# Math 560 - Real Analysis I Notes
# California State University of Long Beach

Notes/Typeset by Hossien Sahebjame
Text used: Real Analysis by Stein Shakarchi
mymathyourmath.com

December 2021

# Contents

# § 1. Rudiments of $\mathbb{R}^d$

In undergraduate analysis we learned about the Riemann integral which limits the class of functions one can integrate over. We extend this notion of "length" in 1-dimension, area in 2, and volume in 3-space into subsets $E$ of $\mathbb{R}^d$. We begin with a point $x \in \mathbb{R}^d$ we denote by

$$x = (x_1, ..., x_d), \qquad x_i \in \mathbb{R}, i = 1, 2, ..., d.$$

Here addition is defined point-wise, that is, for any $x, y \in \mathbb{R}^d$, one has

$$x + y := (x_1 + y_1, ..., x_d + y_d).$$

Similarly, for any scalar $\delta \in \mathbb{R}$, one has

$$\delta x := (\delta x_1, ..., \delta x_d).$$

Define the **norm** of $x \in \mathbb{R}^d$ to be $|x|$ defined via

$$|x| := (x_1^2 + ... + x_d^2)^{\frac{1}{2}}.$$

Thus the "distance" between $x, y \in \mathbb{R}^d$ is given via

$$|x - y|.$$

For a given set $E \subseteq \mathbb{R}^d$ define its **complement** as $E^c$ via

$$E^c := \{x \in \mathbb{R}^d : x \in \mathbb{R}^d \bigwedge x \notin E\}.$$

For two given set $E, F \subseteq \mathbb{R}^d$, define their **distance** to be given via

$$d(E, F) := \inf |x - y|.$$

Here the infimum is taken over all $x \in E, y \in F$.

### § 1.1 Elements of the Topology

Define the **open ball** of radius $r$ centered at $x$ via

$$B_r(x) := \{y \in \mathbb{R}^d : |x - y| < r\}.$$

A subset $E \subseteq \mathbb{R}^d$ is defined to be **open** if for every $x \in E$, there exists and $r > 0$ such that

$$B_r(x) \subseteq E.$$

Then define a set $F \subseteq \mathbb{R}^d$ to be **closed** if $F^c$ is open. It is worthy to note that the open subsets of $\mathbb{R}^d$ are elements of the topology, $\tau_{\mathbb{R}^d}$, on $\mathbb{R}^d$. Furthermore, note that any arbitrary union of open sets is open while any finite union of closed sets is closed. Additionally, arbitrary intersections of closed need be closed while arbitrary intersections of open need not be open. More formally speaking, if $E_\alpha$ is open for every $\alpha \in A$, some indexing set, then

$$\bigcup_{\alpha} E_\alpha$$

is open. A subset $E \subseteq \mathbb{R}^d$ is **bounded** if it is contained in some finite radius ball. A bounded set is **compact** if it is also closed. In $\mathbb{R}^d$, compact is equivalent to closed and bounded. This is more famously known as the Heine-Borel Theorem. A more "topological" definition of compact sets is given any open covering of $E$, that is, for a given collection $\{\mathcal{O}_\alpha : \alpha \in A\}$ of open subsets of $\mathbb{R}^d$ with the property that

$$E \subseteq \bigcup_{\alpha \in A} \mathcal{O}_\alpha,$$

there will always exist some finite subset $B \subseteq A$ such that

$$E \subseteq \bigcup_{\alpha \in B} \mathcal{O}_\alpha.$$

That is, any (arbitrary) open covering of $E$ contains a finite sub-covering of $E$. Define $x \in \mathbb{R}^d$ to be a **limit point** of some $E \subseteq \mathbb{R}^d$ if for every $r > 0$, the intersection

$$B_r(x) \setminus \{x\} \bigcap E$$

is non-trivial. That is, $B_r(x) \setminus \{x\}$ contains points of $E$. Well, at least one point. A point $x \in \mathbb{R}^d$ is said to be an **isolated point** if $x \in E$ and there exists an $r > 0$ such that

$$B_r(x) \bigcap E = \{x\}.$$

A point $x \in \mathbb{R}^d$ is said to be an **interior point** of $E \subseteq \mathbb{R}^d$ if there exists an $r > 0$ such that

$$B_r(x) \subseteq E.$$

The set of all interior points of a given set $E \subseteq \mathbb{R}^d$ is denoted $\text{int}(E)$. We denote the **closure** of $E$ via $\overline{E}$ which consists of $E$ union with all of its limit points. Let the **boundary** of $E$ be denoted via $\partial E$ as

$$\partial E := \{x \in \mathbb{R}^d : x \in \overline{E} \setminus \text{int}(E)\}.$$

Not that the closure of any set is a closed set and that a set is closed if and only if it contains all of its own limit points. Lastly, a set $E \subseteq \mathbb{R}^d$ is **perfect** if $E$ contains no isolated points.

*<back2top>*

### § 1.2 Rectangles

Define a (closed) **rectangle** $R$ in $\mathbb{R}^d$ as the product of $d$ $1-$dimensional closed and bounded intervals. That is,

$$R = \prod_{k=1}^{d} [a_k, b_k].$$

Thus it makes sense for us to define the "length" of these intervals as

$$b_1 - a_1, ..., b_d - a_d.$$

Then the **volume** of $R$, denoted $|R|$ via

$$|R| := \prod_{k=1}^{d} (b_k - a_k).$$

Similarly one can define an **open** rectangle as

$$R = \prod_{k=1}^{d} (a_k, b_k).$$

Furthermore, a rectangle is a **cube** if one has

$$b_m - a_m = b_n - a_n$$

For every $m, n \in \{1, 2, ..., d\}$. Thus if $Q \subseteq \mathbb{R}^d$ is a cube with side length $l$, one has that

$$|Q| = l^d.$$

A union of rectangles is said to be **almost disjoint** if the interiors are disjoint. This leads us to our first result.

> **Lemma 1:** *If a rectangle is the almost disjoint union of finitely many other rectangles say $R = \bigcup_{k=1}^{N} R_k$, then*
>
> $$|R| = \sum_{k=1}^{N} |R_k|.$$

*Proof.* We extend the sides of the $R_1, ..., R_N$ indefinitely. This construction yields finitely many rectangles $\overline{R}_1, ..., \overline{R}_M$ and a partition $J_1, ..., J_N$ of integers between 1 and $M$ such that the unions

$$R = \bigcup_{j=1}^{M} \overline{R}_j$$

and

$$R_k = \bigcup_{j \in J_k} \overline{R}_j$$

are almost disjoint. For the rectangle $R$, we see that

$$|R| = \sum_{j=1}^{M} |\overline{R}_j|.$$

This is because our grid partitions the sides of $R$ and each $\overline{R}_j$ consists of taking products of the intervals in these partitions. Thus when we add the volumes of the $\overline{R}_j$ we sum up the corresponding products of the lengths of these intervals that arise which holds for the $R_1, ..., R_N$. Thus we get

$$\begin{aligned} |R| &= \sum_{j=1}^{M} \overline{R}_j \\ &= \sum_{k=1}^{N} \sum_{j \in J_k} |\overline{R}_j| \\ &= \sum_{k=1}^{N} |R_k|. \end{aligned}$$

as needed. □

Furthermore we have the result

**Lemma 2:** *If $R_1, ..., R_N$ are rectangles and $R \subset \bigcup_{k=1}^{N} R_k$, then*

$$|R| \leq \sum_{k=1}^{N} |R_k|.$$

*Proof.* By proof above, note the sets corresponding to the $J_k$ need not be disjoint anymore. □

This leads us to a big theorem.

**Theorem 3** *Every open subset $\mathcal{O} \subseteq \mathbb{R}$ can be written uniquely as a countable union of disjoint open intervals.*

*Proof.* For each $x \in \mathcal{O}$, let $I_x$ denote the largest interval containing $x$ such that

$$I_x \subset \mathcal{O}.$$

Moreover since $\mathcal{O}$ is open, there exists some $\epsilon > 0$ such that

$$(x - \epsilon, x + \epsilon) \subset \mathcal{O}.$$

So if

$$a_x := \inf\{a < x : (a, x) \subset \mathcal{O}\}$$

and

$$b_x := \sup\{x < b : (x, b) \subset \mathcal{O}\}$$

we then have that

$$a_x < x < b_x.$$

Note that there may exists infinitely many values for $a_x, b_x$. If we now take

$$I_x := (a_x, b_x),$$

then $x \in I_x$ by construction and we even have

$$I_x \subset \mathcal{O}.$$

Thus

$$\mathcal{O} := \bigcup_{x \in \mathcal{O}} I_x.$$

Suppose that for $x, y \in \mathbb{R}$ we have that $I_x, I_y$ intersect. Then

$$x \in I_x \cup I_y \subset \mathcal{O}.$$

By maximality of the intervals we have constructed, we have that

$$I_x \cup I_y \subset I_x$$

and

$$I_x \cup I_y \subset I_y.$$

Forcing

$$I_x = I_y.$$

Thus two distinct intervals in

$$\mathcal{I} := \{I_x\}_{x \in \mathcal{O}}$$

need be disjoint. We must now only show there are countably many distinct intervals in $\mathcal{I}$. By their disjointness and by density, each $I_x$ will always contain a rational number. As the intervals are distinct, they contain distinct rationals. This forces $\mathcal{I}$ to be countable as desired. □

Typically, if $\mathcal{O}$ is open with

$$\mathcal{O} = \bigcup_{j=1}^{\infty} I_j,$$

where the $I_j$ are disjoint open intervals, the measure of $\mathcal{O}$ ought to be

$$\sum_{j=1}^{\infty} |I_j|.$$

Now we generalize the previous theorem to all of $\mathbb{R}^d$:

> **Theorem 4** *Every open subset $\mathcal{O} \subseteq \mathbb{R}^d$ can be written as a countable union of almost disjoint closed cubes.*

*Proof.* By construction we will find a countable collections $\mathcal{Q}$ of closed cubes whose interiors are disjoint. We would like to have

$$\mathcal{O} = \bigcup_{Q \in \mathcal{Q}} Q.$$

First consider the grid in $\mathbb{R}^d$ by taking all closed cubes of unit length with integer vertices. I.e., consider the grid generated by the lattice $\mathbb{Z}^d$. We also use grids formed by cubes of side length $2^{-N}$ obtained by successively bisecting the original grid in half. We keep or toss our cubes by the following rule: If

$$Q \subset \mathcal{O}$$

then we accept $Q$ and if $Q$ intersects both $\mathcal{O}$ and $\mathcal{O}^c$ then we accept it tentatively, and if

$$Q \subset \mathcal{O}^c$$

then we toss it out. As a second step we bisect the tentative ones into $2^d$ cubes of side length 1/2. We repeat this process. By construction, we have that this collection $\mathcal{Q}$ is not only countable but consists of almost disjoint cubes. To see why

$$\mathcal{O} = \bigcup_{Q \in \mathcal{Q}} Q,$$

for any $x \in \mathcal{O}$ there exists a cube with side length $2^{-N}$ containing $x$ properly contained inside of $\mathcal{O}$. This cube has either been accepted or is contained in a cube previously accepted in another stage. This shows the union covers $\mathcal{O}$. □

Thus if $\mathcal{O} = \bigcup_{j=1}^{\infty} R_j$, then it is natural to assign to $\mathcal{O}$ the measure of

$$\sum_{j=1}^{\infty} R_j.$$

The next section will be of great significance as it shows an interesting set of measure 0.

*<back2top>*

### § 1.3 Cantors Set

We define what is know as the **Cantor Set**. We begin with defining the closed unit interval as

$$\mathcal{C}_0 = [0, 1].$$

Then take $\mathcal{C}_1$ to be the set obtained from cutting out the middle third from $\mathcal{C}_0$. Thus we get

$$\mathcal{C}_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

Repeating this process gives us

$$\mathcal{C}_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

We repeat this process indefinitely which gives us a sequence $\mathcal{C}_k$ of decreasing compact sets in the sense that

$$\mathcal{C}_0 \supset \mathcal{C}_1 \supset ... \supset \mathcal{C}_k \supset \mathcal{C}_{k+1} \supset ...$$

The Cantor set is then defined to be the intersection

$$\mathcal{C} = \bigcap_{k=0}^{\infty} \mathcal{C}_k.$$

Clearly $\mathcal{C}$ is non-empty as all end points of the $\mathcal{C}_k$ belong to $\mathcal{C}$. Note that $\mathcal{C} \subset \mathbb{R}$ is both closed and bounded and thus compact. Furthermore the space is totally disconnected with no isolated points.

Naturally one would like to know the size or measure of this Cantor Set. Not to be confused with the cardinality of the set which is the same as the continuum. In terms of measure the Cantor Set is rather small and in fact has measure 0. As $\mathcal{C}_k$ is a disjoint union of $2^k$ intervals each of length $3^{-k}$, then we have that the length of each of the $\mathcal{C}_k$ is just $(\frac{2}{3})^k$ which tends to 0 as $k \to \infty$. It is worth noting that the *Hausdorff dimension* of the cantor set is

$$\frac{\log 2}{\log 3}.$$

*<back2top>*

# § 2. The Exterior Measure

The exterior measure attempts to describe the size of a set using outside approximations. We make this definition precise as follows. If $E$ is *any* subset of $\mathbb{R}^d$, define the **exterior measure** of $E$ as

$$m_*(E) = \inf \sum_{j=1}^{\infty} |Q_j|.$$

Here the infimum is taken over all countable coverings

$$E \subseteq \bigcup_{j=1}^{\infty} Q_j$$

by closed cubes. Note $m_*(E) \in [0, \infty]$. We calculate the exterior measure of a few basic examples.

*Example 1* The exterior measure of a point is just zero. Furthermore, the exterior measure of a collection of points is also zero, even countably infinite. Take $\mathbb{Q}$ for example. We also note the exterior measure of the empty set is 0.

*Example 2* The exterior measure of closed cubes is equal to its volume. To see this, let $Q$ be a closed cube in $\mathbb{R}^d$. Then since $Q$ covers itself we have

$$m_*(Q) \leq |Q|.$$

We must show then that

$$|Q| \leq m_*(Q),$$

which would give us equality. Fix $\epsilon > 0$ and for each $j$, choose an open cube $U_j$ containing $Q_j$ such that

$$|U_j| \leq (1 + \epsilon)|Q_j|.$$

As $\bigcup_j U_j$ is a covering of $Q = \bigcup_j Q_j$ which is compact thus we can find a finite sub-covering. That is, we can find a finite subset $B \subset \mathbb{N}$ such that

$$Q \subseteq \bigcup_{j \in B} U_j.$$

Consider the closures

$$\{\overline{U_j}\}_{j\in B}.$$

Then by Lemma 2, since $Q \subset \bigcup_{j\in B} U_j$ we have that (since closure of an open set has the same measure)

$$|Q| \leq \sum_{j\in B} |U_j|.$$

Then we have

$$\begin{aligned}
|Q| &\leq \sum_{j\in B} |U_j| \\
&\leq \sum_{j\in B} (1+\epsilon)|Q_j| \\
&= (1+\epsilon)\sum_{j\in B} |Q_j| \\
&\leq (1+\epsilon)\sum_{j\in \mathbb{N}} |Q_j| \\
&= (1+\epsilon)m_*(Q).
\end{aligned}$$

as $\epsilon$ was arbitrary we are done.

*Example 3* If $Q$ is an open cube, then we still have that $m_*(Q) = |Q|$. Since $Q \subseteq \overline{Q}$ and their volumes agree, we have

$$m_*(Q) \leq |Q|.$$

To show the reverse inequality, if $Q_0$ is a closed cube inside of $Q$ then any (countable) covering of $Q$ also covers $Q_0$. Hence

$$|Q_0| \leq m_*(Q).$$

We can pick $Q_0$ with volume arbitrarily close to the value of $|Q|$, we get

$$|Q| \leq m_*(Q).$$

We then have equality as needed.

*Example 4* The exterior measure of a rectangle agrees with its volume. By example 2,

$$|R| \leq m_*(R).$$

To get the reverse inequality, consider the collection $\mathcal{Q}$ of all cubes properly contained inside of $R$ and the collection $\mathcal{Q}'$ of all cubes intersecting $R^c$. Then

$$R \subseteq \bigcup_{Q\in\mathcal{Q}\cup\mathcal{Q}'} Q.$$

As $\bigcup_{Q\in\mathcal{Q}} Q \subseteq R$ one gets

$$\sum_{Q\in\mathcal{Q}} |Q| \leq |R|.$$

Moreover there are $\mathcal{O}(k^{d-1})$ cubes in $\mathcal{Q}'$, with volume $k^{-d}$, so that

$$\sum_{Q\in\mathcal{Q}'} |Q| = \mathcal{O}(\frac{1}{k})$$

thus we get

$$\sum_{Q\in\mathcal{Q}\cup\mathcal{Q}'} |Q| \leq |R| + \mathcal{O}(\frac{1}{k}).$$

Sending $k \to \infty$ we get the desired result.

*Example 5* Note $m_*(\mathbb{R}^d) = \infty$. Since any covering of $\mathbb{R}^d$ also covers any cube $Q \subset \mathbb{R}^d$. Thus we get

$$|Q| \leq m_*(\mathbb{R}^d).$$

Since $Q$ can have arbitrary large volume we conclude that $m_*(\mathbb{R}^d) = \infty$.

*Example 6* The cantor set $\mathcal{C}$ has exterior measure zero. To see this, note by construction we have $\mathcal{C} \subset \mathcal{C}_k$ for every $k$. Each $\mathcal{C}_k$ has $2^k$ disjoint intervals each of length $3^{-k}$. Thus

$$0 \leq m_*(\mathcal{C}) \leq \left(\frac{2}{3}\right)^k.$$

This goes to 0 as $k \to \infty$.

*<back2top>*

§ **2.2 Properties of $m_*$**

We note some of the key properties of the exterior measure. First, given a subset $E \subset \mathbb{R}^d$, for any $\epsilon > 0$, there will exists a covering $\{Q_j\}_j$ of $E$, i.e.,

$$E \subset \bigcup_j Q_j,$$

such that

$$\sum_j m_*(Q_j) \leq m_*(E) + \epsilon.$$

That is, $E$ can be approximated from the outside by closed cubes arbitrarily close to $E$.

**Observation 1 (Monotonicity)** *If $E_1 \subset E_2$, then*

$$m_*(E_1) \leq m_*(E_2).$$

*Proof.* If $\{Q_j\}$ is any covering of $E_2$ by closed cubes, then it is also a covering for $E_1$ and thus the infimum for $E_1$ is taken over a larger collection than that of $E_2$. □

**Observation 2 (Countable sub-additivity)** *If $E = \bigcup_j E_j$, then*

$$m_*(E) \leq \sum_j m_*(E_j).$$

*Proof.* We may assume for every $j$ that

$$m_*(E_j) < \infty.$$

Otherwise we get equality. Let $\epsilon > 0$ be given. For each $j$ one is then guaranteed of a collection $\{Q_{k,j}\}_{k\in\mathbb{N}}$ of closed cubes with

$$\sum_{k\in\mathbb{N}} |Q_{k,j}| \leq m_*(E_j) + \frac{\epsilon}{2^j}.$$

Then $E \subset \bigcup_{j,k\in\mathbb{N}} Q_{k,j}$ is a covering of $E$ by closed cubes. Thus we get that

$$\begin{aligned} m_*(E) &\leq \sum_{j,k} |Q_{k,j}| \\ &= \sum_j \sum_k |Q_{k,j}| \\ &\leq \sum_j (m_*(E_j) + \frac{\epsilon}{2^j}) \\ &= \sum_j m_*(E_j) + \epsilon. \end{aligned}$$

We are finished as this holds for any given $\epsilon > 0$. □

**Observation 3 (Approximation by open sets)** *If $E \subset \mathbb{R}^d$, then $m_*(E) = \inf m_*(\mathcal{O})$.* Here the infimum is taken over all open set $\mathcal{O}$ containing $E$.

*Proof.* Let $\mathcal{O}$ be an open cover for $E$. Then by monotonicity we get

$$m_(E) \leq \inf m_*(\mathcal{O}).$$

To see the reverse inequality, let $\epsilon > 0$ be given. We choose our cubes $Q_j$ such that $E \subset \bigcup_j Q_j$, such that

$$\sum_j |Q_j| \leq m_*(E) + \frac{\epsilon}{2}.$$

For each $j$, let $Q_j^0$ denote the open cube containing $Q_j$ such that

$$|Q_j^0| \leq |Q_j| + \frac{\epsilon}{2^{j+1}}.$$

Then $\mathcal{O} = \bigcup_j Q_j^0$ is an open set thus by monotonicity, we get

$$\begin{aligned} m_*(\mathcal{O}) &\leq \sum_j m_*(Q_j^0) \\ &= \sum_j |Q_j^0| \\ &\leq \sum_j \left( |Q_j| + \frac{\epsilon}{2^{j+1}} \right) \\ &\leq \sum_j |Q_j| + \frac{\epsilon}{2} \\ &\leq m_*(E) + \epsilon. \end{aligned}$$

Hence $\inf m_*(\mathcal{O}) \leq m_*(E)$. as needed. $\square$

**Observation 4** *If $E = E_1 \cup E_2$ and $d(E_1, E_2) > 0$, then*

$$m_*(E) = m_*(E_1) + m_*(E_2).$$

*Proof.* By monotonicity, we get that

$$m_*(E) \leq m_*(E_1) + m_*(E_2).$$

We must show the reverse inequality. Choose a $\delta > 0$ such that

$$d(E_1, E_2) > \delta > 0.$$

For $\epsilon > 0$, select a covering of $E$ by closed cubes,

$$E \subset \bigcup_j Q_j$$

such that

$$\sum_j |Q_j| \leq m_*(E) + \epsilon.$$

After subdividing the $Q_j$ we can assume they each have diameter less than $\delta$. Then by our choice of $\delta$, these newly subdivided $Q_j$ intersect either $E_1$ or $E_2$. Let $J_1, J_2$ denote the set of indices $j$ for which $Q_j$ intersects $E_1, E_2$ respectively. Then we get that

$$J_1 \cap J_2 = \emptyset$$

and that

$$E_1 \subset \bigcup_{j \in J_1} Q_j, E_2 \subset \bigcup_{j \in J_2} Q_j.$$

Thus we have

$$\begin{aligned} m_*(E_1) + m_*(E_2) &\leq \sum_{j \in J_1} |Q_j| + \sum_{j \in J_2} |Q_j| \\ &\leq \sum_j |Q_j| \\ &\leq m_*(E) + \epsilon. \end{aligned}$$

We are finished as $\epsilon$ was arbitrary. $\square$

**Observation 5** *If $E$ is a contable union of almoast disjoint cubes, then*

$$m_*(E) = \sum_j |Q_j|.$$

*Proof.* For some fixed $\epsilon > 0$ let $\tilde{Q}_j$ denote the cube properly contained inside of $Q_j$ such that

$$|Q_j| \leq |\tilde{Q}_j| + \frac{\epsilon}{2^j}.$$

Then for every $N$, the cubes

$$\tilde{Q}_1, \tilde{Q}_2, ..., \tilde{Q}_N$$

are disjoint thus have some finite distance between them. Thus we can apply a repeated use of our previous observation to see that

$$\begin{aligned} m_*(\bigcup_j \tilde{Q}_j) &= \sum_{j=1}^{N} |\tilde{Q}_j| \\ &\geq \sum_{j=1}^{N} \left( |Q_j| - \frac{\epsilon}{2^j} \right) \end{aligned}$$

Since $\sum_{j=1}^{N} Q_j \subset E$ we have that

$$m_*(E) \geq \sum_{j=1}^{N} |Q_j| - \epsilon.$$

As $N \to \infty$, we see that

$$\sum_j |Q_j| \leq m_*(E) + \epsilon.$$

This holds for any arbitrary $\epsilon > 0$ thus together with Observation 2, we get equality as desired. $\square$

*<back2top>*

## § 3 The Lebesgue measure

There are a number of ways one can define measurability which all turn out to be the same.

Definition: We say that a subset $E \subset \mathbb{R}^d$ is *Lebesgue measurable* or simply *measurable* if for any $\epsilon > 0$, there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ such that

$$m_*(\mathcal{O} \setminus E) \leq \epsilon.$$

This can be compared to Observation 3 which holds for *all* sets. If $E$ is measurable we define its *Lebesgue measure* $m(E)$ by

$$m_*(E) := m(E).$$

Similarly, for any subset $E \subseteq \mathbb{R}^d$ and any given set $A$, $E$ is (Lebesgue) measurable if one has

$$m(A) = m(A \cap E) + m(A \cap E^c).$$

The Lebesgue measure inherits all Observations of the exterior measure.

*<back2top>*

## § 3.1 Properties of $m$

We immediately inherit a few key properties of the Lebesgue measure.

**Property 1** *Every open set in $\mathbb{R}^d$ is measurable.*

*Proof.* Let $\mathcal{O} \subseteq \mathbb{R}^d$ be open. Then by Theorem 3, one can write $\mathcal{O}$ as a disjoint union of open intervals. That is,

$$\mathcal{O} = \prod_{k=1}^{d} (a_k, b_k).$$

Where each interval is measurable and furthermore a countable union of measurable sets is still measurable thus $\mathcal{O}$ is measurable as needed. □

**Property 2** *If $m_*(E) = 0$, then $E$ is measurable. In particular, if $F$ is a subset of a set of exterior measure $0$, then $F$ is measurable itself.*

*Proof.* To show $F$ is measurable, let $\epsilon > 0$ be fixed. Then there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ with

$$m_*(\mathcal{O}) \leq \epsilon.$$

Since $\mathcal{O} \setminus E \subset \mathcal{O}$, we get

$$m_*(\mathcal{O} \setminus E) \leq \epsilon.$$

and thus $E$ is measurable. □

**Property 3** *A countable union of measurable sets is still measurable.*

*Proof.* Let us assume that $E = \bigcup_j E_j$ where $E_j$ is measurable for each $j$. By measurability, for each $j$ we may find an open set $\mathcal{O}_j$ with

$$E_j \subset \mathcal{O}_j$$

such that

$$m_*(\mathcal{O}_j \setminus E_j) \leq \frac{\epsilon}{2^j}.$$

Then the union $\bigcup_j \mathcal{O}_j$ is open and contains $E$. Then

$$\mathcal{O} \setminus E \subset \bigcup_j (\mathcal{O}_j \setminus E_j).$$

Thus by monotonicity and sub-additivity of the exterior measure, we get

$$\begin{aligned} m_*(\mathcal{O} \setminus E) &\leq \sum_j m_*(\mathcal{O}_j \setminus E_j) \\ &\leq \epsilon. \end{aligned}$$

as desired. □

**Property 4** *Closed sets are measurable*

*Proof.* First, note it suffices to prove that compact sets are measurable. As a matter of fact, any closed set $F$ can be written as a union of compact sets. I.e., one can write

$$F = \bigcap_k F \cap B_k$$

where $B_k$ are balls centered at the origin of radius $k$. Suppose then that $F$ is compact and thus has finite exterior measure. Let $\epsilon > 0$, then by Observation 3 one can find an open set $\mathcal{O}$ containing $F$ such that

$$m_*(\mathcal{O}) \leq m_*(F) + \epsilon.$$

As $F$ is closed, $\mathcal{O} \setminus F$ is open by definition. By a theorem, we can write $\mathcal{O}$ as the countable union of almost disjoint cubes. That is,

$$\mathcal{O} \setminus F = \bigcup_j Q_j.$$

For any fixed $N \in \mathbb{N}$, the finite union

$$\bigcup_{j=1}^{N} Q_j$$

is compact. Thus $d(K, F) > 0$. By observations 4,5,6 we get that

$$\begin{aligned} m_*(\mathcal{O}) &\geq m_*(F) + m_*(K) \\ &= m_*(F) + \sum_j m_*(Q_j). \end{aligned}$$

Hence

$$\begin{aligned} \sum_j m_*(Q_j) &\leq m_*(\mathcal{O}) - m_*(F) \\ &\leq \epsilon. \end{aligned}$$

which holds as $N$ tends to $\infty$. By sub-additivity we get

$$\begin{aligned} m_*(\mathcal{O} - F) &\leq \sum_j m_*(Q_j) \\ &\leq \epsilon. \end{aligned}$$

as we needed. □

We give a Lemma relating closed and compact set.

**Lemma 5** *If $F$ is closed and $K$ is compact and the sets are disjoint, then*

$$d(F, K) > 0.$$

*Proof.* Since $F$ is a closed set, for each $x \in K$, there exists some $\delta_x > 0$ such that

$$d(x, F) > 3\delta_x.$$

Note that by our choice of $\delta_x$, we have that

$$K \subset \bigcup_{x \in K} B_{2\delta_x}(x).$$

By the compactness of $K$, there exists some finite subset $A \subset K$ such that

$$K \subset \bigcup_{x \in A} B_{2\delta_x}(x).$$

Take $\delta$ to be the minimum of the $\delta_x$ for $x \in A$, then

$$d(F, K) \geq \delta > 0.$$

Then if $x \in K$ and if $y \in F$, then for some $x_0 \in A$ we get

$$|x_0 - x| \leq 2\delta_0.$$

and by construction we have

$$\begin{aligned} |y - x| &\geq |y - x_0| - |x_0 - x| \\ &\geq 3\delta_0 - 2\delta_0 \\ &\geq \delta. \end{aligned}$$

as needed. □

**Property 5** *The complement of a measurable set it measurable*

*Proof.* If $E \subset \mathbb{R}^d$ is measurable, then for every $n \in \mathbb{N}$ we may choose an open set $\mathcal{O}_n$ with $E \subset \mathcal{O}_n$ and

$$m_*(\mathcal{O}_n \setminus E) \leq \frac{1}{n}.$$

Since $\mathcal{O}_n$ is open for each $n$, its complement $\mathcal{O}_n^c$ is closed thus measurable. Thus $U = \bigcup_n \mathcal{O}_n^c$ is measurable as well. Note just by construction we have that

$$U \subset E^c.$$

Also note

$$E^c \setminus U \subset \mathcal{O}_n \setminus E.$$

Then by monotonicity, we get that

$$m_*(E^c \setminus U) \leq \frac{1}{n}$$

This gives us a measure zero set $E^c \setminus U$ which is measurable and equal to the union of $U$ and $E^c \setminus U$ thus $E^c$ is measurable. □

**Property 6** *A countable intersection of measurable sets is measurable*

*Proof.* This follows from Properties 3,5 and by set theory since we know that

$$\left(\bigcap_j E_j\right) = \left(\bigcup_j E_j\right)^c.$$

□

Inc conlusion, measurable sets are closed under the basicv operations of set theory. Furhtermore, we have shown closures over *countable* unions and intersections, not just finite. Note measurable sets do not behave nicely with respect to *uncountable* unions and intersections. This leads us to our next big result.

**Theorem 6** *If $E_1, E_2, ...$ are disjoint measurable sets, then we have the following equality*

$$m(E) = \sum_{j=1}^{\infty} m(E_j).$$

*Proof.* We may first assume that the $E_j$ are bounded for each $j$. By Property 5, then we have that for each $j$, $E_j^c$ is also measurable. Next we fix some $\epsilon > 0$ then choose closed subsets $F_j \subset E_j$ such that

$$m(E_j \setminus F_j) < \frac{\epsilon}{2^j}.$$

As the $F_j$ are compact, then for any finite subset $A \subset \mathbb{N}$, we now that $\{F_j\}_{j\in A}$ forms a compact collection of sets. Furthermore, note the $F_j$ are disjoint since they are inside of the $E_j$ which were assumed to be disjoint. Thus we have by addivitiy that

$$m(\bigcup_{j\in A} F_j) = \sum_{j\in A} m(F_j).$$

And since $\bigcup_{j\in A} F_j \subset E$, then by monotonocity we get that

$$\begin{aligned} m(E) &\geq \sum_{j\in A} m(F_j) \\ &\geq \sum_{j\in A} m(E_j) + \epsilon. \end{aligned}$$

The reverse inequality also holds by countable sub-additivity and this case is complete.

Next we show the general case in which the $E_j$ are not-bounded. We select a sequence of cubes $\{Q_j\}_{j\in\mathbb{N}}$ which tends to all of $\mathbb{R}^d$. That is, $Q_j \subset Q_{j+1}$ for all $j$ and that the union of these cubes is all of $\mathbb{R}^d$. So we define a new sequence. Take

$$S_1 := Q_1, S_j := Q_j \setminus Q_{j-1}.$$

The latter for $j \geq 2$.. These $S_j$ are clearly measurable as they are complements of measurable sets. We can then define measurable unions of sets via

$$E_{k,j} = E_k \cap S_j.$$

Then we have that

$$\bigcup_{k,j} E_{k,j},$$

which is a disjoint and bounded union and note that $E_j = \bigcup_{k,j} E_{k,j}$. Using these facts together with what has been proven already, we obtain

$$\begin{aligned} m(E) &= \sum_{k,j} E_{k,j} \\ &= \sum_{k} \sum_{j} m(E_{k,j}) \\ &= \sum_{k} m(E_k). \end{aligned}$$

as desired. □

<*back2top*>

### § 3.2 MCT for sets

Before jumping to the notion of Monotone converging measurable functions we will discuss what it means for monotone converging sets.

If $E_1, E_2, ...$ are measurable increasing sets in the senses that $E_k \subset E_{k+1}$ and if $E = \bigcup_k E_k$, we write $E_k \nearrow E$.

Similarly, if the sets $E_1, E_2, ...$ are decreasing in the senses that $E_k \supset E_{k+1}$ and if $E = \bigcap_k E_k$ we then write $E_k \searrow E$. Moreover, the latter condition requires at least one of the $E_k$, say $E_1$ has finite measure. That is, $m(E_1) < \infty$. This leads use to a major corollary.

**Corollary 7** *If $E_1, E_2, ...$ are measurable subsets of $\mathbb{R}^d$. Then*

- *If $E_k \nearrow E$, then $m(E) = \lim_{N\to\infty} m(E_k)$.*
- *If $E_k \searrow E$ and if $m(E_k) < +\infty$ for some $k$, then $m(E) = \lim_{N\to\infty} m(E_N)$.*

*Proof.* For the proof of (a), we define measurable sets $G_1 = E_1$ for $k = 1$, then for every $k \geq 2$ define

$$G_k := E_k \setminus E_{k-1}$$

By construction, the $G_k$ are measurable, and disjoint. Moreover, note that $E = \bigcup_k G_k$. Then we have

$$\begin{aligned} m(E) &= m(\bigcup_k G_k) \\ &= \sum_{k\in\mathbb{N}} m(G_k) \\ &= \lim_{N\to\infty}\sum_{k=1}^N m(G_k) \\ &= \lim_{N\to\infty} m(\bigcup_{k=1}^N G_k) \\ &= \lim_{N\to\infty} m(\bigcup_{k=1}^N E_k) \\ &= \lim_{N\to\infty} m(E_N). \end{aligned}$$

as desired.

For (b), note since the $E_k$ are decreasing, then

$$E_1 \setminus E_k \subset E_1 \setminus E_{k+1}.$$

Is an increasing sequence. First note from basic set theory, one has

$$\begin{aligned} \bigcup E_1 \setminus E_k &= \bigcup E_1 \cap E_k^c \\ &= E_1 \cap \bigcup E_k^c \\ &= E_1 \cap (\bigcap E_k)^c \\ &= E_1 \setminus \bigcap_k E_k \end{aligned}$$

Then we can compute

$$\begin{aligned} m(E_1 \setminus \bigcap E_k) &= m(E_1) - m(\bigcap E_k) \\ &= m(\bigcup E_1 \setminus E_k) \\ &= \lim_{n\to\infty} m(E_1 \setminus E_k) \\ &= m(E_1) - \lim_{n\to\infty} m(E_n) \end{aligned}$$

and since $m(E_1) < +\infty$, we can subtract it from both sides to obtain the desired result. □

The following theorem will play a central roll in approximations of measurable functions.

**Theorem 8** *Suppose $E \subset \mathbb{R}^d$ is measurable. Then, for every $\epsilon > 0$,*

- *There exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$, and*

$$m(\mathcal{O} \setminus E) \leq \epsilon.$$

- *There exists a closed set $F$ with $F \subset E$ and*

$$m(E \setminus F) \leq \epsilon.$$

- *If $m(E) < +\infty$, then there exists a compact set $K$ with $K \subset E$ and*

$$m(E \setminus K) \leq \epsilon.$$

- *If $m(E) < +\infty$, then there exists a finite union of $F = \bigcup_{j=1}^{N} Q_j$ of closed cubes such that the symmetric difference of $E$ and $F$ is less than or equal to $\epsilon$.*

*Proof.* For (a), this just follows from definitions of measurability. For (b), note that since $E$ is measurable, then so is $E^c$. So there exists an open set containing $E^c$ within $\epsilon$. So there exists some open set $\mathcal{O}$ with $E^c \subset \mathcal{O}$ such that

$$m(\mathcal{O} \setminus E^c) \leq \epsilon.$$

Since $\mathcal{O}$ is open, its complement $\mathcal{O}^c$ is closed and contained inside of $E$. Moreover, one has that $E \setminus \mathcal{O}^c = \mathcal{O} \setminus E^c$ and thus

$$m(E \setminus \mathcal{O}^c) \leq \epsilon.$$

as needed. For (c), we first pick a closed set $F$ such that

$$m(E \setminus F) \leq \epsilon.$$

For each $n \in \mathbb{N}$, let $B_n$ denote the ball centered at the origin of radius $n$. Define compact sets via

$$K_n := F \cap B_n.$$

Then $E \setminus K_n$ is a measurable sequence of decreasing sets to $E \setminus F$. Since $m(E) < +\infty$, we conclude for very large $n \in \mathbb{N}$ that

$$m(E \setminus K_n) \leq \epsilon.$$

For the last part, we choose a family of closed cubes $\{Q_j\}_{j \in \mathbb{N}}$ so that

$$E \subset \bigcup_{j \in \mathbb{N}} Q_j,$$

and that

$$\sum_{j \in \mathbb{N}} |Q_j| \leq m(E) + \frac{\epsilon}{2}.$$

Since $m(E) < +\infty$, the series converges and there exists an $N \in \mathbb{N}$ such that

$$\sum_{j=N+1}^{\infty} |Q_j| \leq \frac{\epsilon}{2}.$$

If there exists some finite subset $A \subset \mathbb{N}$ where $F = \bigcup_{j \in A} Q_j$, then one can compute

$$\begin{aligned} m(E \triangle F) &= m(E \setminus F) + m(F \setminus E) \\ &\leq m\Big( \bigcup_{j=N+1}^{\infty} Q_j \Big) + m\Big( \bigcup_{j=1}^{\infty} Q_j \setminus E \Big) \\ &\leq \sum_{j=N+1}^{\infty} |Q_j| + \sum_{j=1}^{\infty} |Q_j| - m(E) \\ &\leq \epsilon. \end{aligned}$$

as $\epsilon > 0$ was arbitrarily chosen, we are done. □

*<back2top>*

### § 3.3 Invariance Properties

Often times we would like to know which properties are invariant under the measure function, $m$. One key property that is preserved is this notion of translation invariance. That is, one would like to know the measure of the set $\{x + h : x \in E, h \in \mathbb{R}^d\}$, in fact the measure is preserved in the following sense:

$$m(E_h) = m(E),$$

where $E_h := \{x + h : x \in E, h \in \mathbb{R}^d\}$. Suppose we are given some $\delta > 0$, then the set

$$\delta E := \{\delta x : x \in E\},$$

has (Lebesgue) measure

$$m(\delta E) = \delta^d m(E).$$

Furthermore, note that the measure function is reflection invariant as well. That is, if

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

then

$$m(-E) = m(E).$$

*<back2top>*

### § 4 $\sigma-$algebras & Borel sets

Definition: A $\sigma-$*algebras* of sets is a collection of subsets closed under countable unions, intersections, and complements. The collection of all $\sigma-$algebras is of course a $\sigma-$algebra. We will consider a particular collection of $\sigma-$algebras called the *Borel* $\sigma-$*algebra in* $\mathbb{R}^d$. This Borel $\sigma-$algebra will be the "smallest" $\sigma-$algebra containing all of the open subsets of $\mathbb{R}^d$.

*<back2top>*

### § 4.1 Borel Sets

From the view point of the Borel sets, the Lebesgue sets arise from the *completion* of the $\sigma-$algebra of the Borel sets, that is, by adjoining all Borel sets of measure zero.

Starting with open and closed sets, which are the most simplest Borel sets, one could list the Borel sets in order of their complexity. Next in order would be countable intersections of open sets, called $G_\delta$ sets and their respective complemenets, the countable union of closed sets, called the $F_\sigma$ sets. This gives rise to our next Corollary.

**Corollary 9** *A subset $E \subseteq \mathbb{R}^d$ is measurable*

- *if and only if $E$ differs from a $G_\delta$ by a set of measure zero*
- *if and only if $E$ differs from a $F_\sigma$ by a set of measure zero*

*Proof.* Clearly $E$ is measurable whenever it satisfies the first or second. This is because the $F_\sigma$, $G_\delta$, and sets of measure zero are all measurable. On the other hand, if $E$ is measurable, then for each $n \in \mathbb{N}$, we may select an open $\mathcal{O}_n$ of $E$ such that

$$m(\mathcal{O}_n \setminus .E) \leq \frac{1}{n}.$$

Then $S = \bigcup_{n \in \mathbb{N}} \mathcal{O}_n$ is a $G_\delta$ set containing $E$ such that

$$S \setminus E \subset \mathcal{O}_n \setminus E$$

for every $n \in \mathbb{N}$ thus by monotonicity, we get

$$m(S \setminus E) \leq \frac{1}{n},$$

has zero exterior measure and is therefore measurable. For the second implication, we invoke **Theorem 8(ii)** with $\epsilon = \frac{1}{n}$. □

<*back2top*>

## § 5 Construction of non-measurable set.

It turns out that not all subsets of $\mathbb{R}^d$ are in fact measurable. In this section, we give the construction of a non-measurable set $\mathcal{N}$. We will consider the case when $d = 1$, and consider the following subset of $\mathbb{R}$. We place an equivalence relation $\sim$ on [0,1] by putting $x \sim y$ whenever $x - y \in \mathbb{Q}$. We quickly get the three properties of an equivalence relation, that is $\sim$ is reflexive, symmetric, and transitive. Two equivalence classes are either the same or disjoint and we can write [0,1] as the union over all equivalence classes we write as

$$[0, 1] = \bigcup_{\alpha} \mathcal{E}_\alpha.$$

Next, we construct our set $\mathcal{N}$ by selecting one element $x_\alpha \in \mathcal{E}_\alpha$ from each equivalence class. That is, we write

$$\mathcal{N} = \{x_\alpha\}.$$

This leads us to our big theorem of this section:

**Theorem 10** *The set $\mathcal{N}$ is non-measurable.*

*Proof.* We prove this theorem by contradiction. We assume that $\mathcal{N}$ is measurable. Let $\{r_k\}_k$ be an enumeration of $\mathbb{Q} \cap [-1, 1]$. Consider the translates

$$\mathcal{N}_k = \mathcal{N} + r_k.$$

I claim the $\mathcal{N}_k$ are disjoint and that

$$[0, 1] \subset \bigcup_{k \in \mathbb{N}} \mathcal{N} + r_k \subset [-1, 2].$$

To see why the sets are disjoint, consider the intersection

$$\mathcal{N}_k \cap \mathcal{N}_{k'}$$

being non-empty, then there exists rationals $r_k, r_{k'}$ that are distinct and $\alpha, \beta$ with

$$x_\alpha + r_k = x_\beta + r_{k'}.$$

Then we have

$$x_\alpha - x_\beta = r_{k'} - r_k.$$

And thus $\alpha \neq \beta$ and $x_\alpha - x_\beta \in \mathbb{Q}$ hence $x_\alpha \sim x_\beta$ which contradicts $\mathcal{N}$ containing only one representative from each equivalence class. The inclusion

$$\bigcup_{k \in \mathbb{N}} \mathcal{N}_k$$

is immediate by construction. To see why $[0, 1] \subset \bigcup_{k \in \mathbb{N}} \mathcal{N}_k$, let $x \in [0, 1]$. Then $x \sim x_\alpha$ for some $\alpha$ thus

$$x - x_\alpha = r_k$$

for some $k \in \mathbb{N}$. Thus $x \in \mathcal{N}_k$ and we get the first inclusion. If $\mathcal{N}$ is measurable, then so is $\mathcal{N}_k$ for each $k$ and by monotonicity, we have that

$$1 \leq \sum_{k \in \mathbb{N}} m(\mathcal{N}_k) \leq 3.$$

Since $\mathcal{N}_k$ is a translate of $\mathcal{N}$, we have $m(\mathcal{N}) = m(\mathcal{N}_k)$ for every $k$. We conclude that

$$1 \leq \sum_{k \in \mathbb{N}} m(\mathcal{N}) \leq 3.$$

This contradicts our set $\mathcal{N}$ being measurable since it can neither have measure zero or infinite measure. □

## § 5.2 Axiom of Choice

The construction of the set $\mathcal{N}$ is possible based on the following general proposition.

**Proposition 11** Suppose $E$ is a set and $\{E_\alpha\}_{\alpha\in A}$ is a collection of non-empty subsets of $E$. (Here, the indexing set $A$ is not assumed to be countable). Then there exists a "choice" function

$$\alpha \to x_\alpha,$$

such that $x_\alpha \in E_\alpha$ for every $\alpha \in A$.

Definition: In this general form, this assertion is known as the *axiom of choice.* An initial use of this mere fact was used to prove the *well-ordering* principle. We make this a bit more rigorous as follows.

Definition: A set $E$ is *linearly ordered* if there exists a binary relation $\leq$ on $E$ such that

(a) $x \leq x$ ; $\forall x \in E$

(b) If $x, y \in E$ are distinct, then either $x \leq y$ or $y \leq x$ but not both.

(c) If $x \leq y$ and $y \leq z$, then $x \leq z$.

Definition: We say that a set $E$ can be *well-ordered* if it can be linearly ordered in such a way that *every* non-empty subset $A \subset E$ has a smallest element in that ordering. That is an element $x_0 \in A$ such that $x_0 \leq x$ for every other $x \in A$. In general, any set $E$ can be well-ordered. It is in fact nearly obvious that the well-ordering principle implies the axiom of choice: If we well order $E$, then we can choose $x_\alpha$ to be the smallest element of $E_\alpha$. Conversely, the axiom of choice implies the well-ordering principle.

*<back2top>*

§ **6 Measurable functions** For $x \in E$, we begin with the *characteristic function* of a given set $E \subset \mathbb{R}^d$. That is, define

$$\chi_E(x) = \begin{cases} 0 & ; x \notin E \\ 1 & ; x \in E. \end{cases}$$

We then lean over toward the building blocks of the Riemann integral. For the Riemann integral, these will be *step functions.* A **step function** is a finite sum

$$f = \sum_{k=1}^{N} a_k \chi_{R_k}(x)$$

Where each $R_j$ is a rectangle. A **simple function** is a finite sum

$$f = \sum_{k=1}^{N} a_k \chi_{R_k},$$

where each $R_k$ is measurable of finite measure, and the $a_k$ are constants.

*<back2top>*

§ **6.1 Properties of finite- valued**

We begin by considering those subsets of $\mathbb{R}^d$ that take on real-values. That is, for any $x \in E \subset \mathbb{R}^d$, we have

$$-\infty \leq f(x) \leq +\infty,$$

in the case of finite-valued functions $f$, we say that $f$ is finite-valued if the inequality is strict.

$$-\infty < f(x) < +\infty.$$

Definition: A function defined on some measurable subset $E \subset \mathbb{R}^d$, is said to be *measurable* if for every $a \in \mathbb{R}$, the set

$$f^{-1}([-\infty, a)) = \{x \in E : f(x) < a\}$$

is measurable. To simplify this notion we often just write

$$\{x \in E : f(x) < a\} = \{f < a\}.$$

Note first off that there are many equivalent definitions of a measurable function for us now given complemenets. For example, we may require the inverse image of closed intervals need be measurable. In fact, to prove $f$ is measurable if and only if $\{x \in E : f(x) \leq a\} = \{f \leq a\}$ is measurable for every $a \in \mathbb{R}$, note in one direction one has

$$\{f \leq a\} = \bigcap_{k=1}^{\infty} \{f < a + \frac{1}{k}\}.$$

Recall a countable union of measurable sets need be measurable thus for the other direction, observe that

$$\{f < a\} = \bigcup_{k=1}^{\infty} \{f \leq a - \frac{1}{k}\}.$$

Similarly, $f$ is measurable if and only if $\{f \leq a\}$ or $\{f > a\}$ is measurable for every $a \in \mathbb{R}$. This is immediate from the fact that

$$\{f \geq a\} = \{f < a\}^c$$

and in the second case we have that

$$\{f \leq a\} = \{f > a\}^c.$$

Consequently, $-f$ is measurable whenever $f$ is. In particular, one can show if $f$ is finite-valued (that is, $-\infty < f(x) < +\infty$.), then $f$ is measurable if and only if the set

$$\{a < f < b\}$$

is measurable for every $a, b \in \mathbb{R}$. This leads us to our next few properties:

**Property 1** *The finite-valued function $f(x)$ is measurable if and only if $f^{-1}(\mathcal{O})$ is measurable for every open set $\mathcal{O}$, and if and only if $f^{-1}(F)$ is measurable for every closed set $F$.*

**Property 2** *If $f$ is continuous on $\mathbb{R}^d$, then $f$ is measurable. If $f$ is measurable and finite-valued and $\Phi$ is continuous, then $\Phi \circ f$ is continuous.*

By continuity of $\Phi$, we have that

$$\Phi^{-1}((-\infty, a))$$

is an open set, call it $\mathcal{O}$. Hence

$$(\Phi \circ f)^{-1} = f^{-1}(\mathcal{O}) = f^{-1}(\mathcal{O}).$$

is measurable.

**Property 3** *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions, then*

$$\sup_n f_n(x),\ \inf_n f_n(x),\ \lim_{n\to\infty} \sup, f_n(x),\ \lim_{n\to\infty} \inf f_n(x)$$

*are measurable.*

**Property 4** *If $\{f_n\}_{n=1}^{\infty}$ is a collection of measurable functions and*

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

*Then $f(x)$ is measurable.*

**Property 5** *If $f$ and $g$ are measurable, then*

*(a) The integer powers $f^k$ for $k \geq 1$ are measurable.*

*(b) $f + g$, $fg$ are measurable if both $f, g$ are both finite-valued.*

*Proof.* To see (a), note if $k$ is odd, then $\{f^k > a\} = \{f > a^{\frac{1}{k}}\}$, and if $k$ is even and $a \geq 0$, then

$$\{f^k > a\} = \{f > a^{\frac{1}{k}}\} \cup \{f < -a^{\frac{1}{k}}\}.$$

For (b), note that to see why $f + g$ is measurable, we can write

$$\{f + g > a\} = \bigcup_{r\in\mathbb{Q}} \{f > a - r\} \cap \{g > r\}.$$

To see why the product is measurable, note that

$$fg = \frac{1}{4}[(f + g)^2 - (f - g)^2].$$

□

Definition: we shall say two functions $f, g$ defined on a set $E$ are equal *almost everywhere* if the set

$$\{x \in E : f(x) \neq g(x)\}$$

has measure zero.

Note that if $f, g$ are defined almost everywhere on measurable subset of $\mathbb{R}^d$, then the functions $f + g$, $fg$ can only be defined on the intersection of the domains of $f$ and $g$. We summarize this fact with the following property:

**Property 6** *Suppose $f$ is measurable and $f(x) = g(x)$ a.e. on $E$. Then $g$ is measurable.*

*Proof.* To show this, we must show for any measurable set $E \subseteq \mathbb{R}^d$, $g(E) \in \mathcal{M}$. That, is $g^{-1}(E)$ is measurable. Note then that

$$\begin{aligned} g^{-1}(E) &= \{x : g(x) \in E\} \\ &= \{x : g(x) \in E, f(x) = g(x)\} \cup \{x : g(x) \in E, f(x) \neq g(x)\} \end{aligned}$$

Since $f = g$ a.e., this tells us the latter of the two sets is a null set. Moreover, if $A$ is a measurable set in the range of $g$, then we can write

$$\begin{aligned} A &= \{x : g(x) \in E, f(x) = g(x)\} \\ &= \{x : f(x) \in E, f(x) = g(x)\} \\ &= \{x : f(x) \in E\} \cap \{x : f(x) = g(x)\}. \end{aligned}$$

Here the first of the two sets is measurable as $f$ is measurable, and the second set is the complement of a measurable set and is thus measurable. ☐

*<back2top>*

## § 7 Approximation via simple and step functions.

A lot of the theorems in this section are of the same nature. We begin by approximating point-wise, non-negative measurable functions by simple functions. This leads us to our first big theorem of this section.

**Theorem 12** *Suppose $f$ is a non-negative measurable function on $\mathbb{R}^d$. Then there exists an increasing sequence of non-negative simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that converges pointwise to $f$, namely,*

$$\varphi_k \leq \varphi_{k+1} \text{and} \lim_{k\to\infty} \varphi_k(x) = f(x)$$

*for all $x$.*

*Proof.* We first begin with a truncation. For $N \geq 1$, let $Q_N$ denote the cube centered at the origin of side length $N$. Then define

$$F_N(x) = \begin{cases} f(x) & ; x \in Q_N, f(x) \leq N \\ N & ; x \in E, f(x) > N \\ 0 & \text{otherwise} \end{cases} .$$

Then it is clear that

$$F_N \to f$$

as $N \to \infty$. Next, we partition the range of $F_N(x)$. Namely, $[0, N]$, as follows. For fixed $M, N \geq 1$, we define

$$E_{l,M} := \{x \in Q_N : \frac{l}{m} < F_N(x) \leq \frac{l+1}{M}\},$$

for $l \in [0, MN)$. Then we can define

$$F_{N,M}(x) = \sum_l \frac{l}{M} \chi_{E_{l,M}}(x).$$

Where each $F_{N,M}$ is a simple function satisfying

$$0 \leq F_N(x) - F_{N,M}(x) \leq \frac{1}{M},$$

for every $x$. If we choose $N = M = 2^k$ with $k \geq 1$, and let $\varphi_k = F_{2^k,2^k}$, then we see that

$$0 \leq F_M(x) - \varphi_k(x) \leq \frac{1}{2^k}$$

for every $x$, and $\{\varphi_k\}$ is increasing and satisfies the desired properties. □

Note that the result holds for non-negative functions that are extended valued, if the limit $\infty$ is allowed. We now drop the assumption that $f$ be non-negative.

**Theorem 13** *Suppose $f$ is measurable on $\mathbb{R}^d$. Then there exists a sequence of simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that satisfy*

$$|\varphi_k(x)| \leq |\varphi_{k+1}(x)|$$

*and*

$$\lim_{k=1}^{\infty} \varphi_k(x) = f(x),$$

*for all $x$.*

*Proof.* We use the decomposition of $f$ into $f^+$ and $f^-$. That is, we can write

$$f(x) = f^+(x) - f^-(x)$$

where

$$f^+(x) := \max(f(x), 0), f^-(x) := \max(-f(x), 0).$$

Since both functions are non-negative, the previous result gives us two increasing sequences of non-negative simple functions $\{\varphi_k^{(1)}\}_{k=1}^{\infty}$ and $\{\varphi_k^{(2)}\}_{k=1}^{\infty}$ converging point-wise to $f^+$ and $f^-$ respectively. Then if we take

$$\varphi_k(x) := \varphi_k^{(1)}(x) - \varphi_k^{(2)}(x)$$

we see that $\varphi_k \to f$ for every $x$. Finally, the sequence $\{|\varphi_k|\}$ is increasing by definitions of $f^+, f^-$ and properties of $\varphi_k^{(1)}$, $\varphi_k^{(2)}$ imply that

$$|\varphi_k(x)| = \varphi_k^{(1)}(x) + \varphi_k^{(2)}(x)$$

The next step, is to approximate by step functions. Here, in general, the convergence may hold only a.e. □

**Theorem 14** *Suppose $f$ is measurable on $\mathbb{R}^d$. Then there exists a sequence of step functions $\{\psi_k\}_{k=1}^{\infty}$ that converges point-wise to $f(x)$ for almost every $x$.*

*Proof.* By the previous theorem, there are simple functions $\{\varphi_k\}_{k=1}^{\infty}$ such that

$$\lim_{k=1}^{\infty} \varphi_k(x) = f(x)$$

for every $x$. To approximate each $\varphi_k$ by a step function, recall Theorem 8 (iv) which states if $E$ is measurable of finite measure, then for any $\epsilon > 0$ there exists a finite subset $A \subset \mathbb{N}$ such that

$$m(E \triangle \bigcup_{k \in A} Q_k) \leq \epsilon.$$

By considering the grid formed by extending the sides of the cubes, we see there exists almost disjoint rectnalges $\overline{R}_1, ..., \overline{R}_M$ such that

$$\bigcup_{k \in A} Q_k = \bigcup_{j \in B} \overline{R}_j$$

By taking closed rectangles $R_j$ contained in $\overline{R}_j$ and slightly smaller in size, we find a collection of disjoint closed rectangles that satisfy

$$m(E \triangle \bigcup_{j \in B} R_j) \leq 2\epsilon.$$

So then by definition of a simple function that for each $k$, there exists a step function $\psi_k$, and a measurable set $F_k$ such that $m(F_k) < \frac{1}{2^k}$ and that

$$\varphi_k(x) = \psi_k(x)$$

for every $x \notin F_k$. If we define

$$F : \bigcap_{l=1}^{\infty} \bigcup_{k>l}^{\infty} F_k,$$

then $m(F) = 0$ since

$$m(\bigcup_{k>l}^{\infty} F_k) \leq \sum_{k>l} m(F_k) \leq \frac{1}{2^l}.$$

For $x \notin F$, there exists some $k_0 \in \mathbb{N}$ such that

$$x \in \bigcap_{k>k_0} F_k^c,$$

thus for every $k > k_0$, one has

$$\begin{aligned} |f(x) - \psi_k(x)| &\leq |f(x) - \varphi_k(x)| + |\varphi_k(x) - \psi_k(x)| \\ &= |f(x) - \varphi_k(x)| \end{aligned}$$

and since $\varphi_k \to f$ as $n \to \infty$, we conclude that

$$\lim_{k \to \infty} \psi_k(x) = f(x)$$

for every $x \notin F$ as desired. $\square$

<back2top>

## § 8 Littlewood's Three principles.

Although the notion of measurable sets and functions represents new tools, we should not overlook their relation to the older concepts they replaced. Littlewood aptly summarized these connections in the form of three principles that provide a useful intuitive guide in the initial study of the theory.

(a) Every set is nearly a finite union of intervals.

(b) Every function is nearly continuous.

(c) Every convergent sequence is nearly uniformly convergent.

The sets and functions referred to above are of course assumed to be measurable. The catch is in the word "nearly". A precise version of the first principle appears in part (iv) of Theorem 8. An important formulation of the third principle appears in the following result.

**Theorem 15 (Egorov)** *Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence of measurable functions defined on a measurable set $E$ with $m(E) < \infty$ and assume $f_k \to f$ a.e. on $E$. Given $\epsilon > 0$ we can always find a closed set $A_\epsilon \subset E$ such that*

$$m(E \setminus A_\epsilon) \leq \epsilon,$$

*and $f_k \to f$ uniformly on $A_\epsilon$.*

*Proof.* We may assume, without any loss of generality, that $f_k \to f$ for every $x \in E$. For each pair of non-negative integers $n, k$ let

$$E_k^n := \{x \in E : |f_j(x) - f(x)| < \frac{1}{n}\}.$$

Now, fix $n \in \mathbb{N}$ and note that the $E_k^n$ are increasing in set containment. That is,

$$E_k^n \subset E_{k+1}^n$$

for every $k$. Furthermore, note that $E_k^n \nearrow E$. By MCT for measurable sets, there exists some $k_n \in \mathbb{N}$ such that

$$m(E \setminus E_{k_n}^n) \leq \epsilon.$$

By construction then we have that

$$|f_j(x) - f(x)| < \frac{1}{n}$$

whenever $j > k_n$ and $x \in E_{k_n}^n$. Next, we select our $N \in \mathbb{N}$ such that

$$\sum_{n=N}^{\infty} \frac{1}{2^n} < \frac{\epsilon}{2},$$

and let

$$\tilde{A}_\epsilon = \bigcap_{n \geq \mathbb{N}} E_{k_n}^n.$$

Then note

$$\begin{aligned} m(Em \setminus \tilde{A}_\epsilon) &\leq \sum_{n=N}^{\infty} m(E \setminus E^n_{k_n}) \\ &< \frac{\epsilon}{2}. \end{aligned}$$

Next, if $\delta > 0$, we can choose $n \geq N$ such that $\frac{1}{n} < \delta$ and note that $x \in \tilde{A}_\epsilon$ implies $x \in E^n_{k_n}$. So then whenever $j > k_n$, we have

$$|f_j(x) - f(x)| < \delta.$$

Thus $f_k \to f$ uniformly on $A_\epsilon$. Finally, by Theorem 8, we can choose a closed subset $A_\epsilon \subset \tilde{A}_\epsilon$ with

$$m(\tilde{A}_\epsilon \setminus A_\epsilon) < \frac{\epsilon}{2}$$

This results in

$$m(E \setminus \tilde{A}_\epsilon) < \epsilon.$$

as desired. □

The next result attests to the second of three principles.

**Theorem 16 (Lusin)** *Suppose $f$ is measurable and finite valued on $E$ with $E$ having finite measure. Then for every $\epsilon > 0$, there exists a closed set $F_\epsilon$ with*

$$F_\epsilon \subset E$$

*and*

$$m(E \setminus F_\epsilon) \leq \epsilon,$$

*and such that $f|_{F_\epsilon}$ is continuous.*

*Proof.* Let $f_n$ be a sequence of step functions such that

$$f_n \to f$$

a.e. $x$. Then we can find sets $E_n$ such that $m(E_n) < \frac{1}{2^n}$ and $f_n$ is continuous outside of $E_n$. By Egorov's Theorem, we can find a closed set $A_{\frac{\epsilon}{3}} \subset E$ on which $f_n \to f$ uniformly and

$$m(E \setminus A_{\frac{\epsilon}{3}}) \leq \frac{\epsilon}{3}.$$

We then consider the the set

$$F' = A_{\frac{\epsilon}{3}} \setminus \bigcup_{n \geq N} E_n$$

for $N$ so large that

$$\sum_{n \geq N} \frac{1}{2^n} < \frac{\epsilon}{3}.$$

Now for all $n \geq N$, the function $f_n$ is continuous on $F'$, thus $f$ is also continuous on $F'$. We can then approximate $F'$ by a closed set $F_\epsilon$ such that

$$m(F' \setminus F_\epsilon) \leq \epsilon.$$

and we are done. □

<back2top>

§ **Exercises (with solutions):**

*Exercise 1* Suppose $E$ is a given set, and $\mathcal{O}_n$ is the open set

$$\mathcal{O}_n = \{x : d(x, E) < \frac{1}{n}\}.$$

Show If $E$ is compact, then

$$m(E) = \lim_{n\to\infty} m(\mathcal{O}_n).$$

*Proof.* Note that for every $k$, one has

$$\mathcal{O}_k \supset \mathcal{O}_{k+1}$$

And thus $\mathcal{O}_n \searrow E$, Then by MCT for measurable sets, one has that

$$m(E) = \lim_{n\to\infty} m(\mathcal{O}_n)$$

□

*Exercise 2* If $\delta = (\delta_1, \delta_2, ..., \delta_d)$ is a $d-$tuple of positive numbers $\delta_i \in (0, \infty)$. and $E \subset \mathbb{R}^d$, defined $\delta E$ by

$$\delta E := \{(\delta_1 x_1, ..., \delta_d x_d) : (x_1, x_2, ..., x_d) \in E\}$$

Prove $\delta E$ is measurable whenever $E \subset \mathbb{R}^d$ is measurable and that

$$m(\delta E) = \delta_1 \cdot ... \cdot \delta_d m(E).$$

*Proof.* We would like to find for any $\epsilon > 0$ that there exists an open set $\mathcal{O}$ containing $\delta E$ such that

$$m(\mathcal{O} \setminus \delta E) \leq \epsilon.$$

First let us define the product as

$$\Delta = \delta_1 \cdot ... \cdot \delta_d$$

Recall if $Q = \prod_{k=1}^{d} [a_k, b_k]$ is a closed cube with $a_k < b_k$, then

$$m(Q) = \prod_{k=1}^{d} (b_k - a_k).$$

Furthermore, note that if $Q$ is any closed cube, then

$$\delta Q = \prod_{k=1}^{d} [\delta_k a_k, \delta_k b_k]$$

in which case we obtain

$$m(\delta Q) = \prod_{k=1}^{d} \delta_k (b_k - a_k) = \Delta m(Q).$$

Next, notice that if $\mathcal{O}$ is open, then $\delta\mathcal{O}$ is open as well. To see this, for any $x \in \mathcal{O}$ implies there exists an $r > 0$ such that $B_r(x) \subset \mathcal{O}$. Take $\tau := \min\{\delta_1, ..., \delta_d\}$, then if $x \in \mathcal{O}$, we get that $\delta x \in \delta\mathcal{O}$ and $B_{\tau r}(\delta x) \subset \delta\mathcal{O}$. By measurability of $E$, for any given $\epsilon > 0$, there exists some open set, call it $\mathcal{O}$, containing $E$ such that

$$m(\mathcal{O} \setminus E) \leq \frac{\epsilon}{\Delta}.$$

Clearly, $E \subset \mathcal{O}$ implies $\delta E \subset \delta\mathcal{O}$, then we can compute out the measure:

$$\begin{aligned} m(\delta\mathcal{O} \setminus \delta E) &= \inf\left\{\sum_{k=1}^{\infty} m(\delta Q_k) : \mathcal{O} \setminus E \subset \bigcup_{k=1}^{\infty} Q_k\right\} \\ &= \inf\left\{\Delta \sum_{k=1}^{\infty} m(Q_k) : \mathcal{O} \setminus E \subset \bigcup_{k=1}^{\infty} Q_k\right\} \\ &= \Delta m(\mathcal{O} \setminus E) \\ &\leq \Delta \frac{\epsilon}{\Delta} \\ &< \epsilon. \end{aligned}$$

as needed. □

*Exercise 3* Let $B$ be a ball in $\mathbb{R}^d$ of radius $r$. Then prove

$$m(B) = v_d r^d$$

where $v_d = m(B_1)$ is the measure of the unit ball.

*Proof.* Since the Lebesgue measure is translation invariant, one can consider balls centered at the origin. Then by Theorem 4, one can write

$$B_1(0) = \bigcup_{k=1}^{\infty} Q_j$$

where $\{Q_j\}_j$ is a collection of almost disjoint closed cubes. Thus we have

$$v_d = m(B_1(0)) = \sum_{k=1}^{\infty} m(Q_k).$$

Note any dilation of the unit ball is done by taking the $d-$tuple $\delta = (r, r, ..., r)$, then we can compute

$$\begin{aligned} m(B_r(0)) &= m(\delta B_1(0)) \\ &= r^d m(B_1(0)) \\ &= r^d v_d \end{aligned}$$

as desired. □

*Exercise 4 (**Borel-Cantelli Lemma**)* Suppose $\{E_k\}_{k\in\mathbb{N}}$ is a collection of measurable subsets of $\mathbb{R}^d$. Furthermore, suppose that

$$\sum_{k=1}^{\infty} m(E_k) < \infty.$$

Let

$$\begin{aligned} E &= \{x \in \mathbb{R}^d : x \in E_k, \text{for infinitely many } k\} \\ &= \lim_{k\to\infty} \sup(E_k). \end{aligned}$$

Show $E$ is measurable with measure zero.

*Proof.* Since for each $k \in \mathbb{N}$, $E_k$ is measurable, then the countable union

$$A_n := \bigcup_{k \geq n} E_k$$

is measurable as well. Note since we can write

$$E = \bigcap_{n=1}^{\infty} A_n$$

then $E$ is measurable. To show $m(E) = 0$, suppose towards a contradiction that $m(E) = \delta > 0$. We can then define

$$S_N := \bigcap_{k=1}^{N} \bigcup_{n \geq k} E_n = \bigcup_{n \geq N} E_n.$$

Then we have that $S_N \searrow E$ and thus for every $N \in \mathbb{N}$, one has

$$\begin{aligned} \delta &\leq m(S_N) \\ &= m(\bigcup_{n \geq N} E_n) \\ &= \sum_{n=N}^{\infty} m(E_n). \end{aligned}$$

This contradicts the measures being finite. □

*<back2top>*

## § 10. Integration

The general notion of a Lebesgue integral on $\mathbb{R}^d$ will be defined in a step by step manner. At each stage, we will see the definitions of the integral satisfies elementary properties such as linearity and monotonicity. We proceed in four stages, by progressively integrating.

1. Simple functions
2. Bounded functions supported on a set of finite measure
3. Non-negative functions
4. Integrable functions (the general case)

<back2top>

## § 10.1 Simple functions

We assume that all functions are measurable. At the beginning we consider only finite valued functions, then carry on into complex and extended valued. Recall from the previous sections that a simple function is a finite sum

$$\varphi(x) = \sum_{k=1}^{N} a_k \chi_{E_k}(x),$$

where the $a_k$ are constants and the $E_k$ measurable sets of finite measure.

Definition: The *canonical form* of $\varphi$ is the unique decomposition of the simple function above such that the $a_k$ are distinct, and non-zero and the $E_k$ are pair-wise disjoint. Finding the canonical form of $\varphi$ is straightforward: since $\varphi$ can take on only finitely many values, say $c_1, ..., c_M$, we may set

$$F_k := \{x : \varphi(x) = c_k\},$$

and note the $F_k$ are pairwise disjoint. Therefore,

$$\varphi(x) = \sum_{k=1}^{M} c_k \chi_{F_k},$$

is the desired canonical form.

Definition: If $\varphi$ is a simple function with canonical form $\varphi(x) = \sum_{k=1}^{M} c_k \chi_{F_k}$, then we define its *Lebesgue integral* of $\varphi$ by

$$\int_{\mathbb{R}^d} \varphi(x)dx = \sum_{k=1}^{M} c_k m(F_k).$$

If $E \subset \mathbb{R}^d$ is measurable with finite measure, then $\varphi(x)\chi_E(x)$is also a simple function and *define*

$$\int_E \varphi(x)dx = \int \varphi(x)\chi_E(x).$$

To emphasize the use of the Lebesgue measure $m$, one often writes

$$\int_{\mathbb{R}^d} \varphi(x)dm(x).$$

for the Lebesgue integral of $\varphi$.

**Proposition 10.1** *The integral of simple functions defined above satisfies the following properties:*

1. *Independence of representation. If $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ is any representation of $\varphi$, then*

$$\int \varphi = \sum_{k=1}^{N} a_k m(E_k)$$

2. *Linearity. If $\varphi, \psi$ are simple functions and $a, b \in \mathbb{R}$, then*

$$\int (a\varphi + b\psi) = a\int \varphi + b\int \psi.$$

3. *Additivity. If $E, F$ are disjoint subsets of $\mathbb{R}^d$ with finite measure, then*

$$\int_{E\cup F} \varphi = \int_E \varphi + \int_F \varphi.$$

4. *Monotonicity. If $\varphi \leq \psi$ are simple functions, then*

$$\int \varphi \leq \int \psi.$$

5. *Triangle inequality. If $\varphi$ is a simple function, then so is $|\varphi|$, and*

$$\left| \int \varphi \right| \leq \int |\varphi|.$$

*Proof.* For 1., suppose $\varphi = \sum_{k=1}^N a_k \chi_{E_k}$, where the $E_k$ are disjoint, but we do not suppose the $a_k$ are distinct and nonzero. For each distinct non-zero value $a$ amongst $\{a_k\}$, define

$$E'_a := \{\bigcup_k E_k\}$$

where the union is taken over those indices $k$ such that $a_k = a$. Note then that the $E'_a$ are disjoint, and

$$m(E'_a) = \sum m(E_k).$$

Here the sum is taken over the same set of $k'$s. Then we clearly have

$$\varphi(x) = \sum a_k \chi_{E'_a}$$

where the sum is taken over the distinct non-zero values of $\{a_k\}$. Thus we have

$$\begin{aligned} \int \varphi &= \sum a m(E'_a) \\ &= \sum_{k=1}^N a_k m(E_k). \end{aligned}$$

Next suppose that

$$\varphi = \sum_{k=1}^N a_k \chi_{E_k}$$

where the $E_k$ are no longer assumed to be disjoint. We can then "refine" the decomposition

$$\bigcup_{k=1}^N E_k$$

by finding sets $E_1^*, E_2^*, ..., E_n^*$ such that

$$\bigcup_{k=1}^{N} E_k = \bigcup_{j=1}^{n} E_j^*.$$

And the $E_j^*$ are disjoint. For each $k$, we have

$$E_k = \bigcup E_j^*$$

here the union is taken over those $E_j^*$ that are contained in $E_k$. For each $j$ now, let

$$a_j^* = \sum a_k$$

with the sum taken over all $k$ such that $E_k$ contains $E_j^*$. Then we have

$$\varphi = \sum_{j=1}^{n} a_j^* \chi_{E_j^*}$$

however the $E_j^*$ are disjoint, thus we obtain

$$\begin{aligned} \int \varphi &= \sum a_j^* m(E_j^*) \\ &= \sum \sum_{E_k \supset E_j^*} a_k m(E_j^*) \\ &= \sum a_k m(E_k) \end{aligned}$$

and thus 1. is proven. For 2., using any representation of $\varphi$ and $\psi$ works together with the linearity of 1. For additivity, note that if $E \cap F = \emptyset$, then

$$\chi_{E \cup F} = \chi_E + \chi_F$$

this fact together with linearity of the integral gives us 3. For 4., if $\eta \geq 0$ is a non-negative simple function, then its canonical form is everywhere non-negative as well and thus

$$\int \eta \geq 0$$

by definition of the integral. Applying this argument to $\varphi - \psi$ gives us the desired result. Finally, for the triangle inequality, we write $\varphi$ is its canonical form, that is,

$$\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$$

and note that

$$|\varphi| = \sum_{k=1}^{N} |a_k| \chi_{E_k}.$$

Thus we can compute

$$\begin{aligned} \left|\int \varphi\right| &= \left|\sum_{k=1}^{N} a_k m(E_k)\right| \\ &\leq \sum_{k=1}^{N} |a_k| m(E_k) \\ &= \int |\varphi|. \end{aligned}$$

as desired. □

*<back2top>*

### § 10.2 Bounded functions supported on finite measure set

Definition: The *support* of a measurable function $f$ is the set of all points where $f$ does not vanish. This is given as

$$\text{supp}(f) := \{x : f(x) \neq 0\}.$$

We shall say $f$ is supported on a set $E$ if $f(x) = 0$ whenever $x \notin E$. Since $f$ is measurable, so is the set supp$(f)$. We will next be interested in those bounded measurable functions that have

$$m(\text{supp}(f)) < \infty.$$

The key lemma that follows allows us to define the integral for the class of bounded functions supported on sets of finite measure.

**Lemma 17** *Let $f$ be bounded function supported on a set of finite measure $E$. If $\{\varphi_k\}$ is any sequence of simple functions bounded by $M$, supported on $E$ and with $\varphi_k \to f$ for a.e. $x$, then*

1. *The limit $\lim_{n\to\infty} \int \varphi_n$ exists.*
2. *If $f = 0$ a.e., then the limit $\lim_{n\to\infty} \int \varphi_n$ equals 0.*

*Proof.* Since $m(E) < +\infty$, by Egorov's Theorem we are guaranteed the existence of a closed measurable set $A_\epsilon \subset E$ such that

$$m(E \setminus A_\epsilon) \leq \epsilon,$$

and $\varphi_n \to f$ uniformly on $A_\epsilon$. Then let $I_n := \int \varphi_n$ and compute out

$$\begin{aligned} |I_n - I_m| &= \int_E |\varphi_n(x) - \varphi_m(x) \\ &= \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)| dx + \int_{E\setminus A_\epsilon} |\varphi_n(x) - \varphi_m(x)| dx \\ &\leq \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)| dx + 2Mm(E \setminus A_\epsilon) \\ &\leq \int_{A_\epsilon} |\varphi_n(x) - \varphi_m(x)| dx + 2M\epsilon. \end{aligned}$$

By the uniform convergence, one has for all $x \in A_\epsilon$ and large $n, m$ that $|\varphi_n(x) - \varphi_m(x)| < \epsilon$ thus

$$|I_n - I_m| \leq m(E)\epsilon + 2M\epsilon.$$

as $\epsilon > 0$ are arbitrary, we are done. $\square$

We use Lemma 17 to turn to the integration of bounded functions that are supported on sets of finite measure.

Definition: For such functions, define its *Lebesgue integral* to be given via

$$\int f(x)dx = \lim_{n\to\infty} \int \varphi_n(x)dx.$$

Here $\{\varphi_n\}$ is any sequence of simple functions satisfying

$$|\varphi_n(x)| \leq M,$$

each $\varphi_n$ is supported on the support of $f$ and $\varphi_n(x) \to f(x)$ as $n \to \infty$.

If $E \subset \mathbb{R}^d$ with finite measure, and $f$ is bounded with $m(\text{supp}(f)) < \infty$, then it is natural to define

$$\int_E f(x)dx = \int f(x)\chi_E(x)dx.$$

Clearly, if $f$ is simple itself, then $\int f$ as defined above coincides with the integral of simple functions studied earlier. This definition also satisfies certain properties.

**Proposition 10.2** *Suppose $f$ and $g$ are bounded functions supported on a set of finite measure. Then the following properties hold*

1. *Linearity. If $a, b \in \mathbb{R}$, then*
$$\int af + bg = a\int f + b\int g.$$
2. *Additivity. If $E, F$ are disjoint subsets of $\mathbb{R}^d$, then*
$$\int_{E\cup F} f = \int_E f + \int_F f$$
3. *Monotonicity. If $f \leq g$, then*
$$\int f \leq \int g$$
4. *Triangle inequality. $|f|$ is also bounded, supported on a set of finite measure, and*
$$\left|\int f\right| \leq \int |f|.$$

We are now ready to prove one of the biggest convergence theorems, known as the bounded convergence theorem.

**Theorem 18 (Bounded convergence theorem)** *Suppose that* $\{f_n\}_{n\in\mathbb{N}}$ *is a sequence of measurable functions that are all bounded by* $M$*, supported on* $E$ *of finite measure and* $f_n \to f$ *a.e.* $x$ *as* $n \to \infty$*, then* $f$ *is measurable, bounded, supported on* $E$ *for a.e.* $x$ *and*

$$\int |f_n - f| \to 0$$

*as* $n \to \infty$*. Consequently,*

$$\int f_n \to \int f$$

*as* $n \to \infty$*.*

*Proof.* By assumption, on sees that $f$ is bounded by $M$ a.e., and vanishes outside of $E$, except possibly on a set of measure zero. By the triangle inequality, one only needs to prove $\int |f_n - f| \to 0$ as $n \to \infty$. Given some $\epsilon > 0$, by Egorov's Theorem one is guaranteed the existence of a closed subset $A_\epsilon \subset E$ such that

$$m(E \setminus A_\epsilon) \leq \epsilon$$

and $f_n \to f$ uniformly on $A_\epsilon$. Then we know for $n \geq N \in \mathbb{N}$, that

$$|f_n(x) - f(x)| \leq \epsilon.$$

We can thus compute

$$\begin{aligned}\int |f_n(x) - f(x)|dx &\leq \int_{A_\epsilon} |f_n(x) - f(x)|dx + \int_{E\setminus A_\epsilon} |f_n(X) - f(x)|dx \\ &\leq \epsilon m(E) + 2Mm(E \setminus A_\epsilon)\end{aligned}$$

and since $\epsilon > 0$ are arbitrary, this holds for all large $n$ and we are done. $\square$

We note that the above congruence theorem is a statement about the interchange of an integral and a limit. Since its conclusion says

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n$$

A useful observation at this point is the following: if $f \geq 0$ is bounded and supported on a set of finite measure $E$, and if $\int f = 0$, then $f = 0$ a.e. $x$. Indeed, if for each integer $k \geq 1$ we set

$$E_k := \{x \in E : f(x) > \frac{1}{k}\}$$

then, the fact that $\frac{1}{k}\chi_{E_k}(x) \leq f(x)$ implies

$$\frac{1}{k} m(E_k) \leq \int f,$$

by monotonicity of the integral. Thus $m(E_k) = 0$ for all $k$, and since

$$\{x : f(x) > 0\} = \bigcup_{k=1}^{\infty} E_k$$

we see $f = 0$ a.e. $x$.

*<back2top>*

## § 11. Return to Riemann integral

We shall show that Riemann integrable functions are also Lebesgue integrable. This will be formalized in the flowing theorem.

**Theorem 19** *Suppose $f$ is Riemann integrable on $[a,b]$. Then $f$ is measurable, and*

$$\int_{[a,b]}^{\mathcal{R}} f(x)dx = \int_{[a,b]}^{\mathcal{L}} f(x)dx,$$

*where the left handed side is the standard Riemann integral and the right handed side is the standard Lebesgue integral.*

*Proof.* By definition, a Riemann integrable function is bounded. Let's say there exists some $M \in \mathbb{R}$ such that

$$|f(x)| \leq M.$$

Again, by definition of Riemann integrability we can construct two sequence of step functions $\{\varphi_k\}$ and $\{\psi_k\}$ such that

$$|\varphi_k|, |\psi_k| \leq M$$

for all $x \in [a,b]$ and $k \geq 1$,

$$\varphi_1(x) \leq \varphi_2(x) \leq ... \leq f \leq .... \leq \psi_2(x) \leq \psi_1(x),$$

and

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k(x)dx = \lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{R}} f(x)dx. \qquad (*)$$

Several observations are in order here. First, for step functions, it is clear that Riemann and Lebesgue integrals agree. Thus

$$\int_{[a,b]}^{\mathcal{R}} \varphi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)dx$$

and

$$\int_{[a,b]}^{\mathcal{R}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \psi_k(x)dx. \qquad (**)$$

Next, if we let

$$\tilde{\varphi}(x) = \lim_{k\to\infty} \varphi_k(x) \qquad \text{and} \qquad \tilde{\psi}(x) = \lim_{k\to\infty} \psi_k(x),$$

we have

$$\tilde{\varphi} \leq f \leq \tilde{\psi}.$$

Moreover, both $\tilde{\varphi}$ and $\tilde{\psi}$ are measurable as they are limits of step functions. By the bounded convergence theorem, we obtain

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\varphi}(x)dx$$

and

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\psi}(x)dx$$

Thus by $(*)$ and $(**)$, we have

$$\int_{[a,b]}^{\mathcal{L}} (\tilde{\psi}(x) - \tilde{\varphi}(x))dx = 0$$

and since $\psi_k - \varphi_k \geq 0$, we have $\tilde{\psi} - \tilde{\varphi} \geq 0$. Thus $\tilde{\psi} - \tilde{\varphi} = 0$ a.e. $x$. Therefore

$$\tilde{\psi} = \tilde{\varphi} = f$$

a.e. $x$ thus $f$ is measurable. Since $\varphi_k \to f$ a.e., we have that

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} f(x)dx,$$

and by $(*)$ and $(**)$ we conclude that

$$\int_{[a,b]}^{\mathcal{R}} f(x)dx = \int_{[a,b]}^{\mathcal{L}} f(x)dx$$

as needed. □

*<back2top>*

## § 12. Stage three: non-negative functions

We proceed with functions that are measurable and non-negative but not necessarily bounded. It will be important to allow these functions to take on extended real values such as $\infty$. Recall the supremum of a set of positive numbers to be $+\infty$ if the set is unbounded. In the case of such a function $f$, we define its (extended) Lebesgue integral by

$$\int f(x)dx = \sup_g \int g(x)dx,$$

here the supremum is taken over all measurable functions $g$ such that

$$0 \leq g \leq f,$$

and where $g$ is bounded and supported on a finite measure set. With the above definition of the integral, there are only two possibilities; the supremum is either infinite or finite. In the case where

$\int f(x)dx < \infty$, we shall say $f$ is *Lebesgue integrable*. Clearly, if $E \subset \mathbb{R}^d$ is measurable and $f \geq 0$, then $f_{\chi_E}$ is also positive and we define

$$\int_E f(x)dx = \int f(x)\chi_E(x)dx.$$

Some examples are

$$f_a(x) = \begin{cases} |x|^{-a} & ; |x| \leq 1 \\ 0 & ; |x| > 1 \end{cases}.$$

Then $f_a$ is integrable when $a < d$. Another example is given by

$$F_a(x) = \frac{1}{1+|x|^a}$$

is integrable when $a > d$. The following proposition is for the integral of non-negative measurable functions.

**Proposition 10.3** *The integral of non-negative measurable functions enjoys the following properties:*

1. *Linearity. If $f, g \geq 0$ and $a, b \in \mathbb{R}^+$, then*

$$\int (af + bg) = a \int f + b \int g.$$

2. *Additivity. If $E, F$ are disjoint subsets of $\mathbb{R}^d$ and $f \geq 0$, then*

$$\int_{E \cup F} f = \int_E f + \int_F f.$$

3. *Monotonicity. If $0 \leq f \leq g$, then*

$$\int f \leq \int g.$$

4. *If $g$ is integrable and $0 \leq f \leq g$, then $f$ is integrable.*
5. *If $f$ is integrable, then $f(x) < \infty$ a.e. $x$.*
6. *If $\int f = 0$, then $f(x) = 0$ for a.e. $x$.*

We next turn our attention to some important convergence theorems for non-negative measurable functions.

**Lemma 20 (Fatou's Lemma)** *Suppose that $\{f_n\}$ is a sequence of measurable functions with $f_n \geq 0$. If $\lim_{n\to\infty} f_n(x) = f(x)$ for a.e. $x$, then*

$$\int f \leq \liminf_{n\to\infty} \int f_n.$$

*Proof.* Suppose there exists some function bounded and supported on a finite measure set $E$, call it $g$ such that

$$0 \leq g \leq f.$$

If we set $g_n(x) = \min(g(x), f_n(x))$, then $g_n$ is measurable and supported on a $E$. Moreover, $g_n \to g$ a.e. $x$, thus by bounded convergence theorem,

$$\int g_n \to \int g$$

By our construction, we also have $g_n \leq f_n$ thus by monotonicity, we have $\int g_n \leq \int f_n$ and therefore

$$\int g \leq \lim_{n\to\infty} \inf \int f_n$$

by taking the supremum over all $g$ yields the desired result. □

**Corollary 21** *Suppose $f$ is a non-negative measurable function and $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n \leq f$ and $f_n \to f$ for almost every $x$. Then*

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

*Proof.* Since $f_n \leq f$, by monotonicity of the Lebesgue integral we have

$$\int f_n \leq \int f$$

for every $n$. Hence

$$\lim_{n\to\infty} \sup \int f_n \leq \int f$$

Combining this together with the result from Fatou's Lemma, we obtain the desired result. □

We are now ready for the big theorem for monotone increasingly functions.

**Theorem 22 (Monotone Convergence Theorem)** *Suppose $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n \nearrow f$, then*

$$\lim_{n\to\infty} \int f_n = \int f$$

A useful corollary follows.

**Corollary 23** *Consider a series $\sum_{k=1}^{\infty} a_k(x)$ where $a_k(x) \geq 0$ is measurable for every $k \geq 1$, then*

$$\int \sum_{k=1}^{\infty} a_k(x) = \sum_{k=1}^{\infty} \int a_k(x) dx.$$

*If $\sum_{k=1}^{\infty} \int a_k(x)dx$ is finite, then the series $\sum_{k=1}^{\infty} a_k(x)$ converges for a.e. $x$.*

*Proof.* Let

$$f_n(x) = \sum_{k=1}^{n} a_k(x), f(x) = \sum_{k=1}^{\infty} a_k(x).$$

The $f_n$ are measurable for all $n$, and $f_n(x) \leq f_{n+1}(x)$ and $f_n \to f$ as $n \to \infty$. Since

$$\int f_n = \sum_{k=1}^{\infty} \int a_k(x)dx,$$

by MCT we can pull out the integral thus

$$\sum_{k=1}^{\infty} \int a_k(x)dx = \int \sum_{k=1}^{\infty} a_k(x)dx.$$

If the sum

$$\sum \int a_k < \infty$$

then $\sum_{k=1}^{\infty} a_k(x)dx$ is integrable, and thus finite a.e. $x$. $\square$

*<back2top>*

## § 12.2 General Case

Definition: If $f$ is any real-valued measurable function on $\mathbb{R}^d$, we say $f$ is *Lebesgue integrable* if the non-negative measurable function $|f|$ is integrable in the sense of the previous section. If $f$ is Lebesgue integrable, we give meaning to its integral as follows. First, we define

$$f^+ := \max(f(x), 0), f^- := \max(-f(x), 0).$$

So that $f^+, f^-$ are both non-negative and we have $f = f^+ - f^-$ and since $f^\pm \leq |f|$, both $f^+$ and $f^-$ are integrable whenever $f$ is integrable. We can now define its *Lebesgue integral* by

$$\int f = \int f^+ - \int f^-.$$

And note that the definition of the integral is independent of the decomposition of $f$. The integral of Lebesgue integrable functions is linear, additive, monotonic, and satisfies the triangle inequality. The next theorem relates integrability with finite measure sets.

**Proposition 24** *Suppose $f$ is integrable on $\mathbb{R}^d$. Then for every $\epsilon > 0$,*

1. *There exists a finite measure ball $B$ such that*

$$\int_{B^c} |f| < \epsilon.$$

2. *There exists a $\delta > 0$ such that*

$$\int_E |f| < \epsilon.$$

   *whenever $m(E) < \delta$.*

*Proof.* By replacing $f$ with $|f|$, we may assume without any loss of generality that $f \geq 0$. For (1), let $B_N$ denote the ball of radius $N$ centered at the origin. Note that if $f_N(x) = f(x)\chi_{B_N}(x)$, then $f_N$ is measurable and $f_N \leq f_{N+1}$ and

$$\lim_{N\to\infty} f_N = f,$$

by MCT then we have that

$$\lim_{N\to\infty} \int f_N = \int f.$$

In particular, for some large $N$, one has

$$0 \leq \int f - \int f\chi_{B_N}(x) < \epsilon$$

and since $1 - \chi_{B_N} = \chi_{B_N^c}(x)$, this implies that

$$\int_{B_N^c} f < \epsilon.$$

as desired.

For (2), assuming that $f \geq 0$, we take

$$f_N(x) = f(x)\chi_{E_N}$$

where

$$E_N := \{x : f(x) \leq N\}.$$

Once again, $f_N \geq 0$ is measurable and $f_N \leq f_{N+1}$ so given some $\epsilon > 0$, by the MCT there exists an integer $N \in \mathbb{N}$ such that

$$\int (f - f_N) < \frac{\epsilon}{2}.$$

We now choose $\delta > 0$ such that $N\delta < \frac{\epsilon}{2}$. Then if $m(E) < \delta$, then

$$\begin{aligned}
\int_E f &= \int_E (f - f_N) + \int_E f_N \\
&\leq \int (f - f_N) + \int_E f_N \\
&\leq \int (f - f_N) + Nm(E) \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

as needed. □

Next, we state and prove the dominated convergence theorem.

**Theorem 25 (Dominated convergence theorem)** *Suppose* $\{f_n\}$ *is a sequence of measurable functions such that* $f_n \to f$ *a.e.* $x$*, as* $n \to \infty$*. If* $|f_n(x)| \leq g(x)$*, where* $g$ *is integrable, then*

$$\int |f_n - f| \to 0$$

*as* $n \to \infty$ *and consequently,*

$$\int f_n \to \int f$$

*as* $n \to \infty$

*Proof.* For each $N \geq 0$ let

$$E_N := \{x : |x| \leq N, g(x) \leq N\}$$

Given some $\epsilon > 0$ by proposition 24 part 1, we have

$$\int_{E_N^c} g(x)dx < \epsilon.$$

Then the $f_n \chi_{E_N}$ are bounded by $N$, and supported on a set of finite measure thus by the bounded convergence theorem we have

$$\int_{E_N} |f_n - f| < \epsilon$$

as $n \to \infty$ and thus we compute

$$\begin{aligned}
\int .|f_n - f| &= \int_{E_N} |f_n - f| + \int_{E_N^c} |f_n - f| \\
&\leq \int_{E_N} |f_n - f| + 2\int_{E_N^c} g \\
&\leq \epsilon + 2\epsilon \\
&= 3\epsilon.
\end{aligned}$$

as $\epsilon$ was arbitrary, the theorem is proven. □

## § 13. $L^1$ functions

The fact that integrable functions form a vector space is an important result about the algebraic properties of these functions. One fundamental fact is that this space of functions is complete with respect to the appropriate norm.

Definition: For any integrable function $f$ on $\mathbb{R}^d$ we define the *norm* of $f$ to be

$$||f|| = ||f||_{L^1} = ||f||_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)|dx.$$

Note that $||f|| = 0$ if and only if $f = 0$, thus in $L^1$ two functions are equivalent if they agree almost everywhere. Moreover, $L^1(\mathbb{R}^d)$ is a vector space and we have the following proposition:

**Proposition 26** *Suppose* $f, g \in L^1(\mathbb{R}^d)$*, then*

1. $||af||_{L^1(\mathbb{R}^d)} = |a|||f||_{L^1(\mathbb{R}^d)}$
2. $||f+g||_{L^1(\mathbb{R}^d)} \leq ||f||_{L^1(\mathbb{R}^d)} + ||g||_{L^1(\mathbb{R}^d)}$
3. $||f||_{L^1(\mathbb{R}^d)} = 0$ if and only if $f = 0$ a.e. $x$
4. $d(f,g) = ||f-g||_{L^1(\mathbb{R}^d)}$ defines a metric on $L^1(\mathbb{R}^d)$

Definition: A vector space is said to be *complete* if every Cauchy sequence converges. The following is a theorem about the completeness of $L^1$.

**Theorem 27 (Riesz-Fischer)** *The vector space $L^1$ is complete in the metric.*

*Proof.* Suppose $\{f_n\}$ is a Cauchy sequence in the norm. That is,

$$||f_n - f_m|| \to 0$$

as $n \to \infty$. We plan on constructing a subsequence converging to $f$. both point-wise almost everywhere and in the norm. We consider the subsequence $\{f_{n_k}\}$ of $\{f_n\}$ with the following property:

$$||f_{n_{k+1}} - f_{n_k}|| \leq \frac{1}{2^k}$$

for every $k \geq 1$. We are guaranteed such a sequence since $\{f_n\}$ is Cauchy. That is,

$$||f_n - f_m|| \leq \epsilon$$

whenever $m, n \geq N(\epsilon)$. Thus it suffices to take

$$n_k = N(\frac{1}{2^k}).$$

We now consider the following series

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}(f_{n_{k+1}} - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty}|f_{n_{k+1}} - f_{n_k}(x)|$$

Thus by the monotone convergence theorem $g$ is integrable and since $|f| \leq g$, so is $f$. In particular, the series defining $f$ converges almost everywhere and since the partial sums are the $f_{n_k}$, we find that

$$f_{n_k}(x) \to f(x)$$

a.e. $x$. We must now prove $f_{n_k} \to f$ in $L^1$. Note that

$$|f - f_{n_k}| \leq g$$

for every $k$, then by the dominated convergence theorem

$$||f_{n_k} - f||_{L^1} \to 0$$

as $k \to \infty$. Lastly, we must show $f \in L^1$. Recall that the sequence $\{f_n\}$ is Cauchy so given $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that whenever $m, n \geq N$ one has

$$||f_n - f_m|| < \frac{\epsilon}{2}$$

If $n_k$ is chosen such that $n_k > N$ and

$$||f_{n_k} - f|| < \frac{\epsilon}{2},$$

then by triangle inqeuality we have

$$|f_n - f|| \leq ||f_n - f_{n_k}|| + ||f_{n_k} - f|| < \epsilon$$

whenever $n > N$. Thus $\{f_n\}$ has limit $f \in L^1$ as desired. □

Since every sequence that converges in the norm is also a Cauchy sequence in that norm, we obtain the following Corollary:

**Corollary 28** *If $\{f_n\}$ converges to $f$ in $L^1$, then there exists a subsequence $\{f_{n_k}\}$ such that*

$$f_{n_k} \to f$$

*a.e. $x$.*

Definition: We say a family of integrable functions is *dense* in $L^1$ if for any $f \in L^1$ and $\epsilon > 0$ there exists an integrable function $g$ such that

$$||f - g||_{L^1} < \epsilon.$$

We describe those families of integrable functions that are dense in $L^1$ in the following theorem.

**Theorem 29** *The following families of functions are dense in $L^1$:*

1. *The simple functions.*
2. *The step functions.*
3. *The continuous functions of compact support.*

*Proof.* Let $f$ be integrable on $\mathbb{R}^d$ and assume $f$ is real valued. In this case we can write

$$f = f^+ - f^-$$

where $f^+, f^{\geq 0}$, it suffices to prove the theorem for when $f \geq 0$. For (1), by theorem 14 we are guaranteed the existence of non-negative simple functions $\{\varphi_k\}$ that increase to $f$ point wise. By the dominated convergence theorem we have

$$||f - \varphi_k||_{L^1} \to 0 \quad \text{as } k \to \infty,$$

thus there are simple functions arbitrarily close to $f$ in the $L^1$ norm.
For (2), note by (1) it suffices to approximate simple functions by step functions. Recall a simple

function is a finite linear combination of characteristic functions of finite measure $E$. So it suffices to show if $E$ is such a set, then

$$||\chi_E - \psi||_{L^1}$$

is small. As done in theorem 14's proof, there exists an almost disjoint collection of rectangles $\{R_j\}$ such that

$$m(E \triangle \bigcup_{j=1}^{M} \leq 2\epsilon.$$

Thus $\chi_E$ and $\psi = \sum_J \chi_{R_j}$ differ at most on a set of measure $2\epsilon$ thus

$$||\chi_E - \psi||_{L^1} < 2\epsilon.$$

And we are done with the proof of (2).
For (3) it suffices to use (2) when $f$ is the characteristic function of a rectangle. We choose a piece wise linear function defined via

$$g(x) = \begin{cases} 1 & ; x \in [a,b] \\ 0 & ; x \leq a - \epsilon, x \geq b + \epsilon \end{cases}.$$

Thus $g$ is linear on $[a-\epsilon, a]$ and $[b, b+\epsilon]$, then we have

$$||f - g||_{L^1} < 2\epsilon$$

as needed. □

*<back2top>*

### § 13.2 Invariance properties

Definition: If $f$ is a function defined on $\mathbb{R}^d$, the *translation* of $f$ by a vector $h \in \mathbb{R}^d$ is the function $f_h$ defined via

$$f_h(x) = f(x-h).$$

Here we examine some key properties of integrable functions. First thing, there is translation -invariance of the integral. That is, if $f$ is integrable then so it $f_h$, and

$$\int_{\mathbb{R}^d} f(x-h)dx = \int_{\mathbb{R}^d} f(x)dx \qquad (1)$$

The above equality follows from the fact that $mm(E) = m(E_h)$. Thus if $f$ is complex valued, then

$$\int_{\mathbb{R}^d} |f(x-h)|dx = \int_{\mathbb{R}^d} |f(x)|dx,$$

this shows $f_h \in L^1(\mathbb{R}^d)$ and also

$$||f_h|| = ||f||.$$

Thus (1) holds precisely when $f \in L^1(\mathbb{R}^d)$. Similarly to invariance properties of a measurable set, if $f$ is integrable, then so is $f(\delta x)$ and $f(-x)$ and we have

$$\delta^d \int_{\mathbb{R}^d} f(\delta x) = \int_{\mathbb{R}^d} f(x)dx,$$

and

$$\int_{\mathbb{R}^d} f(-x)dx = \int_{\mathbb{R}^d} f(x)dx.$$

Suppose now that $f, g$ are two integrable functions defined on $\mathbb{R}^d$ so that for some fixed $x \in \mathbb{R}^d$, the function defined via

$$y \mapsto f(x-y)g(y)$$

and as a consequence,

$$y \mapsto f(y)g(x-y)$$

is also integrable and we have

$$\int_{\mathbb{R}^d} f(x-y)g(y)dy = \int_{\mathbb{R}^d} f(y)g(x-y)dy.$$

Definition: The integral on the left hand side is denote $(f*g)(x)$ and defined as the *convolution* of $f$ and $g$.

Note that for any given $x \in \mathbb{R}^d$, the statement $f_h \to f$ as $h \to 0$ is the same as the continuity of $f$ at point $x$. It is possible that an integrable function $f$ be discontinuous at every $x$, however there is an overall continuity that some arbitrary $f \in L^1(\mathbb{R}^d)$ enjoys, one that holds in the norm.

**Proposition 30** *Suppose that $f \in L^1(\mathbb{R}^d)$. Then*

$$||f_h - f||_{L^1} \to 0$$

*as $h \to 0$.*

*Proof.* For any given $\epsilon > 0$, one can find a function $g$ such that

$$||f - g|| < \epsilon.$$

Then we can compute

$$f_h - f = (g_h - g) + (f_h - g_h) - (f - g),$$

however, we have

$$||f_h - g_h|| = ||f - g|| < \epsilon$$

while $g$ is continuous with compact support, we have that

$$||g_h - g|| = \int_{\mathbb{R}^d} |g(x-h) - g(x)|dx \to 0, \qquad \text{as } h \to \infty$$

Thus if $|h| < \delta$, where $\delta$ is sufficiently small, then

$$||g_h - g|| < \epsilon$$

and as a consequence,

$$||f_h - f|| < 3\epsilon, \qquad \text{whenever } |h| < \delta$$

and we are finished with the proof as $\epsilon$ was chosen arbitrarily. $\square$

<*back2top*>

## § 14 Fubini's Theorem

In elementary calculus integrals of continuous functions of several variables are often calculated by iterating one-dimensional integrals. In general, we may write $\mathbb{R}^d$ as the product

$$\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, \qquad \text{where } d = d_1 + d_2 \text{ and } d_1, d_2 \geq 1.$$

A point in $\mathbb{R}^d$ then takes the form $(x, y)$ where $x \in \mathbb{R}^{d_1}$, $y \in \mathbb{R}^{d_2}$, with this in mind we formalize the notion of a "slice".

Definition: If $f$ is a function in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, the *slice* of $f$ corresponding to $y \in \mathbb{R}^{d_2}$ is the function $f^y$ of the $x \in \mathbb{R}^{d_1}$ variable, given by

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

Similarly, the slice of $f$ for a fixed $x \in \mathbb{R}^{d_1}$ is $f_x(y) = f(x, y)$.

Definition: In the case of a set $E \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, we define its slices by

$$E^y := \{x \in \mathbb{R}^{d_1} : (x, y) \in E\} \text{and} E_x := \{y \in \mathbb{R}^{d_2} : (x, y) \in E\}$$

Note that $f$ being measurable on $\mathbb{R}^d$ does not guarantee the slice $f^y$ be measurable on $\mathbb{R}^{d_1}$. This leads us to our next big theorem.

**Theorem 31** *Suppose that f(x,y) is integrable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for almost every $y \in \mathbb{R}^{d_2}$,*

1. *The slice $f^y$ is integrable on $\mathbb{R}^{d_1}$*
2. *The function defined by $\int_{\mathbb{R}^{d_1}} f^y(x)dx$ is integrable on $\mathbb{R}^{d_2}$.*
3. *Moreover,*

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f(x, y)dx \right) dy = \int_{\mathbb{R}^d} f.$$

*Proof.* Any finite linear combination of functions that are integrable is still integrable. Thus we let $\{f_k\}_{k=1}^N$ be a sequence of integrable functions. For each $k$, there exists a set $A_k \subset \mathbb{R}^{d_2}$ of measure zero such that $f_k^y$ is integrable on $\mathbb{R}^{d_1}$ whenever $y \notin A_k$. If $A = \bigcup_{k=1}^N A_k$, then $m(A) = 0$ and in $A^c$ the $y-$slice corresponding to any finite linear combination of the $f_k$ is measurable, and also integrable. By linearity of the integral we conclude that any linear combination of the $f_k$'s is integrable. Let $\mathcal{F}$ denote the set of integrable functions on $\mathbb{R}^d$ which satisfy all three conditions of Theorem 31.

Suppose now that $\{f_k\}$ is a sequence of integrable functions such that $f_k \nearrow f$ or $f_k \searrow f$ where $f_k$ is integrable on $\mathbb{R}^d$, then so is $f$. Note it suffices to only consider the case of an increasing sequence. Also, we may replace $f_k$ by $f_k - f_1$ and assume the $f_k$'s are non-negative. Then by the monotone convergence theorem,

$$\lim_{k\to\infty} \int_{\mathbb{R}^d} f_k(x, y)dxdy = \int f(x, y)dxdy.$$

By our assumption, for each $k$, there exists a set $A_k \subset \mathbb{R}^{d_2}$ such that $f_k^y$ is integrable on $\mathbb{R}^{d_1}$ whenever $y \notin A_k$. If $A = \bigcup_{k=1}^{\infty} A_k$, then $m(A) = 0$ in $\mathbb{R}^{d_2}$ and if $y \notin A$, then $f_k^y$ is integrable on $\mathbb{R}^{d_1}$ for every $k$, and by the monotone convergence theorem,

$$g_k(y) = \int_{\mathbb{R}^{d_1}} f_k^y dx \quad \text{increases to a limit} \quad g(y) = \int_{\mathbb{R}^{d_1}} f^y(x)dx.$$

as $k \to \infty$. By assumption, each $g_k$ is integrable, so another application of the monotone convergence theorem yields

$$\int_{R^{d_2}} g_k(y)dy \to \int_{\mathbb{R}^{d_2}} g(y)dy \quad \text{as } k \to \infty.$$

By assumption that the $f_k$ are integrable we have

$$\int_{\mathbb{R}^{d_2}} g_k(y)dy = \int_{\mathbb{R}^d} f_k(x,y)dxdy$$

Combining this with the facts up top we get that

$$\int_{\mathbb{R}^{d_2}} g(y)dy = \int_{\mathbb{R}^d} f_k(x,y)dxdy.$$

Since $f$ is integrable, the right hand side is finite which implies $g$ is integrable. Consequently, $g(y) < \infty$ a e. $y$ hence $f^y$ is integrable for a.e. $y$ and

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f(x,y)dx \right) dy = \int_{\mathbb{R}^d} f(x,y)dxdy$$

This proves $f$ is integrable over $\mathbb{R}^d$, that is, $f \in \mathcal{F}$.

Next, we show any characteristic function of a set $E$ that is a $G_\delta$ set and of finite measure is integrable over $\mathbb{R}^d$. First suppose $E$ is a bounded open cube, so

$$E = Q_1 \times Q_2, \qquad \text{where } Q_1,\ Q_2 \text{ are open cubes in } \mathbb{R}^{d_1},\ \mathbb{R}^{d_2} \text{ respectively}$$

Then for each $y$, the function $\chi_E(x,y)$ is measurable in $x$ and integrable with

$$g(y) = \int_{\mathbb{R}^{d_1}} \chi_E(x,y)dx \begin{cases} |Q_1| & ; y \in Q_2 \\ 0 & ; \text{otherwise} \end{cases}.$$

Consequently, $g = |Q_1|\chi_{Q_2}$ is also measurable and integrable with

$$\int_{\mathbb{R}^{d_2}} g(y)dy = |Q_1||Q_2|.$$

Since we have

$$\int_{\mathbb{R}^d} \chi_E(x,y)dxdy = |E| = |Q_1||Q_2|,$$

we deduce that $\chi_E$ is integrable over $\mathbb{R}^d$. Now suppose $E$ is a subset of the boundary of some closed cube. Then since the boundary has measure zero in $\mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d} \chi_E(x,y)dxdy = 0.$$

Next we note that for a.e. $y$, the slice $E^y$ has measure zero and therefore if $g(y) = \int_{\mathbb{R}^{d_1}} \chi_E(x,y)dx$ we have $g(y) = 0$ a.e. $y$ and thus

$$\int_{\mathbb{R}^{d_2}} g(y)dy = 0$$

thus $\chi_E \in \mathcal{F}$.

Next suppose that $E$ is a finite union of closed cubes with disjoint interiors. Say $E = \bigcup_{k=1}^N Q_k$. If $\tilde{Q}_k$ denotes the interior of the $Q_k$, then we can write $\chi_E$ as a linear combination of the $\tilde{Q}_k$ and $\chi_{A_k}$ where $A_k \subset \partial Q_k$. By our analysis we know $\chi_{Q_k}$ and $\chi_{A_k}$ are integrable over $\mathbb{R}^d$ and since the space of integrable functions is closed under linear combinations, we have that$\chi_E$ is also integrable over $\mathbb{R}^d$.

Next we prove if $E$ is open and of finite measure that $\chi_E \in \mathcal{F}$. By an earlier theorem, we can write $E$ as a countable union of almost disjoint closed cubes

$$E = \bigcup_{k=1}^\infty Q_k.$$

Consequently, if we let

$$f_k = \sum_{j=1}^k \chi_{Q_j}$$

then note $f_k \nearrow f = \chi_E$ which is integrable since $m(E)$ is finite. Therefore by an earlier argument in this proof, $f \in \mathcal{F}$.

Finally, if $E$ is a $G_\delta$ set of finite measure, then $\chi_E$ is integrable. By definition, there exists open sets $\{\tilde{\mathcal{O}}_k\}_{k=1}^\infty$ such that

$$E = \bigcap_{k=1}^\infty \tilde{\mathcal{O}}_k.$$

Since $E$ has finite measure, there exists some open set $\tilde{\mathcal{O}}_0$ of finite measure with $E \subset \tilde{\mathcal{O}}_0$. If we let

$$\mathcal{O}_k = \mathcal{O}_0 \cap \bigcap_{j=1}^k \tilde{\mathcal{O}}_j,$$

then we have a decreasing sequence of finite measure sets $\mathcal{O}_1 \supset \mathcal{O}_2 \supset ...$ with

$$E = \bigcap_{k=1}^\infty \mathcal{O}_k$$

and thus the sequence of functions $\chi_{\mathcal{O}_k}$ decreases to $f = \chi_E$ and since $\chi_{\mathcal{O}_k}$ is integrable over $\mathbb{R}^d$ for every $k$, we conclude that $\chi_E \in \mathcal{F}$ as well.

Next If $E$ has measure 0, then we show $\chi_E \in \mathcal{F}$. Since $E$ is measurable, we may choose a set $G$ of type $G_\delta$ with $E \subset G$ and $m(G) = 0$. Since $\chi_G$ is integrable over $\mathbb{R}^d$, we find that

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}d_1} \chi_G(x,y) \right) dy = \int_{\mathbb{R}^d} \chi_G = 0.$$

Therefore

$$\int_{\mathbb{R}^{d_1}} \chi_G(x,y)dx = 0 \qquad \text{for a.e. } y.$$

Consequently, the slice $G^y$ has measure 0 for a.e. $y$. Noting that $E^y \subset G^y$ yields $m(E^y) = 0$ a.e. $y$ and thus $\int_{\mathbb{R}^{d_1}} \chi_E(x,y)dx = 0$ for a.e. $y$. Therefore,

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} \chi_E(x,y)dx \right) dy = 0 = \int_{\mathbb{R}^d} E,$$

Thus $\chi_E$ is integrable over $\mathbb{R}^d$.

The next stage is if $E$ is any subset of $\mathbb{R}^d$ with finite measure. To prove this, recall first that there exists a set of finite measure $G$ of type $G_\delta$ with $E \subset G$ and $m(G \setminus E) = 0$. Since

$$\chi_E = \chi_G - \chi_{G \setminus E},$$

and integrable functions are closed under linear combinations, we find that $\chi_E \in \mathcal{F}$.

Lastly, we prove if $f$ is integrable, then $f \in \mathcal{F}$. Note first that we can write

$$f = f^+ - f^-$$

Thus by the first step, we may assume $f$ is non-negative. By theorem 12, there exists a sequence of simple functions $\{\varphi_k\}$ that increase to $f$. Since each $\varphi_k$ is a finite linear combination of characteristic functions of sets with finite measure, we have that $\varphi_k \in \mathcal{F}$ by steps 1 and 5, and hence by the second step $f \in \mathcal{F}$. □

§ **14.2 Applications of Fubini**

We begin this section with a theorem.

**Theorem 32** *Suppose $f(x,y)$ is a non-negative measurable functions on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for almost every $y \in \mathbb{R}^{d_2}$:*

1. *The slice $f^y$ is measurable on $\mathbb{R}^{d_1}$.*
2. *The function defined by $\int_{\mathbb{R}^{d_1}} f^y(x)dx$ is measurable on $\mathbb{R}^{d_2}$.*
3. $\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f(x,y)dx \right) dy = \int_{\mathbb{R}^d} f(x,y)dxdy$ *in the extended sense.*

*Proof.* Consider the truncations

$$f_k(x,y) = \begin{cases} f(x,y) & ;\text{if } |(x,y)| < k \text{ and } f(x,y) < k \\ 0 & ;\text{otherwise} \end{cases}.$$

So each $f_k$ is measurable and by part (i) of Fubini's Theorem, there exists a set $E_k \subset \mathbb{R}^{d_2}$ of measure zero such that the slice $f_k^y(x)$ is measurable for all $y \in E_k^c$. Then if we set $E = \bigcup_k E_k$, we find that $f^y(x)$ is measurable for all $y \in E^c$ and all $k$. Moveover, $m(E) = 0$ and since $f_k^y \nearrow f^y$, by the monotone convergence theorem we have that if $y \notin E$, then

$$\int_{\mathbb{R}^{d_1}} f_k(x,y)dx \nearrow \int_{\mathbb{R}^{d_1}} f(x,y)dx, \qquad \text{as } k \to \infty.$$

Again, by Fubini's Theorem, $\int_{\mathbb{R}^{d_1}} f_k(x,y)dx$ is measurable for all $y \in E^c$, hence so is $\int_{\mathbb{R}^{d_1}} f_(x,y)dx$. Another application of the MCT gives us

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f_k(x,y)dx \right) dy \to \int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f(x,y)dx \right) dy.$$

However, by part (iii) of Fubini's theorem we have that

$$\int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f_k(x,y)dx \right) dy = \int_{\mathbb{R}^d} f_k.$$

A final application of the MCT gives that

$$\int_{\mathbb{R}^d} f_k \to \int_{\mathbb{R}^d} f$$

combining the last three lines gives the desired result. □

The following is an immediate result of Theorem 32 applied to the function $\chi_E$

**Corollary 33** *If $E$ is a measurable set in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, then for almost every $y \in \mathbb{R}^{d_2}$ the slice*

$$E^y = \{c \in \mathbb{R}^{d_1} : (x,y) \in E\}$$

*is a measurable subset of $\mathbb{R}^{d_1}$. Moreover $M(E^y)$ is a measurable function of $y$ and*

$$m(E) = \int_{\mathbb{R}^{d_2}} M(E^y)dy.$$

Definition: In relating a set $E$ to its slices $E_x$ and $E^y$, matters are straightforward for the basic sets which arise when we consider $\mathbb{R}^d$ as the product $R^{d_1} \times \mathbb{R}^{d_2}$, these are the *product set* $E = E_1 \times E_2$ with $E_j \subset \mathbb{R}^{d_j}$. This leads us to our next Proposition:

**Proposition 34** *If $E = E_1 \times E_2$ is a measurable subset of $\mathbb{R}^d$ and $m_*(E_2) > 0$, then $E_1$ is measurable.*

*Proof.* By Corollary 33, we know for a.e $y \in \mathbb{R}^{d_2}$, the slice function

$$(\chi_{E_1 \times E_2})^y(x) = \chi_{E_1}(x)\chi_{E_2}(y)$$

is measurable as a function of $x$. In fact, we claim that there exists some $y \in E_2$ such that the above slice function is measurable in $x$. For such a $y$ one would have

$$\chi_{E_1 \times E_2}(x,y) = \chi_{E_1}(x)$$

which would imply $E_1$ is measurable. To prove existence of such a $y$, we use the fact that $m_*(E_2) > 0$. Let $F$ denote the set of $y$ such that $E^y$ is measurable. Then by the previous corollary, $m(F^c) = 0$. However $E_2 \cap F \neq \emptyset$ because $m_*(E_2 \cap F) > 0$. To see this note that

$$E_2 = (E_2 \cap F) \bigcup (E_2 \cap F^c),$$

hence

$$0 < m_*(E_2) = m_*(E_2 \cap F) + m_*(E_2 \cap F^c) = m_*(E_2 \cap F)$$

as $E_2 \cap F^c$ is a subset of a set of measure zero. □

To deal with the converse of this Proposition, we need the following lemma:

**Lemma 35** *If $E_1 \subset \mathbb{R}^{d_1}$ and $E_2 \subset \mathbb{R}^{d_2}$, then*

$$m_*(E_1 \times E_2) \leq m_*(E_1)m_*(E_2),$$

*with the understanding that if one of the sets $E_j$ has exterior measure zero, then $m_*(E_1 \times E_2) = 0$.*

*Proof.* Let $\epsilon > 0$ be given. We know we are guaranteed the existence of two collections of cubes $\{Q_k\}_k$ and $\{\tilde{Q}_j\}$ in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ respectively such that

$$E_1 \subset \bigcup_k Q_k \quad \text{and} \quad E_2 \subset \bigcup_j \tilde{Q}_j$$

and

$$\sum_k |Q_k| \leq m_*(E_1) + \epsilon \quad \text{and} \quad \sum_j |\tilde{Q}_j| \leq m_*(E_2) + \epsilon$$

Since $E_1 \times E_2 \subset \bigcup_{j,k=1}^{\infty} Q_k \times \tilde{Q}_j$, then by sub-additivity, we have

$$\begin{aligned} m_*(E_1 \times E_2) &\leq \sum_{j,k=1}^{\infty} |Q_k \times \tilde{Q}_j| \\ &= \left(\sum_k |Q_k|\right)\left(\sum_j |\tilde{Q}_j|\right) \\ &\leq (m_*(E_1) + \epsilon)(m_*(E_2) + \epsilon). \end{aligned}$$

Let us suppose neither $E_1$ nor $E_2$ has exterior measure zero, then we obtain

$$m_*(E_1 \times E_2) \leq m_*(E_1)m_*(E_2) + \mathcal{O}(\epsilon)$$

and since $\epsilon$ was arbitrary, we conclude that

$$m_*(E_1 \times E_2) \leq m_*(E_1)m_*(E_2).$$

If for instance one has that, without any loss of generality, $m(E_1) = 0$, consider for each positive integer $j$ the set

$$e_2^j = E_2 \cap \{y \in \mathbb{R}^{d_2} : |y| \leq j\}.$$

Then by the above argument, we find that $m_*(E_1 \times E_2^j) = 0$. And since $(E_1 \times E_2^j) \nearrow (E_1 \times E_2)$ as $j \to \infty$, we conclude that $m_*(E_1 \times E_2) = 0$ as well. $\square$

The next proposition relates slices to the Lebesgue measure.

**Proposition 36** *Suppose $E_1$ and $E_2$ are measurable subsets of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ respectively. Then $E = E_1 \times E_2$ is a measruable subset of $\mathbb{R}^d$. Moreover,*

$$m(E) = m(E_1)m(E_2),$$

*with the understanding that if one of the sets has measure zero, then $m(E) = 0$.*

*Proof.* It suffices to prove $E$ is measurable, because then the assertion about $m(E)$ follows from Corollary 33. Since each $E_j$ is measurable, there exists a set $G_j \subset \mathbb{R}^{d_j}$ of type $G_\delta$ with $E_j \subset G_j$ such that

$$m_*(G_j \setminus E_j) = 0$$

Clearly $G = G_1 \times G_2$ is measurable in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and we can write

$$\begin{aligned} G \setminus E &= (G_1 \times G_2) \setminus (E_1 \times E_2) \\ &\subset ((G_1 \setminus E_1) \times G_2) \cup (G_1 \times (G_2 \setminus E_2)) \\ &= \emptyset \end{aligned}$$

We conclude that $m_*(G \setminus E) = 0$ thus $E$ is measurable. □

As a consequence of this Proposition, we have the following Corollary:

**Corollary 37** *Suppose $f$ is a measurable function on $\mathbb{R}^{d_1}$. Then the function $\tilde{f}$ defined via*

$$\tilde{f}(x, y) = f(x)$$

*is measurable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$.*

*Proof.* To see this, we may assume $f$ is real valued. Recall that if $a \in \mathbb{R}$, and $E_1 := \{x \in \mathbb{R}^{d_1} : f(x) < a\}$, then $E_1$ is measurable by definition. Since

$$\{(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} : \tilde{f}(x, y) < a\} = E_1 \times \mathbb{R}^{d_2},$$

the previous proposition shows that $\{\tilde{f}(x, y) < a\}$ is measurable for each $a \in \mathbb{R}$ forcing $\tilde{f}(x, y)$ to be measurable on all of $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. □

Finally, we return to an interpretation of the integral that first arose in calculus. We have in mind the notion that $\int f$ describes the area under the graph of $f$. Here we relate this to the Lebesgue integral and show how it extends to a more general context.

**Corollary 38** *Suppose $f(x)$ is a non-negative function on $\mathbb{R}^d$ and let*

$$\mathcal{A} := \{(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} : 0 \leq y \leq f(x)\},$$

*then*

1. *$f$ is measurable on $\mathbb{R}^d$ if and only if $\mathcal{A}$ is measurable in $\mathbb{R}^{d+1}$.*
2. *If the conditions in (1) hold, then*

$$\int_{\mathbb{R}^d} f(x)dx = m(\mathcal{A}).$$

*Proof.* If $f$ is measurable on $\mathbb{R}^d$, then the previous proposition guarantees the existence of a function

$$F(x, y) = y - f(x)$$

and is measurable in $\mathbb{R}^{d+1}$. Thus we see that

$$\mathcal{A} = \{y \geq 0\} \cap \{F \leq 0\}$$

is measurable.
Conversely, suppose $\mathcal{A}$ is measurable. Note that for each $x \in \mathbb{R}^{d_1}$, the slice

$$\mathcal{A}_x = \{y \in \mathbb{R} : (x, y) \in \mathcal{A}\}$$

is a closed segment, namely $\mathcal{A}_x = [0, f(x)]$. Consequently, by theorem 33 gives us the measurability of $m(\mathcal{A}_x) = f(x)$. Moreover,

$$\begin{aligned} m(\mathcal{A}) &= \int \chi_{\mathcal{A}}(x, y) dx dy \\ &= \int_{\mathbb{R}^{d_1}} m(\mathcal{A}_x) \\ &= \int_{\mathbb{R}^{d_1}} f(x) dx. \end{aligned}$$

as desired. □

We conclude this section with a useful proposition.

**Proposition 39** *If $f$ is a measurable function on $\mathbb{R}^d$, then the function $\tilde{f}(x, y) = f(x-y)$ is measurable on $\mathbb{R}^d \times \mathbb{R}^d$.*

*Proof.* By picking $E = \{z \in \mathbb{R}^d : f(z) < a\}$, we see it suffices to prove that whenever $E$ is a measurable subset of $\mathbb{R}^d$, then $\tilde{E} = \{(x, y) : x - y \in E\}$ is a measurable subset of $\mathbb{R}^d \times \mathbb{R}^d$. Note that if $\mathcal{O}$ is open, then $\tilde{\mathcal{O}}$ is open as well. Taking countable intersections shows that if $E$ is a $G_\delta$ set, then so is $\tilde{E}$. Letting $B_k = \{|y| < k\}$ then define

$$\tilde{E}_k = \tilde{E} \cap B_k$$

then notice that for each $k$, we have $m(\tilde{E}_k) = 0$. Take $\mathcal{O}$ to be open in $\mathbb{R}^d$ and compute out $m(\tilde{\mathcal{O}} \cap B_k)$, where we have $\chi_{\tilde{\mathcal{O}} \cap B_k} = \chi_{\mathcal{O}}(x - y)\chi_{B_k}(y)$:

$$\begin{aligned} m(\tilde{\mathcal{O}} \cap B_k) &= \int \chi_{\mathcal{O}}(x - y)\chi_{B_k}(y) dy dx \\ &= \int \left( \int \chi_{\mathcal{O}}(x - y) dx \right) \chi_{B_k}(y) dy \\ &= m(\mathcal{O}) m(B_k) \end{aligned}$$

by translation invariance of Lebesgue measure. Now if $m(E) = 0$ then there exists a sequence of open sets $\mathcal{O}_n$ such that $E \subset \mathcal{O}_n$ and $m(\mathcal{O}_n) \to 0$ as $n \to \infty$. It follows from above that

$$\tilde{E}_k \subset \tilde{\mathcal{O}}_n \cap B_k$$

and

$$m(\mathcal{O}_n \cap B_k) \to 0$$

in $n$ for each fixed $k$. This shows $m(\tilde{E}_k) = 0$ and hence $m(\tilde{E}) = 0$. The proof is concluded by recalling any measurable set $E$ can be written as the difference of a $G_\delta$ set and a set of measure zero. □

<back2top>

## § 15 A Fourier inversion formula

The question of the inversion of the Fourier transform encompasses in effect the problem at the origin of Fourier analysis. This involves establishing the validity of the inversion formula for a function $f$ in terms of its Fourier transform $\tilde{f}$. That is,

$$\tilde{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \xi} dx, \qquad (***)$$

and

$$f(x) = \int_{\mathbb{R}^d} \tilde{f}(\xi) e^{2\pi i x \xi} d\xi. \qquad (****)$$

The most elegant and useful formulations of Fourier inversions are in terms of the $L^2$ theory. This leads us to our first theorem, or result, of this section.

**Proposition 40** *Suppose $f \in L^1(\mathbb{R}^d)$. Then $\tilde{f}$ defined in $(***)$ is continuous and bounded on $\mathbb{R}^d$.*

*Proof.* Note since

$$|f(x)e^{-2\pi i x \xi}| = |f(x)|,$$

the integral representing $\tilde{f}$ converges for every $\xi$ and

$$\sup_{\xi \in \mathbb{R}^d} |\tilde{f}(\xi)| \leq \int_{\mathbb{R}^d} |f(x)| = ||f||.$$

to verify $\tilde{f}$ is continuous, note for every $x$,

$$f(x)e^{-2\pi i x \xi} \to f(x) e^{-2\pi i x \xi_0}$$

as $\xi \to \xi_0$. Hence $\tilde{f}(\xi) \to \tilde{f}(\xi_0)$ by the dominated convergence theorem. □

One has a bit more, in fact, one has $\tilde{f}(\xi) \to 0$ as $|\xi| \to 0$ but not much can be said about the decrease at infinity. As a consequence, for a general $f \in L^1(\mathbb{R}^d)$, it is not generally true that $\tilde{f} \in L^(\mathbb{R}^d)$. Then the inversion formula $(****)$ becomes problematic.

**Theorem 41** *Suppose $f \in L^(\mathbb{R}^d)$ and assume also that $\tilde{f} \in L^(\mathbb{R}^d)$. Then the formula $(****)$ holds for almost every $x$.*

An immediate consequence follows:

**Corollary 42** *Suppose $\tilde{f}(\xi) = 0$ for all $\xi$. Then $f = 0$ a.e. $x$.*

Now for the "multiplication" formula.

**Lemma 43** *Suppose $f, g \in L^1(\mathbb{R}^d)$. Then*

$$\int_{\mathbb{R}^d} \tilde{f}(\xi) g(\xi) d\xi = \int_{\mathbb{R}^d} f(y) \tilde{g}(y) dy.$$

<*back2top*>

§ **Exercises (with solutions)**

*Exercise 1* Suppose $f$ is integrable on $(-\pi, \pi]$ and extended to $\mathbb{R}$ by making int periodic on $2\pi$. Show that

$$\int_{-\pi}^{\pi} .f(x)dx = \int_I f(x)dx,$$

where $I$ is is any interval in $\mathbb{R}$ of length $2\pi$.

*Proof.* Given that $f$ is $2\pi$ periodic, we have that $f(x) = f(x + 2\pi n)$ for any $n \in \mathbb{N}$. As the hint tells us, observe that for some $k \in \mathbb{N}$, one has

$$I = (a, b] \subset (k\pi, (k + 4)\pi],$$

and define $c = (k + 2)\pi$. Clearly if $c \neq a$, then $c \in (a, b]$. Thus

$$\int_{(a,b]} f(x)dx = \int_{(a,c]} f(x)dx + \int_{(c,b]} f(x)dx.$$

And since $f(x)|_{(c,b]} = f(x - 2\pi)|_{(k\pi,a]}$ we have

$$\int_{(c,b]} f(x)dx = \int_{(k\pi,a]} f(x)dx$$

Thus we can rewrite

$$\int_{(a,b]} f(x)dx = \int_{(k\pi,c]} f(x)dx + \int_{(k\pi,(k+2)\pi]} f(x)dx$$

Finally, we can break up the integral into

$$\int_{(k\pi,(k+2)\pi]} f(x)dx = \int_{(k\pi,(k+1)\pi]} f(x)dx + \int_{((k+1)\pi,(k+2)\pi]} f(x)dx.$$

However by its $2\pi$ periodicity, we have that

$$\int_{(a,b]} f(x)dx = \int_{(k\pi,(k+2)\pi]} f(x)dx = \int_{((k+1)\pi,(k+3)\pi]}$$

And thus this inequality holds for every integer, specially $k = -1$. And therefore

$$\int_{(a,b]} f(x)dx = \int_{(-\pi,\pi]} f(x)dx.$$

as desired. □

*Exercise 2* Suppose $f$ is integrable on $[0, b]$ and

$$g(x) = \int_x^b \frac{f(t)}{t}dt, \qquad \text{for } x \in (0, b]$$

Prove that $g$ is integrable on $[0, b]$ and that

$$\int_0^b g(x)dx = \int_0^b f(t)dt.$$

*Proof.* It suffices to assume $f$ to be a non-negative $L^1$ function. Define the set

$$T(b) := \{(x,t) \in \mathbb{R} \times \mathbb{R} : x \in (0,t], t \in [x,b]\}.$$

Next, define the function

$$F(x,t) = \frac{f(t)}{t}\chi_{T(b)}.$$

Certainly, $F$ is measurable since $\chi_{T(b)}$, $f(t)$, and $\frac{1}{t}$ are all measurable functions. Next, note that

$$\begin{aligned}\int_{[0,b]} g(x)dx &= \int_{[0,b]}\int_{[x,b]} \frac{f(t)}{t}dtdx \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{f(t)}{t}\chi_{T(b)}dtdx\end{aligned}$$

By Tonelli's Theorem then we have

$$\begin{aligned}\int_{[0,b]}\int_{[x,b]} \frac{f(t)}{t}dtdx &= \int_{[0,b]}\int_{[0,t]} \frac{f(t)}{t}dxdt \\ &= \int_{[0,b]} f(t)dt\end{aligned}$$

Thus since $f \in L^1$ we have

$$\int_{[0,b]} g(x)dx = \int_{[0,b]} f(t)dt < \infty$$

implying $g \in L^1$ as desired. □

*Exercise 3* Let $\Gamma \subset \mathbb{R}^d \times \mathbb{R}$, and $\Gamma := \{(x,y) \in \mathbb{R}^d \times \mathbb{R} : y = f(x)\}$ and assume $f$ is measurable on $\mathbb{R}^d$. Show then that $\Gamma \subset \mathbb{R}^{d+1}$ is measurable and $m(\Gamma) = 0$.

*Proof.* First, assume $f$ is almost everywhere finite. Next, partition $\mathbb{R}^d$ into an almost disjoint union of closed unit cubes $\{Q_k\}_{k=1}^{\infty}$ and define

$$\Gamma_k := \{(x,y) \in Q_k \times \mathbb{R} : y = f(x)\}.$$

Next, define the sets

$$F_{k,n}^i := \{x \in Q_k : \frac{i}{2^n} \le f(x) < \frac{i+1}{2^n}\}.$$

Now define

$$E_{k,n}^i := F_{k,n}^i \times [\frac{i}{2^n}, \frac{i+1}{2^n}).$$

Finally, set

$$E_{k,n} = \bigcup_{i=-\infty}^{\infty} E_{k,n}^i$$

where all these sets are measurable since $f$ is measurable. Notice that

$$\Gamma_k \subset E_{k,n}, \qquad \text{for all } n \in \mathbb{N}$$

and that

$$E_{k,n+1} \subset E_{k,n}, \qquad \text{for all } n \in \mathbb{N} \qquad (*)$$

Next, note that

$$m^{d+1}(E_{k,n}) \leq \sum_{i=-\infty}^{\infty} m^{d+1}(E^i_{k,n}) \leq ...$$

$$... \leq \sum_{i=-\infty}^{\infty} m^d(F^i_{k,n}) \cdot m([\frac{i}{2^n}], \frac{i+1}{2^n}) = ...$$

$$... = \frac{1}{2^{n-1}} \sum_{i=-\infty}^{\infty} m^d(F^i_{k,n}) \leq \frac{1}{2^{n-1}} m^d(Q_k) = \frac{1}{2^{n-1}}.$$

By $(*)$, the $E_{k,n}$ are collapsing with finite measure thus $\Gamma_k$ need be measurable as well since

$$m_*(\Gamma_k) \leq \lim_{n\to\infty} m(E_{k,n}) = \lim_{n\to\infty} \frac{1}{2^{n-1}} = 0.$$

Thus for each $k$, $\Gamma_k$ is measurable forcing $\Gamma$ to be measurable since it is a countable union of measurable sets. Lastly, observe that

$$m^{d+1}(\Gamma) \leq \sum_{k=1}^{\infty} m^{d+1}(\Gamma_k) = 0$$

as we needed to show. □

*Exercise 4* If $f$ is integrable on $\mathbb{R}$, show that $F(x) = \int_{-\infty}^{x} f(t)dt$ is uniformly continuous.

*Proof.* By Proposition 24 (2), for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\int_E |f| < \epsilon, \qquad \text{whenever } m(E) < \delta$$

Next, note that

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_{-\infty}^{x} f(t)dt - \int_{-\infty}^{y} f(t)dt \right| \\ &\leq \left| \int_{x}^{y} f(t)dt \right| \\ &\leq \int -x^y |f(t)|dt \end{aligned}$$

Since we have that $f$ is integrable, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\int_x^y |f(t)|dt < \epsilon, \qquad \text{whenever } |x - y| < \delta,$$

thus $F(x)$ is uniformly continuous. □

*Exercise 5. Tchebychev Inequality* Suppose $f \geq 0$ and $f$ is integrable. If $\alpha > 0$ and $E_\alpha = \{f : f(x) > \alpha\}$, prove that

$$m(E_\alpha) \leq \frac{1}{\alpha}\int f.$$

*Proof.* We rewrite our set $E_\alpha$ as

$$E_\alpha := \{x : \frac{f(x)}{\alpha} > 1\}.$$

Then we can compute

$$\begin{aligned} m(E_\alpha) &= \int_{E_\alpha} dx \\ &\leq \int_{E_\alpha} \frac{f(x)}{\alpha} dx \\ &\leq \frac{1}{\alpha}\int f(x)dx. \end{aligned}$$

□

*Exercise 6* Prove if $f$ is integrable on $\mathbb{R}^d$, real-valued, and $\int_E f(x)dx \geq 0$ for every measurable set $E$, then $f(x) \geq 0$ a.e. $x$. As a result, if $\int_E f(x)dx = 0$ for every measurable $E$, then $f(x) = 0$ a.e. $x$.

*Proof.* Consider first the set

$$A = \{x : f(x) < 0\}.$$

Clearly we have that

$$A = \bigcup_{n=1}^{\infty} \{x : f(x) < -\frac{1}{n}\}.$$

Assume towards a contradiction that $m(A) > 0$. We have that

$$m(A) \leq \sum_{n=1}^{\infty} m(\{x : f(x) < -\frac{1}{n}\}).$$

Since $m(A) > 0$ for at least one $n \in \mathbb{N}$, call this set $E$. Then

$$\begin{aligned} \int_E f &\leq \int_E -\frac{1}{n} \\ &= -\frac{1}{n}m(E) \\ &< 0. \end{aligned}$$

a contradiction. Thus the same reasoning will result if $f \leq 0$ so combining these two facts results in given any measurable subset $S \subset \mathbb{R}^d$ such that

$$\int_S f = 0$$

will result in $0 \leq f \leq 0$ a.e. $x$ i.e., $f = 0$ a.e. $x$. □

*Exercise 7* Prove that if $f$ is integrable on $\mathbb{R}^d$ and $\delta > 0$, then $f(\delta x) \to f$ in the $L^1$ norm as $\delta \to 1$.

*Proof.* Since continuous functions of compact support are dense in $L^1(\mathbb{R}^d)$, for any $\epsilon > 0$, there exists a continuous function with compact support $g$ such that

$$||f - g||_{L^1(\mathbb{R}^d)} < \frac{\epsilon}{3}.$$

It follows that since

$$f(\delta x) - f(x) = f(\delta x) - g(\delta x) + g(\delta x) - g(x) + g(x) - f(x),$$

by the triangle inqequality we have

$$||f(\delta x) - f(x)||_{L^1} \leq ||g(\delta x) - g(x)||_{L^1} + ||f(\delta x) - g(\delta x)||_{L^1} + ||f(x) - g(x)||_{L^1}$$

or just

$$||f(\delta x) - f(x)||_{L^1} \leq ||g(\delta x) - g(x)||_{L^1} + ||f(\delta x) - g(\delta x)||_{L^1} + \frac{\epsilon}{3}.$$

By the dilation properties, we see that

$$||f(\delta x) - g(\delta x)||_{L^1} = \frac{1}{|\delta|^d}||f - g||_{L^1} \leq \frac{\epsilon}{3|\delta|^d} \leq \frac{\epsilon}{3}.$$

Note since $g$ is continuous with compact support, then $g(x)$ and $g(\delta x)$ are uniformly continuous and thus $g$ attains its max value, thus there exists a bound for $g$, call it $M$. Next, note that the sets $E$ and $\delta E$ are both compact. Then $E \triangle \delta E$ and $E \cap \delta E$ are compact as well. It will then follow that

$$\begin{aligned} \int_{E \cup \delta E} |g(x) - g(\delta x)| dx &= \int_{E \cap \delta E} |g(x) - g(\delta x)| dx + \int_{E \triangle \delta E} |g(x) - g(\delta x)| dx \\ &\leq \int_{E \cap \delta E} |g(x) - g(\delta x)| dx + 2Mm(E \triangle \delta E) \end{aligned}$$

Next, observe that for every $\alpha \neq 1$, there exists some $\alpha$ such that $\delta x = x + \frac{x}{\alpha}$. Since $E \cap \delta E$ is compact, we can select the diameter

$$r = \max_{x,y \in E \cap \delta E} |d(x, y)|.$$

Certainly now $x, x + \frac{x}{\alpha} \in B_{\frac{|r|}{|\alpha|}}(x)$. Thus, for any $\xi > 0$, we can have for a $\delta$ sufficiently close to 1, we have

$$|x - \delta x \leq \frac{|r|}{|\alpha|} < \xi.$$

Thus for any $\epsilon > 0$, there exists a $\delta$ sufficiently close to 1 such that

$$\int_{E \cap \delta E} |g(x) - g(\delta x)| dx < \frac{\epsilon}{6}.$$

Since $m(E \triangle \delta E) < \infty$, we have for some $\delta$ sufficiently close to 1 that

$$m(E \triangle \delta E) < \frac{\epsilon}{12}.$$

Thus choose $\delta$ close enough to 1 such that

$$||g(x) - g(\delta x)||_{L^1} = \int_{E \cup \delta E} |g(x) - g(\delta x)| dx < \frac{\epsilon}{6} + 2M \frac{\epsilon}{12M}.$$

Finally, for $\delta$ sufficiently close to 1, we have

$$||f(\delta x) - f(x)||_{L^1} \leq ||g(\delta x) - g(x)||_{L^1} + ||f(\delta x) - g(\delta x)||_{L^1} + ||f(x) - g(x)||_{L^1} = \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,$$

as desired. □

<back2top>

## § 17 Differentiation

The differentiation and integration are inverse operations. This fact was understood early on in calculus. Our objective is the formulation and proof of the fundemental theorem of calculus. We shall try to achieve this by answering two questions. The first questions will be stated as follows.

- Suppose $f$ is integrable on $[a, b]$, and $F$ is its indefinite integral $F(x) = \int_a^x f(y)dy$. Does this imply that $F$ is differentiable (at least for almost every $x$) and that $F' = f$?

For the second question, we reverse the order of differentiation and integration.

- What conditions on a function $F$ on $[a, b]$ guarantee that $F'$ exists, that this function is integrable, and that moreover

$$F(b) - F(a) = \int_a^b F'(x)dx?$$

In particular, we shall find that this question is connected to the problem of rectifiability of curves.

<back2top>

### § 17.1 Differentiation of the integral

We begin with the first problem, differentiation of the integral. If $f$ is given on $[a, b]$ and integrable on that interval, we let

$$f(x) = \int_a^x f(y)dy, \qquad a \leq x \leq b.$$

To deal with $F'$, recall the definition of a derivative as the limit of the quotient

$$\frac{F(x+h) - F(x)}{h}, \qquad \text{when } h \text{ tends to } \infty$$

We note that this quotient takes on the form

$$\frac{1}{h} \int_x^{x+h} f(y)dy = \frac{1}{|I|} \int_I f(y)dy,$$

where we use the notation $I = (x, x+h)$ and $|I|$ for length of this interval. We pause to note the above expression is the "average" value of $f$ over $I$, and that in the limit $|I| \to 0$, we may expect these averages to tend to $f(x)$. Reformulating the question slightly, we may ask whether

$$\lim_{|I|\to 0}{}_{x\in I} \frac{1}{|I|}\int_I f(y)dy = f(x)$$

holds for suitable points $x$. In higher dimensions we can pose a similar question. With this in mind, we restate our first problem in the context of $\mathbb{R}^d$ for all $d \geq 1$. So, suppose $f$ is integrable on $\mathbb{R}^d$. Is it true, then that

$$\lim_{m(B)\to 0}{}_{x\in B} \frac{1}{m(B)}\int_B f(y)dy = f(x), \qquad \text{for a.e. } x$$

The limit is taken as the volume of open balls $B$ containing $x$ goes to zero.

Definition: We shall refer to this problem as the *averaging problem*. We remark that if $B$ is any ball of radius $r$ in $\mathbb{R}^d$, then $m(B) = v_d r^d$ where $v_d$ is the measure of the unit ball.

Note of course that in the special case when $f$ is continuous at $x$, the limit does converge to $f(x)$. Indeed, given $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < eps$ whenever $|x - y| < \delta$. Since

$$f(x) - \frac{1}{m(B)}\int_B f(y)dy = \frac{1}{m(B)}\int_B (f(x) - f(y))dy,$$

we find that whenever $B$ is a ball of radius $< \frac{\delta}{2}$ containing $x$, then

$$\left| f(x) - \frac{1}{m(B)}\int_B f(y)dy \right| \leq \frac{1}{m(B)}\int_B |f(x) - f(y)|dy < \epsilon.$$

The averaging problem has an affirmative answer which we will now turn to.

*<back2top>*

### § 17.2 The Hardy-Littlewood maximal function

The maximal function that we consider below arose first in the one-dimensional situation treated by Hardy and Littlewood. The relevant definition is as follows.

Definition: If $f$ is integrable on $\mathbb{R}^d$, we define its *maximal function* $f^*$ by

$$f^*(x) = \sup_{x\in B} \frac{1}{m(B)}\int_B |f(y)|dy, \qquad x \in \mathbb{R}^d.$$

Here the supremum is taken over all balls containing the point $x$. In other words, we replace the limit in the averaging problem with a supremum, and $f$ by its absolute value. This leads us the key properties of $f^*$.

**Theorem 44** *Suppose $f$ is integrable on $\mathbb{R}^d$. Then:*

1. *$f^*$ is measurable.*

2. $f^*(x) < \infty$ *for almost every* $x$.
3. $f^*$ *satisfies*

$$(1) m(\{x \in \mathbb{R}^d : f^*(x) > \alpha\}) \leq \frac{A}{\alpha}||f||_{L^1(\mathbb{R}^d)}$$

*for all* $\alpha > 0$ *where* $A = 3^d$ *and* $||f||_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)|dx$.

*Proof.* The only simple assertion in the theorem is (1). Indeed, the set $E_\alpha = \{x \in \mathbb{R}^d : f^*(x) > \alpha\}$ is open because if $x' \in E_\alpha$, there exists a ball such that $x' \in B$ and

$$\frac{1}{m(B)} \int_B |f(y)|dy > \alpha.$$

Now any point $x$ close enough to $x'$ will also belong to $B$ hence $x \in E_\alpha$ as well. Properties (2) and (3) require a bit more work to prove with (2) being a consequence of (3). This follows once we observe that

$$\{x : f^*(x) = \infty\} \subset \{x : f^*(x) > \alpha\}, \qquad \text{for all } \alpha$$

Taking the limit as $\alpha \to \infty$, the third property yields

$$m(\{x : f^*(x) = \infty\}) = 0$$

as desired. For the proof of (3), let

$$E_\alpha = \{x : f^*(x) > \alpha\},$$

then for each $x \in E_\alpha$, there exists a ball $B_x$ containing $x$ such that

$$\frac{1}{m(B_x)} \int_{B_x} |f(y)|dy > \alpha.$$

Therefore, for each ball $B_x$ we have

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)|dy.$$

Fix a compact set $K \subset E_\alpha$. Note

$$K \subset \bigcup_{x \in E_\alpha} B_x$$

and $K$ is compact, there exists a finite subset $A \subset E_\alpha$ such that

$$K \subset \bigcup_{x \in A} B_x.$$

By the covering lemma, there exists a sub-collection of disjoint balls $B_{i_1}, ..., B_{i_k}$ with

$$m(\bigcup_{x \in A} B_x) \leq 3^d \sum_{j=1}^{k} m(B_{i_j}).$$

Since the balls are disjoint and satisfy (2) and (3), we find that

$$\begin{aligned}
m(K) &\leq m(\bigcup_{x\in A} B_x) \\
&\leq 3^d \sum_{j=1}^{k} m(B_{i_j}) \\
&\leq \frac{3^d}{\alpha} \sum_{j=1}^{k} \int_{B_{i_j}} |f(y)|dy \\
&= \frac{3^d}{\alpha} \int_{\bigcup_{j=1}^{k} B_{i_j}} |f(y)|dy \\
&\leq \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)|dy
\end{aligned}$$

□

This leads us to our next big Lemma:

**Lemma 45** *Suppose* $\mathcal{B} = \{B_1, B_2, ..., B_N\}$ *is a finite collection of balls in* $\mathbb{R}^d$*. Then there exists a disjoint sub-collection* $B_{i_1}, B_{i_2}, ..., B_{i_k}$ *of* $\mathcal{B}$ *that satisfies*

$$m(\bigcup_{l=1}^{N} B_l) \leq 3^d \sum_{j=1}^{k} m(B_{i_j})$$

*Proof.* The argument we give relies on the following observation: Suppose $B, B'$ are a pair of intersecting balls with radius of $B'$ being not greater than that of $B$. Then the ball $B'$ is contained in the ball $\tilde{B}$ that is concentric with $B$ but with 3 times its radius.

We first pick a ball $B_{i_1}$ in $\mathcal{B}$ with maximal radius then delete from $\mathcal{B}$ the ball $B_{i_1}$ and any ball intersecting it. Thus all deleted balls are contained in the ball $\tilde{B}_{i_1}$ concentric with $B_{i_1}$, but with 3 times its radius.

The remaining balls yield a new collection $\mathcal{B}'$, for which we repeat the procedure. We pick $B_{i_2}$ with maximal radius in $\mathcal{B}'$ then delete from $\mathcal{B}'$ any ball intersecting $B_{i_2}$. Continuing this way, we find, after at most $N$ steps, a collection of disjoint balls $B_{i_1}, B_{i_2}, ..., B_{i_k}$.

Finally, to show these disjoint balls satisfy the above inequality, we use the same argument as in the beginning. That is, let $\tilde{B}_{i_j}$ denote the ball concentric with $B_{i_j}$, but with 3 times its radius. Since any ball $B \in \mathcal{B}$ must intersect some ball $B_{i_j}$ for some $j$, and have less than or equal radius than $B_{i_j}$, thus we must have $B \subset \tilde{B}_{i_j}$, thus by monotonicity we have

$$\begin{aligned}
m(\bigcup_{l=1}^{\infty} B_l) &\leq m(\bigcup_{j=1}^{\infty} \tilde{B}_{i_j}) \\
&\leq \sum_{j=1}^{\infty} m(\tilde{B}_{i_j}) \\
&= 3^d \sum_{j=1}^{\infty} m(B_{i_j})
\end{aligned}$$

where in the last step we used the dilation fact about Lebesgue measure. □

<*back2top*>

### § 17.3 Lebesgue differentiation

The estimate obtained for the maximal function now leads to a solution for the averaging problem.

**Theorem 46** *If $f$ is integrable on $\mathbb{R}^d$, then*

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{1}{m(B)}\int_B f(y)dy = f(x), \text{for a.e. } x, (*).$$

*Proof.* It suffices to show for each $\alpha > 0$ that the set

$$E_\alpha = \left\{x : \lim_{m(B)\to 0}\sup_{x\in B}\left|\frac{1}{m(B)}\int_B f(y)dy - f(x)\right| > 2\alpha\right\}$$

has measure zero. Because this then guarantees that the set $E = \bigcup_{n=1}^{\infty} E_{\frac{1}{n}}$ has measure zero and the limit in $(*)$ holds for all points of $E^c$. By theorem 29, for any $\epsilon > 0$, there exists a continuous function $g$ of compact support such that

$$||f - g||_{L^!(\mathbb{R}^d)} < \epsilon.$$

The continuity of $g$ implies that

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{1}{m(B)}\int_B g(y)dy = g(x), \qquad \text{for all } x.$$

Since we may rewrite $\frac{1}{m(B)}\int_B f(y)dy - f(x)$ as

$$\frac{1}{m(B)}\int_B (f(y) - g(y))dy + \frac{1}{m(B)}\int_B g(y)dy - g(x) + g(x) - f(x)$$

We find that

$$\lim_{m(B)\to 0}\sup_{x\in B}\left|\frac{1}{m(B)}\int_B f(y)dy - f(x)\right| \le (f-g)^*(x) + |g(x) - f(x)|$$

Here, $f^*$ denotes the maximal function. Consequently, if

$$F_\alpha = \{x : (f-g)^*(x) > \alpha\} \qquad \text{and} \qquad G_\alpha = \{x : |f(x) - g(x)| > \alpha\},$$

then

$$E \subset (F_\alpha \cup G_\alpha)$$

because if $u_1, u_2 > 0$, then $u_1 + u_2 > 2\alpha$ only if $u_i > \alpha$ for at least one of the $u_i$. On the other hand, Tchebychev's inequality yields

$$m(G_\alpha) \le \frac{1}{\alpha}||f - g||_{L^1(\mathbb{R}^d)},$$

and on the other hand the weak type estimate for the maximal function gives

$$m(F_\alpha) \leq \frac{A}{\alpha}||f-g||_{L^1(\mathbb{R}^d)}.$$

The function $g$ was selected so that $||f-g||_{L^1(\mathbb{R}^d)} < \epsilon$ hence we get

$$m(E_\alpha) \leq \frac{A}{\alpha} + \frac{1}{\alpha}\epsilon$$

and as $\epsilon$ was arbitrarily chosen, we are done. □

Note that an immediate consequence of the theorem yields the fact that $f^*(x) \geq |f(x)|$ for a.e. $x$.
Definition: We say a measurable function $f$ is *locally integrable* if for every ball $B$, the function $f(x)\chi_B(x)$ is integrable. We shall denote by $L^1_{\text{loc}}(\mathbb{R}^d)$ the space of locally integrable functions. For example, the functions $e^{|x|}$ and $|x|^{-\frac{1}{2}}$ are both locally integrable but not integrable on all of $\mathbb{R}^d$.

**Theorem 47** *If $f \in L_{\text{loc}^1(\mathbb{R}^d)}$, then*

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{1}{m(B)}\int_B f(y)dy = f(x), \text{for a.e. } x$$

Our first application of this theorem yields an interesting insight into the nature of measurable sets.
Definition: If $E$ is a measurable set and $x \in \mathbb{R}^d$, we say $x$ is a point of *Lebesgue density* of $E$ if

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{m(B\cap E)}{m(B)} = 1.$$

This says that the small balls around $x$ are almost entirely covered by $E$. More precisely, for $\alpha < 1$ close to 1, and every ball of sufficiently small radius containing $x$, we have

$$m(B\cap E) \geq \alpha m(B).$$

Thus $E$ covers at at least a proportion $\alpha$ of $B$. An application of the bounded convergence theorem to the characteristic function $\chi$ of some measurable set $E$ is the following corollary:

**Corollary 48** *Suppose $E$ is a measurable subset of $\mathbb{R}^d$. Then*

1. *Almost every $x \in E$ is a point of density of $E$*
2. *Almost every $x \notin E$ is not a point of density of $E$.*

We next consider a notion that for integrable functions serves as a useful substitute for point-wise continuity.
Definition: If $f \in L_{\text{loc}^1(\mathbb{R}^d)}$, the *Lebesgue set* of $f$ consists of all points $\overline{x} \in \mathbb{R}^d$ for which $f(\overline{x})$ is finite and

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{1}{m(B)}\int_B |f(y)-f(\overline{x})|dy = 0.$$

At this stage, two simple observations are to be made: First, $\overline{x}$ belongs to the Lebesgue set of $f$ whenever $f$ is continuous at $\overline{x}$. Secondly, if $\overline{x}$ is in the Lebesgue set of $f$, then

$$\lim_{m(B)\to 0}{}_{x\in B}\frac{1}{m(B)}\int_B f(y)dy = f(\overline{x}).$$

This leads us to our next corollary:

**Corollary 49** *If $f$ is locally integrable on $\mathbb{R}^d$, then almost every point belongs to the Lebesgue set of $f$.*

*Proof.* An application of the bounded convergence theorem to the function $|f(y) - r|$ shows that for each rational number $r$, there exists a set $E_r$ of measure zero such that

$$\lim_{m(B)\to 0} {}_{x\in B} \frac{1}{m(B)} \int_B |f(y) - r| dy = |f(x) - r|, \qquad \text{whenever } x \notin E_r$$

So then if $E = \bigcup_{r\in\mathbb{Q}} E_r$, then $m(E) = 0$. Now suppose that $\overline{x} \notin E$ and $f(\overline{x})$ is finite. Given $\epsilon > 0$, there exists a rational number $r$ such that

$$|f(\overline{x}) - r| < \epsilon.$$

Since

$$\frac{1}{m(B)} \int_B |f(y) - f(\overline{x})| dy \leq \frac{1}{m(B)} \int_B |f(y) - r| dy + |f(\overline{x}) - r|$$

we must have

$$\lim_{m(B)\to 0} {}_{x\in B} \frac{1}{m(B)} \int_B |f(y) - f(\overline{x})| dy \leq 2\epsilon$$

thus $\overline{x}$ is in the Lebesgue set of $f$ as desired. □

Definition: A collection of set $\{U_\alpha\}$ is said to *shrink regularly* to $\overline{x}$ (or has *bounded eccentricity* at $\overline{x}$) if there exists a constant $c > 0$such that for each $U_\alpha$, there exists a ball $B$ with

$$\overline{x} \in B, U_\alpha \subset B, \qquad \text{and} \qquad m(U_\alpha) \geq cm(B).$$

Thus $U_\alpha$ is contained in $B$ but its measure is comparable to that of $B$. For example, the set of all open cubes containing $\overline{x}$ shrink regularly to $\overline{x}$. However, in $\mathbb{R}^d$, with $d \geq 2$, the collection of open rectangles does not shrink regularly to $\overline{x}$.

**Corollary 50** *Suppose $f$ is locally integrable on $\mathbb{R}^d$. If $\{U_\alpha\}$ shrinks regularly to $\overline{x}$, then*

$$\lim_{m(U_\alpha)\to 0} {}_{x\in U_\alpha} \int_{U_\alpha} f(y) dy = f(\overline{x}),$$

*for every point $\overline{x}$ in the LEbesgue set of $f$.*

*Proof.* The corollary is proven once we observe that if $\overline{x} \in B$ with $U_\alpha \subset B$, and $m(U_\alpha) \geq cm(B)$, then

$$\frac{1}{m(B)} \int_{U_\alpha} |f(y) - f(\overline{x})| dy \leq \frac{1}{cm(B)} \int_B |f(y) - f(\overline{x})| dy.$$

□

<back2top>

## § 18 Good kernels and approximations

We shall now turn to averages of functions given in terms of convolutions, which can be written as

$$(f*K_\delta)(x) = \int_{\mathbb{R}^d} f(x-y)K_\delta(y)dy.$$

Here $f$ is a general integrable function which we keep fixed, while the $K_\delta$ vary over a specific family of functions, referred to as kernels. We call the $K_\delta$ "good kernels" if

1. $\int_{\mathbb{R}^d} K_\delta(x) = 1.$
2. $\int_{\mathbb{R}^d} |K_\delta(x)|dx \leq A.$
3. For every $\eta > 0$,
$$\int_{|x|\geq\eta} |K_\delta(x)|dx \to 0, \text{as } \delta \to 0$$

Here $A$ is a constant independent of $\delta$. The main use of these kernels is that whenever $f$ is bounded, then $(f*K_\delta)(x) \to f(x)$ as $\delta \to 0$. To obtain a similar conclusion, one also valid at all points of the Lebesgue set of $f$, we need to strengthen somewhat our assumptions on the kernels $K_\delta$. To reflect this sitaution we adopt a different terminology.

Definition: We refer to the resulting narrower class of kernels as *approximations to the identity*. The assumptions are again, that the $K_\delta$ are integrable and satisfy condition (1). But instead of (2) and (3), we assume

1. $|K_\delta(x)| \leq A\delta^{-d}, \text{for all } \delta > 0$
2. $|K_\delta(x)| \leq \frac{A\delta}{|x|^{d+1}}, \text{for all } \delta > 0 \text{ and } x \in \mathbb{R}^d$

The term "approximation to the identity" originates in the fact that the mapping $f \mapsto f*K_\delta$ converges to the identity mapping $f \mapsto f$ as $\delta \to 0$. As $\delta \to 0$, the family of kernels converges to the so-called unit mass at the origin.

Definition: The family of kernels converge to the *Dirac delta* function defined via

$$\mathcal{D}(x) = \begin{cases} \infty & ; \text{if } x = 0 \\ 0 & ; \text{if } x \neq 0 \end{cases}, \quad \text{and} \int \mathcal{D}(x)dx = 1.$$

Since each $K_\delta$ integrates to 1, we may say loosely that

$$K_\delta \to \mathcal{D}, \quad \text{as } \delta \to 0.$$

If we think of the convolution $f*\mathcal{D}$ as $\int f(x-y)\mathcal{D}(y)dy$, the product $f(x-y)\mathcal{D}(y) = 0$ execpt when $y = 0$ and the mass of $\mathcal{D}$ is concentrated at $y = 0$, so we may expect that

$$(f*\mathcal{D})(x) = f(x).$$

Thus $f*\mathcal{D} = f$ and $\mathcal{D}$ plays the role of the identity for convolutions. We now turn to a series of examples of approximations to the identity.

*Example 1* Suppose $\varphi$ is a non-negative bounded function in $\mathbb{R}^d$ that is supported on the unit ball $|x| \leq 1$ and such that

$$\int_{\mathbb{R}^d} \varphi = 1.$$

Then, if we set $K_\delta(x) = \delta^{-1}\varphi(\delta^{-1}x)$, then the family $\{K_\delta\}_{\delta>0}$ is an approximation to the identity.

*Example 2* The Poisson kernel for the upper half plane is given via

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}, \qquad x \in \mathbb{R},$$

where the parameter is now $\delta = y > 0$.

We now turn to a general result about approximations to the identity that highlights the role of the Lebesgue set.

**Theorem 51** *If $\{K_\delta\}_{\delta>0}$ is an approximation to the identity and if $f$ is integrable on $\mathbb{R}^d$, then*

$$(f{*}K_\delta)(x) \to f(x), \text{as } \delta \to 0$$

*for every $x$ in the Lebesgue set of $f$. In particular, the limit holds for a.e. $x$.*

Before proving the Theorem, we state a useful lemma:

**Lemma 52** *Suppose that $f$ is integrable over $\mathbb{R}^d$, and that $x$ is a point of the Lebesgue set of $f$. Let*

$$\mathcal{A}(r) = \frac{1}{r^d}\int_{|y|\le r} |f(x-y) - f(x)|dy, \text{whenever } r > 0.$$

*Then $\mathcal{A}(r)$ is a continuous function of $r > 0$ and*

$$\mathcal{A}(r) \to 0, \text{as } r \to 0.$$

*Moreover, $\mathcal{A}(r)$ is bounded, that is, $\mathcal{A}(r) \le M$ for some $M > 0$ and all $r > 0$.*

*Proof.* The continuity of $\mathcal{A}(r)$ follows by invoking the absolute continuity of an earlier theorem.

The fact that $\mathcal{A}(r) \to 0$ as $r \to 0$ follows since $x$ is in the Lebesgue set of $f$ and the measure of a ball of radius $r$ is $v_d r^d$. This together with the continuity of $\mathcal{A}(r)$ for $0 < r \le 1$ show $\mathcal{A}(r)$ is bounded when $r \in (0,1]$. To prove $\mathcal{A}(r)$ is bounded for $r > 1$, note that

$$\begin{aligned}\mathcal{A}(r) &\le \frac{1}{r^d}\int_{|y|\le r} |f(x-y)dy + \frac{1}{r^d}\int_{|y|\le r} |f(x)|dy \\ &\le r^{-d}||f||_{L^1(\mathbb{R}^d)} + v_d|f(x)|.\end{aligned}$$

and the Lemma is proven. We can begin the proof of Theorem 51 now. □

*Proof.* Since the integral of each kernel $K_\delta$ is equal to 1, we may write

$$(f{*}K_\delta)(x) - f(x) = \int [f(x-y) - f(x)]K_\delta(y)dy.$$

Consequently,

$$|(f{*}K_\delta)(x) - f(x)| \le \int |f(x-y) - f(x)|K_\delta(y)dy.$$

Thus, it suffices to prove the right hand side goes to 0 as $\delta \to 0$. The key is to write the intergal over $\mathbb{R}^d$ as a sum of integrals over annuli as follows:

$$\int |f(x-y) - f(x)||K_\delta(y)|dy = \int_{|y|\leq\delta} + \sum_{k=0}^{\infty} \in_{2^k\delta<|y|\leq 2^{k+1}\delta}$$

By using the second property (2)

$$\begin{aligned}\int_{|y|\leq\delta} |f(x-y) - f(x)||K_\delta(y)|dy &\leq \frac{c}{\delta^d}\int_{|y|\leq\delta} |f(x-y) - f(x)|dy \\ &\leq c\mathcal{A}(\delta)\end{aligned}$$

Next using the second property (3), we have

$$\begin{aligned}\int_{2^k\delta<|y|\leq 2^{k+1}\delta} |f(x-y) - f(x)|dy &\leq \frac{c\delta}{(2^k\delta)^{d+1}}\int_{|y|\leq 2^{k+1}\delta} |f(x-y) - f(x)|dy \\ &\leq \frac{c'}{2^k(2^{k+1}\delta)^d}\int_{|y\leq 2^{k+1}\delta} |f(x-y) - f(x)|dy \\ &\leq c'2^{-k}\mathcal{A}(2^{k+1}\delta)\end{aligned}$$

Putting these together we see that

$$|(f{*}K_\delta)(x) - f(x)| \leq c\mathcal{A}(\delta) + c'\sum_{k=0}^{\infty} 2^{-k}\mathcal{A}(2^{k+1}\delta).$$

Given $\epsilon > 0$, first we choose $N$ so large that

$$\sum_{k\geq N} 2^{-k} < \epsilon.$$

Then, by making $\delta$ sufficiently small, we have by our Lemma

$$\mathcal{A}(2^k\delta) < \frac{\epsilon}{N}, \qquad \text{whenever } k = 0, 1, 2, ..., N-1.$$

Hence recalling $\mathcal{A}$ is bounded, we find that

$$|(f{*}K_\delta)(x) - f(x)| \leq C\epsilon,$$

for all sufficiently small $\delta$ and the theorem is proved. □

In addition to this point wise result, convolutions with approximations to the identity also provide convergence in the $L^1$ norm.

**Theorem 53** *Suppose that $f$ is integrable on $\mathbb{R}^d$ and that $\{K_\delta\}_{\delta>0}$ is an approximation to the identity. Then for each $\delta > 0$, the convolution*

$$(f{*}K_\delta)(x) = \int_{\mathbb{R}^d} f(x-y)K_\delta(y)dy$$

*is integrable, and*

$$||(f{*}K_\delta) - f||_{L^1(\mathbb{R}^d)} \to 0, \text{as } \delta \to 0$$

<back2top>

### § 19 Differentiability of functions

We now take up the second question raised at the beginning of this section, that of finding a broad condition on functions $F$ that guarantees the identity

$$F(b) - F(a) = \int_a^b F'(x)dx.$$

There are two problems that arise from this formulation. First, because of the existence of non-differentiable functions the right hand side may not be meaningful. Second, even if the function $F'$ existed for every $x$, the function $F'$ would not necessarily be (Lebesgue) measurable. To fix this problem, we study functions of bounded variation.

### § 19.2 Functions of bounded variation

Let $\gamma$ be a parameterized curve in the plane given by $z(t) = (x(t), y(t))$, where $t \in [a, b]$. Here $x(t), y(t)$ are continuous real-valued functions on $[a, b]$.
Definition: The curve $\gamma$ is *rectifiable* if there exists an $M < \infty$ such that for any partition $a = t_0 < t_1 < ... < t_N = b$ or $[a, b]$,

$$\sum_{j=1}^{N} |z(t_j) - z(t_{j-1}) \leq M.$$

Definition: By definition, the *length* $L(t)$ of the curve is the supremum over all partitions of the sum on the left hand side, that is,

$$L(\gamma) = \sup_{a=t_0<t_1<...<t_N=b} \sum_{j=1}^{N} |z(t_j) - z(t_{j-1})|$$

If the derivatives of $x(t), y(t)$ exists, then we ask if one has the desired formula

$$L(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}dt.$$

Suppose $F$ is a complex-valued function defined on $[a, b]$ and $a = t_0 < t_1 < ... < t_N = b$ is a partition of the interval. The variation of this partition is defined by

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|$$

Definition: The function $F$ is said to be of *bounded variation* if the variations of $F$ over all partitions are bounded, that is, there exists an $M < \infty$ such that

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \leq M$$

for all partitions of $[a, b]$. In this definition, we do not assume $F$ to be continuous; however when applying this to the case of curves, we will assume $F(t) = z(t) = x(t) + iy(t)$ is continuous. Observe that if $\mathcal{P}'$ is a partition that is a refinment of a partition given by $\mathcal{P}$, then the variation of $F$ on $\mathcal{P}'$ is greater than or equal to the variation of $F$ on $\mathcal{P}$.

**Theorem 54** *A curve parameterized by $(x(t), y(t))$, $t \in [a, b]$, is rectifiable if and only if both $x(t)$ and $y(t)$ are of bounded variation.*

*Proof.* The proof is immediate once we realize that if $F(t) = x(t) + iy(t)$, then

$$F(t_j) - F(t_{j-1}) = (x(t_j) - x(t_{j-1})) + i(y(t_j) - y(t_{j-1}))$$

and if $a, b \in \mathbb{R}$, then

$$|a + bi| \le |a| + |b| \le 2|a + bi|.$$

□

Definition: A real-valued function $F$ defined on $[a, b]$ is said to be *increasing* if $F(t_1) \le F(t_2)$ for whenever $a \le t_1 \le t_2 \le b$. If the inequalities are strict, $F$ is said to be *strictly increasing.*

*Example 1* If $F$ is real-valued, monotonic, and bounded, then $F$ is of bounded variation. Indeed, if for example $F$ is increasing and bounded by $M$, we see that

$$\begin{aligned} \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| &= \sum_{j=1}^{N} F(t_j) - F(t_{j-1}) \\ &= F(b) - F(a) \\ &\le 2M. \end{aligned}$$

*Example 2* If $F$ is differentiable at every point, and $F'$ is bounded, then $F$ is of bounded variation. Indeed, $|F'| \le M$, and the Mean value theorem implies

$$|F(b) - f(a)| \le M|x - y|, \qquad \text{for all } x, y \in [a, b]$$

hence $\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \le M(b - a)$.

Definition: The *total variation* of $f$ on $[a, x]$ is defined by

$$T_F(a, x) = \sup \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|,$$

where the sup is taken over all partitions of $[a, b]$.

*<back2top>*