum_{i=0}^{m-1} f(\xi_{i+1})\phi(x_i) \\ &= -\sum_{i=1}^{m-1} \phi(x_i)[f(\xi_{i+1}) - f(\xi_i)] + f(\xi_m)\phi(b) - f(\xi_1)\phi(a). \end{aligned}$$

Now add and subtract

$$\phi(a)[f(\xi_1) - f(a)] + \phi(b)[f(b) - f(\xi_m)]$$

on the right-hand side. This yields

$$R_\Gamma = -T_R + [f(b)\phi(b) - f(a)\phi(a)],$$

where

$$T_R = \sum_{i=1}^{m-1} \phi(x_i)[f(\xi_{i+1}) - f(\xi_i)] + \phi(a)[f(\xi_1) - f(a)] + \phi(b)[f(b) - f(\xi_m)].$$

This is in fact a Riemann–Stieltjes sum for $\int_a^b \phi\,df$.

From this we deduce $\int_a^b f\,d\phi$ exists if and only if $\int_a^b \phi\,df$ exists.

Moreover,

$$\int_a^b f\,d\phi = [f(b)\phi(b) - f(a)\phi(a)] - \int_a^b \phi\,df,$$

as desired. □

Next, suppose f is bounded and ϕ is increasing on $[a,b]$. For a partition $\Gamma = \{x_i\}_{i=0}^m$ of $[a,b]$, define

$$\begin{aligned} m_i &= \inf_{x\in[x_{i-1},x_i]} f(x), \\ M_i &= \sup_{x\in[x_{i-1},x_i]} f(x), \\ L_\Gamma &= \sum_{i=1}^m m_i[\phi(x_i) - \phi(x_{i-1})], \\ U_\Gamma &= \sum_{i=1}^m M_i[\phi(x_i) - \phi(x_{i-1})]. \end{aligned}$$

Note that

$$L_\Gamma \le R_\Gamma \le U_\Gamma.$$

We call L_Γ and U_Γ the **lower and upper Riemann–Stieltjes sums for** Γ.

Lemma 2.15. *Let f be bounded and ϕ be increasing on $[a,b]$.*

(i) *If Γ' is a refinement of Γ (that is, $\Gamma \subset \Gamma'$), then*

$$L_{\Gamma'} \ge L_\Gamma \quad \textit{and} \quad U_{\Gamma'} \le U_\Gamma.$$

(ii) *For any partitions Γ_1 and Γ_2,*

$$L_{\Gamma_1} \le U_{\Gamma_2}.$$

Proof. For (i), it is enough to check the case that $\Gamma' = \Gamma \cup \{x'\}$. In this case, if $x' \in (x_{i-1}, x_i)$ (where $\Gamma = \{x_k\}$), then

$$\sup_{[x_{i-1},x']} f(x) \le M_i \quad \text{and} \quad \sup_{[x',x_i]} f(x) \le M_i,$$

so

$$\sup_{[x_{i-1},x']} f(x)[\phi(x') - \phi(x_{i-1})] + \sup_{[x',x_i]} f(x)[\phi(x_i) - \phi(x')] \le M_i[\phi(x_i) - \phi(x_{i-1})],$$

giving $U_{\Gamma'} \le U_\Gamma$. A similar argument handles lower sums.

For (ii), note that $\Gamma_1 \cup \Gamma_2$ is a refinement of both Γ_1 and Γ_2, and hence

$$L_{\Gamma_1} \le L_{\Gamma_1\cup\Gamma_2} \le U_{\Gamma_1\cup\Gamma_2} \le U_{\Gamma_2},$$

as desired. □

The following result gives sufficient conditions for the existence of $\int f\,d\phi$.

Theorem 2.16. *Suppose $f \in C([a,b])$ and $\phi \in BV([a,b])$. Then $\int_a^b f\,d\phi$ exists, and*

$$\left|\int_a^b f\,d\phi\right| \le \|f\|\,\|\phi\|_{BV} = \Big[\sup_{[a,b]} |f|\Big] \cdot V[\phi; a, b].$$

Proof. It suffices to consider the case that ϕ is increasing (and non-constant).

In this case,

$$L_\Gamma \leq R_\Gamma \leq U_\Gamma,$$

and hence it suffices to show

$$\lim_{|\Gamma|\to 0} L_\Gamma = \lim_{|\Gamma|\to 0} U_\Gamma.$$

Let $\Gamma = \{x_i\}$ be a partition of $[a,b]$. By uniform continuity of f, for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\Gamma| < \delta \implies M_i - m_i < \tfrac{\varepsilon}{\phi(b)-\phi(a)}.$$

Thus

$$0 \leq U_\Gamma - L_\Gamma = \sum [M_i - m_i](\phi(x_i) - \phi(x_{i-1}) < \varepsilon,$$

and so

$$\lim_{|\Gamma|\to 0} [U_\Gamma - L_\Gamma] = 0.$$

It remains to prove that $\lim_{|\Gamma|\to 0} U_\Gamma$ exists. If not, there would exist $\varepsilon_0 > 0$ and sequences of partitions $\{\Gamma_k\}$, $\{\Gamma_k'\}$ such that

$$|\Gamma_k|, |\Gamma_k'| \to 0 \quad \text{but} \quad U_{\Gamma_k} - U_{\Gamma_k'} > \varepsilon_0.$$

However, this means that for large enough k,

$$L_{\Gamma_k} - U_{\Gamma_k'} > 0,$$

contradicting that $L_\Gamma \leq U_{\Gamma'}$ for any partitions.

The desired bound, i.e.

$$\left| \int_a^b f\, d\phi \right| \leq \|f\| \|\phi\|_{BV}$$

follows from an analogous bound on R_Γ and taking the limit. □

Combining this result with the 'integration by parts' formula, we see that $\int f\, d\phi$ exists if either f or ϕ is continuous and the other is of bounded variation.

We turn to the following mean value theorem for Riemann–Stieltjes integrals.

Theorem 2.17 (Mean value theorem)**.** *Let $f \in C([a,b])$ and ϕ be a bounded increasing function on $[a,b]$. Then there exists $\xi \in [a,b]$ so that*

$$\int_a^b f\, d\phi = f(\xi)[\phi(b) - \phi(a)].$$

Proof. We have

$$(\min f)[\phi(b) - \phi(a)] \le R_\Gamma \le (\max f)[\phi(b) - \phi(a)]$$

for any partition Γ. Since $\int_a^b f\,d\phi$ exists, we therefore have

$$\min f \le \frac{\int_a^b f\,d\phi}{\phi(b) - \phi(a)} \le \max f.$$

The result now follows from the intermediate value theorem. □

We can define Riemann–Stieltjes integrals on open intervals, half-open intervals, infinite intervals, etc. For example, for (a, b) we would set

$$\int_a^b f\,d\phi = \lim_{a' \to a,\ b' \to b} \int_{a'}^{b'} f\,d\phi,$$

where the right-hand side has integrals over $[a', b']$.

2.4. **Further results.** Suppose f is bounded and ϕ is increasing. Then we always have

$$\sup_\Gamma L_\Gamma \le \inf_\Gamma U_\Gamma.$$

Question. If

$$\sup_\Gamma L_\Gamma = \inf_\Gamma U_\Gamma, \tag{2.4}$$

then does $\int_a^b f\,d\phi$ exist? [This is the case, for example, for Riemann integrals.]

Answer. No. Let $[a, b] = [-1, 1]$ and define

$$f(x) = \begin{cases} 0 & x \in [-1, 0) \\ 1 & x \in [0, 1], \end{cases}$$

$$\phi(x) = \begin{cases} 0 & x \in [-1, 0] \\ 1 & x \in (0, 1]. \end{cases}$$

As f and ϕ have a common discontinuity, $\int_a^b f\,d\phi$ does not exist. Depending on whether or not a partition Γ straddles 0, we have $R_\Gamma \in \{0, 1\}$ and in particular does not have a limit.

However, $U_\Gamma \equiv 1$, while $L_\Gamma \in \{0, 1\}$. Thus (2.4) holds.

We do have the following results, the proofs of which we leave as exercises.

Theorem 2.18. *Let f be bounded and ϕ increasing on $[a, b]$. If $\int_a^b f\,d\phi$ exists, then*

$$\lim_{|\Gamma| \to 0} L_\Gamma = \lim_{|\Gamma| \to 0} U_\Gamma = \sup_\Gamma L_\Gamma = \inf_\Gamma U_\Gamma = \int_a^b f\,d\phi.$$

[Hint: given $\varepsilon > 0$, take a sufficiently fine partition and refine it in two ways, first picking points that almost attain the infimum, and second picking points that almost attain the supremum. This will give you good approximations to U_Γ and L_Γ that are close to the value of the integral.]

Theorem 2.19. *Let f be bounded and ϕ increasing and continuous on $[a,b]$. Then*

$$\lim_{|\Gamma|\to 0} L_\Gamma = \sup_\Gamma L_\Gamma, \qquad \lim_{|\Gamma|\to 0} U_\Gamma = \inf_\Gamma U_\Gamma.$$

Moreover, if (2.4) *holds, then $\int_a^b f\,d\phi$ exists and*

$$\sup_\Gamma L_\Gamma = \inf_\Gamma U_\Gamma = \int_a^b f\,d\phi.$$

[Hint: the proof is similar in spirit to that of Theorem 2.8.]

3. Lebesgue measure and outer measure

Reference: Wheeden–Zygmund Chapter 3

3.1. **Lebesgue outer measure; the Cantor set.**

Given $a_k \leq b_k$ ($k = 1, \dots, n$), we define the n-dimensional intervals

$$I = \{x \in \mathbb{R}^n : a_k \leq x_k \leq b_k, \quad k = 1, \dots, n\}$$

and their volumes

$$v(I) = \prod_{k=1}^{n} [b_k - a_k].$$

Definition 3.1. Any set $E \subset \mathbb{R}^n$ may be covered by a countable collection S of intervals I_k (that is, $E \subset \cup_k I_k$). For each such cover S, define

$$\sigma(S) = \sum_{I_k \in S} v(I_k).$$

The **outer measure** of a set $E \subset \mathbb{R}^n$ is defined by

$$|E|_e = \inf \sigma(S) \in [0, \infty]$$

where the infimimum is taken over all such covers S.

Theorem 3.2. *If I is an interval, then $|I|_e = v(I)$.*

Proof. Since I is a cover of itself, we have

$$|I|_e \leq v(I).$$

Conversely, let $S = \{I_k\}$ be a cover of I and let $\varepsilon > 0$. Denote by I_k^* and interval containing I_k in its interior, with

$$v(I_k^*) < (1 + \varepsilon) v(I_k).$$

Since $I \subset \cup_k (I_k^*)^\circ$ (where $\circ$ denotes interior) and I is compact, it follows that

$$I \subset \cup_{k=1}^{N} I_k^* \quad \text{for some} \quad N.$$

Thus

$$v(I) \leq \sum_{k=1}^{N} v(I_k^*) \leq (1 + \varepsilon) \sum_{k=1}^{N} v(I_k) \leq (1 + \varepsilon) \sigma(S).$$

This implies

$$v(I) \leq \sigma(S),$$

and hence upon taking the infimum that

$$v(I) \leq |I|_e.$$

This completes the proof. □

One can check that the boundary of any interval has outer measure zero.

We record a few other properties of outer measure.

Theorem 3.3. *If* $E_1 \subset E_2$ *then* $|E_1|_e \leq |E_2|_e$.

Proof. This follows from the fact that any cover of E_2 is a cover of E_1. □

Theorem 3.4. *If* $E = \cup_k E_k$ *is a countable union, then*

$$|E|_e \leq \sum_k |E_k|_e.$$

Proof. It suffices to assume $|E_k|_e < \infty$ for each k.

Let $\varepsilon > 0$ and for each k choose intervals I_j^k so that

$$E_k \subset \cup_j I_j^k \quad \text{and} \quad \sum_j v(I_j^k) \leq |E_k| + \varepsilon 2^{-k}.$$

Then

$$E \subset \cup_{j,k} I_j^k,$$

and so

$$|E|_e \leq \sum_k \sum_j v(I_j^k) \leq \sum_k [|E_k|_e + \varepsilon 2^{-k}] \leq \sum_k |E_k|_e + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, the result follows. □

Remark 3.5. Any subset of a set with outer measure zero has outer measure zero, and the countable union of sets of outer measure zero has outer measure zero. In particular, since a point has outer measure zero, any countable subset of $\mathbb{R}^n$ has outer measure zero.

On the other hand, there are uncountable subsets with outer measure zero.

Before presenting an example, we introduce the notion of a perfect set.

Definition 3.6. A set C is **perfect** if C is closed and every point in C is a limit point of C. That is, for any $x \in C$, there exists $\{x_k\} \subset C \backslash \{x\}$ so that $x_k \to x$.

We leave the following as an exercise.

Proposition 3.7. *A perfect set is uncountable.*

Example 3.1 (Cantor set). For a closed interval $[a, b]$, define

$$F([a, b]) = [a, \tfrac{2}{3}a + \tfrac{1}{3}b] \cup [\tfrac{1}{3}a + \tfrac{2}{3}b, b].$$

Note $\{a, b\} \subset F([a, b]) \subset [a, b]$. We extend this to disjoint closed intervals $\{I_j\}_{j=1}^n$ via

$$F(\cup_{j=1}^n I_j) = \cup_{j=1}^n F(I_j).$$

Note that $F(I_j)$ are also disjoint, and that $F(\cup I_j)$ contains the endpoints of all the I_j.

Now define a sequence of sets $\{C_k\}$ via

$$C_0 = [0, 1], \quad C_{k+1} = F(C_k) \subset C_k.$$

By construction, C_k is the union of 2^k closed disjoint intervals of length $(\frac{1}{3})^k$.

The set

$$C := \bigcap_{k=0}^{\infty} C_k$$

is called the **Cantor set** (or the Cantor $\frac{1}{3}$ set). Note that C is a closed subset of $[0, 1]$ that contains the endpoints of all of the intervals in each C_k.

As C is covered by the intervals in each C_k, we deduce

$$|C|_e \leq 2^k 3^{-k} \quad \text{for any} \quad k, \quad \text{so that} \quad |C|_e = 0.$$

Moreover, we claim C is perfect (and hence uncountable). Indeed, if $x \in C$ then x belongs to some interval in C_k for each k. Thus, since the length of these intervals approaches 0, x is the limit of the endpoints of these intervals (which belong to C by construction).

We will next construct a function related to the Cantor set that we will use in later sections.

Example 3.2 (Cantor–Lebesgue function). Let C_k be as in the Cantor set construction, and define

$$D_k = [0, 1] \backslash C_k.$$

Then D_k consists of $2^k - 1$ intervals I_j^k (ordered from left to right) removed in the first k stages of the Cantor set construction.

Let f_k be the continuous function on $[0, 1]$ satisfying

- $f_k(0) = f_k(1) = 1$,
- $f_k(x) = j2^{-k}$ on I_j^k, $j = 1, \ldots, 2^k - 1$,
- f_k is linear on each interval of C_k.

Each f_k is increasing, with

$$f_{k+1} = f_k \quad \text{on} \quad I_j^k, \quad j = 1, \ldots, 2^k - 1.$$

Furthermore

$$|f_k - f_{k+1}| < 2^{-k}.$$

Thus

$$\sum_k [f_k - f_{k+1}]$$

converges uniformly on $[0, 1]$, and hence $\{f_k\}$ converges uniformly on $[0, 1]$.

Let $f = \lim_{k\to\infty} f_k$. Then

- $f(0) = f(1) = 1$,
- f is increasing and continuous on $[0, 1]$,
- f is constant on every interval removed in the Cantor set construction.

The function f is called the **Cantor–Lebesgue function.**

We next consider the question of approximating the outer measure of sets.

Theorem 3.8. *Let $E \subset \mathbb{R}^n$. For any $\varepsilon > 0$, there exists an open set G so that*

$$E \subset G \quad \text{and} \quad |G|_e \leq |E|_e + \varepsilon.$$

In particular,

$$|E|_e = \inf\{|G|_e : E \subset G, \quad G \text{ open}\}.$$

Proof. Let $\varepsilon > 0$. Choose intervals I_k with

$$E \subset \bigcup_{k=1}^{\infty} I_k \quad \text{and} \quad \sum_{k=1}^{\infty} v(I_k) \leq |E|_e + \tfrac{1}{2}\varepsilon.$$

Let I_k^* be an interval with $I_k \subset (I_k^*)^\circ$ and

$$v(I_k^*) \leq v(I_k) + \varepsilon 2^{-(k+1)}.$$

The set

$$G = \cup (I_k^*)^\circ$$

is open, contains E, and satisfies

$$|G|_e \leq \sum_{k=1}^{\infty} v(I_k^*) \leq \sum_{k=1}^{\infty} [v(I_k) + \varepsilon 2^{-(k+1)}] \leq |E|_e + \varepsilon,$$

which completes the proof. □

We next need the concept of a G_δ set.

Definition 3.9. A set is called a G_δ set if it is the countable intersection of open sets.

Theorem 3.10. *If $E \subset \mathbb{R}^n$, then there exists a G_δ set H such that*

$$E \subset H \quad \text{and} \quad |E|_e = |H|_e.$$

Proof. By the previous theorem, for each k we may find $G_k \supset E$ so that

$$|G_k|_e \leq |E|_e + \tfrac{1}{k}.$$

Now set

$$H = \cap_{k=1}^{\infty} G_k.$$

Then H is G_δ, contains E, and for each k we have

$$|E|_e \leq |H|_e \leq |G_k|_e \leq |E|_e + \tfrac{1}{k}.$$

This implies $|E|_e = |H|_e$. □

The notion of outer measure is not tied to our choice to define intervals relative to the standard coordinate axes.

Suppose we rotate to new coordinates x', and write I' for an interval with edges parallel to the new coordinate axes. The volume of an interval is invariant under rotation.

Then we may define

$$|E|_e' = \inf \sum v(I_k'),$$

with the infimum taken over all coverings of E by rotated intervals I'.

Theorem 3.11. *We have $|E|_e = |E|_e'$ for all $E \subset \mathbb{R}^n$.*

Proof. First, given any I' and $\varepsilon > 0$, let I_1' be an interval with $I' \subset (I_1')^\circ$ and

$$v(I_1') \le v(I') + \varepsilon.$$

We may write I_1' as a countable union of nonoverlapping intervals I_ℓ. In particular, for each N

$$\sum_{\ell=1}^{N} v(I_\ell) \le v(I_1'), \quad \text{whence} \quad \sum_{\ell=1}^{\infty} v(I_\ell) \le v(I_1') \le v(I') + \varepsilon.$$

Now let $E \subset \mathbb{R}^n$. Given $\varepsilon > 0$, choose $\{I_k\}_{k=1}^\infty$ so that

$$E \subset \cup I_k \quad \text{and} \quad \sum v(I_k) \le |E|_e + \tfrac{1}{2}\varepsilon.$$

For each k, we may (by the argument above) choose $\{I_k, \ell'\}$ so that

$$I_k \subset \cup_\ell I_{k,\ell}' \quad \text{and} \quad \sum_\ell v(I_{k,\ell}') \le v(I_k) + \varepsilon 2^{-(k+1)}.$$

Thus $E \subset \cup_{k,\ell} I_{k,\ell}'$ and

$$\sum_{k,\ell} v(I_{k,\ell}') \le \sum_k v(I_k) + \tfrac{1}{2}\varepsilon \le |E|_e + \varepsilon,$$

which implies $|E|_e' \le |E|_e + \varepsilon$. As ε was arbitrary, we have $|E|_e' \le |E|_e$.

A similar argument proves the reverse inequality. □

3.2. **Lebesgue measurable sets.** Recall the notations

$$A \backslash B = A \cap B^c, \quad B^c = \{x : x \notin B\}.$$

Definition 3.12. A set $E \subset \mathbb{R}^n$ is **(Lebesgue) measurable** if for every $\varepsilon > 0$, there exists an open set G such that

$$E \subset G \quad \text{and} \quad |G \backslash E|_e < \varepsilon.$$

If E is measurable, its outer measure is called its **(Lebesgue) measure** and is denoted by $|E|$. That is,

$$|E| = |E|_e \quad \text{for measurable} \quad E.$$

Remark 3.13. Compare carefully with Theorem 3.8. It is always true that there exists open $G \supset E$ with

$$|G|_e \leq |E|_e + \varepsilon.$$

However, when $E \subset G$, we have

$$G \subset E \cup G \backslash E,$$

which only implies

$$|G|_e \leq |E|_e + |G \backslash E|_e.$$

In particular, we cannot deduce $|G \backslash E|_e < \varepsilon$.

Example 3.3. Every open set is measurable. Indeed, if E is open and we take $G = E$, then $|G \backslash E|_e = |\emptyset|_e = 0$.

Example 3.4. If $|E|_e = 0$, then E is measurable. Indeed given $\varepsilon > 0$, by Theorem 3.8 we may find G so that

$$|G| < \varepsilon.$$

As $G \backslash E \subset G$, we have

$$|G \backslash E|_e < \varepsilon,$$

giving the claim.

Theorem 3.14. *Let $\{E_k\}$ be a countable collection of measurable sets. Then $E := \cup E_k$ is measurable, with*

$$|E| \leq \sum |E_k|.$$

Proof. Let $\varepsilon > 0$. For each k, let G_k be an open set so that

$$E_k \subset G_k \quad \text{and} \quad |G_k \backslash E_k|_e < \varepsilon 2^{-k}.$$

Then $G = \cup G_k$ is open and $E \subset G$.

Moreover,

$$G \backslash E \subset \bigcup [G_k \backslash E_k],$$

so that

$$|G \backslash E|_e \leq |\bigcup G_k \backslash E_k|_e \leq \sum |G_k - E_k|_e < \varepsilon.$$

Thus E is measurable. The subadditivity follows from the analogous property for outer measure. □

Corollary 3.15. *An interval I is measurable, with $|I| = v(I)$.*

Proof. Write I as the union of its (open) interior and its boundary. As the boundary has measure zero, the result follows. □

Our next result is the following:

Theorem 3.16. *Closed sets are measurable.*

We need a few lemmas.

Lemma 3.17. *If $\{I_k\}_{k=1}^N$ is a finite collection of nonoverlapping intervals, then $\cup I_k$ is measurable and*

$$|\cup I_k| = \sum |I_k|.$$

Proof. Measurability follows from the previous theorem. The equality is left as an exercise (cf. Theorem 3.2). □

Recall that the distance between two sets E_1 and E_2 is defined by

$$d(E_1, E_2) = \inf\big\{|x_1 - x_2| : x_1 \in E_1, \quad x_2 \in E_2\big\}.$$

We then have the following lemma.

Lemma 3.18. *If $d(E_1, E_2) > 0$ then $|E_1 \cup E_2|_e = |E_1|_e + |E_2|_e$.*

Proof. It suffices to prove that

$$|E_1|_e + |E_2|_e \le |E_1 \cup E_2|_e.$$

To this end, let $\varepsilon > 0$ and choose intervals $\{I_k\}$ so that

$$E_1 \cup E_2 \subset \bigcup I_k \quad \text{and} \quad \sum |I_k| \le |E_1 \cup E_2|_e + \varepsilon.$$

We may assume that each I_k has diameter less than $d(E_1, E_2)$, for otherwise we may divide each I_k into a finite number of subintervals with this property.

In particular, $\{I_k\}$ splits into $\{I_k^1\}$ and $\{I_k^2\}$, where $\{I_k^j\}_k$ covers E_j.

Then

$$|E_1|_e + |E_2|_e \le \sum_k |I_k^1| + \sum_k |I_k^2| = \sum |I_k| \le |E_1 \cup E_2|_e + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this gives the desired inequality. □

We will use this along with the following fact (which is left as an exercise): if E_1 and E_2 are compact and disjoint, then $d(E_1, E_2) > 0$.

Proof of Theorem 3.16. Suppose F is a compact set.

Given $\varepsilon > 0$, let G be an open set with

$$F \subset G \quad \text{and} \quad |G| < |F|_e + \varepsilon.$$

As $G \backslash F$ is open, there exist nonoverlapping closed intervals I_k so that

$$G \backslash F = \cup I_k$$

(exercise).

Now since

$$G = F \cup \big[\cup_k I_k\big] \supset F \cup \big[\cup_{k=1}^N I_k\big]$$

for every N, and F and $\cup_{k=1}^N I_k$ are disjoint and compact, we deduce

$$|G| \geq \Big| F \cup \big[\cup_{k=1}^N I_k\big] \Big|_e = |F|_e + \Big| \sum_{k=1}^N I_k \Big|_e ,$$

and hence

$$\sum_{k=1}^N |I_k| = \Big| \bigcup_{k=1}^N I_k \Big| \leq |G| - |F|_e \leq \varepsilon$$

for any N. We conclude

$$|G \backslash F|_e \leq \sum |I_k| < \varepsilon,$$

which implies that F is measurable.

Finally, for arbitrary closed F we may write F as a countable union of compact sets:

$$F = \bigcup_k \big[F \cap \{|x| \leq k\}\big],$$

which implies the result. □

Next, we prove:

Theorem 3.19. *If E is measurable then E^c is measurable.*

Proof. For each k, let $G_k \supset E$ be open with $|G_k \backslash E|_e < \frac{1}{k}$.

Since G_k^c is closed, it is measurable.

Now set $H = \cup_k G_k^c$, which is measurable and satisfies $H \subset E^c$.

We may now write $E^c = H \cup Z$, with $Z = E^c \backslash H$.

Then

$$Z \subset E^c \backslash G_k^c = G_k \backslash E,$$

so that $|Z|_e < \frac{1}{k}$ for every k. In particular, $|Z|_e = 0$ and hence is measurable.

Thus $E^c = H \cup Z$ is the union of measurable sets, and hence measurable. □

We record some corollaries:

Theorem 3.20. *The countable intersection of measurable sets is measurable.*

Proof. Indeed, its complement is the countable union of measurable sets. □

Theorem 3.21. *If E_1, E_2 are measurable, then $E_1 \backslash E_2$ is measurable.*

Proof. Indeed, $E_1 \backslash E_2 = E_1 \cap E_2^c$. □

The previous results show that the class of measurable subsets contains the emptyset and is closed under (i) complements, (ii) countable unions, and (iii) countable intersections. Such a class is called a **σ-algebra.**

For example, note that if $\{E_k\}$ are measurable, then

$$\limsup E_k = \cap_{j=1}^\infty \cup_{k=j}^\infty E_k \quad \text{and} \quad \liminf E_k = \cup_{j=1}^\infty \cap_{k=j}^\infty E_k$$

are both measurable.

If $\mathcal{C}_1, \mathcal{C}_2$ are two collections of sets, we say $\mathcal{C}_1$ is **contained in** $\mathcal{C}_2$ if

$$S \in \mathcal{C}_1 \implies S \in \mathcal{C}_2.$$

If $\mathcal{F}$ is a family of σ-algebras Σ, we define

$$\cap_{\Sigma \in \mathcal{F}} \Sigma$$

to be the collection of all sets E that belong to every Σ in $\mathcal{F}$. Then $\cap_{\Sigma \in \mathcal{F}} \Sigma$ is a σ-algebra that is contained in every Σ in $\mathcal{F}$.

Given a collection $\mathcal{C}$ of sets in $\mathbb{R}^n$, consider the family $\mathcal{F}$ of all σ-algebras that contain $\mathcal{C}$, and let

$$\mathcal{E} = \cap_{\Sigma \in \mathcal{F}} \Sigma.$$

Then $\mathcal{E}$ is the smallest σ-algebra containing $\mathcal{C}$. [That is, any σ-algebra containing $\mathcal{C}$ contains $\mathcal{E}$.]

The smallest σ-algebra of subsets of $\mathbb{R}^n$ containing all of the open subsets of $\mathbb{R}^n$ is called the **Borel σ-algebra** of $\mathbb{R}^n$, denoted $\mathcal{B}$. The sets in $\mathcal{B}$ are called Borel subsets of $\mathbb{R}^n$ [they include open sets, closed sets, G_δ sets...].

Theorem 3.22. *Every Borel set is measurable.*

Proof. The collection $\mathcal{M}$ of measurable subsets is a σ-algebra that contains the open sets. □

3.3. **A nonmeasurable set.** Not every set is measurable, as we now show.

We present a construction due to Vitali in the setting of $\mathbb{R}$.

The construction relies on the **axiom of choice**: let $\{E_\alpha : \alpha \in A\}$ be a collection of nonempty disjoint sets, where A is an index set. There exists a set consisting of exactly one element from each E_α $(\alpha \in A)$.

We also need the following lemma:

Lemma 3.23. *Let $E \subset \mathbb{R}$ be measurable, with $|E| > 0$. Then the set*

$$D = \{x - y : x, y \in E\}$$

contains an interval centered at 0.

Proof. Let $\varepsilon > 0$ to be chosen below, and let $G \supset E$ be an open set with $|G| < (1+\varepsilon)|E|$.

Write G as a union of nonoverlapping intervals: $G = \cup I_k$.

Defining $E_k = E \cap I_k$, we have that $E = \cup_k E_k$ and that each E_k is measurable.

Furthermore, $\#(E_k \cap E_j) \leq 1$ for $j \neq k$.

Now, we have

$$|G| = \sum |I_k| \quad \text{and} \quad |E| = \sum |E_k|.$$

As $|G| < (1+\varepsilon)|E|$, we must have

$$|I_{k_0}| < (1+\varepsilon)|E_{k_0}| \quad \text{for some} \quad k_0.$$

Choose $\varepsilon = \frac{1}{3}$ and denote $I_0 = I_{k_0}$, $E_0 = E_{k_0}$. Then we have

$$E_0 \subset I_0 \quad \text{with} \quad |E_0| > \tfrac{3}{4}|I_0|.$$

Now, let d satisfy $|d| < \frac{1}{2}|I_0|$ and consider the set $E_0 + d$. We claim that

$$E_0 \cap [E_0 + d] \neq \emptyset.$$

Indeed, if E_0 and $[E_0 + d]$ are disjoint, then

$$\tfrac{3}{2}|I_0| < 2|E_0| = |E_0| + |[E_0 + d]| = |E_0 \cup [E_0 + d]| \leq |I_0| + |d|,$$

contradicting $|d| < \frac{1}{2}|I_0|$.

This implies that for any $|d| < \frac{1}{2}|I_0|$, there exist $x, y \in E_0$ so that $x - y = d$. Thus

$$D_0 = \{x - y : x, y \in E_0\}$$

contains an interval of length $|\frac{1}{2}|I_0|$ centered at the origin, and hence the same is true for $D \supset D_0$. □

Theorem 3.24. *There exist nonmeasurable sets.*

Proof. Define an equivalence relation on $\mathbb{R}$ as follows:

$$x \sim y \quad \text{if and only if} \quad x - y \in \mathbb{Q}.$$

An equivalence class has the form

$$[x] = \{x + r : r \in \mathbb{Q}\}.$$

For any x, y we have either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.

In particular, $[0] = \mathbb{Q}$ and all other classes are disjoint sets in $\mathbb{R}\backslash\mathbb{Q}$.

The number of distinct classes is uncountable, as each $[x]$ is countable but

$$\bigcup_{x \in \mathbb{R}} [x] = \mathbb{R}$$

is uncountable.

Using the axiom of choice, let E be a set with exactly one element from each equivalence class.

Any two points of E must differ by an irrational number, and thus

$$D = \{x - y : x, y \in E\}$$

cannot contain an interval.

Using the lemma, either E is not measurable or $|E| = 0$.

Suppose $|E| = 0$. Then since E has an element from every class and $[x] = \{x + r : r \in \mathbb{Q}\}$, we have

$$\bigcup_{r \in \mathbb{Q}} [E + r] = \bigcup_{x \in \mathbb{R}} [x] = \mathbb{R}.$$

Thus

$$|\mathbb{R}| = \left|\bigcup_{r\in\mathbb{Q}} [E+r]\right| \leq \sum_{r\in\mathbb{Q}} |E+r| \leq \sum_{r\in\mathbb{Q}} |E| = 0,$$

giving a contradiction. We conclude that E is not measurable. $\square$

Corollary 3.25. *If $A \subset \mathbb{R}$ has $|A|_e > 0$, then A contains a nonmeasurable set.*

Proof. Let E be the nonmeasurable set constructed above and set $E_r = E + r$. Then $\{E_r\}_{r\in\mathbb{Q}}$ are disjoint sets with

$$\bigcup_{r\in\mathbb{Q}} E_r = \mathbb{R}.$$

Hence

$$A = \bigcup_{r\in\mathbb{Q}} [A \cap E_r] \quad \text{and} \quad |A|_e \leq \sum_r |A \cap E_r|_e.$$

If $A \cap E_r$ is measurable, then by the lemma above we must have $|A \cap E_r| = 0$ (since the set of differences of elements in E_r cannot contain an interval).

As $|A|_e > 0$, it follows that there exists $r \in \mathbb{Q}$ such that $A \cap E_r$ is not measurable. $\square$

3.4. **Properties of Lebesgue measure.** We turn to general properties of Lebesgue measure.

The definition of measurable concerns approximation by open sets ‘from without’. We next consider approximation by closed sets ‘from within’.

Lemma 3.26. *A set $E \subset \mathbb{R}^n$ is measurable if and only if for every $\varepsilon > 0$, there exists closed $F \subset E$ such that*

$$|E \backslash F|_e < \varepsilon.$$

Proof. Exercise: use the fact that E is measurable if and only if E^c is measurable, along with the definition of measurable. $\square$

Theorem 3.27. *If $\{E_k\}$ is a countable collection of disjoint measurable sets, then*

$$\left|\bigcup_k E_k\right| = \sum_k |E_k|.$$

Proof. First consider the case that each E_k is bounded.

Let $\varepsilon > 0$ and for each k, let $F_k \subset E_k$ be closed with $|E_k \backslash F_k| < \varepsilon 2^{-k}$.

Then $E_k = F_k \cup [E_k \backslash F_k]$, so

$$|E_k| \leq |F_k| + \varepsilon 2^{-k}.$$

Since the E_k are bounded and disjoint, the F_k are compact and disjoint. Thus, by Lemma 3.18, we have

$$\left|\bigcup_{k=1}^m F_k\right| = \sum_{k=1}^m |F_k| \quad \text{for each} \quad m.$$

As

$$\cup_{k=1}^m F_k \subset \cup_{k=1}^m E_k,$$

we deduce

$$\sum_{k=1}^m |F_k| \leq \left|\bigcup_{k=1}^\infty E_k\right| \quad \text{for any} \quad m.$$

Thus

$$\left|\bigcup_{k=1}^\infty E_k\right| \geq \sum_{k=1}^\infty |F_k| \geq \sum_{k=1}^\infty \big[|E_k| - \varepsilon 2^{-k}\big] = \sum_{k=1}^\infty |E_k| - \varepsilon.$$

We conclude

$$\left|\bigcup_{k=1}^\infty E_k\right| \geq \sum_{k=1}^\infty |E_k|.$$

As the reverse inequality is always true, the theorem holds in this case.

For the general case, let I_j be an increasing sequence of intervals with $\cup I_j = \mathbb{R}^n$. Define

$$S_1 = I_1, \quad S_j = I_j \backslash I_{j-1} \quad \text{for} \quad j \geq 2.$$

The sets

$$E_k^j = E_k \cap S_j$$

are bounded, disjoint, and measurable, with

$$E_k = \bigcup_j E_k^j \quad \text{and} \quad \bigcup_k E_k = \bigcup_{k,j} E_k^j.$$

By the case above,

$$\left|\bigcup_k E_k\right| = \left|\bigcup_{k,j} E_k^j\right| = \sum_{k,j} |E_k^j| = \sum_k \big(\sum_j |E_k^j|\big) = \sum_k |E_k|,$$

as desired. □

We have the following corollaries:

Corollary 3.28. *If $\{I_k\}$ is a sequence of nonoverlapping intervals, then*

$$\big|\cup I_k\big| = \sum |I_k|.$$

Proof. As the I_k° are disjoint, we have

$$|\cup I_k| \geq |\cup I_k^\circ| = \sum |I_k^\circ| = \sum |I_k|.$$

As the reverse inequality is always true, this completes the proof. □

Corollary 3.29. *If $E_2 \subset E_1$ (both measurable) and $|E_2| < \infty$, then*

$$|E_1 \backslash E_2| = |E_1| - |E_2|.$$

Proof. Write $E_1 = E_2 \cup E_1 \backslash E_2$. □

We turn to the next property of Lebesgue measure.

Theorem 3.30. *Let $\{E_k\}$ be a sequence of measurable sets.*

(i) *If $E_k \nearrow E$ then $\lim_{k\to\infty} |E_k| = |E|$.*
(ii) *If $E_k \searrow E$ and $|E_k| < \infty$ for some k, then $\lim_{k\to\infty} |E_k| = |E|$.*

Proof. (i) Without loss of generality we may assume $|E_k| < \infty$ for all k.

We write

$$E = \bigcup_k S_k, \quad \text{where} \quad S_1 = E_1, \quad S_k = E_k \backslash E_{k-1} \quad (k \geq 2).$$

Then

$$|E| = \big|\cup S_k\big| = |E_1| + \sum_{k\geq 2} |E_k \backslash E_{k-1}| = |E_1| + \sum_{k\geq 2} \big(|E_k| - |E_{k-1}|\big) = \lim_{k\to\infty} |E_k|,$$

proving (i).

(ii) Without loss of generality, $|E_1| < \infty$. Now write

$$E_1 = E \cup \big[\cup_{k\geq 1} E_k \backslash E_{k+1}\big].$$

Then

$$|E_1| = |E| + \sum_{k\geq 1} [|E_k| - |E_{k+1}|] = |E| + |E_1| - \lim_{k\to\infty} |E_k|,$$

which implies the desired result. □

Remark 3.31. We need to assume $|E_k| < \infty$ for some k. Indeed, suppose $E_k = \{|x| > k\}$. Then $|E_k| = +\infty$ for each k, but $E_k \searrow \emptyset$.

We close this section with an analogous result about outer measure, which we leave as an exercise.

Theorem 3.32. *If $E_k \nearrow E$ then $\lim_{k\to\infty} |E_k|_e = |E|_e$.*

Hint. Approximate by a measurable set and apply the previous theorem.

3.5. **Characteriziations of measurability.** Measurability was defined in terms of approximation 'from without' by an open set. We also saw that measurability is equivalent to a statement about approximation 'from within' by a closed set.

Here we give some other characterizations. Recall that a G_δ set is a countable intersection of open sets, and an F_σ set is a countable union of closed sets.

Theorem 3.33.

(i) *A set E is measurable if and only if $E = H \backslash Z$, where H is G_δ and $|Z| = 0$.*
(ii) *A set E is measurable if and only if $E = H \cup Z$, where H is F_σ and $|Z| = 0$.*

Proof. It suffices to prove the $\implies$ directions.

Suppose E is measurable. For each k, let $G_k \supset E$ be an open set with
$$|G_k \backslash E| < \tfrac{1}{k}.$$
Then $H = \cap_k G_k$ is G_δ, with
$$E \subset H \quad \text{and} \quad |H \backslash E| \leq \inf_k |G_k \backslash E| = 0.$$
Thus (i) follows with $Z = H \backslash E$.

The result in (ii) follows either from taking complements in (i), or by using approximation from within by closed sets. [The details are left as an exercise.] □

The following characterization is also left as an exercise.

Theorem 3.34. *Suppose $|E|_e < \infty$. Then E is measurable if and only if for any $\varepsilon > 0$ we may write*
$$E = [S \cup N_1] \backslash N_2,$$
where S is a finite union of nonoverlapping intervals and $|N_1|_e, |N_2|_e < \varepsilon$.

Finally, the following characterization becomes important when one wants to introduce abstract measure theory. We rely on Theorem 3.10.

Theorem 3.35 (Carathéodory)**.** *A set E is measurable if and only if for every A,*
$$|A|_e = |A \cap E|_e + |A \backslash E|_e.$$

Proof. $\implies$ Suppose E is measurable and let $A \subset \mathbb{R}^n$.

Let $H \supset A$ be G_δ with $|A|_e = |H|$. Write H as the disjoint union of measurable sets
$$H = [H \cap E] \cup [H \backslash E], \quad \text{so that} \quad |H| = |H \cap E| + |H \backslash E|.$$
Then
$$|A|_e = |H \cap E| + |H \backslash E| \geq |A \cap E|_e + |A \backslash E|_e.$$
As the reverse inequality always holds, we deduce
$$|A|_e = |A \cap E|_e + |A \backslash E|_e.$$

$\impliedby$ Suppose E satisfies the 'splitting' condition above.

First consider the case $|E|_e < \infty$. Then choose a G_δ set $H \supset E$ with $|H| = |E|_e$. Then
$$H = E \cup [H \backslash E]$$
and by hypothesis
$$|H| = |H \cap E|_e + |H \backslash E|_e = |E|_e + |H \backslash E|_e.$$
Thus $|H \backslash E|_e = 0$, so that writing
$$E = H \backslash [H \backslash E]$$

(H G_δ and $H\backslash E$ measure zero) shows that E is measurable.

If $|E|_e = \infty$, then we let $B_k = \{|x| \leq k\}$ and $E_k = E \cap B_k$.

Each E_k has finite outer measure, and $E = \cup_k E_k$.

Let $H_k \supset E_k$ be a G_δ set with $|H_k| = |E_k|_e$. By hypothesis,

$$|H_k| = |H_k \cap E|_e + |H_k\backslash E|_e \geq |E_k|_e + |H_k\backslash E|_e.$$

Thus $|H_k\backslash E| = 0$.

Now $H = \cup H_k$ is measurable, $H \supset E$, and $H\backslash E = \cup H_k\backslash E$.

Thus $|H\backslash E| = 0$, and so (writing $E = H\backslash[H\backslash E]$) we conclude that E is measurable. □

Corollary 3.36. *If E is a measurable subset of A, then*

$$|A|_e = |E| + |A\backslash E|_e.$$

Thus if $|E| < \infty$, then $|A\backslash E|_e = |A|_e - |E|$.

We conclude with a strengthening of Theorem 3.10.

Theorem 3.37. *Let $E \subset \mathbb{R}^n$. There exists a G_δ set $H \supset E$ such that for any measurable M,*

$$|E \cap M|_e = |H \cap M|.$$

Proof. Suppose $|E|_e < \infty$ and let $H \supset E$ be a G_δ set with $|E|_e = |H|$.

If M is measurable, then by Carathéodory,

$$|E|_e = |E \cap M|_e + |E\backslash M|_e \quad \text{and} \quad |H| = |H \cap M| + |H\backslash M|.$$

Because all of these terms are finite and $E\backslash M \subset H\backslash M$, we deduce

$$|E \cap M|_e \geq |H \cap M|.$$

However, the reverse inequality is true because $E \subset H$. Thus $|E \cap M|_e = |H \cap M|$.

If $|E|_e = \infty$, then write $E = \cup E_k$ with $|E_k|_e < \infty$ and $E_k \nearrow E$.

By the case above, for each k there is a G_δ set $U_k \supset E_k$ such that

$$|E_k \cap M|_e \equiv |U_k \cap M| \quad \text{for measurable} \quad M.$$

Set $H_k = \cap_{m=k}^{\infty} U_m$, which is measurable and satisfies $H_k \nearrow H := \cup H_k$.

Note that $E_k \subset H_k \subset U_k$, so that

$$|E_k \cap M|_e \equiv |H_k \cap M| \quad \text{for measurable} \quad M.$$

Now, since $E_k \nearrow E$ and $H_k \nearrow H$, we have

$$E_k \cap M \nearrow E \cap M \quad \text{and} \quad H_k \cap M \nearrow H \cap M.$$

Thus, by Theorem 3.32, we have

$$|E \cap M|_e \equiv |H \cap M| \quad \text{for measurable} \quad M.$$

The set H is not G_σ (it is "$G_{\sigma\delta}$"). To obtain a G_δ set, write

$$H = H_1 \backslash Z, \quad H_1 \ G_\delta, \quad |Z| = 0.$$

Then $E \subset H_1$, and since

$$H_1 \cap M = (H \cap M) \cup (Z \cap M),$$

we have

$$|H_1 \cap M| = |H \cap M| = |E \cap M|_e.$$

This completes the proof. □

3.6. **Lipschitz transformations of $\mathbb{R}^n$.** This proofs in this section were skipped in lecture.

Recall the following:

Definition 3.38. A function $T : \mathbb{R}^n \to \mathbb{R}^n$ is called **Lipschitz** if there exists $c > 0$ so that

$$\text{for all} \quad x, y \in \mathbb{R}^n, \quad |T(x) - T(y)| \leq c|x - y|.$$

Lipschitz functions are automatically continuous.

Theorem 3.39. *Lipschitz maps preserve measurability.*

Proof. (i) We first show that Lipschitz maps preserve the class of F_σ sets. Indeed, since any closed set is a countable union of compact sets, and continuous functions preserve compact sets, we have that T maps closed sets into F_σ sets (cf. $T(\cup E_k) = \cup T(E_k)$). The result follows.

(ii) We next show that Lipschitz maps preseve measure zero sets. Indeed, the image of a set with diameter d has diameter at most cd. Thus, there exists $c' > 0$ so that

$$|T(I)| \leq c'|I|$$

for any interval I (note $T(I)$ is F_σ and hence measurable). Now cover any measure zero set by intervals of arbitrarily small measure to conclude the result.

Now if E is measurable, we may write $E = H \cup Z$ where H is F_σ and $|Z| = 0$. Then measurability of $T(E)$ follows from (i) and (ii). □

Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation (and hence represented by an $n \times n$ matrix, also denoted T).

A parallelepiped

$$P = \left\{ \sum_{k=1}^{n} t_k e_k, \quad t_k \in [0,1] \right\}$$

satisfies $|P| = v(P)$ (exercise), and hence $|P|$ is the absolute value of the $n \times n$ determinant of the matrix whose rows are $\{e_1, \ldots, e_n\}$.

Theorem 3.40. *A linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^n$ *satisfies*

$$|T(E)| = |\det T| \cdot |E|$$

for any measurable set E.

Proof. It is a fact of linear algebra that $|T(I)| = |\det T| \cdot |I|$ when I is an interval.

Now for $E \subset \mathbb{R}^n$ and $\varepsilon > 0$, choose intervals $\{I_k\}$ covering E with

$$\sum |I_k| < |E|_e + \varepsilon.$$

Then

$$|T(E)|_e \leq \sum |T(I_k)| = |\det T| \sum |I_k| < \delta(|E|_e + \varepsilon).$$

It follows that

$$|T(E)|_e \leq |\det T| \cdot |E|_e. \tag{3.1}$$

We wish to show that $|T(E)| = |\det T| \cdot |E|$. It suffices to consider $|\det T| > 0$.

Now choose open $G \supset E$ with $|G \backslash E| < \varepsilon$.

Write G as a union of nonoverlapping intervals $\{I_k\}$. Since the $T(I_k)$ are non-overlapping parallelipipeds, we have

$$|T(G)| = \sum |T(I_k)| = |\det T| \sum |I_k| = |\det T| \cdot |G|.$$

Using $E \subset G$ and (3.1),

$$|\det T| \cdot |E| \leq |\det T| \cdot |G| = |T(G)| \leq |T(E)| + |T(G \backslash E)| \leq |T(E)| + \delta\varepsilon,$$

and hence

$$|\det T| \cdot |E| \leq |T(E)|.$$

Combining with (3.1), we conclude $|T(E)| = |\det T| \cdot |E|$. □

4. Lebesgue measurable functions

Reference: Wheeden–Zygmund Chapter 4

4.1. Properties of measurable functions, I.

Definition 4.1. Let $f : E \to \mathbb{R} \cup \{\pm\infty\}$ for some $E \subset \mathbb{R}^n$. We call f **Lebesgue measurable (on E)** if

$$\forall a \in \mathbb{R} \quad \{x \in E : f(x) > a\} \quad \text{is measurable.}$$

We abbreviate the set appearing above by $\{f > a\}$. Note that

$$E = \{f = -\infty\} \cup \Big[\bigcup_{k=1}^{\infty} \{f > -k\}\Big],$$

so that if f is measurable then measurability of E is equivalent to measurability of $\{f = -\infty\}$.

We shall always assume $\{f = -\infty\}$ is measurable, so that we only consider measurable functions defined on measurable sets.

Example 4.1. If $E = \mathbb{R}^n$ and f is continuous, note that $\{f > a\}$ is always open. Thus continuous functions are measurable.

If E is Borel and $\{f > a\}$ is Borel for every a, then f is measurable. In fact, we call f **Borel measurable**.

Theorem 4.2. *Let $f : E \to \mathbb{R} \cup \{\pm\infty\}$ for some measurable E. Then f is measurable if and only if any of the following statements hold for every $a \in \mathbb{R}$:*

(i) *$\{f \geq a\}$ is measurable.*
(ii) *$\{f < a\}$ is measurable.*
(iii) *$\{f \leq a\}$ is measurable.*

Proof. To see that measurability implies (i), write

$$\{f \geq a\} = \bigcap_{k=1}^{\infty} \{f > a - \tfrac{1}{k}\}.$$

To see (i) implies (ii), note $\{f < a\} = \{f \geq a\}^c$.

To see (ii) implies (iii), write

$$\{f \leq a\} = \cap_{k=1}^{\infty} \{f < a + \tfrac{1}{k}\}.$$

Finally, to see (iii) implies measurability, write $\{f > a\} = \{f \leq a\}^c$. $\square$

The following corollary is left as an exercise:

Corollary 4.3. *If f is measurable then $\{f > -\infty\}$, $\{f < \infty\}$, $\{f = \infty\}$, $\{a \leq f \leq b\}$, $\{f = a\}$, and so on, are all measurable.*

Definition 4.4. For $f : E \to \mathbb{R}$ and $S \subset \mathbb{R}$, we define

$$f^{-1}(S) = \{x \in E : f(x) \in S\}.$$

We call this set the **inverse image** of S under f.

Theorem 4.5. *A function f is measurable if and only if $f^{-1}(G)$ is measurable for every open $G \subset \mathbb{R}$.*

Proof. $\Longleftarrow$: If $G = (a, \infty)$, then $f^{-1}(G) = \{f > a\}$. Thus if $f^{-1}(G)$ is measurable for every open G, we have that f is measurable.

$\Longrightarrow$: Suppose f is measurable and $G \subset \mathbb{R}$ is open. Then G can be written in the form $G = \cup_k (a_k, b_k)$.

As $f^{-1}((a_k, b_k)) = \{a_k < f < b_k\}$, we have that $f^{-1}((a_k, b_k))$ is measurable for each k. Thus, using $f^{-1}(G) = \cup_k f^{-1}((a_k, b_k))$, we conclude that $f^{-1}(G)$ is measurable. □

Remark 4.6. The proof above also shows that f is Borel measurable if and only if $f^{-1}(G)$ is Borel measurable for every open $G \subset \mathbb{R}$.

We also have the following characterization:

Theorem 4.7. *Let $A \subset \mathbb{R}$ be dense. Then f is measurable if $\{f > a\}$ is measurable for all $a \in A$.*

Proof. Let $a \in \mathbb{R}$ and choose $\{a_k\} \subset A$ so that $a_k \searrow a$. Then

$$\{f > a\} = \cup_k \{f > a_k\},$$

and hence the theorem follows. □

Definition 4.8. A property $P(x)$ (for $x \in E$) is said to hold **almost everywhere in** E if the set

$$\{x \in E : P(x) \text{ does not hold}\}$$

has measure zero. We write $P(x)$ holds a.e.

For example, if we say $f = 0$ a.e. in E then we mean

$$|\{x : f(x) \neq 0\}| = 0.$$

Theorem 4.9. *If f is measurable and $g = f$ a.e., then g is measurable and*

$$|\{g > a\}| = |\{f > a\}|$$

for all $a \in \mathbb{R}$.

Proof. Let $Z = \{f \neq g\}$. Note that $|Z| = 0$ and

$$\{g > a\} \cup Z = \{f > a\} \cup Z$$

Thus $\{g > a\} \cup Z$ is measurable, and hence (since Z has measure zero) we have $\{g > a\}$ is measurable. This shows that g is measurable, as well as the desired equality of measures. □

Using the previous theorem, we can extend the definition of measurable functions to include those functions that are only defined almost everywhere.

The composition of measurable functions need not be measurable (see the homework). However, we do have the following:

Theorem 4.10. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be continuous and let f be finite a.e. on $E \subset \mathbb{R}^n$. If f is measurable, then so is $\phi \circ f$.*

Proof. Let us assume that f is finite everywhere in E.

By Theorem 4.5, it is enough to show that

$$\{x : \phi(f(x)) \in G\}$$

is measurable for every open $G \subset \mathbb{R}$.

To see this, note that

$$\{x : \phi(f(x)) \in G\} = [\phi \circ f]^{-1}(G) = f^{-1} \circ \phi^{-1}(G).$$

As ϕ is continuous, we have $\phi^{-1}(G)$ is open. As f is measurable, we therefore have $f^{-1} \circ \phi^{-1}(G)$ is measurable. The result follows. □

Example 4.2. For a measurable function f, we have that $|f|$, $|f|^p$ $(p > 0)$, e^{cf}, and so on, are measurable. In fact, this does not require f to be finite a.e.

One also has that $f^+ = \max\{f, 0\}$ and $f^- = -\min\{f, 0\}$ are measurable whenever f is.

Theorem 4.11. *If f and g are measurable, then $\{f > g\}$ is measurable.*

Proof. Write $\mathbb{Q} = \{r_k\}$, so that

$$\{f > g\} = \cup_k \{f > r_k > g\} = \cup_k \big[\{f > r_k\} \cap \{g < r_k\}\big].$$

This implies the result. □

The following is left as an exercise:

Theorem 4.12. *If f is measurable and $\lambda \in \mathbb{R}$, then $f + \lambda$ and λf are measurable.*

We next consider sums of measurable functions, say $f + g$. Sums are not well-defined if they are of the form $\infty + (-\infty)$ or $(-\infty) + \infty$, so we will consider the simpler case that $f + g$ is well-defined everywhere.

Theorem 4.13. *If f and g are measurable and $f+g$ is well-defined everywhere, then $f+g$ is measurable.*

Proof. By the previous resut, $a-g$ is measurable for any $a \in \mathbb{R}$. As

$$\{f+g>a\} = \{f > a-g\},$$

the result follows from Theorem 4.11. □

The previous two theorems show us that the set of measurable functions on a set E forms a vector space.

In the following, we adopt the convention $0 \cdot \pm\infty = \pm\infty \cdot 0 = 0$.

Theorem 4.14. *If f and g are measurable, then so is fg. If $g \neq 0$ a.e., then f/g is measurable.*

Proof. Recall that F^2 is measurable whenever F is. Thus, if f and g are measurable and finite, so is

$$fg = \tfrac{1}{4}[(f+g)^2 - (f-g)^2].$$

We leave the case of infinite f, g as an exercise, along with the second part of the theorem. □

We turn to the question of taking limit operations.

Theorem 4.15. *If $\{f_k\}$ is a sequence of measurable functions, then* $\sup_k f_k$ *and* $\inf_k f_k$ *are measurable.*

Proof. It suffices to prove the result for $\sup_k f_k$, as $\inf_k f_k = -\sup_k(-f_k)$.

To prove measurability of $\sup_k f_k$, we note

$$\{\sup_k f_k > a\} = \cup_k \{f_k > a\},$$

which completes the proof. □

Theorem 4.16. *If $\{f_k\}$ is a sequence of measurable functions, then* $\limsup f_k$ *and* $\liminf f_k$ *are measurable.*

In particular, if $\lim f_k$ *exists a.e., then it is measurable.*

Proof. This follows from the previous result, since

$$\limsup_{k\to\infty} f_k = \inf_j \sup_{k \geq j} f_k, \quad \liminf f_k = \sup_j \inf_{k\geq j} f_k.$$

This completes the proof. □

Notation. Given a set E, we define the **characteristic function of** E (also called the **indicator function of** E) by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{if } x \notin E. \end{cases}$$

We remark that E is measurable if and only if χ_E is.

A **simple function** is a function of the form

$$f(x) = \sum_{k=1}^{N} a_k \chi_{E_k}(x)$$

for some distinct $\{a_k\}$ and disjoint $\{E_k\}$.

A simple function is measurable if and only if each E_k is. [*Exercise.*]

Simple functions play an important role in the theory of measurable functions.

Theorem 4.17.

(i) *Every function can be written as the limit of a sequence of simple functions.*
(ii) *Every nonnegative function can be written as the increasing limit of a sequence of simple functions.*
(iii) *A measurable function can be written as the limit of a sequence of measurable simple functions.*

Proof. We begin with (ii) and suppose $f \geq 0$.

Let $k \in \mathbb{N}$. We partition $[0,k]$ as follows:

$$[0,k] = \bigcup_{j=1}^{k2^k} [(j-1)2^{-k}, j2^{-k}].$$

Let

$$f_k(x) = \begin{cases} \frac{j-1}{2^k} & f(x) \in [(j-1)2^{-k}, j2^{-k}), \quad j = 1, \ldots, k2^k, \\ k & f(x) \geq k. \end{cases}$$

Each f_k is a simple function, defined where f is.

By passing from f_k to f_{k+1}, each subinterval

$$[(j-1)2^{-k}, j2^{-k}]$$

is divided in half. It follows that $f_k \leq f_{k+1}$.

Note also that $f_k \to f$. Indeed, wherever f is finite, we have

$$0 \leq f - f_k \leq 2^{-k},$$

and $f_k \to \infty$ wherever $f = \infty$. This proves (ii).

To prove (i), we write $f = f^+ - f^-$ and apply part (ii) to f^+ and f^-.

Finally for (iii) we may assume $f \geq 0$ (otherwise, write $f = f^+ - f^-$).

In this case, the sets $\{f \in [(j-1)2^{-k}, j2^{-k})\}$ and $\{f \geq k\}$ are all measurable, and the result follows. □

Remark 4.18. If f is bounded, the simple functions converge to f uniformly (exercise).

4.2. **Semicontinuous functions.**

Definition 4.19. Let $f : E \to \mathbb{R}$ and let $x_0 \in E$ be a limit point of E. The function f is **upper semicontinuous at** x_0 if

$$\limsup_{x\to x_0} f(x) \le f(x_0).$$

We write this as f is **usc** at x_0.

Similarly, f is **lower semicontinuous at** x_0 (written **lsc**) if

$$\liminf_{x\to x_0} f(x) \ge f(x_0).$$

Remark 4.20. If $f(x_0) = \infty$, then f is automatically usc at x_0. Similarly, if $f(x_0) = -\infty$, then f is automatically lsc at x_0.

Remark 4.21. For finite f, we have that f is usc at x_0 if for any $M > f(x_0)$, there exists $\delta > 0$ so that

$$\forall x \in E \quad |x - x_0| < \delta \implies f(x) < M.$$

Similarly f is lsc at x_0 if for any $m < f(x_0)$ there exists $\delta > 0$ so that

$$\forall x \in E \quad |x - x_0| < \delta \implies f(x) > m.$$

Equivalently, f is lsc at x_0 if and only if $-f$ is usc at x_0.

Remark 4.22. One can check that f is continuous at x_0 if and only if $|f(x_0)| < \infty$ and f is both usc and lsc at x_0.

Remark 4.23. Usc functions 'jump up'; lsc functions 'jump down'.

Example 4.3. The following functions are usc on $\mathbb{R}$ but not continuous at $x_0 \in \mathbb{R}$:

$$u_1 = \chi_{[x_0,\infty)}, \quad u_2 = \chi_{\{x_0\}}.$$

We call a function **usc relative to** E if it is usc at every limit point of E that belongs to E (similarly for lsc or continuous).

We have the following characterizations. Recall that $A \subset E$ is **relatively open** (in E) if $A = E \cap G$ for some open $G \subset \mathbb{R}$ (and similarly for relatively closed).

Theorem 4.24.

(i) *A function f is usc relative to E if and only if for all $a \in \mathbb{R}$,*

$$\{x \in E : f(x) \ge a\}$$

is relatively closed; this is equivalent to

$$\{x \in E : f(x) < a\}$$

being relatively open.

(ii) *A function f is lsc relative to E if and only if for all $a \in \mathbb{R}$,*
$$\{x \in E : f(x) \leq a\}$$
is relatively closed; this is equivalent to
$$\{x \in E : f(x) > a\}$$
being relatively open.

Proof. It is enough to prove (i).

$\Longrightarrow$: Suppose f is usc relative to E and let $a \in \mathbb{R}$. Suppose $x_0 \in E$ is a limit point of
$$E_a := \{x \in E : f(x) \geq a\}.$$
Then there exist $\{x_k\} \subset E_a$ so that $x_k \to x_0$.

As f is usc at x_0, we have
$$f(x_0) \geq \limsup_{k\to\infty} f(x_k) \geq a.$$
Thus $x_0 \in E_a$, so that E_a is relatively closed.

$\Longleftarrow$: Suppose $x_0 \in E$ is a limit point of E and f is *not* usc at x_0.

Then $f(x_0) < \infty$ and there exist $M \in \mathbb{R}$ and $x_k \in E$ with
$$|x_k - x_0| < \tfrac{1}{k}, \quad f(x_0) < M \leq f(x_k).$$
Thus
$$\{x \in E : f(x) \geq M\}$$
is not relatively closed in E_a: it does not contain all of its limit points that belong to E. □

We have the following corollary, which we leave as an exercise:

Corollary 4.25. *A finite function f is continuous relative to E if and only if all sets of the form*
$$\{x \in E : f(x) \geq a\} \quad \text{and} \quad \{x \in E : f(x) \leq a\}$$
are relatively closed (where $a \in \mathbb{R}$). This is equivalent to all sets of the form
$$\{x \in E : f(x) > a\} \quad \text{and} \quad \{x \in E : f(x) < a\}$$
being relatively open.

We also have the following:

Corollary 4.26. *If E is measurable and $f : E \to \mathbb{R}$ is usc relative to E, then f is measurable. (Similarly if f is lsc or continuous relative to E).*

Proof. Suppose f is usc relative to E. Since
$$E_a := \{x \in E : f(x) \geq a\}$$
is relatively closed for $a \in \mathbb{R}$, we may write $E_a = E \cap F$ for some closed F. Thus E_a is measurable for all $a \in \mathbb{R}$, and so the result follows from Theorem 4.2. □

Remark 4.27. The previous results imply that if f is usc on $\mathbb{R}^n$, then f is Borel measurable (similarly for lsc or continuous). Indeed, we can write

$$\{f > a\} = \cup_{k=1}^{\infty} \{f \geq a + \tfrac{1}{k}\},$$

and hence $\{f > a\}$ is F_σ (and in particular Borel) for every $a \in \mathbb{R}$.

4.3. **Properties of measurable functions, II.** The following result is known as Egorov's theorem:

Theorem 4.28 (Egorov's theorem). *Let $E \subset \mathbb{R}^n$ be of finite measure.*

Suppose $\{f_k\}$ are measurable functions on E that converge a.e. to a finite limit f.

Then for any $\varepsilon > 0$, there exists a closed set $F \subset E$ such that

$$|E \backslash F| < \varepsilon \quad \textit{and} \quad f_k \to f \quad \textit{uniformly on} \quad F.$$

Roughly speaking: a convergent sequence of measurable functions actually converges uniformly, up to sets of arbitrarily small measure.

To see the necessity of the hypotheses, we consider the following example.

Example 4.4. Let $E = \mathbb{R}^n$ and $f_k = \chi_{\{|x|<k\}}$. Then $f_k \to 1$ on $\mathbb{R}^n$, but $\{f_k\}$ does not converge uniformly outside of any bounded set.

We begin with a lemma.

Lemma 4.29. *Let $E, \{f_k\}, f$ be as in Theorem 4.28.*

For any $\varepsilon > 0$ and $\eta > 0$, there exists a closed set $E \subset F$ and $K > 0$ so that

$$|E \backslash F| < \eta \quad \textit{and} \quad |f(x) - f_k(x)| < \varepsilon \quad \textit{for} \quad x \in F \quad \textit{and} \quad k > K.$$

Proof. Let $\varepsilon, \eta > 0$.

For each m, define

$$E_m = \{x : |f(x) - f_k(x)| < \varepsilon \quad \text{for all} \quad k > m\}.$$

That is,

$$E_m = \bigcap_{k>m} \{x : |f(x) - f_k(x)| < \varepsilon\},$$

so that E_m is measurable.

By construction, $E_m \subset E_{m+1}$.

Moreover, since $f_k \to f$ a.e. in E and f is finite, it follows that

$$E_m \nearrow (E \backslash Z), \quad \text{where} \quad |Z| = 0.$$

Thus (by Theorem 3.30) we have

$$|E_m| \to |E \backslash Z| = |E|.$$

Because $|E| < \infty$, this implies $|E \backslash E_m| \to 0$.

Now choose K so that $|E\backslash E_K| < \frac{1}{2}\eta$, and let $F \subset E_K$ be closed and satisfy

$$|E_K \backslash F| < \tfrac{1}{2}\eta.$$

It follows that $|E\backslash F| < \eta$ and $|f - f_k| < \varepsilon$ in F for any $k > K$.

This completes the proof. □

We can now prove Egorov's theorem.

Proof of Egorov's theorem. Let $\varepsilon > 0$.

Using Lemma 4.29, choose closed sets $F_m \subset E$ and integers $K_{m,\varepsilon}$ such that

$$|E\backslash F_m| < \varepsilon 2^{-m} \quad \text{and} \quad |f - f_k| < \tfrac{1}{m} \quad \text{in} \quad F_m \quad \text{for} \quad k > K_{m,\varepsilon}.$$

The set

$$F = \bigcap_{m=1}^{\infty} F_m$$

is closed and satisfies

$$E\backslash F = E\backslash \Big[\bigcap_{m=1}^{\infty} F_m \Big] = \bigcup_{m=1}^{\infty} E\backslash F_m.$$

Thus

$$|E\backslash F| \leq \sum_m |E\backslash F_m| < \varepsilon.$$

It remains to show that the $\{f_k\}$ converge uniformly on F.

To this end, let $\delta > 0$. Then choose $m_0 > \delta^{-1}$.

As $F \subset F_{m_0}$, we have

$$|f - f_k| < \tfrac{1}{m_0} < \delta$$

on F, provided $k > K_{m_0,\varepsilon}$. This completes the proof. □

We next turn to a result known as Lusin's theorem.

Definition 4.30. A function f defined on a measurable set E has property C on E if for any $\varepsilon > 0$, there exists closed $F \subset E$ so that

(i) $|E\backslash F| < \varepsilon$,
(ii) f is continuous relative to F.

Theorem 4.31 (Lusin's theorem)**.** *Let f be a finite function on a measurable set E. Then f is measurable if and only if f has property C on E.*

Roughly speaking: measurable functions are actually continuous, up to sets of arbitrarily small measure.

We begin with a lemma.

Lemma 4.32. *A simple measurable function (on E) has property C (on E).*

Proof. Let

$$f = \sum_{i=1}^{N} a_i \chi_{E_i}$$

be a simple measurable function on E.

Given $\varepsilon > 0$, choose closed $F_j \subset E_j$ with

$$|E_j \backslash F_j| < \tfrac{\varepsilon}{N}.$$

The set

$$F := \bigcup_{j=1}^{N} F_j$$

is closed, with

$$|E \backslash F| = |\cup E_j \backslash \cup F_j| \leq |\cup E_j \backslash F_j| < \varepsilon$$

(where we use $\cup E_j \backslash \cup F_j \subset \cup E_j \backslash F_j$).

We claim that f is continuous relative to F. To see this, suppose that $\{x_k\} \subset F$ satisfies $x_k \to x_0 \in F$. We need to prove that $f(x_k) \to f(x_0)$.

Suppose x_0 belongs to the set F_j. We claim that there exists k_0 so that for all $k > k_0$, we have $x_k \in F_j$.

If not, then we may find a subsequence $\{x_{k_\ell}\} \subset F \backslash F_j$.

By the pigeonhole principle, we may pass to a further subsequence and assume $\{x_{k_\ell}\} \subset F_{j'}$ for some $j' \neq j$.

However, we must have $x_{k_\ell} \to x_0$ (since the original sequence converges).

This gives a contradiction, because then (since F_j' is closed) we have

$$x_0 \in F_j \cap F_{j'} = \emptyset.$$

Now since f is constant on F_j and $x_k \in F_j$ for $k > k_0$, we can conclude that $f(x_k) \to f(x_0)$, as desired. This completes the proof. □

We can now prove Lusin's theorem.

Proof of Lusin's theorem.

$\implies$: Suppose f is measurable. By Theorem 4.17, there exist measurable simple functions $f_k \to f$.

By Lemma 4.32, each f_k has property C on E. Thus given $\varepsilon > 0$, we may find closed sets $F_k \subset E$ so that

$$|E \backslash F_k| < \varepsilon 2^{-(k+1)} \quad \text{and} \quad f_k \quad \text{is continuous relative to} \quad F_k.$$

We now break into two cases. First, suppose $|E| < \infty$.

Then by Egorov's theorem, there exists closed $F_0 \subset E$ so that

$$|E \backslash F_0| < \tfrac{1}{2}\varepsilon \quad \text{and} \quad f_k \to f \quad \text{uniformly on} \quad F_0.$$

Now let

$$F = F_0 \cap \big(\cap_k F_k\big).$$

Then F is a closed set, each f_k is continuous on F, and $\{f_k\}$ converges uniformly to f on F. Thus (by Theorem 1.8), we have that f is continuous on F. Moreover,

$$|E\backslash F| \leq |E\backslash F_0| + \sum_{k=1}^{\infty} |E\backslash F_k| < \varepsilon,$$

and hence (since ε was arbitrary) we conclude that f has property C on E.

Next, suppose $|E| = +\infty$. Then we write

$$E = \cup_{k=1}^{\infty} E_k, \quad E_k = E \cap \{k-1 \leq |x| < k\}.$$

By the above, we may select closed $F_k \subset E_k$ so that

$$|E_k\backslash F_k| < \varepsilon 2^{-k} \quad \text{and} \quad f \quad \text{is continuous on} \quad F_k.$$

Writing

$$F = \cup_{k=1}^{\infty} F_k,$$

we have

$$|E\backslash F| \leq \sum_k |E_k\backslash F_k| < \varepsilon,$$

with f continuous relative to F. In order to conclude that f has property C on E, we need to verify that F is closed.

To this end, suppose $\{x_n\} \subset F$ satisfies $x_n \to x_0$. Then there exists N and k so that

$$k-1 < x_n < k \quad \text{for all} \quad n \geq N,$$

that is, the tail of the sequence belongs to $F_k \cup F_{k-1}$ for some k. As this is a closed set, it follows that $x_0 \in F_k \cup F_{k-1} \subset F$, as was needed to show.

$\Longleftarrow$ Suppose f has property C on E.

For each k, let $F_k \subset E$ be a closed set such that

$$|E\backslash F_k| < \tfrac{1}{k} \quad \text{and} \quad f \quad \text{is continuous on} \quad F_k.$$

Set $H = \cup_{k=1}^{\infty} F_k$. Then

$$H \subset E \quad \text{and} \quad Z = E\backslash H \quad \text{satisfies} \quad |Z| = 0.$$

Now, for any $a \in \mathbb{R}$, we have

$$\begin{aligned}\{x \in E : f(x) > a\} &= \{x \in H : f(x) > a\} \cup \{x \in Z : f(x) > a\} \\ &= \cup_k \{x \in F_k : f(x) > a\} \cup \{x \in Z : f(x) > a\}.\end{aligned}$$

As $\{x \in Z : f(x) > a\}$ has measure zero, measurability of f follows from that of $\{x \in F_k : f(x) > a\}$.

Indeed, f is continuous on F_k, and hence measurability of the latter set follows from Corollary 4.26. This completes the proof. □

4.4. **Convergence in measure.**

Definition 4.33. Let $\{f_k\}$ and f be measurable functions on a set E that are finite a.e. The sequence $\{f_k\}$ **converges in measure on E to f** if

$$\forall \varepsilon > 0 \quad \lim_{k\to\infty} |\{x \in E : |f(x) - f_k(x)| > \varepsilon\}| = 0.$$

We write $f_k \to^m f$.

Convergence in measure appears in many places throughout analysis. We focus on a few fundamental results.

First, we see that pointwise convergence implies convergence in measure (on sets of finite measure).

Theorem 4.34. *Let f, f_k be measurable and finite a.e. on E. If $f_k \to f$ a.e. on E and $|E| < \infty$, then $f_k \to^m f$ on E.*

Proof. Let $\varepsilon, \eta > 0$ and choose F and K as in Lemma 4.29, that is,

$$|E\backslash F| < \eta \quad \text{and} \quad |f(x) - f_k(x)| \le \varepsilon \quad \text{for} \quad x \in F \quad \text{and} \quad k > K.$$

Then for $k > K$, we have

$$\{x \in E : |f(x) - f_k(x)| > \varepsilon\} \subset E\backslash F.$$

Thus

$$\limsup_{k\to\infty} |\{x \in E : f(x) - f_k(x)| > \varepsilon\}| < \eta.$$

As η was arbitrary, the result follows. □

Note that the conclusion may fail if $|E| = \infty$. Indeed, we can once again take the example $f_k = \chi_{\{|x|<k\}}$.

Convergence in measure does not imply pointwise convergence a.e., even on sets of finite measure.

Example 4.5. Let $\{I_k\}$ be a sequence of subintervals of $[0, 1]$ such that

- each point of $[0, 1]$ belongs to infinitely many I_k,
- $\lim_{k\to\infty} |I_k| = 0$.

For example, we could take $I_1 = [0, 1]$, the next two intervals to be the two halves of $[0, 1]$, the next four intervals to be the four quarters of $[0, 1]$, and so on.

If $f_k = \chi_{I_k}$ then $f_k \to^m 0$. However, $\{f_k(x)\}$ does not converge for any $x \in [0, 1]$.

In the direction of a converse to Theorem 4.34, we have the following.

Theorem 4.35. *If $f_k \to^m f$ on E, then there exists a subsequence f_{k_j} such that $f_{k_j} \to f$ a.e. in E.*

Proof. By definition, for each j there exists k_j so that

$$k \geq k_j \implies |\{|f - f_k| > \tfrac{1}{j}\}| < 2^{-j}.$$

Without loss of generality, we may take k_j to be increasing in j.

Define the sets

$$E_j = \{|f - f_{k_j}| > \tfrac{1}{j}\} \quad \text{and} \quad H_m = \cup_{j=m}^{\infty} E_j.$$

By construction,

$$|E_j| < 2^{-j}, \quad \text{and so} \quad |H_m| \leq 2^{-m+1}.$$

Furthermore,

$$|f - f_{k_j}| \leq \tfrac{1}{j} \quad \text{on} \quad E \backslash E_j.$$

It follows that for $j \geq m$, we have

$$|f - f_{k_j}| \leq \tfrac{1}{j} \quad \text{on} \quad E \backslash H_m,$$

and so $f_{k_j} \to f$ pointwise on $E \backslash H_m$ for any m.

Since $|H_m| \to 0$ as $m \to \infty$, we deduce that $f_k \to f$ a.e. in E, as desired. □

Our last result is a Cauchy criterion for convergence in measure.

Theorem 4.36. *A sequence $\{f_k\}$ converges in measure on E if and only if*

$$\forall \varepsilon > 0 \quad \lim_{k,\ell \to \infty} |\{x \in E : |f_k(x) - f_\ell(x)| > \varepsilon\}| = 0.$$

Proof. $\implies$: This direction follows from the fact that

$$\{|f_k - f_\ell| > \varepsilon\} \subset \{|f_k - f| > \tfrac{1}{2}\varepsilon\} \cup \{|f_\ell - f| > \tfrac{1}{2}\varepsilon\},$$

which is perhaps best proved in the contrapositive.

$\Longleftarrow$: Choose an increasing sequence N_j so that $k, \ell \geq N_j$ implies

$$|\{|f_k - f_\ell| > 2^{-j}\}| < 2^{-j}.$$

Then

$$|f_{N_{j+1}} - f_{N_j}| \leq 2^{-j}$$

except for on a set E_j with $|E_j| < 2^{-j}$.

We set $H_i = \cup_{j=i}^{\infty} E_j$, so that

$$|f_{N_{j+1}}(x) - f_{N_j}(x)| \leq 2^{-j} \quad \text{for} \quad j \geq i \quad \text{and} \quad x \notin H_i.$$

Thus

$$\sum_j [f_{N_{j+1}} - f_{N_j}]$$

converges uniformly outside H_i, and hence $\{f_{N_j}\}$ converges uniformly outside H_i for each i.

As

$$|H_i| \leq \sum_{j \geq i} 2^{-j} = 2^{-i+1},$$

we have $|H_i| \to 0$. Thus $\{f_{N_j}\}$ converges a.e. on E to some f.

In fact, we have that $|f - f_{N_j}| \lesssim 2^{-j}$ outside of each H_i , which implies that $\{f_{N_j}\}$ converges in measure to f.

We wish to upgrade this to $f_k \to^m f$ on E. Thus we let $\varepsilon > 0$ and note that

$$\{|f_k - f| > \varepsilon\} \subset \{|f_k - f_{N_j}| > \tfrac{1}{2}\varepsilon\} \cup \{|f_{N_j} - f| > \tfrac{1}{2}\varepsilon\}$$

for any N_j. Now let $\eta > 0$ and (using the Cauchy criterion) select N_j large enough that

$$|\{|f_k - f_{N_j}| > \tfrac{1}{2}\varepsilon\}| < \tfrac{1}{2}\eta \quad \text{for all large} \quad k.$$

Using $f_{N_j} \to^m f$, we may also choose N_j large enough that

$$|\{|f_{N_j} - f| > \tfrac{1}{2}\varepsilon\}| < \tfrac{1}{2}\eta.$$

Thus

$$|\{|f_k - f| > \varepsilon\}| < \eta \quad \text{for all} \quad k \quad \text{large enough.}$$

As η was arbitrary, this completes the proof. □

5. The Lebesgue integral

Reference: Wheeden–Zygmund Chapter 5

5.1. **The integral of a nonnegative function.** Let $f : E \to \mathbb{R}$ be a nonnegative function on some measurable $E \subset \mathbb{R}^n$. We define the **graph of** f **over** E to be

$$\Gamma(f, E) = \{(x, f(x)) \in \mathbb{R}^{n+1} : x \in E, \quad f(x) < \infty\}.$$

We define the **region under** f **over** E to be

$$R(f, E) = \{(x, y) \in \mathbb{R}^{n+1} : x \in E, \quad 0 \leq y \leq f(x)\}$$

where we understand the last interval to be $[0, \infty)$ if $f(x) = \infty$.

If $R(f, E)$ is measurable (as a subset of $\mathbb{R}^{n+1}$), its measure $|R(f, E)|_{n+1}$ is called the **Lebesgue integral of** f **over** E. We write

$$|R(f, E)|_{n+1} = \int_E f(x)\,dx.$$

We may also write

$$\int_E f\,dx \quad \text{or} \quad \int_E f.$$

If one wishes to emphasize the dimensions, one can write

$$\int_E \cdots \int f(x_1, \cdots, x_n)\,dx_1 \cdots dx_n.$$

So far, we have only defined the integral for nonnegative functions. Existence of the integral is equivalent to measurability of $R(f, E)$ and does not require $|R(f, E)|_{n+1}$ to be finite.

Here is a fundamental result about integrability.

Theorem 5.1. *Let f be nonnegative on a measurable set E. Then $\int_E f$ exists if and only if f is measurable.*

In fact, we will only show the $\Longleftarrow$ direction, saving the other direction for later.

We will need several lemmas.

Lemma 5.2. *Let $E \subset \mathbb{R}^n$ and $a \in [0, \infty]$. Set*

$$E_a = \{(x, y) : x \in E, \quad y \in [0, a]\}$$

(where we understand $y \in [0, a)$ if $a = \infty$).

If $E \subset \mathbb{R}^n$ is measurable, then $E_a \subset \mathbb{R}^{n+1}$ is measurable, with

$$|E_a|_{n+1} = a|E|_n.$$

Here and below we take $0 \cdot \infty = 0$.

Proof. First suppose $a < \infty$. If E is any kind of interval, then the result is immediate.

If E is an open set, then we may write it as a disjoint union of partly open intervals, say $E = \cup I_k$. It follows that $E_a = \cup I_{k,a}$ and hence is measurable. In fact, the $I_{k,a}$ are disjoint and so

$$|E_a| = \sum |I_{k,a}| = \sum a|I_k| = a|E|.$$

Next suppose E is G_δ, with $E = \cap_{k=1}^{\infty} G_k$ and $|E| < \infty$.

We may assume $|G_1| < \infty$ and $G_k \searrow E$ (e.g. by writing $E = G_1 \cap (G_1 \cap G_2) \cap \cdots$).

By Theorem 3.30, we have $|G_k| \to |E|$ as $k \to \infty$. Moreover, by the above, we have $G_{k,a}$ is measurable with $|G_{k,a}| = a|G_k|$.

As $G_{k,a} \searrow E_a$, we deduce that E_a is measurable, with

$$|E_a| = \lim_{k\to\infty} |G_{k,a}| = a \lim_{k\to\infty} |G_k| = a|E|.$$

Now if E is *any* measurable set with $|E| < \infty$, then by Theorem 3.33 we may write $E = H \backslash Z$ where $|Z| = 0$ and H is G_δ (and of finite measure).

Now $E_a = H_a \backslash Z_a$, and hence E_a is measurable, with

$$|E_a| = |H_a| = a|H| = a|E|$$

using the above. This completes the proof of $a \in \mathbb{R}$ and $|E| < \infty$.

If $a \in \mathbb{R}$ and $|E| = \infty$, then the result follows from writing E as a disjoint countable union of finite measure sets.

Finally, if $a = \infty$, then choose $\{a_k\} \subset \mathbb{R}$ with $a_k \nearrow \infty$. The result then follows from the fact that $E_{a_k} \nearrow E_\infty$. □

Lemma 5.3. *If f is a nonnegative measurable function on a measurable set E, then* $|\Gamma(f, E)| = 0$.

Proof. Let $\varepsilon > 0$ and set

$$E_k = \{k\varepsilon \leq f < (k+1)\varepsilon\}, \quad k = 0, 1, 2, \ldots.$$

The sets E_k are disjoint and measurable, with

$$\cup_k E_k = \{f < \infty\}.$$

Thus

$$\Gamma(f, E) = \cup_k \Gamma(f, E_k).$$

By Lemma 5.2, we have

$$|\Gamma(f, E_k)| \leq \varepsilon |E_k|,$$

and thus

$$|\Gamma(f, E)|_e \leq \sum |\Gamma(f, E_k)| \leq \varepsilon \sum |E_k| \leq \varepsilon |E|.$$

When $|E| < \infty$, this implies $|\Gamma(f, E)|_e = 0$, giving the result.

If $|E| = \infty$, we write E as the countable union of disjoint sets of finite measure; then $\Gamma(f, E)$ is the countable union of measure zero sets and hence $|\Gamma(f, E)| = 0$. □

Proof of $\Longleftarrow$ direction of Theorem 5.1. Let f be nonnegative and measurable on E.

Let f_k be simple measurable functions such that $f_k \nearrow f$ (cf. Theorem 4.17).

We then have

$$R(f_k, E) \cup \Gamma(f, E) \nearrow R(f, E),$$

and since $\Gamma(f, E)$ has measure zero, it is enough to prove that each $R(f_k, E)$ is measurable.

Fix k and suppose that

$$f_k = \sum a_j \chi_{E_j}.$$

Then

$$R(f_k, E) = \cup_{j=1}^N E_{j,a_j}.$$

Thus $R(f_k, E)$ is measurable (by Lemma 5.2), and the proof is complete. □

We record the following corollary:

Corollary 5.4. *If f is a nonnegative measurable simple function of the form*

$$f = \sum_{j=1}^N a_j \chi_{E_j},$$

then

$$\int_{\cup E_j} f = \sum_{j=1}^N a_j |E_j|.$$

Proof. First, note that $R(f, E) = \cup_{j=1}^N E_{j,a_j}$. As the E_j are measurable and disjoint, so are E_{j,a_j}. Thus by definition of the integral and Lemma 5.2,

$$\int_{\cup E_j} f = \sum_{j=1}^N |E_{j,a_j}| = \sum_{j=1}^N a_j |E_j|.$$

This completes the proof. □

5.2. **Properties of the integral.** We turn to the following theorem.

Theorem 5.5.

(i) *If f and g are measurable and $0 \leq g \leq f$ on E, then*

$$\int_E g \leq \int_E f.$$

In particular, $\int_E \inf f \leq \int_E f$.

(ii) *If f is nonnegative and measurable on E and $\int_E f$ is finite, then f is finite a.e. on E.*
(iii) *Let $E_1 \subset E_2$ be measurable. If f is nonnegative and measurable on E_2, then*

$$\int_{E_1} f \leq \int_{E_2} f.$$

Proof. Items (i) and (iii) follow from the observations that

$$R(g,E) \subset R(f,E) \quad \text{and} \quad R(f,E_1) \subset R(f,E_2).$$

We turn to (ii). Without loss of generality, assume $|E| > 0$. Suppose $f = \infty$ on some $E_1 \subset E$ with $|E_1| > 0$.

Then, using (i) and (iii), we have

$$\int_E f \geq \int_{E_1} f \geq \int_{E_1} a = a|E_1| \quad \text{for all} \quad a \in \mathbb{R},$$

which contradicts that $\int_E f$ is finite. □

We turn to the following convergence result.

Theorem 5.6 (Monotone convergence theorem for nonnegative functions)**.** *Suppose $\{f_k\}$ is a sequence of nonnegative measurable functions such that $f_k \nearrow f$ on E. Then*

$$\int_E f_k \to \int_E f.$$

Proof. First observe that f is measurable (by Theorem 4.15).

Next, since $R(f_k,E) \cup \Gamma(f,E) \nearrow R(f,E)$ and $\Gamma(f,E)$ has measure zero, we deduce

$$|R(f_k,E)| \to |R(f,E)|,$$

which gives the result. □

We next show countable additivity of the integral.

Theorem 5.7. *Suppose f is nonnegative and measurable on E, where E is the countable union of disjoint measurable sets E_j. Then*

$$\int_E f = \sum_j \int_{E_j} f.$$

Proof. The sets $R(f,E_j)$ are disjoint and measurable, and

$$R(f,E) = \cup_j R(f,E_j).$$

Thus the result follows from Theorem 3.27. □

We now record some theorems that are corollaries of these results.

The first provides an alternate definition of the integral that is similar in spirit to the definition of the Riemann integral.

Theorem 5.8. *Let f be nonnegative and measurable on E. Then*

$$\int_E f = \sup\Big(\sum_j \big[\inf_{x\in E_j} f(x)\big]|E_j|\Big),$$

where the supremum is taken over all decompositions $E = \cup_j E_j$ into the disjoint union of finitely many measurable sets.

Proof. Consider such a decomposition $E = \cup_{j=1}^N E_j$. Let

$$g = \sum_{j=1}^N a_j \chi_{E_j}, \quad a_j := \inf_{y\in E_j} f(y).$$

Then by the results above,

$$\sum_{j=1}^N a_j |E_j| = \int_E g \leq \int_E f.$$

As this decomposition was arbitrary, we deduce

$$\sup \sum_j [\inf_{E_j} f]|E_j| \leq \int_E f.$$

We turn to the reverse inequality.

As in the proof of Theorem 4.17, for each $k \geq 1$ we introduce

$$\{E_j^k : j = 0, \dots, k2^k\}$$

by $E_0^k = \{f \geq k\}$ and

$$E_j^k = \{(j-1)2^{-k} \leq f < j2^{-k}\} \quad \text{for} \quad j \geq 1.$$

Then the simple functions

$$f_k = \sum_j [\inf_{E_j^k} f]\chi_{E_j^k}$$

satisfy $0 \leq f_k \nearrow f$. Thus, by the monotone convergence theorem

$$\sum_j [\inf_{E_j^k} f]|E_j^k| = \int_E f_k \to \int_E f.$$

Thus

$$\sup_j [\inf_{E_j} f]|E_j| \geq \int_E f,$$

which completes the proof. □

This result immediately implies the following:

Theorem 5.9. *If f is nonnegative on E and $|E| = 0$, then $\int_E f = 0$.*

We turn to an improvement of Theorem 5.5(i).

Theorem 5.10. *If f and g are measurable on E and $0 \le g \le f$ a.e. on E, then $\int_E g \le \int_E f$.*

In particular, if f, g are nonnegative and measurable on E and $f = g$ a.e., then $\int_E f = \int_E g$.

Proof. We can write $E = A \cup Z$, where A and Z are disjoint and $Z = \{g > f\}$ has measure zero.

Thus,

$$\int_E f = \int_A f + \int_Z f = \int_A f \ge \int_A g = \int_E g.$$

The result follows. □

In light of the previous result, we may consider integrals $\int_E f$ for measurable functions f that are only defined a.e. on E.

Theorem 5.11. *Let f be nonnegative and measurable on E. Then*

$$\int_E f = 0 \iff f = 0 \quad \textit{a.e. in} \quad E.$$

Proof. $\Longleftarrow$: If $f = 0$ a.e. in E, then by Theorem 5.10 we have

$$\int_E f = \int_E 0 = 0.$$

$\Longrightarrow$: Suppose $f \ge 0$ is measurable on E and $\int_E f = 0$. Then for any $\alpha > 0$,

$$\alpha |\{x \in E : f(x) > \alpha\}| = \int_{\{f>\alpha\}} \alpha \le \int_{\{f>\alpha\}} f \le \int_E f = 0.$$

It follows that

$$|\{f > \alpha\}| = 0 \quad \text{for all} \quad \alpha > 0.$$

Writing

$$\{f > 0\} = \cup_k \{f > \tfrac{1}{k}\},$$

the result follows. □

The proof of the theorem above also establishes the following useful inequality:

Corollary 5.12 (Tchebyshev's Inequality)**.** *Let f be nonnegative and measurable on E. For any $\alpha > 0$,*

$$|\{x \in E : f(x) > \alpha\}| \le \tfrac{1}{\alpha} \int_E f.$$

We turn to linearity properties of the integral.

Theorem 5.13 (Linearity, I)**.** *If $f \ge 0$ is measurable on E and $c \ge 0$, then*

$$\int_E cf = c \int_E f.$$

Proof. If f is a simple function, then so is cf and hence the result follows from the formula for integrating simple functions.

For general f, choose simple measurable $0 \leq f_k \nearrow f$. Then $cf_k \nearrow cf$ and

$$\int_E cf = \lim_{k\to\infty} \int_E cf_k = \lim_{k\to\infty} c \int_E f_k = c \int_E f,$$

giving the result. □

Theorem 5.14 (Linearity, II). *If f and g are nonnegative and measurable on E then*

$$\int_E (f+g) = \int_E f + \int_E g.$$

Proof. Suppose

$$f = \sum_{i=1}^N a_i \chi_{A_i} \quad \text{and} \quad g = \sum_{j=1}^M b_j \chi_{B_j}$$

are simple functions. Then

$$f + g = \sum_{i,j} (a_i + b_j) \chi_{A_i \cap B_j}$$

is simple and

$$\begin{aligned} \int_E (f+g) &= \sum_i a_i \sum_j |A_i \cap B_j| + \sum_j b_j \sum_i |A_i \cap B_j| \\ &= \sum_i a_i |A_i| + \sum_j b_j |B_j| = \int_E f + \int_E g. \end{aligned}$$

Now for general f, g, we choose simple measurable $f_k \nearrow f$ and $g_k \nearrow g$. Then $f_k + g_k$ are simple and $f_k + g_k \nearrow f + g$. Thus

$$\int_E (f+g) = \lim \int_E (f_k + g_k) = \lim_k \Big(\int_E f_k + \int_E g_k \Big) = \int_E f + \int_E g,$$

giving the result. □

Corollary 5.15. *Suppose f and g are measurable on E with $0 \leq f \leq g$. If $\int_E f$ is finite, then*

$$\int_E (g - f) = \int_E g - \int_E f.$$

Proof. We have

$$\int_E f + \int_E (g-f) = \int_E g,$$

and hence (since $\int_E f$ is finite) the result follows from subtraction. □

We turn to the following additivity result:

Theorem 5.16. *Suppose f_k are nonnegative and measurable on E. Then*

$$\int_E \sum_{k=1}^{\infty} f_k = \sum_{k=1}^{\infty} \int_E f_k.$$

Proof. The functions $F_N = \sum_{k=1}^{N} f_k$ are nonnegative, measurable, and increase to $\sum_{k=1}^{\infty} f_k$. Thus (by the monotone convergence theorem and finite linearity)

$$\int_E \sum_{k=1}^{\infty} f_k = \lim_{N\to\infty} \int_E F_N = \lim_{N\to\infty} \sum_{k=1}^{N} \int_E f_k = \sum_{k=1}^{\infty} \int_E f_k,$$

which implies the result. □

Monotone convergence allows us to interchange integration and passage to a limit.

We consider other situations in which we can make this interchange. Mere convergence of f_k to f is not enough:

Example 5.1. Let $E = [0,1]$. For $k \geq 1$ let f_k be defined as follows:

For $x \in [0, \frac{1}{k}]$, the graph of f_k consists of the isosceles triangle with height k and base $[0, \frac{1}{k}]$.

For $x \in [\frac{1}{k}, 1]$, $f_k(x) = 0$.

Then $f_k \to 0$ on $[0,1]$, but

$$\int_0^1 f_k = \tfrac{1}{2} k \cdot \tfrac{1}{k} = \tfrac{1}{2}$$

for all k. Thus $\lim \int_0^1 f_k \neq \int_0^1 \lim f_k$.

In the positive direction, we have the following convergence results.

Theorem 5.17 (Fatou's lemma)**.** *If $\{f_k\}$ is a sequence of nonnegative functions on E, then*

$$\int_E \liminf_{k\to\infty} f_k \leq \liminf_{k\to\infty} \int_E f_k.$$

Proof. The integral on the left exists, since the integrand is nonnegative and measurable.

Define the functions

$$g_k = \inf_{n \geq k} f_n.$$

Then $g_k \nearrow \liminf f_k$ and $0 \leq g_k \leq f_k$.

Therefore by Theorem 5.6 (monotone convergence) and Theorem 5.10, we have
$$\int_E g_k \to \int_E \liminf f_k \quad \text{and} \quad \int_E g_k \leq \int_E f_k,$$
so that
$$\int_E \liminf f_k = \lim \int_E g_k \leq \liminf \int_E f_k,$$
which gives the result. □

Corollary 5.18. *Suppose f_k are nonnegative and measurable on E and $f_k \to f$ a.e. on E. If $\int_E f_k \leq M$ for all k, then $\int_E f \leq M$.*

Proof. By Fatou's lemma,
$$\int_E \liminf f_k \leq \liminf \int_E f_k \leq M.$$
Since $\liminf f_k = f$ a.e. in E, the result follows. □

Finally, we have the following:

Theorem 5.19 (Lebesgue dominated convergence theorem for nonnegative functions)**.**

Let $\{f_k\}$ be nonnegative measurable functions on E such that $f_k \to f$ a.e. on E.

Suppose there exists a measurable function ϕ such that $f_k \leq \phi$ a.e. for all k and $\int_E \phi$ is finite. Then
$$\int_E f_k \to \int_E f.$$

Proof. By Fatou's lemma,
$$\int_E f = \int_E \liminf f_k \leq \liminf \int_E f_k.$$
Thus, it suffices to prove
$$\int_E f \geq \limsup \int_E f_k.$$
For this, we apply Fatou's lemma to the nonnegative function $\phi - f_k$, which yields
$$\int_E \liminf(\phi - f_k) \leq \liminf \int_E (\phi - f_k).$$
As $f_k \to f$ a.e., the integrand on the left equals $\phi - f$ a.e. Thus, by linearity,
$$\int_E \liminf(\phi - f_k) = \int_E \phi - \int_E f.$$
On the other hand,
$$\liminf \int_E (\phi - f_k) = \int_E \phi - \limsup \int_E f_k.$$

Hence

$$-\int_E f \leq -\limsup \int_E f_k,$$

giving the desired inequality. □

5.3. **The integral of arbitrary measurable functions.** To define the integral of an arbitrary measurable function f on a set E, we break into positive and negative parts:

$$f = f^+ - f^-,$$

each of which are measurable. We then define

$$\int_E f = \int_E f^+ - \int_E f^-,$$

provided at least one of these integrals is finite. In this case we say that the integral $\int_E f$ exists.

This agrees with the original definition in the case that $f = f^+$.

As before, we can make sense of this definition even when f is only defined a.e.

If $\int_E f$ exists and is finite, we say that f is **Lebesgue integrable**, or simply **integrable**. We write $f \in L(E)$. That is,

$$L(E) = \left\{ f : \int_E f \quad \text{is finite} \right\}.$$

We have the following **triangle inequality**: if $\int_E f$ exists, then

$$\left| \int_E f \right| \leq \int_E f^+ + \int_E f^- = \int_E (f^+ + f^-) = \int_E |f|.$$

Theorem 5.20. *Let f be measurable on E. Then f is integrable if and only if $|f|$ is.*

Proof. By the triangle inequality, $|f| \in L(E) \implies f \in L(E)$.

Suppose $f \in L(E)$. Then

$$\int_E f^+ - \int_E f^-$$

is finite, and hence (since at least one is finite by definition) both are finite. Thus

$$\int_E |f| = \int_E f^+ + \int_E f^-$$

is finite, i.e. $|f| \in L(E)$. □

Many properties of $\int_E f$ follow from results already established for non-negative f.

Theorem 5.21. *If $f \in L(E)$ then f is finite a.e. in E.*

Proof. This follows from the fact that $|f| \in L(E)$ (and Theorem 5.5(ii)). □

Theorem 5.22.

(i) *If both* $\int_E f$ *and* $\int_E g$ *exist and* $f \leq g$ *a.e. in* E*, then* $\int_E f \leq \int_E g$*. In particular, if* $f = g$ *a.e. in* E *then* $\int_E f = \int_E g$*.*
(ii) *If* $\int_{E_2} f$ *exists and* $E_1 \subset E_2$ *is measurable, then* $\int_{E_1} f$ *exists.*

Proof. For (i), note that $f \leq g$ implies $0 \leq f^+ \leq g^+$ and $0 \leq g^- \leq f^-$.
Thus

$$\int_E f^+ \leq \int_E g^+ \quad \text{and} \quad \int_E f^- \geq \int_E g^-.$$

The desired inequality follows from subtraction of these two inequalities.

For (ii), we note that at least one of $\int_{E_2} f^+$ or $\int_{E_2} f^-$ is finite. Thus at least one of $\int_{E_1} f^+$ or $\int_{E_1} f^-$ is finite, and hence $\int_{E_1} f$ exists. □

Theorem 5.23. *If* $\int_E f$ *exists and* $E = \cup_k E_k$ *is a disjoint union of measurable sets, then*

$$\int_E f = \sum_k \int_{E_k} f.$$

Proof. Each $\int_{E_k} f$ exists by the previous theorem.

We write $f = f^+ - f^-$ and use countable additivity for nonnegative functions to write

$$\int_E f = \sum \int_{E_k} f^+ - \sum \int_{E_k} f^-.$$

At least one of these sums is finite, and hence

$$\int_E f = \sum \Big(\int_{E_k} f^+ - \int_{E_k} f^- \Big) = \sum \int_{E_k} f,$$

which completes the proof. □

We leave the following as exercises:

Theorem 5.24. *If* $|E| = 0$ *or if* $f = 0$ *a.e. in* E*, then* $\int_E f = 0$*.*

Theorem 5.25. *If* $\int_E f$ *is defined, then so is* $\int_E (-f)$*, and*

$$\int_E (-f) = -\int_E f.$$

Theorem 5.26. *If* $\int_E f$ *exists and* $c \in \mathbb{R}$*, then* $\int_E (cf)$ *exists, and*

$$\int_E (cf) = c \int_E f.$$

Theorem 5.27. *If* $f, g \in L(E)$*, then* $f + g \in L(E)$*, and*

$$\int_E (f+g) = \int_E f + \int_E g.$$

Remark 5.28. It is not difficult to prove $f+g \in L(E)$ [it follows from the triangle inequality]. To prove the equality, one must consider all the possible sign combinations of f, g.

Remark 5.29. The preceding show that for $f_k \in L(E)$ and $a_k \in \mathbb{R}$,

$$\int_E \sum_{k=1}^N a_k f_k = \sum_{k=1}^N a_k \int_E f_k.$$

Corollary 5.30. *Let f, ϕ be measurable on E, with $f \geq \phi$ and $\phi \in L(E)$. Then*

$$\int_E [f-\phi] = \int_E f - \int_E \phi.$$

Proof. Note that $\int_E f$ exists, since $f^- < \phi^-$ (and hence $\int_E f^-$ is finite).

Since $f - \phi \geq 0$, we have that $\int_E (f-\phi)$ exists.

If $f \in L(E)$, then the result follows by linearity.

If $f \notin L(E)$, then we must have $\int_E f = +\infty$.

As $\phi \in L(E)$, we also have $f - \phi \notin L(E)$, and hence (since $f - \phi \geq 0$) $\int_E f - \phi = +\infty$. Thus the result follows in this case as well. □

It is an interesting question to ask when $fg \in L(E)$. For now, we give only a simple sufficient condition.

Theorem 5.31. *Let $f \in L(E)$ and let g be a measurable function on g such that $|g| \leq M < \infty$ a.e. on E. Then $fg \in L(E)$.*

Proof. Since $|fg| \leq M|f|$ a.e., it follows that

$$\int_E |fg| \leq \int_E M|f| = M \int_E |f|.$$

Thus $fg \in L(E)$. □

Similarly, we have the following:

Corollary 5.32. *If $f \in L(E)$ and $f \geq 0$ and there exist $\alpha, \beta \in \mathbb{R}$ so that $\alpha \leq g \leq \beta$ a.e. in E, then*

$$\alpha \int_E f \leq \int_E fg \leq \beta \int_E f.$$

As before, we will be interested that guarantee

$$\int_E f_k \to \int_E f$$

in the case that $f_k \to f$. In particular, we can prove extensions of the results we established in the case of nonnegative functions.

Theorem 5.33 (Monotone convergence theorem)**.** *Let $\{f_k\}$ be a sequence of measurable functions on E.*

(i) *If* $f_k \nearrow f$ *a.e. on* E *and there exists* $\phi \in L(E)$ *so that* $f_k \geq \phi$ *on* E *for all* k*, then* $\int_E f_k \to \int_E f$.
(ii) *If* $f_k \searrow f$ *a.e. on* E *and there exists* $\phi \in L(E)$ *so that* $f_k \leq \phi$ *on* E *for all* k*, then* $\int_E f_k \to \int_E f$.

Proof. We focus on (i), leaving (ii) as an exercise.

We may assume that $f_k \nearrow f$ everywhere on E. Thus

$$0 \leq f_k - \phi \nearrow f - \phi$$

on E, so that by the monotone convergence theorem for nonnegative functions we have

$$\int_E (f_k - \phi) \to \int_E (f - \phi).$$

Thus, using Corollary 5.30, we deduce

$$\int_E f_k - \int_E \phi \to \int_E f - \int_E \phi,$$

and since $\phi \in L(E)$ the result follows. □

Theorem 5.34 (Uniform convergence theorem)**.** *Let* $f_k \in L(E)$ *and let* $f_k \to f$ *uniformly on* E*, where* $|E| < \infty$*. Then* $f \in L(E)$ *and*

$$\int_E f_k \to \int_E f.$$

Proof. As

$$|f| \leq |f_k| + |f - f_k|$$

and $f_k \to f$ uniformly on E, we have

$$|f| \leq |f_k| + 1$$

on E for all large k, and hence (since $|E| < \infty$) $f \in L(E)$. Thus

$$\begin{aligned} \left| \int_E f - \int_E f_k \right| = \left| \int_E (f - f_k) \right| &\leq \int_E |f - f_k| \\ &\leq |E| \cdot \sup_{x \in E} |f(x) - f_k(x)| \to 0 \quad \text{as} \quad k \to \infty, \end{aligned}$$

which completes the proof. □

Theorem 5.35 (Fatou's lemma)**.** *Let* $\{f_k\}$ *be a sequence of measurable functions on* E*. If there exists* $\phi \in L(E)$ *such that* $f_k \geq \phi$ *on* E *for all* k*, then*

$$\int_E \liminf_{k \to \infty} f_k \leq \liminf_{k \to \infty} \int_E f_k.$$

Proof. Apply Fatou's lemma for nonnegative functions to the sequence $f_k - \phi$. □

Corollary 5.36. *Let $\{f_k\}$ be a sequence of measurable functions on E. If there exists $\phi \in L(E)$ such that $f_k \le \phi$ on E for all k, then*

$$\int_E \limsup_{k\to\infty} f_k \ge \limsup_{k\to\infty} \int_E f_k.$$

Proof. Use Fatou's lemma and the fact that $\liminf(-f_k) = -\limsup f_k$. □

Theorem 5.37 (Lebesgue's dominated convergence theorem)**.** *Let $\{f_k\}$ be a sequence of measurable functions on E such that $f_k \to f$ a.e. on E. If there exists $\phi \in L(E)$ such that $|f_k| \le \phi$ a.e. in E for all k, then $\int_E f_k \to \int_E f$.*

Proof. We have $-\phi \le f_k \le \phi$, and hence

$$0 \le f_k + \phi \le 2\phi$$

a.e. in E. As $2\phi \in L(E)$, we have by the dominated convergence theorem for nonnegative functions that

$$\int_E f_k + \phi \to \int_E f + \phi.$$

The result follows. □

Corollary 5.38 (Bounded convergence theorem)**.** *Let $\{f_k\}$ be a sequence of measurable functions such that $f_k \to f$ a.e. in E. If $|E| < \infty$ and $|f_k| \le M < \infty$ a.e. in E, then $\int_E f_k \to \int_E f$.*

Proof. Take $\phi \equiv M$ and use the dominated convergence theorem. □

Remark 5.39. To extend the notion of Lebesgue integrability to complex-valued functions, we define

$$\int f_1 + i f_2 = \int f_1 + i \int f_2.$$

5.4. **Riemann–Stieltjes and Lebesgue integrals.** This section will be mostly skipped in lecture.

Let f be a measurable function on a set E. We define the **distribution function** of f by

$$\omega(\alpha) = \omega_{f,E}(\alpha) = |\{x \in E : f(x) > \alpha\}|.$$

Here $\alpha \in \mathbb{R}$. This is a decreasing function of α. Note that if we assume that f is finite a.e. and $|E| < \infty$, then

$$\lim_{\alpha\to\infty} \{f > \alpha\} = \{f = \infty\}, \quad \text{so that} \quad \lim_{\alpha\to\infty} \omega(\alpha) = 0$$

and

$$\lim_{\alpha\to-\infty} \omega(\alpha) = |E| < \infty.$$

Thus ω is bounded, and furthermore ω is of bounded variation with variation equal to $|E|$.

In what follows, we let f denote a measurable function, finite a.e. on E, with $|E| < \infty$. We write

$$\omega(\alpha) = \omega_{f,E}(\alpha), \quad \{f > \alpha\},$$

and so on.

Lemma 5.40. *If $\alpha < \beta$ then $|\{\alpha < f \leq \beta\}| = \omega(\alpha) - \omega(\beta)$.*

Proof. This follows from the facts that

$$\{f > \beta\} \subset \{f > \alpha\}, \quad \{\alpha < f \leq \beta\} = \{f > \alpha\} \backslash \{f > \beta\},$$

and $|\{f > \beta\}| < \infty$ (cf. Corollary 3.29). □

We denote

$$\omega(\alpha+) = \lim_{\varepsilon \searrow 0} \omega(\alpha + \varepsilon) \quad \text{and} \quad \omega(\alpha-) = \lim_{\varepsilon \searrow 0} \omega(\alpha - \varepsilon).$$

Lemma 5.41. *The following hold:*

- $\omega(\alpha+) = \omega(\alpha)$ *(i.e. ω is continuous from the right)*
- $\omega(\alpha-) = |\{f \geq \alpha\}|$.

Proof. For $\varepsilon_n \searrow 0$, we get that

$$\{f > \alpha + \varepsilon_n\} \nearrow \{f > \alpha\} \quad \text{and} \quad \{f > \alpha - \varepsilon_n\} \searrow \{f \geq \alpha\}.$$

As these sets have finite measure, we deduce

$$\omega(\alpha + \varepsilon_n) \to \omega(\alpha) \quad \text{and} \quad \omega(\alpha - \varepsilon_n) \to |\{f \geq \alpha\}|.$$

This completes the proof. □

Thus ω is a decreasing function that is continuous from the right. It may have jumps $\omega(\alpha-) - \omega(\alpha)$ or intervals of constancy. We can characterize these situations as follows.

Corollary 5.42. *The following hold:*

(a) $\omega(\alpha-) - \omega(\alpha) = |\{f = \alpha\}|$. *Thus ω is continuous at α if and only if $|\{f = \alpha\}| = 0$.*

(b) *ω is constant on (α, β) if and only if*

$$|\{\alpha < f < \beta\}| = 0.$$

Proof. (a) follows from the fact that

$$|\{f \geq \alpha\}| = |\{f > \alpha\}| + |\{f = \alpha\}|.$$

For (b), we use

$$\omega(\alpha) - \omega(\beta-) = |\{f > \alpha\}| - |\{f \geq \beta\}| = |\{\alpha < f < \beta\}|.$$

This is zero if and only if ω is constant on $[\alpha, \beta)$; using right continuity, this is equivalent to being constant on (α, β). □

We now relate the Lebesgue integral to a Riemann–Stieltjes integral:

Theorem 5.43. *If $a < f \le b$ on E (with $a, b, |E|$ finite), then*

$$\int_E f = -\int_a^b \alpha d\omega(\alpha).$$

Proof. The integral on the left exists because $a, b, |E| < \infty$. The integral on the right exists because $\alpha \mapsto \alpha$ is continuous and $\omega \in BV$.

Now partition $[a, b]$ as $\{\alpha_j\}_{j=0}^k$ and set

$$E_j = \{\alpha_{j-1} < f \le \alpha_j\}.$$

Note E is the disjoint union of the E_j. Thus

$$\int_E f = \sum_{j=1}^k \int_{E_j} f,$$

and

$$\sum_{j=1}^k \alpha_{j-1}|E_j| \le \int_E f \le \int_{j=1}^k \alpha_j |E_j|.$$

However, we have just seen that

$$|E_j| = \omega(\alpha_{j-1}) - \omega(\alpha_j),$$

and hence the sums above are Rieman–Stieltjes sums for $-\int_a^b \alpha d\omega(\alpha)$. Sending the mesh of the partition to zero now yields the claim. □

More generally, if f is measurable on E, then

$$\int_{\{a<f\le b\}} f = -\int_a^b \alpha d\omega(\alpha).$$

In fact, if either $\int_E f$ or $\int_{-\infty}^{\infty} \alpha d\omega(\alpha)$ are finite, then

$$\int_E f = -\int_{-\infty}^{\infty} \alpha d\omega(\alpha).$$

We leave the proof as an exercise.

We call two measurable functions f, g on a set E **equidistributed** (or **equimeasurable**) if

$$\omega_{f,E}(\alpha) = \omega_{g,E}(\alpha) \quad \text{for all} \quad \alpha.$$

We may think of f, g as being rearrangements of eachother. We have the following:

Corollary 5.44. *If f, g are equimeasurable on E and $f \in L(E)$, then $g \in L(E)$ with*

$$\int_E f = \int_E g.$$

Remark 5.45. We now see the difference between Riemann and Lebesgue integration The Riemann integral is defined using partitioning of the domain, while the Lebesgue integral uses partitioning of the range.

In fact, let $f \geq 0$ be measurable and finite a.e. on E, with $|E| < \infty$. Let $\Gamma = \{\alpha_j\}$ be a partition of $[0, \infty)$ by a countable number of points $\alpha_j \to \infty$.

Let $E_k = \{\alpha_k \leq f < \alpha_{k+1}\}$ and $Z = \{f = +\infty\}$. Then

$$|Z| = \infty \quad \text{and} \quad |E| = \sum |E_k|.$$

Define

$$s_\Gamma = \sum \alpha_k |E_k| \quad \text{and} \quad S_\Gamma := \sum \alpha_{k+1} |E_k|.$$

We have the following:

Theorem 5.46. *Let $f \geq 0$ be measurable and finite a.e. on E, with $|E| < \infty$. Then*

$$\int_E f = \lim_{|\Gamma| \to 0} s_\Gamma = \lim_{|\Gamma| \to 0} S_\Gamma.$$

Proof. Without loss of generality, suppose f is finite everywhere.

Given Γ, let ϕ_Γ and ψ_Γ be defined by $\phi_\Gamma = \alpha_k$ in E_k and $\psi_\Gamma = \alpha_{k+1}$ in E_k. Then

$$0 \leq \phi_\Gamma \leq f \leq \psi_\Gamma,$$

and so

$$s_\Gamma = \int_E \phi_\Gamma \leq \int f \leq \int_E \psi_\Gamma = S_\Gamma.$$

If $s_\Gamma < \infty$, then we have

$$0 \leq S_\Gamma - s_\Gamma = \sum (\alpha_{k+1} - \alpha_k)|E_k| \leq |\Gamma||E|,$$

so that $S_\Gamma < \infty$ and $S_\Gamma - s_\Gamma \to 0$ as $|\Gamma| \to \infty$. This implies the result when $\int f < \infty$.

If $\int f = \infty$ then we deduce $S_\Gamma = \infty$ (and $s_\Gamma = \infty$), which gives the result. □

Next, we turn to the following result:

Theorem 5.47. *If $a < f \leq b$ on E (with $|E| < \infty$) and ϕ is continuous on $[a, b]$, then*

$$\int_E \phi(f) = -\int_a^b \phi(\alpha) d\omega(\alpha).$$

Proof. First note that $\phi(f) \in L(E)$, and that (as ϕ is continuous) the Rieman–Stieltjes integral exists.

We write $f = \lim f_k$, where $a < f_k \leq b$ is simple;. In particular, we form partitions $\{\alpha_j^k\}$ of $[a, b]$ with mesh size tending to zero and set

$$f_k(x) = \alpha_j^k \quad \text{for} \quad \alpha_{j-1}^k < f(x) \leq \alpha_j^k.$$

Then $\phi(f_k) \to \phi(f) \in E$. As the $\phi(f_k)$ are uniformly bounded and $|E| < \infty$, the bounded convergence theorem implies

$$\int_E \phi(f_k) \to \int_E \phi(f).$$

However, using that $\phi(f_k)$ is simple, we use Lemma 5.40 to deduce

$$\int_E \phi(f_k) = -\sum_j \phi(\alpha_j^k)[\omega(\alpha_j^k) - \omega(\alpha_{j-1}^k)],$$

giving

$$\int_E \phi(f_k) \to -\int_a^b \phi(\alpha)\, d\omega(\alpha).$$

This completes the proof. □

We also have the following extension: if $\phi(f) \in L(E)$ then

$$\int_E \phi(f) = -\int_{-\infty}^{\infty} \phi(\alpha)\, d\omega,$$

which we leave as an exercise.

In fact if ϕ is continuous and nonnegative then we can write

$$\int_E \phi(f) = -\int_{-\infty}^{\infty} \phi(\alpha)\, d\omega(\alpha)$$

without restricting either side to be finite.

In particular, for any continuous ϕ,

$$\int_E |\phi(f)| = -\int_{-\infty}^{\infty} |\phi(\alpha)|\, d\omega(\alpha).$$

We apply this to the special class of functions $\phi(\alpha) = |\alpha|^p$, $0 < p < \infty$, which gives

$$\int_E |f|^p = -\int_{-\infty}^{\infty} |\alpha|^p d\omega(\alpha).$$

For nonnegative f, this yields

$$\int_E f^p = -\int_0^{\infty} \alpha^p d\omega(\alpha), \tag{5.1}$$

and in general

$$\int_E |f|^p = -\int_0^{\infty} \alpha^p d\omega_{|f|}(\alpha).$$

For $\phi \geq 0$, we may denote by $L_\phi(E)$ the class of measurable functions f such that $\phi(f) \in L(E)$. When $\phi(\alpha) = |\alpha|^p$ ($p \in (0, \infty)$), we write $L_\phi(E) = L^p(E)$. In particular, $L(E) = L^1(E)$.

To complete this section, we continue from (5.1) above. First observe the L^p version of Tchebyshev's inequality (which we leave as an exercise):

$$\omega(\alpha) \leq \tfrac{1}{\alpha^p} \int_{\{f>\alpha\}} f^p, \quad \alpha > 0.$$

Thus for $f \in L^p$ we have $\alpha^p \omega(\alpha)$ bounded. In fact:

Lemma 5.48. *For $f \in L^p$, $\alpha^p \omega(\alpha) \to 0$ as $\alpha \to \infty$.*

Proof. This will follow from Tchebyshev's inequality, once we prove

$$\lim_{\alpha\to\infty}\int_{\{f>\alpha\}} f^p = 0.$$

To this end, let $\alpha_k \to \infty$ and define $f_k = f$ when $f > \alpha_k$, $f_k = 0$ elsewhere.

Then

$$\int_{\{f>\alpha_k\}} f^p = \int_E f_k^p.$$

Since f is finite a.e., we have $f_k \to 0$ a.e.

Moreover, $0 \le f_k^p \le |f|^p \in L(E)$. Thus, the result follows from the dominated convergence theorem. □

Finally, we have the following:

Theorem 5.49. *If $f \ge 0$ and $f \in L^p$, then*

$$\int_E f^p = p\int_0^\infty \alpha^{p-1}\omega(\alpha)\,d\alpha.$$

Proof. First let $0 < a < b < \infty$. Using the integration by parts formula for Riemann–Stieltjes integrals and the fact that $\alpha \mapsto \alpha^p$ is continuously differentiable on $[a, b]$, we find

$$-\int_a^b \alpha^p\,d\omega(\alpha) = -b^p\omega(b) + \alpha^p\omega(a) + p\int_a^b \alpha^{p-1}\omega(\alpha)\,d\alpha.$$

By the lemma above, $b^p\omega(b) \to 0$ as $b \to \infty$, while $\alpha^p\omega(a) \to 0$ follows from $|E| < \infty$. Thus the result follows from sending $a \to 0$ and $b \to \infty$. □

5.5. **Riemann and Lebesgue integrals.** This section will be mostly skipped in lecture.

In the following, we denote the Riemann integral by $(R)\int$ and the Lebesgue integral by $\int$.

Theorem 5.50. *If f is bounded and Riemann integrable on $[a, b]$, then $f \in L([a, b])$ and*

$$\int_a^b f = (R)\int_a^b f.$$

Proof. Let Γ_k be a sequence of partitions of $[a, b]$ with mesh size tending to zero.

For each k, define two simple functions ℓ_k, u_k on $[a, b)$ by taking the lower and upper bounds on each semi-open interval $[x_i^k, x_{i+1}^k]$ (where $\Gamma_k = \{x_i^k\}$).

The functions ℓ_k, u_k are bounded and measurable on $[a, b)$. If L_k, U_k denote the lower/upper Riemann sums of f, then

$$\int_a^b \ell_k = L_k, \quad \int_a^b u_k = U_k.$$

We have $\ell_k \leq f \leq -K$, and if we let Γ_{k+1} be a refinement of Γ_k then ℓ_k is increasing and u_k decreasing.

Writing $\ell = \lim \ell_k$ and $u = \lim u_k$, we have ℓ, u measurable and $\ell \leq f \leq u$.

By the bounded convergence theorem,

$$L_k \to \int_a^b \ell \quad \text{and} \quad U_k \to \int_a^b u.$$

However, because f is Riemann integrable we have

$$L_k, U_k \to (R)\int_a^b f.$$

Thus

$$(R)\int_a^b f = \int_a^b \ell = \int_a^b u.$$

Using that $u - \ell \geq 0$, we deduce $l = f = u$ a.e. in $[a, b]$. Thus f is measurable and $\int f = (R)\int f$. □

Compare this with the Dirichlet function $f(x) = 1$ for $x \in \mathbb{Q} \cap [0, 1]$ and $f(x) = 0$ otherwise. This function is bounded, Lebesgue integrable ($\int f = 0$), but not Riemann integrable.

Here is a useful result:

Theorem 5.51. *Let $f \geq 0$ on $[a, b]$ and Riemann integrable (hence bounded) on each interval $[a + \varepsilon, b]$, where $\varepsilon > 0$. If*

$$I := \lim_{\varepsilon \to 0}(R)\int_{a+\varepsilon}^b f$$

exists and is finite, then $f \in L[a, b]$ and $\int_a^b f = I$.

Proof. The result follows from the monotone convergence theorem, since $\int_{a+\varepsilon}^b f = (R)\int_{a+\varepsilon}^b f$ for each $\varepsilon > 0$. □

On the other hand, one can construct a function f whose improper Riemann integral exists and is finite, but which is not integrable. (The function must not be nonnegative...)

We conclude with the following characterization of Riemann integrable functions:

Theorem 5.52. *A bounded function is Riemann integrable on $[a, b]$ if and only if it is continuous a.e. on $[a, b]$.*

Proof. $\Longrightarrow$: Let f be bounded and Riemann integrable.

Let Γ_k, ℓ_k, u_k be as above. Let Z be the set of measure zero outside of which $\ell = f = u$.

We will show that if x is not a partitioning point of any Γ_k and $x \notin Z$, then f is continuous at x.

If not, then there exists $\varepsilon > 0$ depending on x (but not k) so that $u_k(x) - \ell(k) \geq \varepsilon$. This implies $u(x) - \ell(x) \geq \varepsilon$, which contradicts $x \notin Z$.

$\Longleftarrow$: Let f be bounded and continuous a.e. on $[a, b]$. Let $\{\Gamma'_k\}$ be a sequence of partitions with mesh size tending to zero and define $\ell'_k, u'_k, L'_k, U'_k$ as above.

Because Γ'_{k+1} need not be a refinement of Γ'_k, ℓ'_k and u'_k may not be monotone. However, by continuity, $\ell'_k \to f$ and $u'_k \to f$ a.e.

Thus, by the bounded convergence theorem,

$$\int_a^b \ell'_k, \int_a^b u'_k \to \int_a^b f.$$

Since $L'_k = \int_a^b \ell'_k$ and $U_k = \int_a^b u'_k$, it follows that f is Riemann integrable. $\square$

6. L^p CLASSES

Reference: Wheeden–Zygmund Chapter 8

6.1. **Definition of L^p.**

Let E be a measurable subset of $\mathbb{R}^n$ and $0 < p < \infty$. We define

$$L^p(E) = \{f : \int_E |f|^p < \infty\}$$

and

$$\|f\|_p = \|f\|_{L^p(E)} = \left(\int_E |f|^p\right)^{\frac{1}{p}}.$$

We define $L^\infty(E)$ as follows. We define

$$\operatorname*{ess\,sup}_E f = \inf\{\alpha : |\{x \in E : f(x) > \alpha\}| = 0\},$$

unless $|\{x \in E : f(x) > \alpha\}| > 0$ for all α, in which case we set $\operatorname{ess\,sup}_E f = \infty$.

The **essential supremum** is the smallest number M such that $f(x) \leq M$ a.e. in E.

A function is **essentially bounded** (or **bounded**) on E if $\operatorname{ess\,sup}_E |f|$ is finite. The set of essentially bounded functions on E is denoted $L^\infty(E)$, and we write

$$\|f\|_\infty = \|f\|_{L^\infty(E)} = \operatorname*{ess\,sup}_E |f|.$$

Theorem 6.1. *If $|E| < \infty$ then $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.*

Proof. Let $M = \|f\|_\infty$. For $M' < M$, the set $A := \{|f| > M'\}$ has positive measure. Moreover,

$$\|f\|_p \geq \left(\int_A |f|^p\right)^{1/p} \geq M'|A|^{1/p}.$$

As $|A|^{1/p} \to 1$ when $p \to \infty$, we find

$$\liminf_{p\to\infty} \|f\|_p \geq M',$$

which then implies

$$\liminf_{p\to\infty} \|f\|_p \geq M.$$

On the other hand,

$$\|f\|_p \leq \left(\int_E M^p\right)^{1/p} = M|E|^{1/p},$$

showing $\limsup_{p\to\infty} \|f\|_p \leq M$. This completes the proof. □

This can fail for $|E| = \infty$ (consider e.g. $f(x) \equiv c$).

Theorem 6.2. *If $0 < p_1 < p_2 \leq \infty$ and $|E| < \infty$, then $L^{p_2} \subset L^{p_1}$.*

Proof. Exercise. For $p_2 < \infty$, split f into the sets where $|f| \leq 1$ and $|f| > 1$. ☐

This also can fail if $|E| = \infty$. Consider e.g. $f(x) = x^{-1/p_1}$ on $(1, \infty)$. Then $f \in L^{p_2} \backslash L^{p_1}$ for $p_1 < p_2 < \infty$.

A function can belong to all L^{p_1} with $p_1 < p_2$ but not belong to L^{p_2}. Consider e.g. x^{-1/p_2} on $(0, 1)$, which belongs to L^{p_1} for $p_1 < p_2$ but not to L^{p_2}. Similarly, $\log(1/x)$ is in $L^{p_1}(0, 1)$ for $p_1 < \infty$ but not in L^∞.

If $f \in L^{p_1} \cap L^\infty$ then $f \in L^{p_2}$ for all $p_2 > p_1$. [*Exercise.*]

The spaces L^p are vector spaces, i.e. closed under addition and scalar multiplication. [*Exercise.*]

6.2. **Hölder and Minkowski inequalities.**

Theorem 6.3 (Young's inequality)**.** *Let $y = \phi(x)$ be continuous, real-valued, and strictly increasing for $x \geq 0$, with $\phi(0) = 0$. Writing $x = \psi(y)$ for the inverse of ϕ, then for $a, b > 0$ we have*

$$ab \leq \int_0^a \phi(x)\, dx + \int_0^b \psi(y)\, dy.$$

Equality holds if and only if $b = \phi(a)$.

Proof. One can draw a picture, interpret the integrals as areas under curves, and the result follows. ☐

Set $\phi(x) = x^\alpha$ for some $\alpha > 0$, and hence $\psi(y) = y^{\alpha^{-1}}$. Then Young's inequality says

$$ab \leq \frac{1}{1+\alpha} a^{1+\alpha} + \frac{1}{1+\alpha^{-1}} b^{1+\alpha^{-1}}.$$

Setting $p = 1 + \alpha$ and $p' = 1 + \alpha^{-1}$, this yields

$$ab \leq \tfrac{a^p}{p} + \tfrac{b^{p'}}{p'}$$

for $a, b \geq 0$, $1 < p < \infty$, and $\frac{1}{p} + \frac{1}{p'} = 1$.

Two numbers p, p' satisfying

$$\tfrac{1}{p} + \tfrac{1}{p'} = 1$$

and $p, p' > 1$ are called **conjugate exponent pairs**. In particular, $p' = \frac{p}{p-1}$ and $2' = 2$.

We write $1' = \infty$ and $\infty' = 1$.

Theorem 6.4 (Hölder's inequality)**.** *For $1 \leq p \leq \infty$,*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g_{L^{p'}}.$$

Proof. The case $p \in \{1, \infty\}$ is straightforward, so consider $1 < p < \infty$.

It suffices to consider the case $0 < \|f\|_p, \|g\|_{p'} < \infty$. In this case, define

$$\tilde{f} = \frac{f}{\|f\|_p} \quad \text{and} \quad \tilde{g} = \frac{g}{\|g\|_{p'}}.$$

Then

$$\int_E |\tilde{f}\tilde{g}| \leq \int_E \tfrac{|\tilde{f}|^p}{p} + \tfrac{|\tilde{g}|^{p'}}{p'} = \tfrac{1}{p} + \tfrac{1}{p'} = 1,$$

and rearranging yields the desired inequality. □

When $p = p' = 2$, Hölder's inequality is called the Cauchy–Schwarz inequality:

$$\int_E |fg| \leq \left(\int |f|^2\right)^{\frac{1}{2}} \left(\int |g|^2\right)^{\frac{1}{2}}.$$

In fact, one has the following 'duality' between L^p and $L^{p'}$.

Theorem 6.5. *Let f be real-valued and measurable on E and $1 \leq p \leq \infty$. Then*

$$\|f\|_p = \sup \int_E fg,$$

where the supremum is taken over all real-valued g such that $\|g\|_{p'} \leq 1$ and $\int_E fg$ exists.

Proof. Let us prove this result in the simple case of $f \geq 0$, $1 < p < \infty$ and $0 < \|f\|_p < \infty$, leaving other cases as exercises (or see Wheeden–Zygmund).

By dividing both sides of the equality by $\|f\|_p$, we may assume $\|f\|_p = 1$. Now let $g = f^{p/p'}$. Then one can verify $\|g\|_{p'} = 1$ and $\int_E fg = 1$, which yields the result in this case. □

Another classical inequality for L^p functions is the following:

Theorem 6.6 (Minkowski's inequality)**.** *For $1 \leq p \leq \infty$,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. The cases $p \in \{1, \infty\}$ are straightforward and left as an exercise.

For $1 < p < \infty$, we write

$$\|f + g\|_p^p = \int |f + g|^{p-1}|f + g| \leq \int |f + g|^{p-1}|f| + \int |f + g|^{p-1}|g|.$$

Now, apply Hölder's inequality (noting $p' = \frac{p}{p-1}$) to get

$$\int |f + g|^{p-1}|g| \leq \|f + g\|_p^{p-1}\|g\|_p,$$

and similarly to get

$$\int |f + g|^{p-1}|f| \leq \|f + g\|_p^{p-1}\|f\|_p.$$

Thus
$$\|f+g\|_p^p \leq \|f+g\|_p^{p-1}(\|f\|_p + \|g\|_p),$$
which implies the result. □

Remark 6.7. Minkowski's inequality fails when $p \in (0,1)$: let $f = \chi_{(0,\frac{1}{2})}$ and $g = \chi_{(\frac{1}{2},1)}$. Then $\|f+g\|_p = 1$ but $\|f\|_p + \|g\|_p = 2 \cdot 2^{-1/p} < 1$.

6.3. ℓ^p **classes.** A sequence $a = \{a_k\}$ belongs to ℓ^p if
$$\|a\|_{\ell^p} = \|a\|_p = \Big(\sum_k |a_k|^p\Big)^{1/p} < \infty.$$
This is the definition for $0 < p < \infty$; for $p = \infty$ we set
$$\|a\|_{\ell^\infty} = \sup_k |a_k|.$$

For ℓ^p spaces we have $\ell^{p_1} \subset \ell^{p_2}$ whenever $0 < p_1 < p_2 \leq \infty$. [Exercise.]

One can also construct sequences belonging to ℓ^{p_2} but not ℓ^{p_1} for any $p_1 < p_2$ [exercise].

One can also prove analogues of Hölder's and Minkowski's inequality, i.e.
$$\|ab\|_1 \leq \|a\|_p \|b\|_{p'}, \quad \|a+b\|_p \leq \|a\|_p + \|b\|_p$$
for suitable ranges of exponents.

6.4. **Banach and metric space properties.** A **Banach space** is a normed vector space such that the space is complete with respect to the metric induced by the norm.

Theorem 6.8. *For $1 \leq p \leq \infty$, L^p is a Banach space with norm $\|f\|_p = \|f\|_{L^p}$.*

Remark 6.9. Elements of L^p are identified as equivalence classes of functions that are equal a.e.

Proof. The results we have established so far show that $f \mapsto \|f\|_p$ is a norm and L^p is a vector space. It therefore remains to show that L^p is complete.

Let $\{f_k\}$ be Cauchy in L^p. If $p = \infty$, then
$$|f_k - f_m| \leq \|f_k - f_m\|_\infty$$
a.e. and hence $\{f_k\}$ converges uniformly a.e. to a bounded limit f; it follows that $f_k \to f$ in L^∞.

If $1 \leq p < \infty$, then Tchebyshev's inequality implies
$$|\{|f_k - f_m| > \varepsilon\}| \leq \varepsilon^{-p} \int |f_k - f_m|^p,$$
and hence $\{f_k\}$ is Cauchy in measure. Thus there exists f such that $f_k \to f$ a.e. (cf. Chapter 4). Now for any $\varepsilon > 0$, there exists K such that
$$\|f_k - f_j\|_p < \varepsilon \quad \text{for} \quad k, j > K.$$

Sending $j \to \infty$, we obtain by Fatou's lemma that $\|f - f_k\|_p < \varepsilon$ for $k > K$. Noting that

$$\|f\|_p \le \|f - f_k\|_p + \|f_k\|_p < \infty,$$

it follows that $f \in L^p(E)$, which completes the proof. □

A metric space is **separable** if it has a countable dense subset. Note that L^∞ is *not* separable, since there exist an uncountable set of functions a distance one apart (e.g. $f_t = \chi_{(0,t)}$ in $L^\infty((0,1))$).

Theorem 6.10. *For $1 \le p < \infty$, L^p is separable.*

Sketch of proof. First consider the case $L^p(\mathbb{R}^n)$.

Consider a class of dyadic cubes in $\mathbb{R}^n$ and let D be the set of all finite linear combinations of characteristic functions of these cubes, wth rational coefficients. This is a countable subset of L^p.

To see that D is dense in L^p, we approximate more and more general functions.

First, we can approximate characteristic functions of open sets (since every open set is a countable union of nonoverlapping dyadic cubes).

We can then approximate characteristic functions of G_δ sets, and thus measurable sets of finite measure.

This lets us reach simple functions whose supports have finite measure, which in turn lets us reach nonnegative functions in L^p and finally arbitrary functions in L^p.

To handle $E \subset \mathbb{R}^n$, just work with the restrictions of functions in D to E. □

Recall that we showed Minkowski's inequality fails for $p \in (0,1)$, so that $\|\cdot\|_p$ fails to be a norm. Still we have the following:

Theorem 6.11. *For $0 < p < 1$, L^p is a complete separable metric space with distance*

$$d(f,g) = \|f - g\|_{L^p}^p.$$

Proof. To show that d is a metric, we need to verify the triangle inequality. This follows from the inequality

$$(a+b)^p \le a^p + b^p \quad \text{for} \quad a, b \ge 0, \quad p \in (0,1).$$

To see this, one can divide by a (say) and reduce the inequality to $(1+t)^p \le 1 + t^p$ for $t > 0$, which can be proved with calculus.

Thus

$$|f - g|^p \le |f - h|^p + |h - g|^p,$$

which gives the triangle inequality upon integrating. The proofs that L^p is complete and separable are the same as those for $p \ge 1$. □

We have analogous results for ℓ^p spaces:

Theorem 6.12. *For $p \in [1, \infty]$, ℓ^p is a Banach space. For $p \in [1, \infty)$, ℓ^p is separable, while ℓ^∞ is not separable.*

For $p \in (0, 1)$, ℓ^p is a complete separable metric space with distance $d(a, b) = \|a - b\|_p^p$.

The proofs are left to the reader. We only point out an example to show that ℓ^∞ is not separable: consider the sequences $a = \{a_k\}$ such that $a_k \in \{0, 1\}$. The number of such sequences is uncountable and $\|a - a'\|_{\ell^\infty} = 1$ for any two different such sequences.

We turn to the following continuity property:

Theorem 6.13 (Translations are continuous in L^p)**.** *For $f \in L^p(\mathbb{R}^n)$ with $1 \le p < \infty$, we have*

$$\lim_{|h| \to 0} \|f(x + h) - f(x)\|_p = 0.$$

Proof. Let C_p be the set of $f \in L^p$ so that the conclusion of the theorem holds.

We first note that (a) C_p is closed under finite linear combinations and (b) C_p is closed under strong L^p limits. In fact, these are both consequences of Minkowski's inequality, e.g. if $C_p \ni f_k \to f$ in L^p then we have

$$\begin{aligned} &\|f(x + h) - f(x)\|_p \\ &\quad \le \|f(x + h) - f_k(x + h)\|_p + \|f_k(x + h) - f_k(x)\|_p + \|f_k - f\|_p \\ &\quad = \|f_k(x + h) - f_k(x)\|_p + 2\|f_k - f\|_p, \end{aligned}$$

which implies the result.

Now, the characteristic function of a cube belongs to C_p. As finite linear combinations of characteristic functions of cubes are dense in L^p (cf. the proof of separability of L^p), we have that (a) and (b) imply $L^p \subset C_p$. This completes the proof. □

Remark 6.14. Translation is also continuous in L^p for $p \in (0, 1)$, but it fails for $p = \infty$. Indeed, consider $\chi_{(0,\infty)}$.

6.5. L^2 **and orthogonality.** We can define an **inner product** on $L^2(E)$ by

$$\langle f, g \rangle = \int_E f \bar{g}.$$

Indeed, by Cauchy–Schwarz,

$$|\langle f, g \rangle| \le \|f\|_2 \|g\|_2.$$

In the following we often denote $\|f\|_2$ by $\|f\|$ and omit reference to the set E.

The product $\langle \cdot, \cdot \rangle$ satisfies the properties of an inner product (e.g. linearity in the first variable) and $\|f\| = \sqrt{\langle f, f \rangle}$

If $\langle f, g\rangle = 0$, then we call f and g **orthogonal**. A set $\{\phi_\alpha\}_{\alpha\in A}$ is **orthogonal** if any two of its elements are orthogonal and **orthonormal** if it is orthogonal and $\|\phi_\alpha\| = 1$ for all $\alpha \in A$.

By convention, we always assume that orthogonal sets consist only of nonzero elements.

Theorem 6.15. *Any orthogonal system $\{\phi_\alpha\}$ in L^2 is countable.*

Proof. Suppose $\{\phi_\alpha\}$ is orthonormal. For $\alpha \neq \beta$, we find (using orthogonality)

$$\|\phi_\alpha - \phi_\beta\|^2 = \|\phi_\alpha\|^2 + \|\phi_\beta\|^2 = 2,$$

so that $\|\phi_\alpha - \phi_\beta\| = \sqrt{2}$. Because L^2 is separable, this implies that $\{\phi_\alpha\}$ must be countable. [To see this, argue by contradiction.] □

A collection $\{\psi_k\}_{k=1}^N \subset L^2$ is **linearly independent** if

$$\sum_{k=1}^N a_k\psi_k = 0 \implies a_k \equiv 0$$

An infinite collection of functions is linearly independent if each finite subcollection is.

Theorem 6.16. *If $\{\psi_k\}$ is orthogonal, then it is linearly independent.*

Proof. If

$$\sum a_k\psi_k = 0$$

then taking inner products with ψ_ℓ implies $a_\ell = 0$. □

The **span** of a set $\{\psi_k\}$ is the collection of all finite linear combinations of the ψ_k.

The **Gram-Schmidt** algorithm takes as input a linearly independent set of vectors and produces an orthogonal set of vectors with the same span as the original vectors. It works by taking in $\{\psi_k\}$ and defining

$$\begin{aligned}
\phi_1 &= \psi_1,\\
\phi_2 &= \psi_2 - \frac{\langle\psi_2,\phi_1\rangle}{\langle\phi_1,\phi_1\rangle}\phi_1,\\
\phi_3 &= \psi_3 - \frac{\langle\psi_3,\phi_1\rangle}{\langle\phi_1,\phi_1\rangle}\phi_1 - \frac{\langle\psi_3,\phi_2\rangle}{\langle\phi_2,\phi_2\rangle}\phi_2,
\end{aligned}$$

and so on.

An orthogonal system $\{\phi_k\}$ is **complete** if $\langle f, \phi_k\rangle = 0$ for all k implies $f = 0$.

A set $\{\psi_k\}$ is a **basis** for L^2 if its span is dense in L^2. Noting that any countable dense set in L^2 is a basis, we deduce that L^2 has an orthogonal basis (cf. Gram–Schmidt).

Theorem 6.17. *Any orthogonal basis in L^2 is complete. In particular, there exists a complete orthonormal basis for L^2.*

Proof. Let $\{\psi_k\}$ be an orthonormal basis for L^2. Suppose now that $\langle f, \psi_k\rangle = 0$ for all k. Then

$$\langle f, f\rangle = \langle f, f - \sum_{k=1}^N a_k\psi_k\rangle \quad \text{for all} \quad N \quad \text{and all} \quad a_k.$$

By Cauchy–Schwarz,

$$|\langle f, f\rangle| \leq \|f\| \cdot \|f - \sum_{k=1}^N a_k\psi_k\|.$$

As the term on the right-hand side can be made arbitrarily small, we deduce $f = 0$. □

6.6. **Fourier series and Parseval's formula.** Let $\{\phi_k\}$ be an orthonormal set in L^2. For $f \in L^2$, we define the **Fourier coefficients** of f (with respect to $\{\phi_k\}$) by

$$c_k = \langle f, \phi_k\rangle = \int_E f\bar{\phi}_k.$$

We define the **Fourier series** of f (with respect to $\{\phi_k\}$) by

$$S[f] = \sum_k c_k\phi_k.$$

We abbreviate this by writing $f \sim \sum_k c_k\phi_k$. We define the partial Fourier series by

$$s_N = \sum_{k=1}^N c_k\phi_k.$$

Theorem 6.18. *Let $\{\phi_k\}$ be an orthonormal set in L^2 and $f \in L^2$.*

(i) *Given N, the best L^2 approximation to f using the ϕ_k is given by the partial Fourier series.*

(ii) *(Bessel's inequality) We have $c := \{c_k\} \in \ell^2$ and*

$$\|c\|_{\ell^2} \leq \|f\|_{L^2},$$

where $\{c_k\}$ are the Fourier coefficients of f.

Proof. Fix N and $\gamma := (\gamma_1, \cdots, \gamma_N)$ and consider linear combinations of the form

$$F = F(\gamma) = \sum_{k=1}^N \gamma_k\phi_k.$$

By orthonormality,

$$\|F\|^2 = \sum_{k=1}^N |\gamma_k|^2.$$

Thus, recalling $c_k := \langle f, \phi_k\rangle$, we can write

$$\begin{aligned}\|f - F\|^2 &= \langle f - \sum \gamma_k\phi_k, f - \sum \gamma_k\phi_k\rangle \\ &= \|f\|^2 - \sum_{k=1}^N [\bar\gamma_k c_k + \gamma_k \bar c_k] + \sum_{k=1}^N |\gamma_k|^2 \\ &= \|f\|^2 + \sum_{k=1}^N |c_k - \gamma_k|^2 - \sum_{k=1}^N |c_k|^2.\end{aligned}$$

It follows that

$$\min_\gamma \|f - F(\gamma)\|^2 = \|f\|^2 - \sum_{k=1}^N |c_k|^2$$

and

$$\text{argmin}_\gamma \|f - F(\gamma)\|^2 = (c_1, \cdots, c_N).$$

This proves (i). Furthermore (evaluating at $\gamma = (c_1, \ldots, c_N)$) we can deduce

$$\sum_{k=1}^N |c_k|^2 = \|f\|^2 - \|f - S_N\|^2,$$

which yields Bessel's inequality upon sending $N \to \infty$. □

If equality holds in Bessel's inequality (i.e. $\|c\|_{\ell^2} = \|f\|_{L^2}$), we say f satisfies **Parseval's formula**. From the proof of Bessel's inequality, we deduce the following:

Theorem 6.19. *Parseval's formula holds if and only if $S[f]$ converges to f in L^2.*

We can also use Fourier coefficients to define L^2 functions.

Theorem 6.20 (Riesz–Fischer)**.** *Let $\{\phi_k\}$ be an orthonormal set in L^2 and $\{c_k\} \in \ell^2$. There exists an $f \in L^2$ such that $S[f] = \sum c_k\phi_k$ and f satisfies Parseval's formula.*

Proof. Write $t_N = \sum_{k=1}^N c_k\phi_k$. For $M < N$, orthonormality implies

$$\|t_N - t_M\|^2 = \sum_{k=M+1}^N |c_k|^2.$$

Thus $\{c_k\} \in L^2$ implies $\{t_N\}$ is Cauchy and hence converges to some $f \in L^2$. Now observe for $N \geq k$

$$\int f\bar\phi_k = \int (f - t_N)\bar\phi_k + \int t_N\bar\phi_k = \int (f - t_N)\bar\phi_k + c_k$$

which tends to c_k as $N \to \infty$ by Cauchy–Schwarz and the fact that $t_N \to f$ in L^2. Thus $S[f] = \sum c_k\phi_k$ and $t_N = s_N(f)$. In particular, Parseval's formula follows from the fact that $t_N \to f$ in L^2. □

This result does not guarantee uniqueness. However, one does have uniqueness if the set $\{\phi_k\}$ is *complete*. Indeed, if f and g have the same Fourier coefficients then $f-g$ is perpendicular to each ϕ_k.

We have the following related result:

Theorem 6.21. *An orthonormal system $\{\phi_k\}$ is complete if and only if Parseval's formula holds for every $f \in L^2$.*

Proof. If $\{\phi_k\}$ is complete and $f \in L^2$, then Bessel's inequality implies that the Fourier coefficients $\{c_k\}$ are in ℓ^2. Thus (by Riesz–Fischer) there exists $g \in L^2$ with $S[g] = \sum c_k\phi_k$ and $\|g\|^2 = \sum |c_k|^2$. Because f, g have the same Fourier coefficients and $\{\phi_k\}$ is complete, we get $f = g$ a.e. Thus $\|f\|^2 = \|g\|^2 = \sum |c_k|^2$.

Conversely, if $\langle f, \phi_k\rangle = 0$ for all k and $\|f\|^2 = \sum |\langle f, \phi_k\rangle|^2$, then $\|f\| = 0$ which shows that the $\{\phi_k\}$ are complete. □

Suppose $\{\phi_k\}$ is a complete orthonormal set in L^2 and $f, g \in L^2$. Let $\{\hat{f}_k\}$ and $\{\hat{g}_k\}$ be the Fourier coefficients of f, g. A consequence of Parseval's theorem is the following:

$$\langle f, g\rangle = \sum_k \hat{f}_k \overline{\hat{g}_k}.$$

[Exercise.]

Two metric spaces (X_1, d_1) and (X_2, d_2) are (linearly) isometric if there exists a surjective linear map $T : X_1 \to X_2$ such that

$$d_1(f, g) = d_2(Tf, Tg)$$

for all $f, g \in X_1$.

Theorem 6.22. *All spaces $L^2(E)$ are linearly isometric with ℓ^2 (and hence with each other).*

Proof. Let $\{\phi_k\}$ be a complete orthonormal set in $L^2(E)$. Define $T : L^2(E) \to \ell^2$ by $Tf = \{\langle f, \phi_k\rangle\}$. This maps into ℓ^2 by Bessel's inequality and onto ℓ^2 by Riesz–Fischer. Furthermore it is an isometry by Parseval's formula. □

6.7. **Hilbert spaces.** A **Hilbert space** over $\mathbb{C}$ is a vector space over $\mathbb{C}$ with an inner product that is complete with respect to the metric induced by the inner product.

That is, if (f, g) denotes the inner product, then the norm is defined by $\|f\| = \sqrt{(f, f)}$ and the metric is defined by $d(f, g) = \|f - g\|$.

Recall that the Cauchy–Schwarz inequality holds for any inner product space:

$$|(f, g)| \leq \|f\|\,\|g\| \quad \text{for all} \quad f, g \in H.$$

This is clear for $g = 0$, while for $g \neq 0$ we find $\lambda = -(f, g)\|g\|^{-2}$ and rearrange the inequality

$$0 \leq (f + \lambda g, f + \lambda g).$$

Note that any Hilbert space is also a Banach space.

A Hilbert space is infinite dimensional if it cannot be spanned by a finite number of elements. Two fundamental examples of Hilbert spaces are L^2 and ℓ^2. In fact:

Theorem 6.23. *All separable infinite dimensional Hilbert spaces are linearly isometric with ℓ^2 (and hence with each other).*

Proof. Given a separable Hilbert space H, we may (by Gram–Schmidt) find an infinite orthonormal set $\{e_k\}$ whose span is dense in H. In fact, $\{e_k\}$ is complete, since if $(f, e_k) \equiv 0$ then

$$\left\| f - \sum_{k=1}^{N} a_k e_k \right\|^2 = \|f\|^2 + \sum_{k=1}^{N} |a_k|^2 \geq \|f\|^2.$$

In particular if f were non-zero, the span of $\{e_k\}$ could not be dense.

Bessel's inequality and the Riesz–Fischer theorem hold for $\{e_k\}$. Indeed, for $f \in H$ we set $c_k = (f, e_k)$ and have

$$0 \leq \left\| f - \sum_{k=1}^{N} c_k e_k \right\|^2 = \|f\|^2 - \sum_{k=1}^{N} |c_k|^2,$$

which yields Bessel's inequality upon sending $N \to \infty$. Thus $\{c_k\} \in \ell^2$. The Riesz–Fischer theorem is proved essentially like it was for L^2 and relies on the fact that H is complete.

Finally, the mapping $f \mapsto \{(f, e_k)\}$ yields a linear isometry from H to ℓ^2 (for all the same reasons as before, namely Bessel's inequality, Riesz–Fischer, and Parseval). □

7. Repeated integration

Reference: Wheeden–Zygmund Chapter 6

We return to the theory of Lebesgue integration and consider the question of repeated integration.

For a continuous function f on an interval $I = [a,b] \times [c,d]$, one has

$$\iint_I f(x,y)\,dx\,dy = \int_a^b \left[\int_c^d f(x,y)\,dy \right] dx,$$

with similar formulas in higher dimensions. We first consider extensions of this to the case of Lebesgue integration.

7.1. Fubini's theorem.

We write $x = (x_1, \cdots, x_n)$ for an element of an n-dimensional interval $I_1 = \prod_{i=1}^n [a_i, b_i]$, and similarly let y be a point of an m-dimensional interval $I_2 = \prod_{i=1}^m [c_i, d_i]$.

We may have $I_1 = \mathbb{R}^n$ or $I_2 = \mathbb{R}^m$.

The product $I = I_1 \times I_2$ is an $(n+m)$-dimensional interval containing points of the form (x, y).

A function f on I will be written $f(x,y)$, and its integral $\int_I f$ denoted by $\iint_I f(x,y)\,dx\,dy$.

Theorem 7.1 (Fubini's theorem). *Let $f(x,y) \in L(I)$, with $I = I_1 \times I_2$.*

(i) *For a.e. $x \in I_1$, $y \mapsto f(x,y)$ is measurable and integrable on I_2.*
(ii) *The function $x \mapsto \int_{I_2} f(x,y)\,dy$ is measurable and integrable on I_1, with*

$$\iint_I f(x,y)\,dx\,dy = \int_{I_1} \left[\int_{I_2} f(x,y)\,dy \right] dx.$$

It is enough to consider the case $I_1 = \mathbb{R}^n$ and $I_2 = \mathbb{R}^m$ [for otherwise we may set $f = 0$ outside I]. We drop I_1, I_2, I from the notation. We write $L(dx), L(dy), L(\,dx\,dy)$, and so on.

The strategy of proof is to build up an increasing class of functions for which the result holds.

We say a function $f \in L(dx\,dy)$ for which Fubini's theorem is true has property F.

Lemma 7.2. *Any finite linear combination of functions with property F has property F.*

Proof. This follows from the fact that measurability/integrability are preserved under finite linear combinations. □

Lemma 7.3. *Let $\{f_k\}$ have property F. If $f_k \nearrow f$ or $f_k \searrow f$ and $f \in L(dx\,dy)$, then f has property f.*

Proof. Let us treat the case $f_k \nearrow f$.

By assumption, for each k there exists $Z_k \subset \mathbb{R}^n$ with $|Z_k|_{\mathbb{R}^n} = 0$ and such that $f_k(x,y) \in L(dy)$ for $x \notin Z_k$.

Let $Z = \cup_k Z_k$, so that $|Z|_{\mathbb{R}^n} = 0$. Then for $x \notin Z$, we have by the monotone convergence theorem (in y)

$$h_k(x) = \int f_k(x,y)\,dy \nearrow h(x) = \int f(x,y)\,dy.$$

By assumption, we have $h_k \in L(dx)$ and $f_k \in L(dx\,dy)$, with

$$\iint f_k(x,y)\,dx\,dy = \int h_k(x)\,dx.$$

Thus, again using the monotone convergence theorem we have

$$\iint f(x,y)\,dx\,dy = \int h(x)\,dx.$$

As $f \in L(dx\,dy)$ (by assumption), we have that $h \in L(dx)$, giving that h is finite a.e. (i.e. $y \mapsto f(x,y)$ is integrable for a.e. x). This completes the proof. □

Now let us prove some special cases of Fubini's theorem.

Lemma 7.4. *If $E = \cap_{k=1}^{\infty} G_k \subset \mathbb{R}^{n+m}$ is G_δ and $|G_1| < \infty$, then χ_E has property F.*

Proof. We proceed in several cases.

Case 1. Let $E = J_1 \times J_2$ be a product of bounded open intervals in $\mathbb{R}^n$ and $\mathbb{R}^m$. Then $|E| = |J_1|\,|J_2|$.

For each x, $y \mapsto \chi_E(x,y)$ is measurable, and

$$h(x) := \int \chi_E(x,y)\,dy \implies h(x) = \begin{cases} |J_2| & x \in J_1 \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\int h(x)\,dx = |J_1|\,|J_2|,$$

while

$$\iint \chi_E(x,y)\,dx\,dy = |E| = |J_1|\,|J_2|,$$

giving the lemma in case 1.

Case 2. If E is a subset of the boundary of an interval in $\mathbb{R}^{n+m}$, then for a.e. x the set $\{y : (x,y) \in E\}$ has $\mathbb{R}^m$-measure zero.

Thus $h(x) = \int \chi_E(x,y)\,dy$ satisfies $h = 0$ a.e. and so $\int h(x)\,dx = 0$. As $\iint \chi_E(x,y)\,dx\,dy = 0$, the result follows in the case.

Case 3. If E is a partly open interval then cases 1 and 2 imply χ_E has property F.

Case 4. Let $E \subset \mathbb{R}^{n+m}$ be open and finite measure. Write $E = \cup I_j$, where I_j are disjoint partly open intervals.

Writing $E_k = \cup_{j=1}^k I_j$, we have $\chi_{E_k} = \sum_{j=1}^k \chi_{I_j}$, so that χ_{E_k} has property F by case 3 and the first lemma above.

As $\chi_{E_k} \nearrow \chi_E$, we deduce that χ_E has property F by the second lemma.

Case 5. Now let $E = \cap_{k=1}^\infty G_k$ be G_δ. We may assume $G_k \searrow E$ (by redefining $\tilde{G}_k = \cap_{j=1}^k G_j$, say), so that $\chi_{G_k} \searrow \chi_E$. Now the lemma follows from case 4 and the second lemma above. □

Lemma 7.5. *If $Z \subset \mathbb{R}^{n+m}$ has measure zero, then χ_Z has property F. Thus for a.e. $x \in \mathbb{R}^n$, the set $\{y : (x, y) \in Z\}$ has $\mathbb{R}^m$-measure zero.*

Proof. Let $H \supset Z$ be a G_δ set with $|H| = 0$. Writing $H = \cap G_k$, we may assume G_1 has finite measure. Thus, by the previous lemma

$$\int \left[\int \chi_H(x, y)\, dy \right] dx = \iint \chi_H(x, y)\, dx\, dy = 0.$$

Thus implies

$$|\{y : (x, y) \in H\}| = \int \chi_H(x, y)\, dy = 0 \quad \text{for a.e.} \quad x.$$

As $Z \subset H$, this implies $|\{y : (x, y) \in Z\}| = 0$ for a.e. x.

It follows that for a.e. x, $y \mapsto \chi_Z(x, y)$ is measurable and $\int \chi_Z(x, y)\, dy = 0$.

Thus

$$\int \left[\int \chi_Z(x, y)\, dy \right] dx = 0,$$

which gives the lemma, since $\iint \chi_Z(x, y)\, dx\, dy = |Z| = 0$. □

Lemma 7.6. *If $E \subset \mathbb{R}^{n+m}$ is measurable with finite measure, then χ_E has property F.*

Proof. We write $E = H \backslash Z$ with H G_δ and $|Z| = 0$. If $H = \cap G_k$ then we may assume $|G_1| < \infty$. As $\chi_E = \chi_H - \chi_Z$, the lemma follows from the results above. □

Now we can complete the proof of Fubini's theorem.

Proof of Fubini's theorem. Let $f \in L(dx\, dy)$. We will show that f has property F.

Writing $f = f^+ - f^-$, we may assume by the lemma above that $f \geq 0$.

For $f \geq 0$, there exist measurable simple functions $f_k \nearrow f$ with $f_k \geq 0$.

As each $f_k \in L(dx\, dy)$, by the second lemma above it suffices to show that each f_k has property F.

However, each f_k has the form $f = \sum_j v_j \chi_{E_j}$ for some finite measure sets E_j, and hence the result follows. □

Fubini's theorem shows that for $f \in L(\mathbb{R}^{n+m})$, the function $y \mapsto f(x,y)$ is measurable for almost every $x \in \mathbb{R}^n$. In fact, we don't need $f \in L(\mathbb{R}^{n+m})$:

Theorem 7.7. *Let $f = f(x,y)$ be measurable on $\mathbb{R}^{n+m}$. Then for a.e. $x \in \mathbb{R}^n$, $y \mapsto f(x,y)$ is measurable on $\mathbb{R}^m$. In particular, if $E \subset \mathbb{R}^{n+m}$ is measurable then*

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

is measurable in $\mathbb{R}^m$ for a.e. $x \in \mathbb{R}^n$.

Proof. The two statements are equivalent if $f = \chi_E$ for some measurable $E \subset \mathbb{R}^{n+m}$.

In the case that $f = \chi_E$ write $E = H \cup Z$ where $H \in F_\sigma$ and $|E|_{n+m} = 0$.

Then $E_x = H_x \cup Z_x$ where $H_x \in F_\sigma$ (in $\mathbb{R}^m$) and $|Z_x|_m = 0$ for a.e. x by the results above.

Thus E_x is measurable for a.e. x.

Now for f measurable function on $\mathbb{R}^{n+m}$ and $a \in \mathbb{R}$, define $E(a) = \{(x,y) : f(x,y) > a\}$. Then since $E(a)$ is measurable in $\mathbb{R}^{n+m}$, we have

$$E(a)_x = \{y : f(x,y) \in E(a)\}$$

is measurable for a.e. x. The exceptional set depends on $a \in \mathbb{R}$.

The union Z of all exceptional sets over $a \in \mathbb{Q}$ still has $\mathbb{R}^n$-measure zero. For $x \notin Z$, we have

$$\{y : f(x,y) > a\}$$

is measurable for all rational a, and hence for all $a \in \mathbb{R}$. □

The following can be deduced from the results above by extending functions by zero. It is left as an exercise.

Theorem 7.8. *Let f be measurable on $E \subset \mathbb{R}^{n+m}$. Let $E_x = \{y : (x,y) \in E\}$.*

(i) *For a.e. $x \in \mathbb{R}^n$, $y \mapsto f(x,y)$ is measurable on E_x.*
(ii) *If $f \in L(E)$ then for a.e. $x \in \mathbb{R}^n$, the function $y \mapsto f(x,y)$ is integrable on E_x. Moreover, $x \mapsto \int_{E_x} f(x,y)\,dy$ is integrabile and*

$$\iint_E f(x,y)\,dy\,dy = \int_{\mathbb{R}^n} \left[\int_{E_x} f(x,y)\,dy \right] dx.$$

7.2. **Tonelli's theorem.** Fubini's theorem says finiteness of a multiple integral implies finiteness of the iterated integrals. The converse is false.

Example 7.1. Let I be the unit square in $\mathbb{R}_2$. Let I_1 be the square of sidelength $1/2$ in the lower left corner of I. Let I_2 be the cube of sidelength $\frac{1}{4}$ touching the top right corner of I_1. Let I_3 be the cube of sidelength $\frac{1}{8}$ touching the top right corner of I_2, and so on.

Subdivide each I_k into for equal subsquares, I_k^j, labeled by starting in the bottom left quadrant and proceeding counterclockwise.

For each k, let $f = |I_k|^{-1}$ on the interiors of I_k^1 and I_k^3 and $f = -|I_k|^{-1}$ on the interiors of I_k^2 and I_k^4. Let $f = 0$ on the rest of I.

By construction,

$$\int_0^1 f(x,y)\,dx = 0 \quad \text{for all} \quad y$$

and

$$\int_0^1 f(x,y)\,dy = 0 \quad \text{for all} \quad x.$$

However,

$$\iint_I |f(x,y)|\,dx\,dy = \sum_k \iint_{I_k} |f(x,y)|\,dx\,dy = \sum_k 1 = \infty.$$

Thus finiteness of the iterated integral does not imply finiteness of the multiple integral.

For nonnegative f, we do have the following:

Theorem 7.9 (Tonelli's theorem)**.** *Let $f(x,y)$ be nonnegative and measurable on an interval $I = I_1 \times I_2$. Then for almost every $x \in I_1$, $y \mapsto f(x,y)$ is measurable on I_2. Moreover, $x \mapsto \int_{I_2} f(x,y)\,dy$ is measurable on I_1 and*

$$\iint_I f(x,y)\,dx\,dy = \int_{I_1}\left[\int_{I_2} f(x,y)\,dy\right]dx.$$

Proof. We will use Fubini's theorem.

For $k = 1, 2, \ldots$ define $f_k(x,y) = 0$ if $|(x,y)| > k$ and $f_k(x,y) = \min\{k, f(x,y)\}$ if $|(x,y)| \le k$.

Then $f_k \ge 0$ and $f_k \nearrow f$ on I. Moreover $f_k \in L(I)$ (since f_k is bounded and compactly supported).

Thus Fubini's theorem applies to each f_k.

Measurability of $\int_{I_2} f(x,y)\,dy$ follows from its analogue for f_k.

Further, by monotone convergence, $\int_{I_2} f_k(x,y)\,dy \nearrow \int_{I_2} f(x,y)\,dy$. (Measurability follows from Theorem 7.8.)

Using monotone convergence once again, we have

$$\iint_I f_k(x,y)\,dx\,dy \to \iint_I f(x,y)\,dx\,dy,$$
$$\int_{I_1}\left[\int_{I_2} f_k(x,y)\,dy\right]dx \to \int_{I_1}\left[\int_{I_2} f(x,y)\,dy\right]dx.$$

As $f_k \in L$, the result follows. □

Remark 7.10. Note that the roles of x and y may be interchanged, so that for $f \geq 0$ measurable we have

$$\iint_I f(x,y)\,dx\,dy = \int_{I_1}\int_{I_2} f(x,y)\,dy\,dx = \int_{I_2}\int_{I_1} f(x,y)\,dx\,dy.$$

In particular, finiteness of any one of the three integrals implies that of the other two.

Thus, finiteness of one of these integrals for $|f|$ implies that f is integrable and all of these integrals are equal.

Tonelli's theorem implies that

$$\iint_I f(x,y)\,dx\,dy = \int_{I_1}\left[\int_{I_2} f(x,y)\,dy\right], dx$$

even if $\iint_I f = \pm\infty$ (i.e. if $\iint_I f$ merely exists). This follows from considering $f^{\pm}$ and applying Tonelli's theorem [exercise].

We record one application of Fubini's theorem:

Theorem 7.11. *Let $f \geq 0$ be defined on a measurable set $E \subset \mathbb{R}^n$. If $R(f,E)$ (the region under f over E) is measurable in $\mathbb{R}^{n+1}$, then f is measurable.*

Proof. For $y \in [0,\infty)$,

$$\{x \in E : f(x) \geq y\} = \{x : (x,y) \in R(f,E)\}.$$

As $R(f,E)$ is measurable, it follows that $\{x \in E : f(x) \geq y\}$ is measurable (in $\mathbb{R}^n$) for almost all such y (as measured in $\mathbb{R}^1$).

Thus $\{f(x) \geq y\}$ is measurable for all y in a dense subset of $(0,\infty)$. For $y < 0$, we simply have $\{x \in E : f(x) \geq y\} = E$, which is measurable. Thus f is measurable. □

8. Differentiation

Reference: Wheeden–Zygmund Chapter 7

The main topic of this chapter is an analogue of the fundamental theorem of calculus for the Lebesgue integral.

8.1. The indefinite integral.

Let $A \subset \mathbb{R}^n$ be measurable. We define the **indefinite integral** of $f : A \to \mathbb{R}$ to be

$$F(E) = \int_E f,$$

where $E \subset A$ is measurable. The function F is a **set function**, i.e. a real-valued function on a σ-algebra Σ of measurable sets such that

(i) $F(E) < \infty$ for all $E \subset \Sigma$,
(ii) if $E = \cup_k E_k$ is a union of disjoint $E_k \in E$ then $F(E) = \sum_k F(E_k)$.

Recall that

$$\text{diam}\,(E) := \sup\{|x - y| : x, y \in E\}.$$

A set function F is **continuous** if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \text{diam}\,(E) < \delta \implies |F(E)| < \varepsilon.$$

Example 8.1. Let $F(E) = 1$ whenever E is measurable and $0 \in E$, and let $F(E) = 0$ otherwise. Then F is not continuous.

A set function F is **absolutely continuous** (with respect to Lebesgue measure) if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : |E| < \delta \implies |F(E)| < \varepsilon.$$

Absolutely continuous set functions are automatically continuous; however, the converse is false.

Example 8.2. Let $A = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ and $D = \{(x, x) : x \in [0, 1]\}$. Consider the σ-algebra of measurable $E \subset A$ such that $E \cap D$ is 'linearly' measurable, and let $F(E)$ be the linear measure of $E \cap D$. Then F is continuous, but not absolutely continuous: there are sets E containing a fixed segment of D with arbitrarily small $\mathbb{R}^2$-measure.

Theorem 8.1. *If $f \in L(A)$ then its indefinite integral is absolutely continuous.*

Proof. Without loss of generality, assume $f \geq 0$ (otherwise consider $f^{\pm}$).

For any k we may write $f = g + h$, where $g = \min\{f, k\}$.

Now, let $\varepsilon > 0$. Choose k large enough that [with h as above] we have

$$0 \leq \int_A h < \tfrac{1}{2}\varepsilon,$$

and hence $0 \leq \int_E h < \frac{1}{2}\varepsilon$ for every measurable $E \subset A$.

[This uses the fact that $\int_{f>k}[f - k] \leq \int_{f>k} f \to 0$ as $k \to \infty$.]

As $0 \le g \le k$, we have $0 \le \int_E g \le k|E| < \frac{1}{2}\varepsilon$ if $|E|$ is small enough.

Thus

$$0 \le \int_E f < \varepsilon \quad \text{for} \quad |E| \quad \text{small enough.}$$

□

In fact, if $F(E)$ is an absolutely continuous set function, then there exists an integrable function f such that $F(E) = \int_E f$ for measurable sets E. This is known as the Radon–Nikodym theorem.

8.2. **Lebesgue differentiation theorem.** In this section we let Q denote an n-dimensional cube with edges parallel to the coordinate axes.

Theorem 8.2. *Let $f \in L(\mathbb{R}^n)$. Then its indefinite integral is differentiable with derivative f almost everywhere, in the following sense:*

$$\lim_{Q \searrow x} \frac{1}{|Q|} \int_Q f(y)\, dy = f(x).$$

Here $Q \searrow x$ means we take the limit over any sequence Q_k of cubes containing x with $|Q_k| \to 0$.

Remark 8.3. In the case of $n = 1$, this is equivalent to

$$\lim_{h \to 0} \frac{1}{2h} \int_{x-h}^{x+h} f(y)\, dy = f(x),$$

which is essentially equivalent to $\frac{d}{dx} \int_a^x f(y)\, dy = f(x)$.

Remark 8.4. If f is continuous, the theorem is proven as follows:

$$\left| \frac{1}{|Q|} \int_Q f(y)\, dy - f(x) \right| = \left| \frac{1}{|Q|} \int_Q [f(y) - f(x)]\, dy \right| \le \sup_{y \in Q} |f(y) - f(x)| \to 0$$

as $Q \searrow x$.

The strategy will then be to approximate $f \in L(\mathbb{R}^n)$ by continuous functions. We begin with the following:

Lemma 8.5. *For $f \in L(\mathbb{R}^n)$, there exists a sequence C_k of continuous functions with compact support so that*

$$\int_{\mathbb{R}^n} |f - C_k|\, dx \to 0 \quad \textit{as} \quad k \to \infty.$$

Proof. Let A be the set of $f \in L(\mathbb{R}^n)$ such that the theorem holds.

To begin, note that (1) A is closed under finite linear combinations.

Next, we show that (2) if $\{f_k\} \subset A$ and $\int |f - f_k| \to 0$ then $f \in A$. To see this, note that f is necessarily integrable (by the triangle inequality). Now, given $\varepsilon > 0$, choose k_0 so that

$$\int |f - f_{k_0}| < \tfrac{1}{2}\varepsilon.$$

Now choose a continuous function C with compact support such that

$$\int |f_{k_0} - C| < \tfrac{1}{2}\varepsilon.$$

Thus $\int |f - C| < \varepsilon$. It follows that $f \in A$.

Now we prove the lemma (i.e. $L(\mathbb{R}^n) \subset A$). Writing $f = f^+ - f^-$, we can use (1) to reduce to the case $f \geq 0$.

Thus there exist nonnegative simple functions $f_k \nearrow f$. In particular, $f_k \in L(\mathbb{R}^n)$ and

$$\int |f - f_k| \to 0.$$

Thus, by (2), we may assume that $f \in L(\mathbb{R}^n)$ is a nonnegative simple function.

Using (1) again, we can reduce to $f = \chi_E$ with $|E| < \infty$.

Let $\varepsilon > 0$ and choose open $G \supset E$ with $|G \backslash E| < \varepsilon$. Then

$$\int |\chi_G - \chi_E| = |G \backslash E| < \varepsilon,$$

and hence we may assume that $f = \chi_G$ for some open G with $|G| < \infty$.

Now write $G = \cup I_k$ where I_k are disjoint partly open intervals.

Set $f_N = \chi_{\cup_{k=1}^N I_k}$. Then

$$\int |f - f_N| = \sum_{k=N+1}^{\infty} |I_k| \to 0 \quad \text{as} \quad N \to \infty,$$

since $\sum_{k=1}^{\infty} |I_k| = |G| < \infty$.

Therefore by (2) it is enough to show that each $f_N \in A$. But by (1), this reduces to proving that $\chi_I \in A$ for any interval I.

Indeed, if I^* is an interval contaniing I in its interior, with $|I^* \backslash I| < \varepsilon$, then we define C to be a continuous function taking values in $[0, 1]$, equal to 1 on I and 0 outside I^*. Then

$$\int |\chi_I - C| \leq |I^* \backslash I| < \varepsilon,$$

showing that $\chi_I \in A$. This completes the proof. □

Another natural object of study will be the **Hardy–Littlewood maximal function**

$$f^*(x) = \sup \tfrac{1}{|Q|} \int_Q f(y)\, dy,$$

where the supremum is over all Q with center x.

Note that:

- $0 \le f^*(x) \le \infty$
- $(f+g)^* = f^* + g^*$
- $(cf)^* = |c| f^*$.

If $f^*(x_0) > \alpha$ for some $x_0 \in \mathbb{R}^n$ and $\alpha > 0$ then because indefinite integrals are absolutely continuous, we have that $f^*(x) > \alpha$ for x near x_0. This proves lower semicontinuity (and hence measurability) of f^*.

We leave as an exercise that f^* is not integrable unless $f = 0$ a.e. However, we will be able to show that f^* is in "weak $L(\mathbb{R}^n)$", which means

$$\exists C > 0 : |\{|f| > \alpha\}| \le \tfrac{C}{\alpha} \quad \text{for all} \quad \alpha > 0.$$

(Any function in $L(\mathbb{R}^n)$ is in weak $L(\mathbb{R}^n)$ by Tchebyshev's inequality. The function $|x|^{-n}$ is in weak $L(\mathbb{R}^n)$ but not $L(\mathbb{R}^n)$.)

Lemma 8.6 (Hardy–Littlewood)**.** *If $f \in L(\mathbb{R}^n)$, then f^* is in weak $L(\mathbb{R}^n)$. In fact, there exists c (independent of f, α) so that*

$$|\{f^* > \alpha\}| \le \tfrac{c}{\alpha} \int_{\mathbb{R}^n} |f|\, dx.$$

for all $\alpha > 0$.

To prove this, we need the following simple form of the Vitali covering lemma:

Lemma 8.7 (Vitali)**.** *Let $E \subset \mathbb{R}^n$ with $|E| < \infty$. Let K be a collection of (open) cubes covering E. There exists $\beta = \beta(n) > 0$ and $\{Q_j\}_{j=1}^N \subset K$ so that*

$$\sum_{j=1}^N |Q_j| \ge \beta |E|.$$

Proof. Without loss of generality, we may assume E is compact (e.g. by approximating from within by a closed set).

By compactness, we may assume $K_1 := K$ is a finite collection of cubes. Let Q_1 be a cube of largest sidelength.

Write $K_1 = K_2 \cup K_2'$, where K_2 contains the cubes in K_1 disjoint from Q_1. Let Q_1^* be the cube concentric with Q_1 with thrice the sidelength. Then every cube in K_2' is contained in Q_1.

Let Q_2 be the largest cube in K_2, and repeat this construction (writing $K_2 = K_3 \cup K_3'$ and defining Q_2^*.

This process terminates after finitely many steps (once $K_N = \emptyset$) and yields $\{Q_j\}_{j=1}^N \subset K$ and $\{Q_j^*\}_{j=1}^N$ such that

$$E \subset \cup_{j=1}^N Q_j^*.$$

Thus

$$|E| \leq \sum_{j=1}^N |Q_j^*| = 3^n \sum_{j=1}^N |Q_j|.$$

The result follows.

□

Remark 8.8. One can prove Lemma 8.7 without assuming that E is measurable, but the proof is more complicated. There are also more refined versions of Vitali covering lemmas that have many interesting applications in analysis (e.g. proving a.e. differentiability of monotone and BV functions; see below).

Proof of Lemma 8.6. Suppose $f \in L(\mathbb{R}^n)$ and f has compact support.

Using the definition of f^*, we can show that there exists $c_1 = c_1(f)$ such that

$$f^*(x) \leq c_1 |x|^{-n} \quad \text{for large enough} \quad |x|.$$

Indeed, suppose $f = 0$ for $|x| > R$. Then for $|x| > 2R$, any cube that contains x that intersects $\{|x| \leq R\}$ must have radius at least $|x| - R \geq \frac{1}{2}|x|$. Thus

$$f^*(x) \leq c_0 |x|^{-n} \int |f|\, dy \leq c_1 |x|^{-n}.$$

This proves that $\{f^* > \alpha\}$ has finite measure for every $\alpha > 0$.

Now let $\alpha > 0$ and define

$$E = \{f^* > \alpha\}.$$

For $x \in E$, there exists a cube Q_x with center x such that

$$|Q_x| < \tfrac{1}{\alpha} \int_{Q_x} |f|.$$

As the collection of $\{Q_x\}_{x \in E}$ covers E, the Vitali lemma implies that there exist $\beta > 0$ and $x_1, \ldots, x_N \in E$ so that $Q_{x_1}, \ldots, Q_{x_N}$ are disjoint and

$$|E| < \tfrac{1}{\beta} \sum_{j=1}^N |Q_{x_j}|.$$

Thus

$$|E| < \tfrac{1}{\beta} \sum_{j=1}^N \tfrac{1}{\alpha} \int_{Q_{x_j}} |f| \leq \tfrac{1}{\beta\alpha} \int |f|.$$

This proves the result (with $c = \beta^{-1}$) in this case.

Now given arbitrary $f \in L(\mathbb{R}^n)$ we may assume $f \geq 0$ (since replacing f with $|f|$ does not change f^*).

Let f_k be a sequence of integrable functions with compact support such that $0 \leq f_k \nearrow f$.

By the above, there exists a constant c independent of k and $\alpha > 0$ such that

$$|\{x \in \mathbb{R}^n : f_k^*(x) > \alpha\}| \leq \tfrac{c}{\alpha} \int f_k \leq \tfrac{c}{\alpha} \int f.$$

As $f_k^* \nearrow f^*$, it follows that

$$|\{x \in \mathbb{R}^n : f^*(x) > \alpha\}| \leq \tfrac{c}{\alpha} \int f,$$

which completes the proof. □

Finally we can prove the Lebesgue differentiation theorem.

Proof of Theorem 8.2. For $f \in L(\mathbb{R}^n)$ there exists a sequence of continuous, integrable C_k so that

$$\int |f - C_k| \to 0.$$

Write $F(Q) = \int_Q f$ and $F_k(Q) = \int_Q C_k$. For any k,

$$\begin{aligned}\limsup_{Q \searrow x} \left| \frac{F(Q)}{|Q|} - f(x) \right| &\leq \limsup_{Q \searrow x} \left| \frac{F(Q)}{|Q|} - \frac{F_k(Q)}{|Q|} \right| \\ &\quad + \limsup_{Q \searrow x} \left| \frac{F_k(Q)}{|Q|} - C_k(x) \right| + |C_k(x) - f(x)|.\end{aligned}$$

Because C_k is continuous, the second term on the RHS tends to zero. Moreover,

$$\left| \frac{F(Q)}{|Q|} - \frac{F_k(Q)}{|Q|} \right| \leq \tfrac{1}{|Q|} \int_Q |f - C_k| \leq (f - C_k)^*(x),$$

and thus for every k

$$\limsup_{Q \searrow x} \left| \frac{F(Q)}{|Q|} - f(x) \right| \leq (f - C_k)^*(x) + |f(x) - C_k(x)|.$$

Let $\varepsilon > 0$ and define E_ε to be the set on which the LHS of the above is greater than ε. In particular, by the above,

$$E_\varepsilon \subset \{(f - C_k)^*(x) > \tfrac{1}{2}\varepsilon\} \cup \{|f - C_k(x)| > \tfrac{1}{2}\varepsilon\}.$$

By the maximal function estimate and Tchebyshev, we find

$$|E_\varepsilon| \leq c\tfrac{2}{\varepsilon} \int |f - C_k| + \tfrac{2}{\varepsilon} \int |f - C_k| \to 0 \quad \text{as} \quad k \to \infty.$$

Here we use that c is independent of k. Thus $|E_\varepsilon| = 0$.

Now let E be the set where

$$\limsup_{Q\searrow x}\left|\frac{F(Q)}{|Q|}-f(x)\right|$$

is positive. Then $E=\cup_k E_{\varepsilon_k}$ for some sequence $\varepsilon_k\searrow 0$, and hence $|E|=0$. Thus

$$\lim_{Q\searrow x}\frac{F(Q)}{|Q|}=f(x)\quad\text{for a.e.}\quad x,$$

which completes the proof. □

One can extend the Lebesgue differentiation theorem to functions that are merely locally integrable — this means that the function is integrable over any bounded measurable subset of $\mathbb{R}^n$.

The Lebesgue differentiation theorem implies that any measurable set E, almost every point of E is a 'point of density' for E — this means that

$$\lim_{Q\searrow x}\frac{|E\cap Q|}{|Q|}=1$$

for a.e. $x\in E$.

8.3. **Further results.** While we will not pursue these topics further, it is worth mentioning some additional related results. The proofs can be found in Wheeden–Zygmund. They rely on a stronger version of the Vitali covering lemma.

- Finite monotone increasing functions are differentiable (with non-negative derivative) almost everywhere.
- Functions of bounded variation are differentiable a.e. with integrable derivatives.
- If $V(x)=V(f;[a,x])$ for some $f\in BV([a,b])$, then $V'(x)=|f'(x)|$ for a.e. x.

A function f is called **absolutely continuous** on $[a,b]$ if for any $\varepsilon>0$, there exists $\delta>0$ such that for any collection $\{[a_i,b_i]\}$ of nonoverlapping subintervals of $[a,b]$,

$$\sum(b_i-a_i)<\delta \implies \sum|f(b_i)-f(a_i)|<\varepsilon.$$

We write $f\in AC([a,b])$.

- If $f\in AC([a,b])$ then $f\in BV([a,b])$.
- A function f is absolutely continuous on $[a,b]$ if and only if f' exists a.e. in (a,b), $f'\in L(a,b)$, and

$$f(x)-f(a)=\int_a^x f'\quad\text{for}\quad a\le x\le b.$$