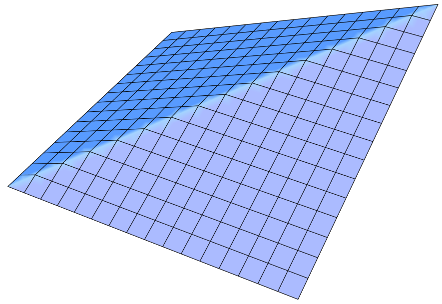


PMATH 450: Lebesgue Integration and Fourier Analysis

Prof. Mahya Ghandehari • Winter 2015 • University of Waterloo
TR 12:30–2:20PM in PHY 150

Transcribed, LaTeXed, and edited by Rob Zimmerman

Note: These are personal lecture notes, not official course materials. They may contain errors (which by default you should attribute to me, Rob Zimmerman, rather than the course instructor). The permanent home of these — and all of my other lecture notes — is `https://rob-zimmerman.github.io/coursenotes`.

Contents

1 Measure Theory Foundations **3**
Riemann Integration and the Need for Measure 3
Lebesgue Outer Measure 7
Lebesgue Measurable Sets 10
Sigma-Algebras and Borel Sets 11
Properties of Lebesgue Measure 16
Riemann Integrability and Discontinuities 22

2 Measurable Functions and Integration **25**
Measurable Functions and Closure Properties 26
Simple Functions and Approximation 29
Lebesgue Integrals of Nonnegative Functions 30
Monotone Convergence Theorem and Fatou's Lemma 34
Integrable Functions and Almost Everywhere Equivalence 38
Dominated Convergence Theorem 41

3 L^p Spaces and Hilbert Spaces **44**
L^p and L^∞ Spaces 44
Hölder and Minkowski Inequalities 47
Inner Products and Hilbert Spaces 56
Orthogonal Series and Bessel's Inequality 62
Orthonormal Bases and Parseval Expansions 65
Existence of Orthonormal Bases 69

4 Fourier Analysis on the Circle **72**
The Circle Group and Trigonometric Polynomials 72
Fourier Coefficients on $L^1(\mathbb{T})$ 77
Dirichlet Kernels and Partial Sums 83
Divergence and the Uniform Boundedness Principle 87
Convolutions 93
Summability Kernels and the Fejér Kernel 97
Pointwise Convergence of Fourier Series 105

5 Fourier Analysis on $\mathbb{R}$ and Applications **113**
Fourier Transform on $\mathbb{R}$ 113
Basic Properties and the Riemann–Lebesgue Lemma 115
Summability Kernels and Fourier Inversion 118
Plancherel and the Inverse Transform 120
The Heat Equation 125
Translation-Invariant Subspaces of $L^2(\mathbb{R})$ 127

Appendix: Lecture and Tutorial Index **129**

Lecture 1 — Tuesday, January 06, 2015

1. Measure Theory Foundations

Riemann Integration and the Need for Measure

To begin with, we briefly review Riemann integration. Let $f : [a, b] \to \mathbb{R}$ be given. Partition $[a, b]$ by

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$

Define

$$U(f, P) = \sum_{i=1}^{n} \sup_{[x_{i-1}, x_i]} f \cdot (x_i - x_{i-1})$$

and

$$L(f, P) = \sum_{i=1}^{n} \inf_{[x_{i-1}, x_i]} f \cdot (x_i - x_{i-1}),$$

the **upper and lower Riemann sums**.

Fact 1.1 For any two partitions P and Q, $U(f, P) \geq L(f, Q)$.

Proof. Let $P : a = x_0 < x_1 < \cdots < x_n = b$ and $Q : a = y_0 < y_1 < \cdots < y_m = b$ be two partitions of $[a, b]$. Let R be the common refinement of P and Q, which is the partition obtained by taking the union of the points in P and Q. Then $U(f, P) \geq U(f, R)$ and $L(f, Q) \leq L(f, R)$, since the upper sum decreases and the lower sum increases when we refine the partition. Therefore, $U(f, P) \geq U(f, R) \geq L(f, R) \geq L(f, Q)$. $\square$

Thus $\inf_P U(f, P) \geq \sup_Q L(f, Q)$. We say that f is **Riemann integrable** if equality holds.

Theorem 1.2 If f is continuous on $[a, b]$, then f is Riemann integrable.

Figure 1 shows the geometric comparison behind the upper and lower sums: refinement makes the two stepwise estimates close in on the same area.

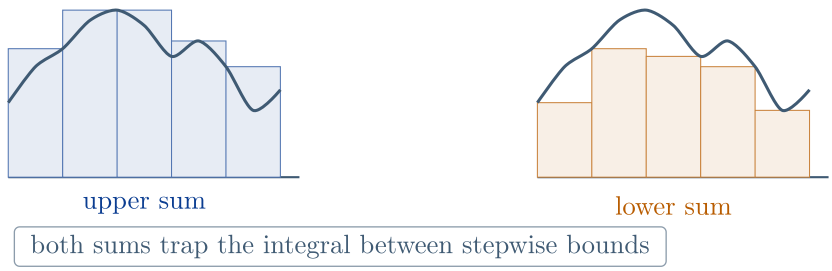


FIGURE 1

Notation 1.3 Let $E \subseteq \mathbb{R}$. Then $\mathbb{1}_E$ denotes the **characteristic function** of E. That is,

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Example 1.4 $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ is not Riemann integrable.

Proof. Let $P : 0 = x_0 < x_1 < x_2 < \cdots < x_n = 1$ be a partition of $[0,1]$. Since $\mathbb{Q}$ and its complement are both dense in $\mathbb{R}$, every nondegenerate subinterval contains rational and irrational points. Therefore $\sup_{[x_{i-1},x_i]} \mathbb{1}_{\mathbb{Q}} = 1$ and $\inf_{[x_{i-1},x_i]} \mathbb{1}_{\mathbb{Q}} = 0$ for every i. Hence,

$$U(\mathbb{1}_{\mathbb{Q}}, P) = 1 \neq 0 = L(\mathbb{1}_{\mathbb{Q}}, P).$$

□

Let $\{x_1, x_2, \dots\} = \mathbb{Q} \cap [0,1]$. For any $n \in \mathbb{N}$, define

$$f_n(x) = \begin{cases} 1 & \text{if } x = x_j \text{ and } 1 \le j \le n \\ 0 & \text{otherwise} \end{cases}.$$

Then $f_n(x) \to \mathbb{1}_{\mathbb{Q}\cap[0,1]}(x)$ for every $x \in [0,1]$, and every f_n is Riemann integrable with

$$\lim_{n\to\infty} \int f_n(x)\, \mathrm{d}x = 0.$$

The pointwise limit is not Riemann integrable, so the Riemann theory cannot even formulate an interchange of limit and integral here. This is precisely the kind of failure that Lebesgue integration is designed to repair; later we will see that the limit is Lebesgue integrable with integral 0.

If $E_1, \dots, E_n$ are disjoint intervals of $\mathbb{R}$ and

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{E_i}$$

for $a_i \in \mathbb{R}$, then

$$\int f(x)\,\mathrm{d}x = \sum_{i=1}^{n} a_i \cdot (\text{length of } E_i).$$

For a function $f : [a, b] \to \mathbb{R}$, partition an interval containing $f([a, b])$ as

$$A = y_0 < y_1 < \cdots < y_n = B.$$

For $i = 0, \dots, n-1$, let $E_i = f^{-1}([y_i, y_{i+1}))$; then

$$f \approx \sum_{i=0}^{n-1} y_i \mathbb{1}_{E_i}.$$

In other words, $f(x)$ is close to

$$\sum_{i=0}^{n-1} y_i \mathbb{1}_{E_i}(x).$$

When the Lebesgue integral is defined, we want

$$\int_a^b \left(\sum_{i=0}^{n-1} y_i \mathbb{1}_{E_i}(x) \right) \mathrm{d}x = \sum_{i=0}^{n-1} y_i (\text{length of } E_i).$$

This suggests the approximation

$$\int_a^b f(x)\,\mathrm{d}x \approx \int \sum_{i=0}^{n-1} y_i \mathbb{1}_{E_i}(x)\,\mathrm{d}x.$$

However, $f^{-1}([y_i, y_{i+1}))$ is not necessarily an interval; we need a better concept of "length" for a general arbitrary set. This motivates the concept of a measure.

We ideally want to have a function $m : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ which satisfies the following properties:

1. $m(I) = \text{length}(I)$ for any interval I.

2. If $E_1, E_2, \ldots$ are disjoint subsets of $\mathbb{R}$, then
$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(E_i)$$
(**countable additivity**).

3. Let $E \subseteq \mathbb{R}$ and $y \in \mathbb{R}$. Then the translation of E with y is $E + y = \{e + y \mid e \in E\}$. So we want $m(E + y) = m(E)$ (**translation invariance**).

4. If $A \subseteq B$, then $m(A) \leq m(B)$ (**monotonicity**).

However, there are storm clouds on the horizon.

Proposition 1.5 There is no function $m : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ satisfying Properties (1)–(4) above.

Proof. We will construct a countable collection of sets which cannot satisfy (1)–(4) for any function m. On $[0, 1)$, define $x \sim y$ if $x - y \in \mathbb{Q}$. This is an equivalence relation.

Using the Axiom of Choice, choose a set $N \subseteq [0, 1)$ containing exactly one representative from each equivalence class. For $r \in \mathbb{Q} \cap [0, 1)$, let
$$N_r = \{r + n \mid n \in N \cap [0, 1 - r)\} \cup \{-1 + r + n \mid n \in N \cap [1 - r, 1)\}.$$

If $m : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ satisfies (1)–(4), then translation invariance and finite additivity give $m(N_r) = m(N)$. Moreover, the sets N_r are pairwise disjoint and
$$\bigcup_{r \in \mathbb{Q} \cap [0,1)} N_r = [0, 1).$$

Therefore,
$$m([0, 1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} m(N).$$

The left-hand side is 1, while the right-hand side is 0 if $m(N) = 0$ and ∞ if $m(N) > 0$. Either way, we have a contradiction. $\square$

Lecture 2 — Thursday, January 08, 2015

Lebesgue Outer Measure

We want to define the Lebesgue measure. Before that, though, we define the Lebesgue outer measure.

Notation 2.1

1. If $a < b$, $a, b \in \mathbb{R}$, then
$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$$
is an **open interval**. With $a = b$, the empty set $\emptyset$ is also considered an open interval.
2. $\mathcal{P}(\mathbb{R})$ is the collection of all subsets of $\mathbb{R}$ and is called the **power set** of $\mathbb{R}$.
3. Let E be a subset of $\mathbb{R}$. The sequence of open intervals $\{I_n\}_{n=1}^{\infty}$ is called an **open cover for E** if $E \subseteq \bigcup_{i=1}^{\infty} I_i$.

Definition 2.2 The **Lebesgue outer measure** is the function $\lambda^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ defined as follows: for every $E \subseteq \mathbb{R}$,

$$\lambda^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \ell(I_i) \mid \{I_i\}_{i=1}^{\infty} \text{ is an open cover of } E \right\}$$

As Figure 2 suggests, the infimum searches through every countable open cover, not merely the first convenient one we happen to draw.

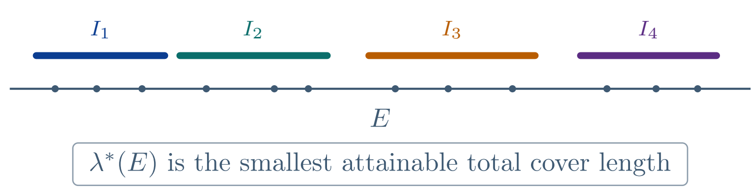


Figure 2

Proposition 2.3 The following hold:

1. $\lambda^*(E) \geq 0$ for every $E \subseteq \mathbb{R}$ (**nonnegativity**).
2. $\lambda^*(\emptyset) = 0$ and $\lambda^*(\{x\}) = 0$.
3. If $A, B \subseteq \mathbb{R}$ and $A \subseteq B$, then $\lambda^*(A) \leq \lambda^*(B)$ (**monotonicity**).
4. $\lambda^*(E) = \lambda^*(E + x)$ for all $E \subseteq \mathbb{R}$ and $x \in \mathbb{R}$ (**translation invariance**).
5. If $A_1, A_2, \dots \subseteq \mathbb{R}$, then
$$\lambda^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \lambda^*(A_i)$$
(**σ-subadditivity**).
6. $\lambda^*(\text{interval}) = \ell(\text{interval})$ for any interval in $\mathbb{R}$.

Proof. (1) is obvious.

For (2), observe that $\{x\} \subseteq (x - \varepsilon, x + \varepsilon)$. Then $\lambda^*(\{x\}) \leq 2\varepsilon$ by definition of λ^*. This holds for every $\varepsilon > 0$; therefore, $\lambda^*(\{x\}) = 0$. The same argument applies for $\emptyset$.

For (3), first let $\mathcal{C}(A)$ denote the collection of all open covers of A; that is,

$$\mathcal{C}(A) = \{\{I_i\}_{i=1}^{\infty} \mid \{I_i\}_{i=1}^{\infty} \text{ is an open cover for } A\}$$

If $A \subseteq B$, then $\mathcal{C}(B) \subseteq \mathcal{C}(A)$. Therefore,

$$\inf\left\{\sum_{i=1}^{\infty} \ell(I_i) \mid \{I_i\} \in \mathcal{C}(A)\right\} \leq \inf\left\{\sum_{i=1}^{\infty} \ell(I_i) \mid \{I_i\} \in \mathcal{C}(B)\right\},$$

which is the same as $\lambda^*(A) \leq \lambda^*(B)$.

For (4), let $x \in \mathbb{R}$. If $\{I_i\}_{i=1}^{\infty}$ is an open cover of E, then $\{I_i + x\}_{i=1}^{\infty}$ is an open cover of $E + x$, and $\ell(I_i + x) = \ell(I_i)$. Hence

$$\lambda^*(E + x) \leq \sum_{i=1}^{\infty} \ell(I_i + x) = \sum_{i=1}^{\infty} \ell(I_i).$$

Taking the infimum over all open covers of E gives $\lambda^*(E + x) \leq \lambda^*(E)$. Replacing x by $-x$ yields

$\lambda^*(E) \leq \lambda^*(E+x)$, so $\lambda^*(E+x) = \lambda^*(E)$.

For (5), if

$$\sum_{i=1}^{\infty} \lambda^*(A_i) = \infty,$$

then we are done, so assume otherwise. Pick $\varepsilon > 0$. For $n \in \mathbb{N}$, let $\{I_{i,n}\}_{i=1}^{\infty}$ be an open cover for A_n such that

$$\lambda^*(A_n) + \frac{\varepsilon}{2^n} > \sum_{i=1}^{\infty} \ell(I_{i,n}).$$

Observe that the collection $\{I_{i,n}\}_{i,n=1}^{\infty}$ is an open cover for $\bigcup_{i=1}^{\infty} A_i$. Then

$$\lambda^*(\bigcup A_i) \leq \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \ell(I_{i,n}) = \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \ell(I_{i,n}) \leq \sum_{n=1}^{\infty}[\lambda^*(A_n) + \varepsilon/2^n] = \sum_{n=1}^{\infty} \lambda^*(A_n) + \varepsilon,$$

which is true for all $\varepsilon > 0$. So

$$\lambda^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \lambda^*(A_i).$$

For (6), we want to prove $\lambda^*(\text{interval}) = \ell(\text{interval})$. We will prove it for closed intervals $[a,b]$ where $a, b \in \mathbb{R}$, which amounts to showing $\lambda^*([a,b]) = b-a$. For the upper bound, $[a,b] \subseteq (a-\varepsilon/2, b+\varepsilon/2)$ gives $\lambda^*([a,b]) \leq b-a+\varepsilon$ for every $\varepsilon > 0$, and hence $\lambda^*([a,b]) \leq b-a$. For the reverse inequality, we want to show that if $\{I_n\}_{n=1}^{\infty}$ is an open cover for $[a,b]$, then

$$\sum_{n=1}^{\infty} \ell(I_n) \geq b-a.$$

Now, $[a,b]$ is compact, so that it admits a finite subcover of $\{I_n\}_{n=1}^{\infty}$, say $I_1, \ldots, I_N$. We know

$$\sum_{n=1}^{\infty} \ell(I_n) \geq \ell(I_1) + \cdots + \ell(I_N).$$

So assume our open cover is finite, say $\{I_n\}_{n=1}^{N}$. If $a \in \bigcup_{n=1}^{N} I_n$, then there exists n_1 such that $a \in I_{n_1} = (a_{n_1}, b_{n_1})$. If $b \in I_{n_1}$, then

$$\sum_{n=1}^{N} \ell(I_n) \geq \ell(I_{n_1}) \geq b-a,$$

so we are done. Otherwise, $b \notin I_{n_1}$, so $b_{n_1} \in [a,b]$. Choose I_{n_2} with $b_{n_1} \in I_{n_2} = (a_{n_2}, b_{n_2})$. Since $b_{n_1} \notin I_{n_1}$, the interval I_{n_2} is distinct from I_{n_1}. Continue in this way: at step j, choose I_{n_j}

containing $b_{n_{j-1}}$. Since the cover is finite and the right endpoints strictly move to the right until b is covered, this process terminates in a finite number of steps. Thus,

$$\sum_{n=1}^{N} \ell(I_n) \geq \sum_{j=1}^{k} \ell(I_{n_j}) = (b_{n_1} - a_{n_1}) + \cdots + (b_{n_k} - a_{n_k}) \geq b_{n_k} - a_{n_1} \geq b - a.$$

Hence, $\lambda^*([a, b]) = b - a$. □

Example 2.4 $\lambda^*(\mathbb{Q}) = 0$. In general, if A is countable, then $\lambda^*(A) = 0$ since

$$A = \{x_1, x_2, \dots\} = \bigcup_{n=1}^{\infty} \{x_n\}$$

and

$$\lambda^*(A) \leq \sum_{n=1}^{\infty} \lambda^*(\{x_n\}) = \sum_{n=1}^{\infty} 0 = 0$$

by σ-subadditivity.

Lebesgue Measurable Sets

Definition 2.5 We say that $A \subseteq \mathbb{R}$ is **Lebesgue measurable** if, for every $E \subseteq \mathbb{R}$,

$$\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \setminus A).$$

Remark 2.6 Because $E = (E \cap A) \cup (E \setminus A)$, we always have

$$\lambda^*(E) \leq \lambda^*(E \cap A) + \lambda^*(E \setminus A)$$

by σ-subadditivity. It is the reverse inequality that is the nontrivial part of the definition of measurability.

Figure 3 depicts Carathéodory's criterion: a measurable set A cuts every test set E into two pieces without creating any extra outer measure.

Notation 2.7 Let $\mathcal{L}(\mathbb{R})$ denote the collection of all Lebesgue measurable subsets of $\mathbb{R}$.

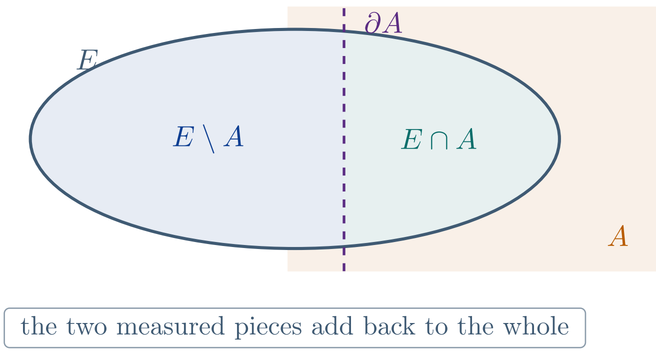


Figure 3

Example 2.8 Here are some examples of Lebesgue measurable sets:

1. $\emptyset \in \mathcal{L}(\mathbb{R})$.
2. $\{x\} \in \mathcal{L}(\mathbb{R})$ for $x \in \mathbb{R}$. Why? Let $E \subseteq \mathbb{R}$. Obviously, $\lambda^*(E \cap \{x\}) \leq \lambda^*(\{x\}) = 0$ because $E \cap \{x\} \subseteq \{x\}$. Also,

$$\lambda^*(E) \leq \lambda^*(E \setminus \{x\}) + \lambda^*(\{x\}) = \lambda^*(E \setminus \{x\}).$$

Hence, $\lambda^*(E) = \lambda^*(E \cap \{x\}) + \lambda^*(E \setminus \{x\})$. In general, if A is such that $\lambda^*(A) = 0$, then A is measurable.

Lecture 3 — Tuesday, January 13, 2015

Sigma-Algebras and Borel Sets

Recall that the outer measure $\lambda^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ is translation invariant, satisfies $\lambda^*(\text{interval}) = \ell(\text{interval})$, is σ-subadditive, and is monotone. We defined a set $A \subseteq \mathbb{R}$ to be measurable when $\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \setminus A)$ for every $E \subseteq \mathbb{R}$.

Theorem 3.1

1. $\emptyset, \mathbb{R} \in \mathcal{L}(\mathbb{R})$.

2. If $A \in \mathcal{L}(\mathbb{R})$, then $A^c \in \mathcal{L}(\mathbb{R})$.
3. If $A_1, A_2, \dots \in \mathcal{L}(\mathbb{R})$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{L}(\mathbb{R})$.

Proof. For (1), let $E \subseteq \mathbb{R}$. Then

$$\lambda^*(E) = \lambda^*(E) + \lambda^*(\emptyset) = \lambda^*(E \cap \emptyset) + \lambda^*(E \setminus \emptyset),$$

so $\emptyset \in \mathcal{L}(\mathbb{R})$. Similarly,

$$\lambda^*(E) = \lambda^*(E \cap \mathbb{R}) + \lambda^*(E \setminus \mathbb{R}),$$

so $\mathbb{R} \in \mathcal{L}(\mathbb{R})$.

For (2), let $E \subseteq \mathbb{R}$. Because $A \in \mathcal{L}(\mathbb{R})$, we have

$$\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \setminus A) = \lambda^*(E \cap A) + \lambda^*(E \cap A^c),$$

so $A^c \in \mathcal{L}(\mathbb{R})$.

For (3), we first check that $\mathcal{L}(\mathbb{R})$ is closed under finite unions. If $A, B \in \mathcal{L}(\mathbb{R})$ and $E \subseteq \mathbb{R}$, then

$$\begin{aligned}\lambda^*(E) &= \lambda^*(E \cap A) + \lambda^*(E \setminus A)\\ &= \lambda^*(E \cap A) + \lambda^*((E \setminus A) \cap B) + \lambda^*(E \setminus (A \cup B))\\ &\geq \lambda^*(E \cap (A \cup B)) + \lambda^*(E \setminus (A \cup B)).\end{aligned}$$

The reverse inequality follows from subadditivity, so $A \cup B$ is measurable. Induction now gives closure under finite unions.

Let $A = \bigcup_{n=1}^{\infty} A_n$ and disjointify the sequence by setting

$$B_1 = A_1 \quad \text{and} \quad B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k, \qquad n \geq 2.$$

Each B_n is measurable, the sets B_n are pairwise disjoint, and $\bigcup_{n=1}^{\infty} B_n = A$. Repeatedly applying measurability gives, for every N and every $E \subseteq \mathbb{R}$,

$$\lambda^*(E) = \sum_{n=1}^{N} \lambda^*(E \cap B_n) + \lambda^*\left(E \setminus \bigcup_{n=1}^{N} B_n\right).$$

Since $E \setminus A \subseteq E \setminus \bigcup_{n=1}^{N} B_n$, we may let $N \to \infty$ to obtain

$$\lambda^*(E) \geq \sum_{n=1}^{\infty} \lambda^*(E \cap B_n) + \lambda^*(E \setminus A) \geq \lambda^*(E \cap A) + \lambda^*(E \setminus A).$$

Subadditivity gives the reverse inequality. Hence A is measurable. $\square$

Definition 3.2 Let X be any set. Let μ be a collection of some subsets of X. We say μ is an **algebra** on X if the following hold:

1. $\emptyset, X \in \mu$.
2. If $A \in \mu$, then $A^c \in \mu$.
3. If $A_1, \ldots, A_n \in \mu$, then $\bigcup_{i=1}^{n} A_i \in \mu$.

The collection μ is called a ***σ*-algebra on *X*** if it is an algebra closed under countable unions; that is, if $A_1, A_2, \dots \in \mu$, then $\bigcup_{n=1}^{\infty} A_n \in \mu$.

Example 3.3 Examples of σ-algebras on $\mathbb{R}$ include the following:

1. $\{\emptyset, \mathbb{R}\}$.
2. $\mathcal{L}(\mathbb{R})$.
3. The **Borel *σ*-algebra**, which is the smallest σ-algebra on $\mathbb{R}$ containing all open sets, and denoted by
$$\mathcal{B}(\mathbb{R}) := \bigcap \{\mathcal{M} \mid \mathcal{M} \text{ is a } \sigma\text{-algebra and } \mathcal{M} \text{ contains all open sets}\}.$$
4. $\bigcap_{\alpha \in A} \mathcal{M}_\alpha$, where $\{\mathcal{M}_\alpha\}_{\alpha \in A}$ is a family of σ-algebras on $\mathbb{R}$.

Definition 3.4 Every subset of $\mathbb{R}$ belonging to $\mathcal{B}(\mathbb{R})$ is called a **Borel set**.

Example 3.5 Examples of Borel sets include the following:

1. Open sets, closed sets (hence compact sets), countable sets, and $(a, b] = (a, b) \cup \{b\}$ (hence all intervals).
2. Let $\mathcal{G}$ denote the family of open sets, and let $\mathcal{F}$ denote the family of closed sets. Let
$$\mathcal{G}_\delta := \{\text{all countable intersections of open sets}\}$$
and
$$\mathcal{F}_\sigma := \{\text{all countable unions of closed sets}\}.$$
Then $\mathcal{G}_\delta, \mathcal{F}_\sigma \subseteq \mathcal{B}(\mathbb{R})$.

We now look into the relationship between $\mathcal{L}(\mathbb{R})$ and $\mathcal{B}(\mathbb{R})$.

Proposition 3.6 The following hold:

1. Every open interval (a, b) with $a < b$ is Lebesgue measurable.
2. Every interval $(a, +\infty)$ or $(-\infty, b)$ is Lebesgue measurable.
3. Every open set is Lebesgue measurable.

Proof. For (1), let $E \subseteq \mathbb{R}$ and $\varepsilon > 0$. By definition of outer measure, there exist open intervals $\{I_n\}_{n=1}^\infty$ such that
$$E \subseteq \bigcup_{n=1}^\infty I_n \qquad \text{and} \qquad \sum_{n=1}^\infty \ell(I_n) < \lambda^*(E) + \varepsilon.$$
Fix n, and define
$$I_n^- := I_n \cap (-\infty, a], \qquad I_n^0 := I_n \cap (a, b), \qquad \text{and} \qquad I_n^+ := I_n \cap [b, \infty).$$
Each of these sets is an interval (possibly empty), and
$$\ell(I_n^-) + \ell(I_n^0) + \ell(I_n^+) = \ell(I_n).$$
The family $\{I_n^0\}_{n=1}^\infty$ covers $E \cap (a, b)$, and the family $\{I_n^-, I_n^+\}_{n=1}^\infty$ covers $E \setminus (a, b)$. Therefore,
$$\lambda^*(E \cap (a, b)) \leq \sum_{n=1}^\infty \ell(I_n^0)$$

and

$$\lambda^*(E \setminus (a,b)) \leq \sum_{n=1}^{\infty} \left(\ell(I_n^-) + \ell(I_n^+)\right).$$

Adding these inequalities gives

$$\lambda^*(E \cap (a,b)) + \lambda^*(E \setminus (a,b)) \leq \sum_{n=1}^{\infty} \ell(I_n) < \lambda^*(E) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\lambda^*(E \cap (a,b)) + \lambda^*(E \setminus (a,b)) \leq \lambda^*(E).$$

The reverse inequality always holds by subadditivity because $E = (E \cap (a,b)) \cup (E \setminus (a,b))$. Hence

$$\lambda^*(E) = \lambda^*(E \cap (a,b)) + \lambda^*(E \setminus (a,b)),$$

so (a,b) is Lebesgue measurable.

For (2),

$$(a, +\infty) = \bigcup_{n=1}^{\infty} (a, a+n) \in \mathcal{L}(\mathbb{R}).$$

For (3), let O be an open set in $\mathbb{R}$. Then $O = \bigcup_{n=1}^{\infty} (a_n, b_n)$, where $a_n, b_n \in \mathbb{R} \cup \{-\infty, \infty\}$. Every (a_n, b_n) is Lebesgue measurable by (1) and (2), and hence their countable union is as well. □

By definition of Borel σ-algebra, $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{L}(\mathbb{R})$. We can summarize everything so far:

Theorem 3.7 The set of Lebesgue measurable sets $\mathcal{L}(\mathbb{R})$ is a σ-algebra which contains the Borel sets $\mathcal{B}(\mathbb{R})$ and all $E \subseteq \mathbb{R}$ such that $\lambda^*(E) = 0$. The restriction $\lambda^*|_{\mathcal{L}(\mathbb{R})} =: \lambda$ is called the **Lebesgue measure** and satisfies the following properties:

1. $\lambda(E) = \lambda(E + x)$ for all $E \in \mathcal{L}(\mathbb{R})$ and $x \in \mathbb{R}$ (**translation invariance**).
2. If $A, B \in \mathcal{L}(\mathbb{R})$ and $A \subseteq B$, then $\lambda(A) \leq \lambda(B)$ (**monotonicity**).
3. $\lambda(I) = \ell(I)$ for any interval I
4. If $A_1, A_2, \dots \in \mathcal{L}(\mathbb{R})$ and $A_i \cap A_j = \emptyset$ for $i \neq j$, then

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

(**countable additivity**).

Proof. All statements except (4) have already been established: $\mathcal{L}(\mathbb{R})$ is a σ-algebra, $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{L}(\mathbb{R})$, null sets are measurable, and $\lambda = \lambda^*|_{\mathcal{L}(\mathbb{R})}$ inherits (1), (2), and (3) from λ^*. So it remains to prove (4). Let

$$A = \bigcup_{n=1}^{\infty} A_n.$$

By countable subadditivity of outer measure,

$$\lambda(A) = \lambda^*(A) \leq \sum_{n=1}^{\infty} \lambda^*(A_n) = \sum_{n=1}^{\infty} \lambda(A_n).$$

For the reverse inequality, first note that if $B, C \in \mathcal{L}(\mathbb{R})$ are disjoint, then measurability of B gives

$$\lambda(B \cup C) = \lambda^*((B \cup C) \cap B) + \lambda^*((B \cup C) \setminus B) = \lambda(B) + \lambda(C).$$

By induction, finite additivity holds for every finite family of pairwise disjoint Lebesgue measurable sets. Therefore, for every $N \geq 1$,

$$\lambda(A) \geq \lambda\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \lambda(A_n).$$

Letting $N \to \infty$ gives

$$\lambda(A) \geq \sum_{n=1}^{\infty} \lambda(A_n)$$

, and hence equality holds. □

The regularity results below will show that every Lebesgue measurable set differs from a Borel set by a null set. The inclusion is strict: the Cantor set is a null Borel set with continuum cardinality, so all of its subsets are Lebesgue measurable, whereas the Borel σ-algebra itself has only continuum cardinality. Thus some subsets of the Cantor set are Lebesgue measurable but not Borel.

Properties of Lebesgue Measure

We now collect several useful properties of the Lebesgue measure λ.

Proposition 3.8 The following hold:

1. If $A \subseteq B$, $A, B \in \mathcal{L}(\mathbb{R})$, and $\lambda(A) < \infty$, then $\lambda(B \setminus A) = \lambda(B) - \lambda(A)$.
2. If $A \subseteq \mathbb{R}$ is compact, then $\lambda(A) < \infty$.
3. λ is **$\boldsymbol{\sigma}$-subadditive**. That is, if $A_1, A_2, \dots \in \mathcal{L}(\mathbb{R})$, then
$$\lambda\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \lambda(A_i).$$
Note that $A_1, A_2, \dots$ do not need to be pairwise disjoint!
4. **Continuity of measure**: if $A_1, A_2, \dots \in \mathcal{L}(\mathbb{R})$ and $A_1 \subseteq A_2 \subseteq \dots$, then
$$\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \lambda(A_n).$$
5. **Downwards continuity**: let $A_1, A_2, \dots \in \mathcal{L}(\mathbb{R})$ and $A_1 \supseteq A_2 \supseteq \dots$, then
$$\lambda\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \lambda(A_n)$$
if $\lambda(A_1) < \infty$.

Proof. For (1), since $A \subseteq B$, we have $B = A \cup (B \setminus A)$, where A and $B \setminus A$ are disjoint. By countable additivity,
$$\lambda(B) = \lambda(A) + \lambda(B \setminus A).$$
Rearranging gives $\lambda(B \setminus A) = \lambda(B) - \lambda(A)$.

For (2), if A is compact, then A is bounded by the Heine–Borel theorem, so there exist $m < M$ with $A \subseteq [m, M]$. By monotonicity and $\lambda([m, M]) = M - m$,
$$\lambda(A) \leq \lambda([m, M]) = M - m < \infty.$$

For (3),
$$\lambda\left(\bigcup_{i=1}^{\infty} A_i\right) = \lambda^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \lambda^*(A_i) = \sum_{i=1}^{\infty} \lambda(A_i),$$
where we used countable subadditivity of λ^*.

Lecture 4 — Thursday, January 15, 2015

Proof continued. For (4), let $A = \bigcup_{n=1}^{\infty} A_n$ and define $B_1 = A_1$, $B_2 = A_2 \setminus A_1, \ldots, B_n = A_n \setminus A_{n-1}$. Then $\{B_i\}_{i=1}^{\infty}$ is a collection of pairwise disjoint sets in $\mathcal{L}(\mathbb{R})$, so that

$$\lambda\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \lambda(B_i) = \lambda(A).$$

Thus

$$\lim_{n\to\infty} \lambda(A_n) = \sum_{i=1}^{\infty} \lambda(B_i) = \lambda(\bigcup_{i=1}^{\infty} B_i) = \lambda(A).$$

For (5), let $B_n := A_1 \setminus A_n$. Then $B_1 \subseteq B_2 \subseteq \cdots$, and

$$\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n.$$

By (4),

$$\lambda\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \lambda(B_n).$$

Also, since $A_n \subseteq A_1$ and $\lambda(A_1) < \infty$,

$$\lambda(B_n) = \lambda(A_1) - \lambda(A_n).$$

Therefore

$$\lambda(A_1) - \lambda\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \big(\lambda(A_1) - \lambda(A_n)\big),$$

so $\lambda(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \lambda(A_n)$. □

Remark 4.1 The condition $\lambda(A_1) < \infty$ in (5) is necessary. For example, if $A_n = [n, \infty)$, then $\{A_n\}_{n=1}^{\infty}$ forms a decreasing chain, but $\lambda(A_n) = \infty$. Then $\bigcap_{n=1}^{\infty} A_n = \emptyset$, and therefore

$$\lambda\left(\bigcap_{n=1}^{\infty} A_n\right) = 0,$$

so

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

Proposition 4.2 (Outer regularity) If $E \in \mathcal{L}(\mathbb{R})$, then

$$\lambda(E) = \inf\{\lambda(U) : E \subseteq U,\ U \text{ open}\}.$$

Proof. For every open $U \supseteq E$, monotonicity gives $\lambda(E) \leq \lambda(U)$, so

$$\lambda(E) \leq \inf\{\lambda(U) : E \subseteq U,\ U \text{ open}\}.$$

If $\lambda(E) = \infty$, this inequality forces every such U to have infinite measure, and equality follows immediately. Assume henceforth that $\lambda(E) < \infty$. Conversely, let $\varepsilon > 0$. By the definition of outer measure, there exist open intervals $\{I_n\}_{n=1}^{\infty}$ such that

$$E \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} \ell(I_n) < \lambda^*(E) + \varepsilon = \lambda(E) + \varepsilon.$$

Set $U := \bigcup_{n=1}^{\infty} I_n$. Then U is open, $E \subseteq U$, and

$$\lambda(U) = \lambda^*(U) \leq \sum_{n=1}^{\infty} \ell(I_n) < \lambda(E) + \varepsilon.$$

Hence

$$\inf\{\lambda(U) : E \subseteq U,\ U \text{ open}\} \leq \lambda(E) + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ yields the reverse inequality. □

Proposition 4.3 If $E \in \mathcal{L}(\mathbb{R})$, then

$$\lambda(E) = \sup\{\lambda(K) \mid K \subseteq E,\ K \text{ compact}\}.$$

Proof. Let $E_n := E \cap [-n, n]$. Then $E_n \nearrow E$, so by continuity from below,

$$\lambda(E) = \lim_{n\to\infty} \lambda(E_n) = \sup_n \lambda(E_n).$$

Thus it suffices to treat bounded measurable sets. Fix n and $\varepsilon > 0$, and set

$$F_n := [-n, n] \setminus E_n.$$

By Proposition 4.2, choose an open set $O \supseteq F_n$ such that

$$\lambda(O) < \lambda(F_n) + \varepsilon.$$

Define

$$K_n := [-n, n] \setminus O.$$

Then K_n is compact and $K_n \subseteq E_n$. Also,

$$E_n \setminus K_n = E_n \cap O \subseteq O \setminus F_n,$$

so

$$\lambda(E_n \setminus K_n) \leq \lambda(O \setminus F_n) = \lambda(O) - \lambda(F_n) < \varepsilon.$$

Hence

$$\lambda(K_n) = \lambda(E_n) - \lambda(E_n \setminus K_n) > \lambda(E_n) - \varepsilon.$$

Since ε is arbitrary,

$$\lambda(E_n) = \sup\{\lambda(K) : K \subseteq E_n,\ K \text{ compact}\}.$$

Taking the supremum over n gives the claim for E. □

Corollary 4.4 If $E \in \mathcal{L}(\mathbb{R})$ with $\lambda(E) < \infty$ and $\varepsilon > 0$, then there exist an open set U and a compact set K such that

$$K \subseteq E \subseteq U \qquad \text{and} \qquad \lambda(U \setminus K) < \varepsilon.$$

Proof. By Proposition 4.2, choose an open $U \supseteq E$ with

$$\lambda(U) < \lambda(E) + \frac{\varepsilon}{2}.$$

Then

$$\lambda(U \setminus E) = \lambda(U) - \lambda(E) < \frac{\varepsilon}{2}.$$

By Proposition 4.3, choose a compact $K \subseteq E$ with

$$\lambda(E \setminus K) < \frac{\varepsilon}{2}.$$

Therefore,

$$\lambda(U \setminus K) \leq \lambda(U \setminus E) + \lambda(E \setminus K) < \varepsilon.$$

□

Corollary 4.5 For every $E \in \mathcal{L}(\mathbb{R})$, there is a Borel set B such that

$$\lambda(E \triangle B) = 0.$$

Proof. Set $E_m := E \cap [-m, m]$. For every $m, n \geq 1$, outer regularity gives an open set $U_{m,n} \supseteq E_m$ such that

$$\lambda(U_{m,n} \setminus E_m) < \frac{1}{n}.$$

Let $B_m := \bigcap_{n=1}^{\infty} U_{m,n}$. Then B_m is Borel, contains E_m, and

$$\lambda(B_m \setminus E_m) \leq \lambda(U_{m,n} \setminus E_m) < \frac{1}{n}$$

for every n, so $\lambda(B_m \setminus E_m) = 0$. The Borel set $B := \bigcup_{m=1}^{\infty} B_m$ contains $E = \bigcup_m E_m$, and

$$B \setminus E \subseteq \bigcup_{m=1}^{\infty} (B_m \setminus E_m),$$

a null set. Hence $\lambda(E \triangle B) = 0$. □

Riemann Integrability and Discontinuities

Theorem 4.6 (Characterization of Riemann integrability) A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if

$$\lambda(\{x \in [a, b] \mid f \text{ is not continuous at } x\}) = 0.$$

Proof. Let $E = \{x \in [a, b] \mid f \text{ is not continuous at } x\}$.

To begin with, we need to show that E is Lebesgue measurable in the first place. We do this by verifying several claims. For each $k \in \mathbb{N}$, let

$$E_k = \left\{ x \in [a, b] \;\middle|\; \begin{array}{l} \text{for every } \delta > 0, \text{ there exist } y, z \in [a, b] \cap [x - \delta, x + \delta] \\ \text{such that } |f(y) - f(z)| \geq 1/k \end{array} \right\}.$$

1. *Claim 1:* $\bigcup_{k=1}^{\infty} E_k \subseteq E$. *Verification.* If $z \in E_k$, then f is not continuous at z. Indeed, if f were continuous at z, then for $\varepsilon = 1/(2k)$ there would be a $\delta > 0$ such that $|z - x| < \delta$ implies $|f(z) - f(x)| < 1/(2k)$. Hence $|f(x) - f(y)| < 1/k$ for all $x, y \in [z - \delta, z + \delta]$, contradicting $z \in E_k$. Therefore $\bigcup_{k=1}^{\infty} E_k \subseteq E$.
2. *Claim 2:* $E = \bigcup_{k=1}^{\infty} E_k$. *Verification.* Conversely, if $z \in E$, then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ we can find y with $|y - z| < \delta$ and $|f(y) - f(z)| \geq \varepsilon_0$. Choose k with $1/k \leq \varepsilon_0$. Then $z \in E_k$. Hence $E \subseteq \bigcup_{k=1}^{\infty} E_k$, and together with Claim (1) we get
$$E = \bigcup_{k=1}^{\infty} E_k.$$
3. *Claim 3:* $E \in \mathcal{L}(\mathbb{R})$. *Verification.* We first show that each E_k is closed in $[a, b]$. Let $x \in [a, b] \setminus E_k$. Then there exists $\delta_0 > 0$ such that
$$|f(y) - f(z)| < 1/k \quad \text{for all } y, z \in [a, b] \cap [x - \delta_0, x + \delta_0].$$
Now let $u \in [a, b]$ satisfy $|u - x| < \delta_0/2$. If $y, z \in [a, b] \cap [u - \delta_0/2, u + \delta_0/2]$, then $y, z \in [a, b] \cap [x - \delta_0, x + \delta_0]$, so $|f(y) - f(z)| < 1/k$. Hence $u \notin E_k$. Therefore $[a, b] \setminus E_k$ is open in $[a, b]$, and E_k is closed in $[a, b]$.

   Since $[a, b]$ is compact, each closed subset E_k is compact, hence Borel, and therefore Lebesgue measurable. By Claim (2),

$$E = \bigcup_{k=1}^{\infty} E_k,$$

so E is a countable union of measurable sets and thus $E \in \mathcal{L}(\mathbb{R})$.

We can now prove the main statement.

Reverse implication. Assume first that $\lambda(E) = 0$. We prove that f is Riemann integrable by making the difference between its upper and lower Riemann sums arbitrarily small. Let $\varepsilon > 0$, and choose $C \geq 0$ such that $|f| \leq C$. Since $\lambda^*(E) = 0$, there is a countable open cover of E by intervals whose total length is less than $\varepsilon/(4C+1)$. Choose $k \in \mathbb{N}$ so large that $(b-a)/k < \varepsilon/2$. Because E_k is compact, finitely many intervals $I_1, \ldots, I_N$ from this cover suffice to cover E_k. Let

$$X = [a, b] \setminus \bigcup_{n=1}^{N} I_n.$$

If $z \in X$, then $z \notin E_k$, so we can choose a relative open interval V_z containing z such that the oscillation of f on $\overline{V_z}$ is less than $1/k$. Compactness gives a finite subcover $V_1, \ldots, V_m$ of X. The intervals

$$I_1, \ldots, I_N, V_1, \ldots, V_m$$

form an open cover of $[a, b]$, so this cover has a Lebesgue number. Choose a partition $P = \{a = x_0 < \cdots < x_n = b\}$ whose mesh is smaller than that number and which includes every endpoint of the I_j lying in $[a, b]$. Each partition interval is contained in one member of the displayed cover. If it is contained in some V_j, its oscillation is less than $1/k$. Otherwise its containing member must be one of the I_j, so the interval lies in $\bigcup_{j=1}^{N} I_j$ and has oscillation at most $2C$. Because every endpoint of every I_j in $[a, b]$ belongs to P, the total length of all partition intervals of the second kind is at most $\sum_{j=1}^{N} \ell(I_j)$. Therefore

$$U(f, P) - L(f, P) \leq 2C \sum_{j=1}^{N} \ell(I_j) + \frac{b-a}{k} < \frac{2C\varepsilon}{4C+1} + \frac{\varepsilon}{2} < \varepsilon.$$

Since ε was arbitrary, f is Riemann integrable.

Forward implication. Assume that f is Riemann integrable. It is enough to show that $\lambda(E_k) = 0$ for every $k \in \mathbb{N}$, because

$$\lambda(E) \leq \sum_{k=1}^{\infty} \lambda(E_k) = 0$$

once this is known. Fix $k \in \mathbb{N}$ and $\varepsilon > 0$. Choose a partition P of $[a, b]$ such that $U(f, P) -$

$L(f,P) < \varepsilon/k$. Write

$$P = \{x_0 = a < x_1 < \cdots < x_n = b\}.$$

Then

$$\frac{\varepsilon}{k} > U(f,P) - L(f,P) = \sum_{i=0}^{n-1}(M_i - m_i)(x_{i+1} - x_i).$$

Let $J = \{i \mid M_i - m_i \geq 1/k\}$. Every point of E_k that is not a partition point lies in the interior of a unique partition interval, and that interval must have oscillation at least $1/k$. Thus

$$E_k \subseteq \bigcup_{i\in J}(x_i, x_{i+1}) \cup \{x_0, \dots, x_n\}.$$

Moreover,

$$\frac{\varepsilon}{k} \geq \sum_{i\in J}(M_i - m_i)(x_{i+1} - x_i) \geq \frac{1}{k}\sum_{i\in J}(x_{i+1} - x_i).$$

Thus

$$\sum_{i\in J}\ell([x_i, x_{i+1}]) < \varepsilon,$$

By monotonicity,

$$\lambda(E_k) \leq \sum_{i\in J}\lambda((x_i, x_{i+1})) + \lambda(\{x_0, \dots, x_n\}) < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\lambda(E_k) = 0$. $\square$

Many functions are not Riemann integrable. For example, the characteristic function $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ is discontinuous at every point of $[0,1]$, and hence is not Riemann integrable.

Example 4.7 Every bounded function on $\mathbb{R}$ with finitely many discontinuities is Riemann integrable on every compact interval.

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and suppose that f has only finitely many discontinuities, say at points $x_1, \dots, x_m \in \mathbb{R}$. To prove Riemann integrability, fix any compact interval $[a,b]$ and consider the restriction $f|_{[a,b]}$.

Its set of discontinuities in $[a,b]$ is contained in

$$\{x_1, \dots, x_m\} \cap [a,b],$$

which is a finite set. Every finite set has Lebesgue measure 0, so

$$\lambda\big(\{x \in [a,b] \mid f|_{[a,b]} \text{ is not continuous at } x\}\big) = 0.$$

By Theorem 4.6, $f|_{[a,b]}$ is Riemann integrable on $[a,b]$. Since $[a,b]$ was arbitrary, f is Riemann integrable on every compact interval.

□

2. Measurable Functions and Integration

Definition 4.8 A real-valued function f is said to be **Lebesgue measurable** (or simply **measurable**) if

$$f^{-1}((\alpha,\infty)) = \{x \mid f(x) > \alpha\}$$

is measurable for every $\alpha \in \mathbb{R}$. A complex-valued function f is called measurable if both its real and imaginary parts are measurable.

Example 4.9 Let E be a measurable set in $\mathbb{R}$, and let $\mathbb{1}_E$ be its characteristic function. Then $\mathbb{1}_E$ is measurable. Indeed, if $\alpha \geq 1$, then $\mathbb{1}_E^{-1}((\alpha,\infty)) = \emptyset$. If $0 < \alpha \leq 1$, then $\mathbb{1}_E^{-1}((\alpha,\infty)) = E$. If $\alpha < 0$, then $\mathbb{1}_E^{-1}((\alpha,\infty)) = \mathbb{R}$. All three sets are measurable.

Example 4.10 Let $E_1,\dots,E_n \in \mathcal{L}(\mathbb{R})$ be pairwise disjoint, and let $a_1,\dots,a_n \in \mathbb{R}$. The function

$$\sum_{i=1}^{n} a_i \mathbb{1}_{E_i}$$

is called a **simple function**. It takes the value 0 outside $\bigcup_{i=1}^n E_i$.

Proposition 4.11 Simple functions are measurable.

Proof. Let $f = \sum_{i=1}^n a_i \mathbb{1}_{E_i}$ be a simple function. For every $\alpha \in \mathbb{R}$, we have Set $E_0 := \mathbb{R} \setminus \bigcup_{i=1}^n E_i$ and $a_0 := 0$. Then

$$f^{-1}((\alpha,\infty)) = \bigcup_{\substack{0 \leq i \leq n \\ a_i > \alpha}} E_i,$$

which is a finite union of measurable sets, and hence measurable. Including E_0 is necessary when $\alpha < 0$, because f vanishes outside the displayed support. $\square$

Lecture 5 — Tuesday, January 20, 2015

Measurable Functions and Closure Properties

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable if $f^{-1}((\alpha, \infty)) \in \mathcal{L}(\mathbb{R})$ for every $\alpha \in \mathbb{R}$.

Proposition 5.1 Let $f : \mathbb{R} \to \mathbb{R}$. The following are equivalent:

1. f is Lebesgue measurable.
2. $f^{-1}((\alpha, \infty)) \in \mathcal{L}(\mathbb{R})$ for all $\alpha \in \mathbb{R}$.
3. $f^{-1}((-\infty, \alpha)) \in \mathcal{L}(\mathbb{R})$ for all $\alpha \in \mathbb{R}$.
4. $f^{-1}((-\infty, \alpha]) \in \mathcal{L}(\mathbb{R})$ for all $\alpha \in \mathbb{R}$.

Proof. (1) *implies* (2). This is the definition.

(2) *implies* (3). Let $\alpha \in \mathbb{R}$. Then

$$f^{-1}((-\infty, \alpha)) = \bigcup_{n=1}^{\infty} f^{-1}((-\infty, \alpha - 1/n]) = \bigcup_{n=1}^{\infty} f^{-1}((\alpha - 1/n, \infty))^c,$$

which is measurable.

(3) *implies* (4). Let $\alpha \in \mathbb{R}$. Then

$$f^{-1}((-\infty, \alpha]) = \{x \mid f(x) \le \alpha\} = \bigcap_{n=1}^{\infty} \{x \mid f(x) < \alpha + 1/n\} = \bigcap_{n=1}^{\infty} f^{-1}((-\infty, \alpha + 1/n)),$$

which is a countable intersection of measurable sets, and hence measurable.

(4) *implies* (1). Let $\alpha \in \mathbb{R}$. Then

$$f^{-1}((\alpha,\infty)) = \{x \mid f(x) > \alpha\} = \{x \mid f(x) \leq \alpha\}^c = f^{-1}((-\infty,\alpha])^c,$$

which is measurable. □

Corollary 5.2 A function $f : \mathbb{R} \to \mathbb{R}$ is measurable if and only if $f^{-1}(B) \in \mathcal{L}(\mathbb{R})$ for every $B \in \mathcal{B}(\mathbb{R})$.

Proof. Suppose first that f is measurable, and let

$$\mathcal{M} = \{B \subseteq \mathbb{R} \mid f^{-1}(B) \in \mathcal{L}(\mathbb{R})\}.$$

Preimages preserve complements and countable unions, so $\mathcal{M}$ is a σ-algebra. It contains every interval (α,∞) by measurability, and these intervals generate $\mathcal{B}(\mathbb{R})$. Hence $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}$.

Conversely, if the inverse image of every Borel set is measurable, then this holds in particular for every open interval (α,∞). Thus f is measurable by definition. □

Proposition 5.3 Let $f, g : \mathbb{R} \to \mathbb{R}$ be measurable, let $\phi : \mathbb{R} \to \mathbb{R}$ continuous, and let $c \in \mathbb{R}$. Then the following functions are measurable:

1. cf.
2. $f + g$.
3. $\phi \circ f$.
4. $|f|$, f^+, and f^-, where $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$.

Proof. For (1), let $\alpha \in \mathbb{R}$.

We have $(cf)^{-1}((\alpha,\infty)) = \{x \mid cf(x) > \alpha\}$. If $c > 0$, this is $\{x \mid f(x) > \alpha/c\} = f^{-1}((\alpha/c,\infty)) \in \mathcal{L}(\mathbb{R})$, since f is measurable. If $c < 0$, this is

$$\{x \mid f(x) < \alpha/c\} = f^{-1}((-\infty,\alpha/c)) \in \mathcal{L}(\mathbb{R}),$$

since f is measurable. If $c = 0$, this is $\{x \mid 0 > \alpha\}$, which is either $\emptyset$ or $\mathbb{R}$, both of which are measurable.

For (2), let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}(f+g)^{-1}((\alpha,\infty)) &= \{x \mid f(x)+g(x) > \alpha\} \\ &= \{x \mid f(x) > \alpha - g(x)\} \\ &= \{x \mid \text{there exists } r \in \mathbb{Q} \text{ such that } f(x) > r \text{ and } r > \alpha - g(x)\} \\ &= \bigcup_{r\in\mathbb{Q}} \{x \mid f(x) > r\} \cap \{x \mid g(x) > \alpha - r\} \\ &= \bigcup_{r\in\mathbb{Q}} (f^{-1}((r,\infty)) \cap g^{-1}((\alpha - r,\infty))),\end{aligned}$$

which is a countable union of measurable sets, and hence measurable.

For (3), let $\alpha \in \mathbb{R}$. Then $(\phi \circ f)^{-1}((\alpha,\infty)) = f^{-1}(\phi^{-1}((\alpha,\infty)))$. Since ϕ is continuous, $\phi^{-1}((\alpha,\infty)) \in \mathcal{B}(\mathbb{R})$ (exercise!), so $f^{-1}(\phi^{-1}((\alpha,\infty))) \in \mathcal{L}(\mathbb{R})$ by Corollary 5.2.

For (4), let $\phi(x) = |x|$, which is continuous. Then by (3), $|f|$ is measurable. Thus $f^+ = (|f|+f)/2$ and $f^- = (|f|-f)/2$ are both measurable. □

Definition 5.4 A function $f : \mathbb{R} \to [-\infty,\infty]$ is called an **extended real function**. We say that $f : \mathbb{R} \to [-\infty,\infty]$ is **Lebesgue measurable** if $f^{-1}((\alpha,\infty)) \in \mathcal{L}(\mathbb{R})$ for all $\alpha \in \mathbb{R}$ and $f^{-1}(\{-\infty\}), f^{-1}(\{\infty\}) \in \mathcal{L}(\mathbb{R})$.

Proposition 5.5 Let $f_1, f_2, \ldots$ be a sequence of extended real-valued measurable functions. Then the following are also measurable:

1. $\sup_{n\in\mathbb{N}} f_n$.
2. $\inf_{n\in\mathbb{N}} f_n$.
3. $\limsup_{n\to\infty} f_n$.
4. $\liminf_{n\to\infty} f_n$.
5. f, where $f_n \to f$ pointwise.

Proof. For (1), let $g = \sup_{n\in\mathbb{N}} f_n$. Let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} g^{-1}((\alpha,\infty)) &= \{x \mid g(x) > \alpha\} \\ &= \{x \mid \sup_{n\in\mathbb{N}} f_n(x) > \alpha\} \\ &= \{x \mid \text{there exists } n \in \mathbb{N} \text{ such that } f_n(x) > \alpha\} \\ &= \bigcup_{n\in\mathbb{N}} \{x \mid f_n(x) > \alpha\} \\ &= \bigcup_{n\in\mathbb{N}} f_n^{-1}((\alpha,\infty)) \in \mathcal{L}(\mathbb{R}) \end{aligned}$$

since f_n is measurable for all n.

For (2), let $h = \inf_{n\in\mathbb{N}} f_n$. Then

$$h^{-1}((\alpha,\infty)) = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} f_n^{-1}((\alpha + 1/m, \infty)),$$

which is measurable. In both the supremum and infimum cases, the fibres over ∞ and $-\infty$ are obtained from the sets above by countable unions, intersections, and complements. Thus both extended real functions are measurable.

For (3), define $g_n = \sup_{k\geq n} f_k$ for every $n \in \mathbb{N}$. Each g_n is measurable by (1), and $\limsup_{n\to\infty} f_n = \lim_{n\to\infty} g_n = \inf_{n\in\mathbb{N}} g_n$ is measurable by (2).

For (4), use the same argument with infima and suprema interchanged.

For (5), let $f = \lim_{n\to\infty} f_n$. Then $f = \liminf_{n\to\infty} f_n$ is measurable by (4). □

Simple Functions and Approximation

Theorem 5.6 Let $f : \mathbb{R} \to \mathbb{R}^{\geq 0}$ be a function. Then f is measurable if and only if there exists a sequence of simple functions $\{\phi_n\}_{n\in\mathbb{N}}$ such that $\phi_n \leq \phi_{n+1}$ and $\phi_n \to f$ pointwise.

Proof. *Reverse implication.* The pointwise limit of measurable functions is measurable by Proposition 5.5, and simple functions are measurable by Proposition 4.11.

Forward implication. Fix $n \in \mathbb{N}$. Define $E_{k,n} = f^{-1}([k/2^n, (k+1)/2^n))$ for $k = 0, 1, \ldots, n2^n - 1$, which is measurable because f is measurable. Let $E_n = \bigcup_{k=0}^{n2^n-1} E_{k,n}$ and define

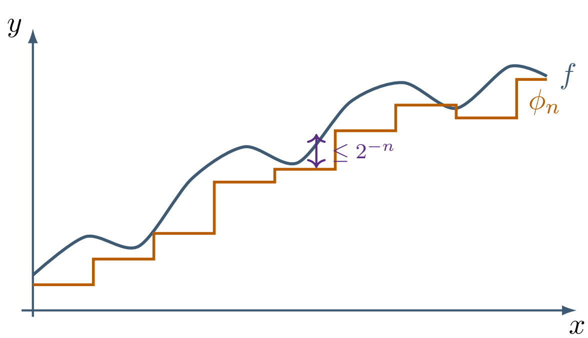
$$\phi_n(x) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\mathbb{1}_{E_{k,n}} + n\mathbb{1}_{E_n^c},$$

which is a simple function. Then $\phi_n \leq \phi_{n+1} \leq f$, and $\lim_{n\to\infty} \phi_n(x) = f(x)$ for every $x \in \mathbb{R}$. Indeed, fix $x \in \mathbb{R}$. There exists $N \in \mathbb{N}$ such that $f(x) \leq N$, so $x \in E_n$ for all $n \geq N$ (since $E_n = \{x \mid f(x) < n\}$). From the way we defined ϕ_n, we have

$$|\phi_n(x) - f(x)| \leq \frac{1}{2^n} \xrightarrow{n\to\infty} 0.$$

$\square$

Figure 4 shows the construction at one fixed resolution: the simple function rounds f downward, with a vertical error no larger than the mesh size.



measurable functions are built from simple functions below

FIGURE 4

Lebesgue Integrals of Nonnegative Functions

Definition 5.7 If

$$\phi = \sum_{i=1}^{n} a_i \mathbb{1}_{E_i},$$

is a nonnegative simple function, represented with $a_i \geq 0$, and $A \in \mathcal{L}(\mathbb{R})$, then the **Lebesgue**

integral over A is

$$\int_A \phi \, \mathrm{d}\lambda := \sum_{i=1}^{n} a_i \lambda(E_i \cap A),$$

with the convention that $0 \cdot \infty = 0$.

Example 5.8 $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ is not Riemann integrable, but it is very Lebesgue integrable:

$$\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}\cap[0,1]} \, \mathrm{d}\lambda = \lambda(\mathbb{Q} \cap [0,1]) = 0.$$

Definition 5.9 If $f : \mathbb{R} \to [0,\infty]$, then for $E \in \mathcal{L}(\mathbb{R})$, the **Lebesgue integral** over E is

$$\int_E f \, \mathrm{d}\lambda = \sup_{\substack{0 \le \phi \le f \\ \phi \text{ simple}}} \int_E \phi \, \mathrm{d}\lambda.$$

Lecture 6 — Thursday, January 22, 2015

Proposition 6.1 Let $A \in \mathcal{L}(\mathbb{R})$, and let $f, g : \mathbb{R} \to [0,\infty]$ be measurable.

1. If $f \le g$ on A, then
$$\int_A f \, \mathrm{d}\lambda \le \int_A g \, \mathrm{d}\lambda.$$
2. If $\emptyset \subseteq B \subseteq A$ and $B \in \mathcal{L}(\mathbb{R})$, then
$$\int_B f \, \mathrm{d}\lambda = \int_A f \cdot \mathbb{1}_B \, \mathrm{d}\lambda.$$
3. The two integrations defined above coincide for simple functions.

Proof. For (1), if $\phi \le f$ and ϕ is simple, then $\phi \le g$.

Therefore the admissible class

$$\{\phi : 0 \le \phi \le f, \ \phi \text{ simple}\}$$

is contained in
$$\{\psi : 0 \le \psi \le g,\ \psi \text{ simple}\}.$$
Taking suprema of the corresponding simple integrals gives
$$\int_A f \,\mathrm{d}\lambda \le \int_A g \,\mathrm{d}\lambda.$$

For (2), let
$$\mathcal{S}_B := \{\phi : 0 \le \phi \le f,\ \phi \text{ simple on } B\}$$
and
$$\mathcal{S}_A := \{\psi : 0 \le \psi \le f\mathbb{1}_B,\ \psi \text{ simple on } A\}.$$
If $\phi \in \mathcal{S}_B$, define $\widetilde{\phi} := \phi\mathbb{1}_B$. Then $\widetilde{\phi} \in \mathcal{S}_A$ and
$$\int_A \widetilde{\phi} \,\mathrm{d}\lambda = \int_B \phi \,\mathrm{d}\lambda.$$
Conversely, if $\psi \in \mathcal{S}_A$, then $\psi = \psi\mathbb{1}_B$, and the restriction $\psi|_B$ belongs to $\mathcal{S}_B$ with
$$\int_A \psi \,\mathrm{d}\lambda = \int_B \psi \,\mathrm{d}\lambda.$$
Hence the two sets of simple integral values are identical, so their suprema are equal:
$$\int_B f \,\mathrm{d}\lambda = \int_A f\mathbb{1}_B \,\mathrm{d}\lambda.$$

For (3), let s be a nonnegative simple function. In the definition
$$\int_A s \,\mathrm{d}\lambda = \sup\left\{\int_A \phi \,\mathrm{d}\lambda : 0 \le \phi \le s,\ \phi \text{ simple}\right\},$$
the admissible family contains $\phi = s$, so the supremum is at least the integral of s previously defined for simple functions. On the other hand, if $0 \le \phi \le s$ are simple, monotonicity of the simple integral gives
$$\int_A \phi \,\mathrm{d}\lambda \le \int_A s \,\mathrm{d}\lambda$$
for the old simple function definition. Thus the supremum cannot exceed that value. Therefore both definitions coincide for simple functions.

□

Lecture 7 — Tuesday, January 27, 2015

Definition 7.1 Let $f : \mathbb{R} \to [-\infty, \infty]$ be a function. We say that f is **Lebesgue integrable** if the following two conditions hold:

1. f is measurable; that is, f^+ and f^- are measurable.
2. $$\int f^+ \, d\lambda < \infty \qquad \text{and} \qquad \int f^- \, d\lambda < \infty.$$

In that case, we define
$$\int f \, d\lambda := \int f^+ \, d\lambda - \int f^- \, d\lambda.$$
If $\emptyset \neq A \in \mathcal{L}(\mathbb{R})$, then f is integrable over A if $f \cdot \mathbb{1}_A$ is integrable.

Proposition 7.2 If $B \subseteq A$ are Lebesgue measurable and f is integrable on B, then
$$\int_B f \, d\lambda = \int_A f \cdot \mathbb{1}_B \, d\lambda.$$

Proof. Write $f = f^+ - f^-$. Then
$$(f\mathbb{1}_B)^+ = f^+\mathbb{1}_B \qquad \text{and} \qquad (f\mathbb{1}_B)^- = f^-\mathbb{1}_B.$$
By Proposition 6.1(2) applied to the nonnegative measurable functions f^+ and f^-,
$$\int_B f^+ \, d\lambda = \int_A f^+\mathbb{1}_B \, d\lambda \qquad \text{and} \qquad \int_B f^- \, d\lambda = \int_A f^-\mathbb{1}_B \, d\lambda.$$
Therefore,
$$\begin{aligned}\int_B f \, d\lambda &= \int_B f^+ \, d\lambda - \int_B f^- \, d\lambda \\ &= \int_A f^+\mathbb{1}_B \, d\lambda - \int_A f^-\mathbb{1}_B \, d\lambda\end{aligned}$$

$$= \int_A (f^+ - f^-)\mathbb{1}_B \, \mathrm{d}\lambda$$
$$= \int_A f\mathbb{1}_B \, \mathrm{d}\lambda.$$

□

Monotone Convergence Theorem and Fatou's Lemma

Theorem 7.3 (Monotone convergence theorem) Let $f_1, f_2, \dots : \mathbb{R} \to [0,\infty]$ be measurable functions and $f_1 \le f_2 \le \dots$ pointwise. Then

$$\int_A \lim_{n\to\infty} f_n \, \mathrm{d}\lambda = \lim_{n\to\infty} \int_A f_n \, \mathrm{d}\lambda$$

for $A \in \mathcal{L}(\mathbb{R})$. In particular,

$$\text{if } \sup_{n\in\mathbb{N}} \int_A f_n \, \mathrm{d}\lambda < \infty, \text{ then } \lim_{n\to\infty} f_n < \infty \text{ almost everywhere on } A.$$

The picture in Figure 5 captures both monotonicities in the theorem: the graphs rise pointwise, and the shaded areas rise with them.

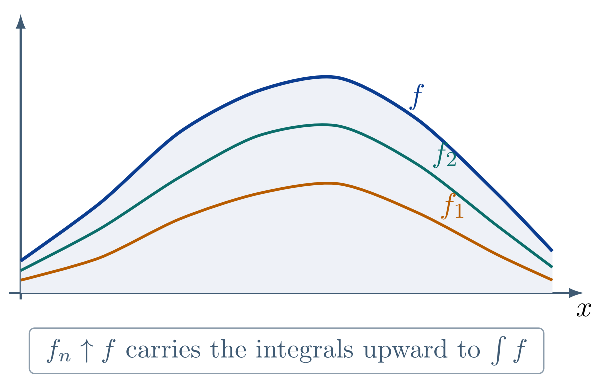


FIGURE 5

Proof. Fix $A \in \mathcal{L}(\mathbb{R})$ and write $f = \lim_{n\to\infty} f_n$. Since

$$0 \le f_1 \le f_2 \le \dots,$$

we have

$$0 \leq \int_A f_1 \,\mathrm{d}\lambda \leq \int_A f_2 \,\mathrm{d}\lambda \leq \dots$$

by Proposition 6.1(1). Thus

$$\lim_{n\to\infty} \int_A f_n \,\mathrm{d}\lambda = \sup_n \int_A f_n \,\mathrm{d}\lambda.$$

The function f is measurable, and $f_n \leq f$ for every n. Monotonicity therefore gives

$$\sup_{n\in\mathbb{N}} \int_A f_n \,\mathrm{d}\lambda \leq \int_A f \,\mathrm{d}\lambda.$$

It remains to prove the reverse inequality

$$\int_A f \,\mathrm{d}\lambda \leq \lim_{n\to\infty} \int_A f_n \,\mathrm{d}\lambda.$$

Let $0 \leq \phi \leq f$ be simple, and let $0 < \delta < 1$. Define

$$A_n = \{x \in A \mid f_n(x) > \delta\phi(x)\}.$$

Then $A_1 \subseteq A_2 \subseteq \dots$, and their union contains $\{x \in A \mid \phi(x) > 0\}$. Moreover,

$$\int_A f_n \,\mathrm{d}\lambda \geq \int_{A_n} f_n \,\mathrm{d}\lambda \geq \int_{A_n} \delta\phi \,\mathrm{d}\lambda,$$

because $\delta\phi(x) \leq f_n(x)$ on A_n. Write

$$\phi = \sum_{i=1}^m a_i \mathbb{1}_{E_i}.$$

Then

$$\int_{A_n} \delta\phi \,\mathrm{d}\lambda = \delta \sum_{i=1}^m a_i \lambda(E_i \cap A_n)$$

for every $n \in \mathbb{N}$. Now,

$$\lim_{n\to\infty} \int_A f_n \,\mathrm{d}\lambda \geq \delta \sum_{i=1}^m a_i \lambda(E_i \cap A)$$

because the union of the A_n contains every $E_i \cap A$ for which $a_i > 0$. Hence

$$\lim_{n\to\infty} \int_A f_n \,\mathrm{d}\lambda \geq \delta \int_A \phi \,\mathrm{d}\lambda.$$

Letting $\delta \to 1$, we have

$$\lim_{n\to\infty} \int_A f_n \,\mathrm{d}\lambda \geq \int_A \phi \,\mathrm{d}\lambda$$

for all simple functions ϕ such that $0 \leq \phi \leq f$. Hence

$$\lim_{n\to\infty} \int_A f_n \, \mathrm{d}\lambda \geq \int_A f \, \mathrm{d}\lambda.$$

Finally, if the common value of these integrals is finite and

$$N = \{x \in A \mid f(x) = \infty\},$$

then $m\mathbb{1}_N \leq f$ for every m. Consequently,

$$m\lambda(N) \leq \int_A f \, \mathrm{d}\lambda < \infty$$

for every m, so $\lambda(N) = 0$. □

Proposition 7.4 Let $f, g : \mathbb{R} \to [0, \infty]$ be measurable, $A \in \mathcal{L}(\mathbb{R})$, and $c \geq 0$. Then:

1. $$\int_A (f+g) = \int_A f + \int_A g \qquad \text{and} \qquad \int_A cf = c\int_A f.$$
2. If $f_1, f_2, \cdots : \mathbb{R} \to [0, \infty]$ are measurable, then
$$\int_A \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_A f_n.$$

Proof. For (1), if f, g are measurable and nonnegative, then there exists $\phi_n \to f$, $0 \leq \phi_n \nearrow f$ simple and then there exists $\psi_n \to g$, $0 \leq \psi_n \nearrow g$ simple and therefore $\phi_n + \psi_n \to f + g$ with $\{\phi_n + \psi_n\}$ an increasing sequence of nonnegative simple functions. Then, by the monotone convergence theorem,

$$\int f + \int g = \lim_{n\to\infty} \int \phi_n + \lim_{n\to\infty} \int \psi_n = \lim_{n\to\infty} \left(\int \phi_n + \int \psi_n \right).$$

By definition of the Lebesgue integral for simple functions, we know $\int \phi_n + \int \psi_n = \int (\phi_n + \psi_n)$ for every n. Passing to the limit gives

$$\int f + \int g = \int (f+g).$$

Similarly, if $c \geq 0$ and $0 \leq \phi_n \nearrow f$ are simple, then $0 \leq c\phi_n \nearrow cf$ and

$$\int c\phi_n = c \int \phi_n.$$

By the monotone convergence theorem,

$$\int cf = c \int f.$$

For (2), let

$$g_n = \sum_{i=1}^{n} f_i,$$

for all $n \in \mathbb{N}$. Then

$$\lim_{n\to\infty} g_n = \lim_{n\to\infty} \sum_{i=1}^{n} f_i,$$

and

$$0 \leq g_1 \leq g_2 \leq \dots.$$

By the monotone convergence theorem,

$$\lim_{n\to\infty} \int_A g_n = \int_A \sum_{i=1}^{\infty} f_i = \lim_{n\to\infty} \sum_{i=1}^{n} \int_A f_i = \sum_{i=1}^{\infty} \int_A f_i,$$

because $g_n = f_1 + \cdots + f_n$. □

Theorem 7.5 (Fatou's lemma) Let $f_n : \mathbb{R} \to [0, \infty]$ be measurable. Then

$$\int_A \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_A f_n$$

for any $A \in \mathcal{L}(\mathbb{R})$.

Proof. Let $n \in \mathbb{N}$ and let $g_n = \inf_{k\geq n} f_k$. Then $\liminf_{n\to\infty} f_n = \lim_{n\to\infty} g_n$. Now,

$$0 \leq g_1 \leq g_2 \leq \dots,$$

so by the monotone convergence theorem,

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

Now, $g_n \le f_k$ for all $k \ge n$, so $\int_A g_n \le \int_A f_k$ for all $k \ge n$. Hence

$$\int_A g_n \le \inf_{k \ge n} \int_A f_k$$

for every n. Hence,

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

□

Remark 7.6 Let $f_n = n\mathbb{1}_{[0,1/n]}$. Then $f_n \to 0$ pointwise, but

$$\lim_{n\to\infty} \int f_n = 1 \neq \int \lim_{n\to\infty} f_n = 0.$$

Lecture 8 — Thursday, January 29, 2015

Integrable Functions and Almost Everywhere Equivalence

Definition 8.1 We say f is integrable on A if $f \cdot \mathbb{1}_A$ is integrable. If $f : \mathbb{R} \to \mathbb{C}$, we say f is integrable if $\mathrm{Im}(f)$ and $\mathrm{Re}(f)$ are both integrable; in that case, we write

$$\int f = \int \mathrm{Re}(f) + i \int \mathrm{Im}(f).$$

Fact 8.2 Let $f : \mathbb{R} \to [-\infty, \infty]$ be integrable. Then $\lambda(f^{-1}(\{-\infty, \infty\})) = 0$.

Proof. Let $A = f^{-1}(\{-\infty, \infty\})$, which is a measurable subset of $\mathbb{R}$. Let

$$g_n = n\mathbb{1}_A \le |f|.$$

Then

$$n\lambda(A) = \int g_n \le \int |f| < \infty$$

because f is integrable. Since this holds for all $n \in \mathbb{N}$, we must have $\lambda(A) = 0$. □

Definition 8.3 Let $f, g : \mathbb{R} \to [-\infty, \infty]$. We say $f = g$ **almost everywhere** if

$$\lambda(\{x \mid f(x) \neq g(x)\}) = 0.$$

Fact 8.4 Let $f : \mathbb{R} \to [-\infty, \infty]$ be integrable, and let $f = g$ almost everywhere for some other function $g : \mathbb{R} \to [-\infty, \infty]$. Then g is integrable as well, with

$$\int f = \int g.$$

Proof. Let $N = \{x \mid f(x) \neq g(x)\}$. This is a measurable null set. Lebesgue measure is complete, so every subset of N is measurable. For each $\alpha \in \mathbb{R}$,

$$\{g > \alpha\} = \big(\{f > \alpha\} \setminus N\big) \cup \big(\{g > \alpha\} \cap N\big),$$

and the same decomposition works for the fibres over ∞ and $-\infty$. Thus g is measurable.

Every nonnegative measurable function supported on N has integral zero: any simple function below it is also supported on N, so its simple integral is zero. Since $f^{\pm} = g^{\pm}$ off N, it follows that

$$\int f^+ \,\mathrm{d}\lambda = \int g^+ \,\mathrm{d}\lambda \qquad \text{and} \qquad \int f^- \,\mathrm{d}\lambda = \int g^- \,\mathrm{d}\lambda.$$

The two integrals on the left are finite because f is integrable. Hence g is integrable, and subtracting the displayed equalities gives

$$\int f \,\mathrm{d}\lambda = \int g \,\mathrm{d}\lambda.$$

□

Definition 8.5 Let $f_n : \mathbb{R} \to \mathbb{R}$ be measurable and $f : \mathbb{R} \to \mathbb{R}$. We say $f_n \to f$ **almost everywhere** if there exists $N \in \mathcal{L}(\mathbb{R})$ such that $\lambda(N) = 0$ and $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in \mathbb{R} \setminus N$ (that is, f_n converges pointwise everywhere outside of N).

Example 8.6 If $f_n \to f$ almost everywhere and f_n is measurable, then f is measurable.

Proof. Let N be such that $\lambda(N) = 0$ and $f_n \to f$ pointwise on $\mathbb{R} \setminus N$. Then f is measurable on $\mathbb{R} \setminus N$ by Proposition 5.5(5). Since N is measurable, f is measurable on $\mathbb{R}$. □

Proposition 8.7 Let $f, g : \mathbb{R} \to \mathbb{R}$ be integrable and let $c \in \mathbb{R}$. Then the following hold:

1. cf is integrable and
$$\int cf = c \int f.$$
2. $f + g$ is integrable with
$$\int (f+g) = \int f + \int g.$$
3. $|f|$ is integrable, and
$$\left| \int f \right| \leq \int |f|.$$

Proof. For (1), if $c \geq 0$, then $(cf)^+ = cf^+$ and $(cf)^- = cf^-$, so
$$\int (cf)^+ = c \int f^+ < \infty \qquad \text{and} \qquad \int (cf)^- = c \int f^- < \infty.$$
Hence cf is integrable, and
$$\int cf = \int (cf)^+ - \int (cf)^- = c \left(\int f^+ - \int f^- \right) = c \int f.$$
If $c < 0$, then $(cf)^+ = |c| f^-$ and $(cf)^- = |c| f^+$. Consequently,
$$\int cf = |c| \int f^- - |c| \int f^+ = c \int f.$$

For (2), we know that if f and g are measurable, then $f + g$ is measurable too. Since $(f+g)^+ \leq f^+ + g^+$, we have
$$\int (f+g)^+ \leq \int f^+ + \int g^+ < \infty$$
and similarly, since $(f+g)^- \leq f^- + g^-$,
$$\int (f+g)^- \leq \int f^- + \int g^- < \infty.$$
Thus $f + g$ is integrable. We want to show
$$\int (f+g) = \int f + \int g.$$

Now

$$(f+g)^+ - (f+g)^- = f+g = (f^+ - f^-) + (g^+ - g^-),$$

so

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+,$$

and

$$\int (f+g)^+ + \int f^- + \int g^- = \int (f+g)^- + \int f^+ + \int g^+$$

and hence

$$\int (f+g) = \int (f+g)^+ - \int (f+g)^- = \int f^+ - \int f^- + \int g^+ - \int g^- = \int f + \int g.$$

For (3), f being integrable implies that f^+, f^- are measurable and hence $|f| = f^+ + f^-$ is measurable as well. Also,

$$\int |f| = \int f^+ + \int f^- < \infty$$

implies that $|f|$ is integrable. Furthermore,

$$\left|\int f\right| = \left|\int f^+ - \int f^-\right| \le \left|\int f^+\right| + \left|\int f^-\right| = \int f^+ + \int f^- = \int (f^+ + f^-) = \int |f|.$$

□

Corollary 8.8 Let $f : \mathbb{R} \to \mathbb{R}$. Then f is integrable if and only if $|f|$ is integrable.

Warning It may be tempting to think that if $f+g$ is integrable, then so are f and g individually. However, this is not necessarily true. For example, consider the function $f = \mathbb{1}_E + \mathbb{1}_{[0,1]\setminus E}$ where $[0,1] \supset E \notin \mathcal{L}(\mathbb{R})$. Then $f = 1$ is integrable on $[0,1]$, but $\mathbb{1}_E$ is not even measurable, let alone integrable.

Dominated Convergence Theorem

We will use Fatou's lemma here. Recall its statement: if $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions with $f_n : \mathbb{R} \to [0,\infty]$ and $A \in \mathcal{L}(\mathbb{R})$, then

$$\int_A \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int_A f_n.$$

Theorem 8.9 (Dominated convergence theorem) Let $f_n : E \to \mathbb{R}$ be measurable, where $E \in \mathcal{L}(\mathbb{R})$. Let $g : \mathbb{R} \to [0, \infty)$ be integrable on E, and suppose that g dominates $\{f_n\}$ on E almost everywhere; that is, $|f_n| \leq g$ almost everywhere on E for every $n \in \mathbb{N}$. If $f_n \to f$ pointwise almost everywhere on E, then f is integrable on E and

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

Proof. We prove this for $E = \mathbb{R}$; the general case follows by extending the functions by zero outside E. Define

$$N = \{x \in \mathbb{R} \mid \lim_{n\to\infty} f_n(x) \neq f(x) \text{ or the limit does not exist}\} \cup \bigcup_{n=1}^{\infty} \{x \in \mathbb{R} \mid |f_n(x)| > g(x)\},$$

which is the set of points that are up to no good. By assumption, $\lambda(N) = 0$. Let $A = \mathbb{R} \setminus N$. We have $|f_n| \leq g$ and $|f| \leq g$ on A. Therefore,

$$\int_A |f_n| \,\mathrm{d}\lambda \leq \int_{\mathbb{R}} g \,\mathrm{d}\lambda < \infty \qquad \text{and} \qquad \int_A |f| \,\mathrm{d}\lambda \leq \int_{\mathbb{R}} g \,\mathrm{d}\lambda < \infty.$$

The restrictions of the f_n to A are measurable, and their pointwise limit $f|_A$ is measurable by Proposition 5.5(5). Hence $f_n \mathbb{1}_A$ and $f \mathbb{1}_A$ are integrable. They agree almost everywhere with f_n and f, respectively, so Fact 8.4 shows that every f_n and f is integrable and that

$$\int_{\mathbb{R}} f_n \,\mathrm{d}\lambda = \int_A f_n \,\mathrm{d}\lambda \qquad \text{and} \qquad \int_{\mathbb{R}} f \,\mathrm{d}\lambda = \int_A f \,\mathrm{d}\lambda.$$

Lecture 9 — Tuesday, February 03, 2015

Proof continued. Since $|f_n| \leq g$ on A, the functions $g + f_n$ and $g - f_n$ are nonnegative on A for every n.

Apply Fatou's lemma to $h_n := g + f_n$:

$$\int_A \liminf_{n\to\infty} h_n \le \liminf_{n\to\infty} \int_A h_n.$$

Because $f_n \to f$ pointwise on A, we have $h_n \to g + f$, so

$$\int_A (g+f) \le \liminf_{n\to\infty} \left(\int_A g + \int_A f_n \right) = \int_A g + \liminf_{n\to\infty} \int_A f_n.$$

Subtracting $\int_A g$ from both sides gives

$$\int_A f \le \liminf_{n\to\infty} \int_A f_n.$$

Now apply Fatou's lemma to $k_n := g - f_n$:

$$\int_A \liminf_{n\to\infty} k_n \le \liminf_{n\to\infty} \int_A k_n.$$

Again $k_n \to g - f$, so

$$\int_A (g-f) \le \liminf_{n\to\infty} \left(\int_A g - \int_A f_n \right) = \int_A g - \limsup_{n\to\infty} \int_A f_n.$$

Subtracting $\int_A g$ and multiplying by -1 give

$$\limsup_{n\to\infty} \int_A f_n \le \int_A f.$$

Therefore,

$$\limsup_{n\to\infty} \int_A f_n \le \int_A f \le \liminf_{n\to\infty} \int_A f_n.$$

Hence

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

Since $\int_{\mathbb{R}} f_n = \int_A f_n$ and $\int_{\mathbb{R}} f = \int_A f$, this also gives

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n = \int_{\mathbb{R}} f.$$

$\square$

3. L^p Spaces and Hilbert Spaces

L^p and L^∞ Spaces

Definition 9.1 Let $A \in \mathcal{L}(\mathbb{R})$ with $\lambda(A) > 0$. Let $f : A \to \mathbb{C}$ be a measurable function. For $1 \leq p < \infty$, define

$$\|f\|_{L^p(A)} = \left(\int_A |f|^p \right)^{1/p}.$$

When A is clear from context, we write $\|f\|_p$ instead of $\|f\|_{L^p(A)}$.

Note that if $f = g$ almost everywhere on A, then $\|f - g\|_p = 0$. We thus define an equivalence relation $\sim$ by $f \sim g$ when $f = g$ almost everywhere. Observe that if $f \sim g$, then $\|f\|_p = \|g\|_p$.

Definition 9.2 We define $L^p(A)$ as the set of equivalence classes of measurable functions $f : A \to \mathbb{C}$ such that $\|f\|_p < \infty$, where two functions are considered equivalent if they are equal almost everywhere.

Example 9.3 $C([a, b])$, the set of continuous functions on $[a, b]$, is a subset of $L^p([a, b])$ for $1 \leq p < \infty$. For example, $C([1, 2]) \subset L^1([1, 2])$, but $C(\mathbb{R}) \not\subset L^1(\mathbb{R})$.

Even in two dimensions, changing p changes the geometry of the norm. Figure 6 shows the three standard unit balls that foreshadow this distinction in function spaces.

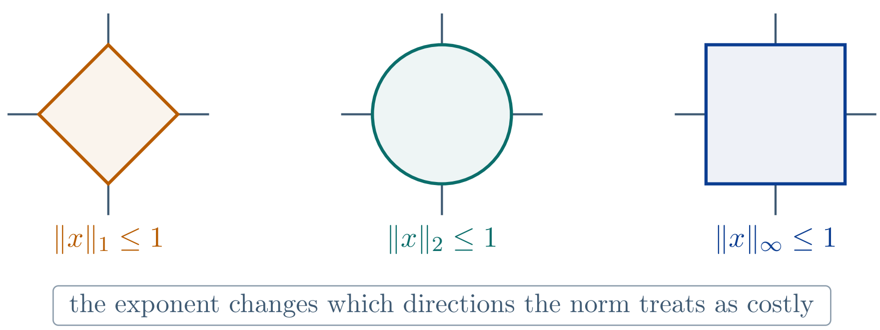


Figure 6

Proposition 9.4 Let $1 \leq p < \infty$ and let $E \in \mathcal{L}(\mathbb{R})$ with $\lambda(E) > 0$. Then the following hold:

1. $L^p(E)$ is a vector space.
2. $\|\cdot\|_p$ satisfies the following:
   (a) $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere
   (b) $\|cf\|_p = |c| \cdot \|f\|_p$ for every $c \in \mathbb{C}$.
   (c) $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Proof. For (1), if $f \in L^p(E)$ and $c \in \mathbb{C}$, then $cf \in L^p(E)$ because cf is measurable and

$$\|cf\|^p = \int_E |cf|^p = \int_E |c|^p|f|^p = |c|^p \int_E |f|^p < \infty.$$

If $f, g \in L^p(E)$, then $f + g$ is measurable and

$$\|f + g\|^p = \int_E |f + g|^p \leq \int_E 2^p(|f|^p + |g|^p) = 2^p \left[\int_E |f|^p + \int_E |g|^p\right] < \infty.$$

The inequality follows from the convexity of $x \mapsto x^p$, namely $((a + b)/2)^p \leq (a^p + b^p)/2$. Thus, $f + g \in L^p(E)$, proving (1).

For (2), (2b) is proven above. For (2a),

$$\|f\|_p = 0 \iff \int_E |f|^p = 0$$

If the integral is zero, let

$$A_n = \{x \in E \mid |f(x)|^p \geq 1/n\}.$$

Then $0 \geq \lambda(A_n)/n$, so every A_n is null. Since $\{f \neq 0\} = \bigcup_{n=1}^\infty A_n$, we have $f = 0$ almost everywhere. The converse is immediate.

We will handle (2c) in the next lecture after we have established Hölder's inequality. $\square$

Definition 9.5 Let $E \in \mathcal{L}(\mathbb{R})$ with $\lambda(E) > 0$ and let $f : E \to \mathbb{C}$. Define the **essential supremum**

$$\|f\|_\infty = \inf \left\{ \sup_{x \in A} |f(x)| : A \subseteq E,\ \lambda(E \setminus A) = 0 \right\}.$$

We define $L^\infty(E)$ as the set of equivalence classes of measurable functions $f : E \to \mathbb{C}$ such that $\|f\|_\infty < \infty$, where two functions are considered equivalent if they are equal almost everywhere.

Example 9.6 If $f = 0$ almost everywhere, then $\|f\|_\infty = 0$, and conversely. If E is a compact interval and f is continuous on E, then $\|f\|_\infty = \sup_E |f(x)|$.

Proposition 9.7 $L^\infty(E)$ is a vector space, and $\|\cdot\|_\infty$ is a norm on $L^\infty(E)$.

Proof. If $f, g \in L^\infty(E)$, choose sets N_f, N_g such that $\lambda(N_f) = \lambda(N_g) = 0$ and $|f(x)| \leq \|f\|_\infty + \varepsilon$ on $E \setminus N_f$ and $|g(x)| \leq \|g\|_\infty + \varepsilon$ on $E \setminus N_g$. Then on $E \setminus (N_f \cup N_g)$,

$$|f + g| \leq |f| + |g| \leq \|f\|_\infty + \|g\|_\infty + 2\varepsilon \qquad \text{and} \qquad |cf| \leq |c|(\|f\|_\infty + \varepsilon),$$

so $f + g, cf \in L^\infty(E)$. For the norm axioms, nonnegativity is immediate, $\|f\|_\infty = 0$ if and only if $f = 0$ almost everywhere (from the preceding example), $\|cf\|_\infty = |c|\|f\|_\infty$, and

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

from the pointwise bound above after removing a set of Lebesgue measure zero. Thus $\|\cdot\|_\infty$ is a norm. $\square$

Lecture 10 — Thursday, February 05, 2015

Hölder and Minkowski Inequalities

Theorem 10.1 (Hölder's inequality) Let p, q be conjugate indices (that is, $1/p + 1/q = 1$), $f \in L^p(E)$ and $g \in L^q(E)$. Then

$$\int_E |fg| \leq \|f\|_p \|g\|_q.$$

In particular, $fg \in L^1(E)$.

Proof. *Case: $p = 1, q = \infty$ (or vice versa).* Define $A = \{x \in E : |g(x)| > \|g\|_\infty\}$. We first show that $\lambda(A) = 0$. Let $A_n = \{x \in E : |g(x)| > \|g\|_\infty + 1/n\}$, so $A = \bigcup_{n=1}^{\infty} A_n$. It suffices to show that $\lambda(A_n) = 0$ for each n. Let $B \subseteq E$ satisfy $\lambda(E \setminus B) = 0$ and $\sup_{x \in B} |g(x)| \leq \|g\|_\infty + 1/n$. Then $A_n \subseteq E \setminus B$, so $\lambda(A_n) = 0$. Thus $\lambda(A) = 0$. Therefore, for $f \in L^1(E)$ and $g \in L^\infty(E)$,

$$\int_E |fg| = \int_{E \setminus A} |fg| \leq \|g\|_\infty \int_E |f| = \|g\|_\infty \|f\|_1.$$

Case: $1 < p, q < \infty$. If $\|f\|_p = 0$, then $f = 0$ almost everywhere, so $fg = 0$ almost everywhere and hence $\int |fg| = 0$, so we are done. Similarly, if $\|g\|_q = 0$, then we are done. Therefore, we may assume that $\|f\|_p \neq 0, \|g\|_q \neq 0$. It is enough to show

$$\int_E \frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q} \leq 1.$$

Let $a = |f|/\|f\|_p$ and $b = |g|/\|g\|_q$. Young's inequality gives

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Therefore

$$\int_E \frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q} \leq \int_E \left(\frac{1}{p} \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g|^q}{\|g\|_q^q} \right) = \frac{1}{p} \frac{\|f\|_p^p}{\|f\|_p^p} + \frac{1}{q} \frac{\|g\|_q^q}{\|g\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

$\square$

Theorem 10.2 (Minkowski's inequality) Let $1 \le p \le \infty$, $f, g \in L^p(E)$. Then

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

Proof. *Case:* $p = \infty$. We have $|f+g| \le |f| + |g|$ by the usual triangle inequality, and moreover $|f| \le \|f\|_\infty$ and $|g| \le \|g\|_\infty$ almost everywhere. Therefore, we have

$$|f+g| \le \|f\|_\infty + \|g\|_\infty$$

almost everywhere, and taking the essential supremum on both sides gives

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

Case: $p = 1$. The pointwise triangle inequality gives

$$\|f+g\|_1 = \int_E |f+g| \le \int_E |f| + \int_E |g| = \|f\|_1 + \|g\|_1.$$

Case: $1 < p < \infty$. It is enough to show that

$$\int_E |f+g|^p \le \|f+g\|_p^{p-1}(\|f\|_p + \|g\|_p).$$

We have

$$\int_E |f+g|^p = \int_E |f+g|^{p-1}|f+g| \le \int_E |f+g|^{p-1}|f| + \int_E |f+g|^{p-1}|g|.$$

Let q be such that $1/p + 1/q = 1$. Then

$$\|(f+g)^{p-1}\|_q = \left(\int_E |f+g|^{(p-1)q}\right)^{1/q} = \left(\int_E |f+g|^p\right)^{1/q} = \|f+g\|_p^{p/q} < \infty$$

because $L^p(E)$ is a vector space. Then by Hölder's inequality,

$$\int_E |f+g|^{p-1}|f| \le \|f+g\|_p^{p/q}\|f\|_p = \|f+g\|_p^{p-1}\|f\|_p.$$

Similarly,

$$\int_E |f+g|^{p-1}|g| \le \|f+g\|_p^{p-1}\|g\|_p.$$

Therefore,

$$\|f+g\|_p^p \leq \|f+g\|_p^{p-1}(\|f\|_p + \|g\|_p),$$

which implies that $\|f+g\|_p \leq \|f\|_p + \|g\|_p$. (If $\|f+g\|_p = 0$, the inequality is immediate.) $\square$

Corollary 10.3 For $1 \leq p \leq \infty$, $L^p(E)$ is a normed vector space.

Definition 10.4 A complete normed vector space is called a **Banach space**.

Theorem 10.5 Let $1 \leq p \leq \infty$. Then $L^p(E)$ is a Banach space.

Proof. We want to show that $L^p(E)$ is complete. Our strategy is to take a Cauchy sequence $\{f_n\}$ and find a subsequence converging to some $f \in L^p(E)$. The Cauchy property will then force the entire sequence to converge to f. Let's get into it!

Case: $1 \leq p < \infty$. Let $\{f_n\}$ be a Cauchy sequence in $L^p(E)$. This means that $\|f_n - f_m\|_p \to 0$ as $m, n \to \infty$. Let $\{f_{n_k}\}$ be a subsequence such that $\|f_{n_{k+1}} - f_{n_k}\|_p < 1/2^k$. Let

$$g_k = \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}|$$

and

$$g = \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|.$$

For every finite k, Minkowski's inequality gives

$$\|g_k\|_p \leq \sum_{i=1}^{k} \|f_{n_{i+1}} - f_{n_i}\|_p \leq \sum_{i=1}^{\infty} \frac{1}{2^i} = 1.$$

Now $g_k \nearrow g$ pointwise. By the monotone convergence theorem,

$$\int_E |g|^p = \lim_{k\to\infty} \int_E |g_k|^p \leq 1,$$

so $g \in L^p(E)$. In particular, g is finite outside a set A with $\lambda(A) = 0$. Hence

$$f = f_{n_1} + \sum_{i=1}^{\infty}(f_{n_{i+1}} - f_{n_i})$$

is well-defined on A^c as the pointwise limit of

$$h_k = f_{n_1} + \sum_{i=1}^{k}(f_{n_{i+1}} - f_{n_i}).$$

Set $f(x) = 0$ on A. Then f is measurable and $f_{n_k} \to f$ almost everywhere. The inequality $|f| \le |f_{n_1}| + g$ and Minkowski's inequality show that $f \in L^p(E)$. Moreover, $|f_{n_k} - f|^p \to 0$ almost everywhere and $|f_{n_k} - f|^p \le |g|^p$. The dominated convergence theorem therefore gives

$$\lim_{k\to\infty} \int_E |f_{n_k} - f|^p = 0.$$

Thus $\|f_{n_k} - f\|_p \to 0$. Since $\{f_n\}$ is Cauchy, the triangle inequality now gives $\|f_n - f\|_p \to 0$.

Lecture 11 — Tuesday, February 10, 2015

Proof continued. *Case:* $p = \infty$. Let $\{f_n\}_{n=1}^{\infty}$ be a $\|\cdot\|_\infty$-Cauchy sequence in $L^\infty(E)$. For $k \in \mathbb{N}$, set

$$A_k = \{x \in E : |f_k(x)| > \|f_k\|_\infty\}.$$

We have already proved $\lambda(A_k) = 0$. For $m, n \in \mathbb{N}$, set

$$B_{n,m} = \{x \in E : |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\},$$

and $\lambda(B_{n,m}) = 0$. Let

$$A = \left(\bigcup_k A_k\right) \cup \left(\bigcup_{n,m} B_{n,m}\right).$$

Then $\lambda(A) = 0$. For all $x \in A^c$, we have $|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty$, so $\{f_n\}$ is uniformly Cauchy on A^c. Hence $f_n \to f$ uniformly on A^c. Define $f(x) = 0$ for $x \in A$. Then

$$\|f_n - f\|_\infty \le \sup_{x \in A^c} |f_n(x) - f(x)| \longrightarrow 0.$$

Also, $f \in L^\infty(E)$ because f is measurable and there exists N such that $\|f_N - f\|_\infty < 1$. Therefore,

$$\|f\|_\infty \le \|f - f_N\|_\infty + \|f_N\|_\infty < 1 + \|f_N\|_\infty < \infty.$$

$\square$

Remark 11.1 We have proved the following:

1. If $f \in L^\infty(E)$, then $\lambda(\{x \in E : |f(x)| > \|f\|_\infty\}) = 0$.
2. If $f_n \xrightarrow{\|\cdot\|_p} f$, then there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \longrightarrow f$ pointwise almost everywhere.
3. $f_n \xrightarrow{\|\cdot\|_\infty} f$ if and only if there exists A such that $\lambda(A) = 0$ and $f_n \to f$ uniformly on A^c (we really only proved one direction, but the other is trivial).

Warning Norm convergence does not force pointwise convergence of the chosen representatives. For example, let $f_n(x) = x^n$ on $[0,1]$. For every $1 \le p < \infty$,

$$\|f_n\|_p = (np+1)^{-1/p} \longrightarrow 0,$$

but $f_n(1) = 1$ for every n, so this representative does not converge pointwise to the zero function.

Let us record some facts about L^p spaces.

Proposition 11.2 Let $[a,b]$ be a compact interval with $a < b$, and let $1 \le p_1 < p_2 < \infty$. Then

$$L^\infty([a,b]) \subsetneq L^{p_2}([a,b]) \subsetneq L^{p_1}([a,b]).$$

Proof. Let $1 \le p_1 < p_2 < \infty$ and $f \in L^{p_2}([a,b])$. By Hölder's inequality with exponents

$$r = \frac{p_2}{p_1} \qquad \text{and} \qquad r' = \frac{p_2}{p_2 - p_1},$$

we have

$$\int_a^b |f|^{p_1} = \int_a^b |f|^{p_1} \cdot 1 \le \left(\int_a^b |f|^{p_2}\right)^{p_1/p_2} (b-a)^{1-p_1/p_2} < \infty.$$

Hence $f \in L^{p_1}([a,b])$, so $L^{p_2}([a,b]) \subseteq L^{p_1}([a,b])$. Also, if $f \in L^\infty([a,b])$, then

$$\|f\|_{p_1}^{p_1} = \int_a^b |f|^{p_1} \leq (b-a)\|f\|_\infty^{p_1},$$

so $L^\infty([a,b]) \subseteq L^{p_1}([a,b])$ for every finite p_1.

To prove strictness, fix $1 \leq p_1 < p_2 < \infty$ and define

$$h(x) := (x-a)^{-1/p_2} \mathbb{1}_{(a,(a+b)/2]}(x).$$

Then

$$\int_a^b |h|^{p_2}\,\mathrm{d}x = \int_a^{(a+b)/2} (x-a)^{-1}\,\mathrm{d}x = \infty,$$

so $h \notin L^{p_2}([a,b])$, while

$$\int_a^b |h|^{p_1}\,\mathrm{d}x = \int_a^{(a+b)/2} (x-a)^{-p_1/p_2}\,\mathrm{d}x < \infty$$

because $p_1/p_2 < 1$. Thus $h \in L^{p_1}([a,b]) \setminus L^{p_2}([a,b])$.

Finally, for each finite $p \geq 1$, the function

$$u_p(x) := (x-a)^{-1/(2p)} \mathbb{1}_{(a,(a+b)/2]}(x)$$

belongs to $L^p([a,b])$ since

$$\int_a^{(a+b)/2} (x-a)^{-1/2}\,\mathrm{d}x < \infty,$$

but $u_p \notin L^\infty([a,b])$. Therefore all inclusions in the chain are strict.

$\square$

Proposition 11.3 Let $[a,b]$ be a compact interval. For $1 \leq p < \infty$, let $C([a,b])$ be the set of continuous functions on $[a,b]$. Then the following hold:

1. $C([a,b])$ is $\|\cdot\|_p$-dense in $L^p([a,b])$. In other words, for every $f \in L^p([a,b])$ and every $\varepsilon > 0$, there exists $h \in C([a,b])$ such that $\|f-h\|_p < \varepsilon$.
2. $C([a,b]) \subsetneq L^\infty([a,b])$, and $C([a,b])$ is $\|\cdot\|_\infty$-closed in $L^\infty([a,b])$ (in particular, $C([a,b])$ is not $\|\cdot\|_\infty$-dense in $L^\infty([a,b])$).

Proof. For (1), let $f \in L^p([a,b])$ and $\varepsilon > 0$. We first treat real-valued f; the complex case follows by applying the same argument to the real and imaginary parts.

Step 1: reduction to a bounded function. For $M > 0$, define

$$f_M(x) := \max\{-M, \min\{f(x), M\}\}.$$

Then $|f_M| \leq |f|$ and $|f - f_M|^p \to 0$ pointwise as $M \to \infty$. Since $|f - f_M|^p \leq |f|^p \in L^1([a,b])$, the dominated convergence theorem gives

$$\|f - f_M\|_p \to 0.$$

Choose M so that $\|f - f_M\|_p < \varepsilon/3$.

Step 2: approximation of f_M by simple functions. Choose $n \in \mathbb{N}$ and define

$$s_n(x) := 2^{-n} \lfloor 2^n f_M(x) \rfloor .$$

Since f_M is bounded, s_n takes only finitely many values, so s_n is simple. Moreover,

$$|f_M(x) - s_n(x)| \leq 2^{-n} \quad \text{for all } x \in [a,b].$$

Hence

$$\|f_M - s_n\|_p \leq (b-a)^{1/p} 2^{-n}.$$

Choose n large enough so that $(b-a)^{1/p}2^{-n} < \varepsilon/3$, and set $s := s_n$. Write

$$s = \sum_{j=1}^{m} a_j \mathbb{1}_{E_j}$$

for measurable sets E_j.

Step 3: approximation of each indicator by a continuous cutoff. Fix

$$\delta := \left(\frac{\varepsilon}{3 \sum_{j=1}^{m} (|a_j| + 1)} \right)^p .$$

By Corollary 4.4 (applied to $E_j \subseteq [a,b]$), for each j there exist a compact set K_j and an open set U_j with

$$K_j \subseteq E_j \subseteq U_j \qquad \text{and} \qquad \lambda(U_j \setminus K_j) < \delta.$$

Since $[a,b]$ is normal, there exists $\eta_j \in C([a,b])$ such that $0 \leq \eta_j \leq 1$, $\eta_j = 1$ on K_j, and $\eta_j = 0$

on $[a,b] \setminus U_j$. Then

$$|\eta_j - \mathbb{1}_{E_j}| \leq \mathbb{1}_{U_j \setminus K_j},$$

so

$$\|\eta_j - \mathbb{1}_{E_j}\|_p^p \leq \lambda(U_j \setminus K_j) < \delta.$$

Define

$$h := \sum_{j=1}^m a_j \eta_j \in C([a,b]).$$

Using the triangle inequality,

$$\|h - s\|_p \leq \sum_{j=1}^m |a_j| \, \|\eta_j - \mathbb{1}_{E_j}\|_p \leq \delta^{1/p} \sum_{j=1}^m |a_j| < \varepsilon/3.$$

Therefore,

$$\|f - h\|_p \leq \|f - f_M\|_p + \|f_M - s\|_p + \|s - h\|_p < \varepsilon.$$

Hence $C([a,b])$ is dense in $L^p([a,b])$.

For (2), first note that every continuous function on a compact interval is bounded, so $C([a,b]) \subseteq L^\infty([a,b])$.

Now let $\{f_n\} \subseteq C([a,b])$ and assume $f_n \xrightarrow{\|\cdot\|_\infty} f$ in $L^\infty([a,b])$. By Remark 11.1(3), there is a set $A \subseteq [a,b]$ with $\lambda([a,b] \setminus A) = 0$ such that $f_n \to f$ uniformly on A. Define $g = f$ on A. For $x \in [a,b] \setminus A$, choose a sequence $\{x_m\} \subseteq A$ with $x_m \to x$, and define

$$g(x) := \lim_{m \to \infty} g(x_m).$$

This is well-defined: if $x_m, y_m \in A$ both converge to x, then uniform convergence on A and uniform continuity of each continuous approximant imply

$$\lim_{m \to \infty} g(x_m) = \lim_{m \to \infty} g(y_m).$$

Hence g is continuous on $[a,b]$, and $g = f$ almost everywhere on $[a,b]$. Therefore the L^∞ class of f belongs to $C([a,b])$, so $C([a,b])$ is $\|\cdot\|_\infty$-closed in $L^\infty([a,b])$.

Finally, the characteristic function $\mathbb{1}_{[a,(a+b)/2]}$ belongs to $L^\infty([a,b])$ but does not agree almost everywhere with any continuous function. So $C([a,b]) \subsetneq L^\infty([a,b])$, and in particular $C([a,b])$ is not $\|\cdot\|_\infty$-dense in $L^\infty([a,b])$.

□

Proposition 11.4 Let $[a, b]$ be a compact interval. For $1 \leq p < \infty$, $L^p([a, b])$ is separable.

Proof. By Proposition 11.3(1), $C([a, b])$ is dense in $L^p([a, b])$. Next, by the Stone–Weierstrass theorem, the space $\mathbb{C}[x]$ of polynomials in x with complex coefficients is $\| \cdot \|_\infty$-dense in $C([a, b])$. Indeed, the real version gives the complex case by writing $f \in C([a, b])$ as $f = f_1 + if_2$ with f_1, f_2 real-valued and continuous:

$$\|f - (p_1 + ip_2)\|_\infty = \|(f_1 - p_1) + i(f_2 - p_2)\|_\infty \leq \|f_1 - p_1\|_\infty + \|f_2 - p_2\|_\infty < \varepsilon$$

for some $p_1, p_2 \in \mathbb{R}[x]$ such that $\|f_1 - p_1\|_\infty < \varepsilon/2$ and $\|f_2 - p_2\|_\infty < \varepsilon/2$.

Next, let $\mathbb{Q}[x]$ be the space of polynomials on $[a, b]$ with rational coefficients. Then $\mathbb{Q}[x]$ is $\| \cdot \|_\infty$-dense in $\mathbb{R}[x]$. To see this, let $p(x) = a_n x^n + \cdots + a_0$, where $a_n, \ldots, a_0 \in \mathbb{R}$. Choose $r_n, \ldots, r_0 \in \mathbb{Q}$ such that $|a_i - r_i| < \varepsilon$, and set $Q(x) = r_n x^n + \cdots + r_0$. If $c = \max(|a|, |b|)$, then

$$\|p - Q\|_\infty \leq |a_n - r_n| c^n + \cdots + |a_0 - r_0|.$$

Letting the rational approximations improve proves the claim.

Finally, the set of polynomials with rational real and imaginary coefficients, $\mathbb{Q}[x] + i\mathbb{Q}[x]$, are $\| \cdot \|_\infty$-dense in $\mathbb{C}[x]$. To see this, let $p(x) = a_n x^n + \cdots + a_0$, where $a_n, \ldots, a_0 \in \mathbb{C}$. Write $a_i = r_i + is_i$ with $r_i, s_i \in \mathbb{R}$. Choose $r'_n, \ldots, r'_0, s'_n, \ldots, s'_0 \in \mathbb{Q}$ such that $|r_i - r'_i| < \varepsilon$ and $|s_i - s'_i| < \varepsilon$, and set

$$q(x) = (r'_n + is'_n)x^n + \cdots + (r'_0 + is'_0).$$

Then

$$\begin{aligned}
\|p - q\|_\infty &\leq |a_n - (r'_n + is'_n)| c^n + \cdots + |a_0 - (r'_0 + is'_0)| \\
&= |(r_n - r'_n) + i(s_n - s'_n)| c^n + \cdots + |(r_0 - r'_0) + i(s_0 - s'_0)| \\
&\leq \sqrt{(r_n - r'_n)^2 + (s_n - s'_n)^2} c^n + \cdots + \sqrt{(r_0 - r'_0)^2 + (s_0 - s'_0)^2} \\
&< \varepsilon\sqrt{2}(c^n + \cdots + 1).
\end{aligned}$$

Therefore $\mathbb{Q}[x] + i\mathbb{Q}[x]$ is $\| \cdot \|_p$-dense in $C([a, b])$, and hence $\| \cdot \|_p$-dense in $L^p([a, b])$. The set $\mathbb{Q}[x] + i\mathbb{Q}[x]$ is countable, so $L^p([a, b])$ is separable.

□

Lecture 12 — Thursday, February 12, 2015

Theorem 12.1 $L^\infty([a,b])$ is not separable.

Proof. Assume $a=0, b=1$. Let $\alpha = (\alpha_i)_{i=1}^\infty \in \{0,1\}^\mathbb{N}$ be a binary sequence and let

$$f_\alpha = \sum_{n=1}^\infty \alpha_n \mathbb{1}_{[\frac{1}{n+1},\frac{1}{n}]}.$$

If $\alpha \neq \beta$ are distinct binary sequences, then there exists $n_0 \in \mathbb{N}$ such that $\alpha_{n_0} \neq \beta_{n_0}$, which implies $\|f_\alpha - f_\beta\|_\infty = 1$. Suppose for a contradiction that $L^\infty([0,1])$ is separable. Let $\{u_n\}_{n=1}^\infty$ be a countable dense subset of $L^\infty([0,1])$. Since $\{u_n\}$ is dense in $L^\infty([0,1])$, for all $\alpha \in \{0,1\}^\mathbb{N}$, there exists $n_\alpha \in \mathbb{N}$ such that $\|f_\alpha - u_{n_\alpha}\|_\infty < 1/2$. Now, if $\alpha \neq \beta$, then $n_\alpha \neq n_\beta$. Then the map $\{0,1\}^\mathbb{N} \to \mathbb{N}$ defined by $\alpha \mapsto n_\alpha$ is injective. Then $|\{0,1\}^\mathbb{N}| \leq |\mathbb{N}|$, a contradiction. $\square$

Inner Products and Hilbert Spaces

The space L^2 is special among the spaces L^p: its norm comes from an inner product, and this additional structure gives projection, orthogonality, and Fourier analyses.

Definition 12.2 Let X be a $\mathbb{C}$-vector space (or $\mathbb{R}$-vector space). A map $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ is called an **inner product** if it satisfies the following:

1. $\langle v,v\rangle \geq 0$ (**nonnegativity**).
2. $\langle v,v\rangle = 0$ if and only if $v=0$.
3. $\langle v,w\rangle = \overline{\langle w,v\rangle}$ (**conjugate symmetry**).
4. $\langle \alpha v_1 + v_2, w\rangle = \alpha\langle v_1,w\rangle + \langle v_2,w\rangle$.

We say that $\langle\cdot,\cdot\rangle$ is **linear in the first coordinate** and **antilinear in the second coordinate**, because $\langle v, \alpha w_1 + w_2\rangle = \overline{\alpha}\langle v,w_1\rangle + \langle v,w_2\rangle$.

Example 12.3 In $\mathbb{C}^2$, define $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1\overline{y_1} + x_2\overline{y_2}$. This is an inner product on $\mathbb{C}^2$.

Proposition 12.4 Let X be a vector space with an inner product. The formula

$$\|v\| = \sqrt{\langle v, v\rangle}$$

defines a nonnegative, definite, and homogeneous function on X.

Proof. By nonnegativity, $\|v\|$ is a nonnegative real number, and definiteness gives $\|v\| = 0$ if and only if $v = 0$. If α is a scalar, then

$$\|\alpha v\|^2 = \langle \alpha v, \alpha v\rangle = |\alpha|^2\langle v, v\rangle = |\alpha|^2\|v\|^2,$$

so $\|\alpha v\| = |\alpha|\|v\|$. □

Proposition 12.5 (Cauchy–Schwarz inequality) Let $v, w \in X$, where X is a vector space with an inner product $\langle \cdot, \cdot \rangle$. Then

$$|\langle v, w\rangle| \leq \|v\|\|w\|.$$

Equality holds if and only if v and w are linearly dependent.

Proof. If $w = 0$, the inequality is immediate. Otherwise, set

$$\alpha = \frac{\langle v, w\rangle}{\|w\|^2}.$$

Then

$$0 \leq \|v - \alpha w\|^2 = \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2},$$

which proves the inequality. Equality holds precisely when $\|v - \alpha w\| = 0$, or equivalently $v = \alpha w$. Together with the case where one vector is zero, this is exactly linear dependence. Conversely, direct substitution shows that linearly dependent vectors give equality. □

Proposition 12.6 Let X be a vector space with inner product $\langle \cdot, \cdot \rangle$. Then $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ is indeed a norm on X.

Proof. Nonnegativity and definiteness follow from the inner product axioms:

$$\|x\| = \sqrt{\langle x, x\rangle} \geq 0 \qquad \text{and} \qquad \|x\| = 0 \iff \langle x, x\rangle = 0 \iff x = 0.$$

For homogeneity,

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = |\alpha|^2 \langle x, x\rangle = |\alpha|^2 \|x\|^2,$$

so $\|\alpha x\| = |\alpha|\,\|x\|$. Triangle inequality follows from Cauchy–Schwarz:

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + 2\,\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Hence $\|x+y\| \leq \|x\| + \|y\|$. □

Definition 12.7 Let X be a vector space with inner product $\langle \cdot, \cdot \rangle$ and $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ as before. If X is complete with respect to $\| \cdot \|$, we call X a **Hilbert space**.

Example 12.8 Let E be a Lebesgue measurable subset of $\mathbb{R}$ with $\lambda(E) > 0$. Then $L^2(E)$ is a Hilbert space with inner product

$$\langle f, g\rangle = \int_E f\overline{g}$$

for $f, g \in L^2(E)$. Note that the norm induced by this inner product is the L^2 norm:

$$\sqrt{\langle f, f\rangle} = \left(\int_E |f|^2\right)^{1/2} = \|f\|_2.$$

Fact 12.9 For a positive-measure Lebesgue set E, the space $L^p(E)$ is a Hilbert space only when $p = 2$.

Proof. Every norm induced by an inner product satisfies the parallelogram identity

$$\|f+g\|^2 + \|f-g\|^2 = 2\|f\|^2 + 2\|g\|^2.$$

Because Lebesgue measure is nonatomic, we may choose disjoint measurable sets $A, B \subseteq E$ of the same finite positive measure m. Set $f = \mathbb{1}_A$ and $g = \mathbb{1}_B$. For $1 \le p < \infty$, the parallelogram identity would give

$$2(2m)^{2/p} = 4m^{2/p},$$

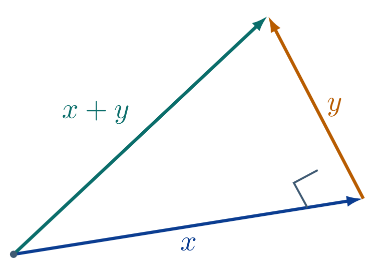
which forces $p = 2$. For $p = \infty$, its two sides are 2 and 4, respectively. Thus no $p \neq 2$ can arise from an inner product, while the preceding example shows that $L^2(E)$ does. $\square$

Definition 12.10 Two vectors v, w in a Hilbert space H are called **orthogonal** if $\langle v, w\rangle = 0$; we write $v \perp w$. A set $S \subseteq H$ is called an **orthogonal set** if for every distinct pair $v, w \in S$, $v \perp w$. A set $S \subseteq H$ is called **orthonormal** if S is orthogonal and $\|v\| = 1$ for every $v \in S$.

Theorem 12.11 (Pythagorean theorem) Let $x, y \in \mathcal{H}$ with $x \perp y$. Then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. More generally, if $\{x_n\}_{n=1}^N$ is an orthogonal set, then

$$\left\|\sum_{n=1}^N x_n\right\|^2 = \sum_{n=1}^N \|x_n\|^2.$$

The familiar right triangle survives in any Hilbert space, as summarized in Figure 7; only the algebraic notions of length and perpendicularity are needed.



$x \perp y$ makes the squared lengths add

FIGURE 7

Proof. Observe that if $x \perp y$, then $\langle x, y\rangle = 0$ and $\langle y, x\rangle = 0$. Therefore,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.$$

The general case follows by induction on N. ☐

Proposition 12.12 If $\{x_n\}_{n=1}^{\infty}$ is an orthogonal set in a Hilbert space $\mathcal{H}$, then

$$\sum_{n=1}^{\infty} x_n$$

is convergent if and only if

$$\sum_{n=1}^{\infty} \|x_n\|^2 < \infty.$$

Proof. *Forward implication.* Assume that

$$\sum_{n=1}^{\infty} x_n$$

converges in $\mathcal{H}$. Let

$$S_n := \sum_{i=1}^{n} x_i.$$

Then $\{S_n\}$ is Cauchy in $\mathcal{H}$. For $n > m$,

$$\|S_n - S_m\|^2 = \left\| \sum_{i=m+1}^{n} x_i \right\|^2 = \sum_{i=m+1}^{n} \|x_i\|^2,$$

where the last equality is the finite Pythagorean theorem for orthogonal families. Since $\|S_n - S_m\| \to 0$ as $m, n \to \infty$, the tails $\sum_{i=m+1}^{n} \|x_i\|^2$ tend to 0. Hence the scalar series

$$\sum_{i=1}^{\infty} \|x_i\|^2$$

is Cauchy in $\mathbb{R}$, and therefore convergent.

Reverse implication. Assume

$$\sum_{i=1}^{\infty} \|x_i\|^2 < \infty.$$

With the same partial sums S_n, for $n > m$ we again have

$$\|S_n - S_m\|^2 = \sum_{i=m+1}^{n} \|x_i\|^2.$$

Because the numerical series converges, its tails tend to 0, so $\|S_n - S_m\| \to 0$. Thus $\{S_n\}$ is Cauchy in $\mathcal{H}$. Since $\mathcal{H}$ is complete, $S_n \to x$ for some $x \in \mathcal{H}$, so

$$\sum_{n=1}^{\infty} x_n$$

converges.

Finally, if $S_n \to x$, then by continuity of the norm,

$$\|x\|^2 = \lim_{n\to\infty} \|S_n\|^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \|x_i\|^2 = \sum_{i=1}^{\infty} \|x_i\|^2.$$

□

Definition 12.13 We define the space ℓ^2 to be the set of all sequences $\beta = (\beta_i)_{i=1}^{\infty}$ of complex numbers such that $\sum_{i=1}^{\infty} |\beta_i|^2 < \infty$. In other words,

$$\ell^2 = \{(\beta_i)_{i=1}^{\infty} : \beta_i \in \mathbb{C} \text{ and } \|\beta\|_2^2 = \sum_{i=1}^{\infty} |\beta_i|^2 < \infty\}.$$

Proposition 12.14 Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal set. Let $\beta \in \ell^2$. Then there exists $x \in \mathcal{H}$ such that $\|x\| = \|\beta\|_2$ and $\langle x, e_i \rangle = \beta_i$.

Proof. Choose

$$x = \sum_{i=1}^{\infty} \beta_i e_i,$$

which is convergent because

$$\sum_{i=1}^{\infty} |\beta_i|^2 < \infty$$

by Proposition 12.12 and

$$\|x\|^2 = \sum_{i=1}^{\infty} |\beta_i|^2 = \|\beta\|_2^2.$$

For each fixed j, continuity of the inner product gives

$$\langle x, e_j \rangle = \lim_{N\to\infty} \left\langle \sum_{i=1}^{N} \beta_i e_i, e_j \right\rangle = \beta_j.$$

□

Lecture 13 — Thursday, February 19, 2015

Orthogonal Series and Bessel's Inequality

Remark 13.1 There are two useful relationships between Riemann and Lebesgue integration. First, suppose that $[a, b]$ is a compact interval and that f is a bounded function on $[a, b]$. If f is Riemann integrable, then $f \in L^1([a, b])$ and

$$\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f.$$

Thus the Riemann and Lebesgue integrals agree in this setting. Second, if f is nonnegative and Riemann integrable on $\mathbb{R}$, possibly as an improper Riemann integral, then $f \in L^1(\mathbb{R})$ and the Lebesgue integral of f agrees with its Riemann integral.

Recall that

$$\ell^2 = \{(a_n)_{n=1}^{\infty} : a_n \in \mathbb{C}, \|a\|_2 < \infty\}$$

and let $\{x_n\}_{n=1}^{\infty} \subseteq \mathcal{H}$ be an orthogonal sequence in a Hilbert space $\mathcal{H}$. The orthogonal series convergence criterion (Proposition 12.12) says that

$$\sum_{n=1}^{\infty} x_n$$

converges if and only if
$$\sum_{n=1}^{\infty} \|x_n\|^2 < \infty.$$
If $x = \sum_{n=1}^{\infty} x_n$, then
$$\|x\| = \left(\sum_{n=1}^{\infty} \|x_n\|^2 \right)^{1/2}.$$

Lemma 13.2 Let $\eta \in \mathcal{H}$ and define $T_\eta : \mathcal{H} \to \mathbb{C}$ by $\xi \mapsto \langle \xi, \eta \rangle$. Then T_η is linear and continuous.

Proof. Linearity follows from linearity of the inner product in the first variable. To prove continuity, suppose that $\xi_n \to \xi$. Then
$$|T_\eta(\xi_n) - T_\eta(\xi)| = |\langle \xi_n, \eta \rangle - \langle \xi, \eta \rangle| = |\langle \xi_n - \xi, \eta \rangle| \leq \|\xi_n - \xi\| \|\eta\| \to 0.$$
$\square$

Proposition 13.3 Let $\mathcal{H}$ be a Hilbert space, and let $\{e_i : i \in \mathbb{N}\}$ be an orthonormal subset of $\mathcal{H}$. If $(\beta_i) \in \ell^2$, then there exists $x \in \mathcal{H}$ such that $\langle x, e_i \rangle = \beta_i$ for all i and
$$\|x\| = \left(\sum_{i=1}^{\infty} |\beta_i|^2 \right)^{1/2} = \|(\beta_i)\|_2.$$

Proof. Consider the series
$$\sum_{i=1}^{\infty} \beta_i e_i.$$
Since $\{e_i\}_{i=1}^{\infty}$ is orthogonal, Proposition 12.12 shows that this series converges, because
$$\sum_{i=1}^{\infty} \|\beta_i e_i\|^2 = \sum_{i=1}^{\infty} |\beta_i|^2 \|e_i\|^2 = \sum_{i=1}^{\infty} |\beta_i|^2 = \|(\beta_i)\|_2^2 < \infty.$$
Set
$$x = \sum_{i=1}^{\infty} \beta_i e_i \in \mathcal{H}$$

so that $\|x\| = \|(\beta_i)\|_2$. Let $j \in \mathbb{N}$ be arbitrary. Since T_{e_j} is continuous by Lemma 13.2, we have

$$\langle x, e_j \rangle = \left\langle \lim_{N\to\infty} \sum_{i=1}^{N} \beta_i e_i, e_j \right\rangle = \lim_{N\to\infty} \left\langle \sum_{i=1}^{N} \beta_i e_i, e_j \right\rangle = \lim_{N\to\infty} \sum_{i=1}^{N} \beta_i \langle e_i, e_j \rangle = \beta_j.$$

□

Proposition 13.4 (Bessel's inequality) Let $\mathcal{H}$ be a Hilbert space, and let $\{e_i : i \in \mathbb{N}\}$ be an orthonormal subset of $\mathcal{H}$. If $x \in \mathcal{H}$, then $(\langle x, e_i \rangle)_{i\in\mathbb{N}} \in \ell^2$, and

$$\|(\langle x, e_i \rangle)_{i\in\mathbb{N}}\|_2 \leq \|x\|.$$

Proof. Let $x \in \mathcal{H}$ be arbitrary. It is enough to show that

$$\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 < \infty.$$

Fix $N \in \mathbb{N}$ and set

$$x_N = \sum_{i=1}^{N} \langle x, e_i \rangle e_i \in \mathcal{H}.$$

Then

$$\langle x - x_N, x_N \rangle = \langle x, x_N \rangle - \langle x_N, x_N \rangle = \langle x, \sum_{i=1}^{N} \langle x, e_i \rangle e_i \rangle - \sum_{i=1}^{N} |\langle x, e_i \rangle|^2 = 0.$$

Furthermore,

$$\|x\|^2 = \|(x - x_N) + x_N\|^2 = \|x - x_N\|^2 + \|x_N\|^2,$$

so $\|x_N\|^2 \leq \|x\|^2$. That is,

$$\sum_{i=1}^{N} |\langle x, e_i \rangle|^2 \leq \|x\|^2$$

for every N. Therefore,

$$\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2,$$

as desired. □

Orthonormal Bases and Parseval Expansions

Definition 13.5 Let $\mathcal{H}$ be a Hilbert space. An orthonormal set $S \subseteq \mathcal{H}$ is called a **basis** for $\mathcal{H}$ if $\overline{\text{span}(S)} = \mathcal{H}$ (i.e., $\text{span}(S)$ is dense in $\mathcal{H}$).

Example 13.6 The space ℓ^2 is a vector space: if $(a_n), (b_n) \in \ell^2$, then $(a_n) + (b_n) = (a_n + b_n) \in \ell^2$, and if $c \in \mathbb{C}$, then $c(a_n) = (ca_n) \in \ell^2$. It has the inner product

$$\langle (a_n), (b_n) \rangle = \sum_{n=1}^{\infty} a_n \overline{b_n}.$$

The norm $\|\cdot\|_2$ comes from this inner product:

$$\|(a_n)\|_2 = \sqrt{\langle (a_n), (a_n) \rangle} = \sqrt{\sum_{n=1}^{\infty} |a_n|^2},$$

so ℓ^2 is a Hilbert space. For $i \in \mathbb{N}$, define $e_i = (0, 0, \dots, 1, 0, \dots) \in \ell^2$, with the 1 in the ith coordinate. Then $\{e_i : i = 1, 2, \dots\}$ is an orthonormal basis for ℓ^2.

Fact 13.7 Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. There is no orthonormal subset S of $\mathcal{H}$ such that $\text{span}(S) = \mathcal{H}$.

Proof. Suppose such an orthonormal set exists. Let $\{x_n\}_{n=1}^{\infty} \subseteq S$, and let

$$x = \sum_{n=1}^{\infty} \frac{1}{2^n} x_n,$$

which exists because

$$\sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right)^2 \|x_n\|^2 = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right)^2 < \infty.$$

Then $x \in \mathcal{H}$, but $x \notin \text{span}(S)$. Indeed, if

$$x = \sum_{j=1}^{N} a_j y_j$$

for some $y_j \in S$ and $a_j \in \mathbb{C}$, then there exists i such that $x_i \notin \{y_1, \ldots, y_N\}$. For this i,

$$\langle x, x_i \rangle = \left\langle \sum_{n=1}^{\infty} \frac{1}{2^n} x_n, x_i \right\rangle = \frac{1}{2^i}$$

by continuity of $\langle \cdot, \cdot \rangle$. On the other hand,

$$\langle x, x_i \rangle = \left\langle \sum_{j=1}^{N} a_j y_j, x_i \right\rangle = 0,$$

which is a contradiction. □

Definition 13.8 An orthonormal set S in a Hilbert space $\mathcal{H}$ is called **complete** if it has the property that if $\langle x, e \rangle = 0$ for every $e \in S$, then $x = 0$.

Theorem 13.9 Let $S = \{e_i : i \in \mathbb{N}\}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. The following are equivalent:

1. S is a basis for $\mathcal{H}$.
2. S is complete.
3. $x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i$ for every $x \in \mathcal{H}$.

Proof. (1) *implies* (2). Assume that S is a basis for $\mathcal{H}$. Let $x \in \mathcal{H}$ and suppose $\langle x, e_i \rangle = 0$ for every $i \in \mathbb{N}$. We prove that $x = 0$. Let $\varepsilon > 0$. Since $\operatorname{span}(S)$ is dense in $\mathcal{H}$, there exists a finite linear combination

$$\sum_{i=1}^{N} a_i e_i$$

such that

$$\left\| x - \sum_{i=1}^{N} a_i e_i \right\| < \varepsilon.$$

Then

$$\left\langle x, \sum_{i=1}^{N} a_i e_i \right\rangle = \sum_{i=1}^{N} \overline{a_i} \langle x, e_i \rangle = 0.$$

Thus

$$\langle x, x\rangle = \|x\|^2 = \left\langle x, x - \sum_{i=1}^{N} a_i e_i + \sum_{i=1}^{N} a_i e_i \right\rangle = \left| \left\langle x, x - \sum_{i=1}^{N} a_i e_i \right\rangle \right|$$

By the Cauchy–Schwarz inequality,

$$\left| \left\langle x, x - \sum_{i=1}^{N} a_i e_i \right\rangle \right| \leq \|x\| \left\| x - \sum_{i=1}^{N} a_i e_i \right\| \leq \varepsilon \|x\|.$$

Then either $\|x\| = 0$, in which case $x = 0$, or

$$\|x\| \leq \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\|x\| = 0$, and therefore $x = 0$. Thus (2) holds.

(2) *implies* (3). Assume that S is complete. Let $x \in \mathcal{H}$. We show that

$$x = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i.$$

Since $x \in \mathcal{H}$, the series

$$y = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i$$

is convergent by Bessel's inequality. Moreover,

$$\langle x - y, e_j\rangle = \langle x, e_j\rangle - \langle y, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle = 0.$$

By completeness, $x - y = 0$, so $x = y$. Hence (3) holds.

(3) *implies* (1). Assume that for every $x \in \mathcal{H}$,

$$x = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i.$$

It remains to show that $\operatorname{span}(S)$ is dense in $\mathcal{H}$. For every $x \in \mathcal{H}$, we have

$$x = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i.$$

Let

$$x_n = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

The asserted series identity means precisely that its partial sums x_n converge to x. Since each x_n belongs to $\operatorname{span}(S)$, this span is dense in $\mathcal{H}$, and S is a basis. $\square$

Lecture 14 — Tuesday, February 24, 2015

Fact 14.1 If $x_n \to x$ in a Banach space, then $\|x_n\| \to \|x\|$.

Proof. By (reverse) triangle inequality,

$$\big| \|x_n\| - \|x\| \big| \leq \|x_n - x\|.$$

If $x_n \to x$, then $\|x_n - x\| \to 0$, so $\|x_n\| \to \|x\|$. $\square$

Remark 14.2 Let $S = \{e_n : n \in \mathbb{N}\}$ be an orthonormal basis for $\mathcal{H}$. Let $x \in \mathcal{H}$. By Theorem 13.9, we have

$$x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i.$$

Then

$$\|x\|^2 = \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2$$

where the sum is understood as the limit of its partial sums. Thus

$$x_N = \sum_{i=1}^{N} \langle x, e_i \rangle e_i \to x$$

where $\|x_N\| \to \|x\|$. But

$$\|x_N\| = \sqrt{\sum_{i=1}^{N} |\langle x, e_i \rangle|^2}$$

and

$$\|x\| = \lim_{N\to\infty} \|x_N\| = \sqrt{\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2}.$$

We now prove that this expansion is unique. Suppose $(\alpha_i), (\beta_i) \in \ell^2$ and

$$\sum_{i=1}^{\infty} \alpha_i e_i = \sum_{i=1}^{\infty} \beta_i e_i,$$

so

$$\sum_{i=1}^{\infty} (\alpha_i - \beta_i) e_i = 0.$$

Fix $j \in \mathbb{N}$. For each $N \geq j$, define

$$u_N := \sum_{i=1}^{N} (\alpha_i - \beta_i) e_i.$$

Then $u_N \to 0$ in $\mathcal{H}$, and by orthonormality,

$$\langle u_N, e_j\rangle = \alpha_j - \beta_j \text{ for } N \geq j.$$

The map $T_{e_j}(v) := \langle v, e_j\rangle$ is continuous by Lemma 13.2, so

$$\alpha_j - \beta_j = \lim_{N\to\infty} \langle u_N, e_j\rangle = \left\langle \lim_{N\to\infty} u_N, e_j \right\rangle = \langle 0, e_j\rangle = 0.$$

Hence $\alpha_j = \beta_j$. Since j was arbitrary, $\alpha_i = \beta_i$ for every $i \in \mathbb{N}$. Therefore the expansion of x in the orthonormal basis S is unique.

Existence of Orthonormal Bases

Theorem 14.3

1. Every Hilbert space admits an orthonormal basis.
2. If $\mathcal{H}$ is a separable Hilbert space, then every orthonormal subset of $\mathcal{H}$ is countable.

Proof. For (1), we apply Zorn's lemma. Let $\mathcal{O} = \{\text{all orthonormal subsets of } \mathcal{H}\}$, which is nonempty because it contains the empty set. Partially order $\mathcal{O}$ by inclusion. If $\mathcal{C}$ is a chain in $\mathcal{O}$, then $S = \bigcup_{C \in \mathcal{C}} C$ is again an orthonormal set: any two of its elements already belong to a common member of the chain. Hence every chain in $\mathcal{O}$ has an upper bound in $\mathcal{O}$. By Zorn's lemma, there is a maximal element $S_0 \in \mathcal{O}$.

It remains to prove that S_0 is an orthonormal basis for $\mathcal{H}$. Since S_0 is orthonormal, it suffices to show that it is complete. Let $x \in \mathcal{H}$ and assume $\langle x, e \rangle = 0$ for every $e \in S_0$. If $x \neq 0$, then $S_0 \cup \{x/\|x\|\}$ is orthonormal, contradicting the maximality of S_0. Hence $x = 0$, so S_0 is complete and therefore is an orthonormal basis.

For (2), let S be an orthonormal subset of $\mathcal{H}$. Since $\mathcal{H}$ is separable, there is a countable dense subset D of $\mathcal{H}$. For each $e \in S$, there exists $d \in D$ such that $\|e - d\| < 1/2$. If $e, e' \in S$ are distinct, then the reverse triangle inequality gives

$$\|d - d'\| \geq \|e - e'\| - \|e - d\| - \|e' - d'\| > 1 - 1/2 - 1/2 = 0,$$

so $d \neq d'$. Thus the map $e \mapsto d$ is an injection from S to D, so S is countable. □

Corollary 14.4 If $\mathcal{H}$ is an infinite-dimensional separable Hilbert space, then $\mathcal{H}$ is isometrically isomorphic to ℓ^2.

Proof. By Theorem 14.3(1), $\mathcal{H}$ admits an orthonormal basis S. By Theorem 14.3(2), every orthonormal subset of a separable Hilbert space is countable, so we may write $S = \{e_n\}_{n=1}^{\infty}$. Define

$$U : \mathcal{H} \to \ell^2, \qquad U(x) := (\langle x, e_n \rangle)_{n=1}^{\infty}.$$

The map U is linear by linearity of the inner product in its first variable. Because $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis, Theorem 13.9 gives

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 = \|U(x)\|_2^2,$$

so U is an isometry.

To prove surjectivity, let $\beta = (\beta_n)_{n=1}^{\infty} \in \ell^2$. Since $\sum_{n=1}^{\infty} |\beta_n|^2 < \infty$ and $\{e_n\}$ is orthonormal, Proposition 12.12 implies that

$$x := \sum_{n=1}^{\infty} \beta_n e_n$$

converges in $\mathcal{H}$. For each fixed j, continuity of $v \mapsto \langle v, e_j \rangle$ gives

$$\langle x, e_j \rangle = \left\langle \sum_{n=1}^{\infty} \beta_n e_n, e_j \right\rangle = \beta_j.$$

Hence $U(x) = \beta$, so U is onto.

Therefore U is a linear surjective isometry from $\mathcal{H}$ onto ℓ^2; that is, an isometric isomorphism.

$\square$

Example 14.5 Let $\mathcal{H} = L^2([0,1])$, and for $n \in \mathbb{Z}$ let $\chi_n(x) = e^{2\pi i n x}$. Then $S = \{\chi_n : n \in \mathbb{Z}\}$ is orthonormal in $L^2([0,1])$. To see this, let $n, m \in \mathbb{Z}$ be arbitrary. Then

$$\langle \chi_n, \chi_m \rangle = \int_0^1 e^{2\pi i n x} e^{-2\pi i m x} \, \mathrm{d}x = \int_0^1 e^{2\pi i (n-m) x} \, \mathrm{d}x.$$

If $m = n$, then $\langle \chi_n, \chi_m \rangle = 1$. If $m \neq n$, then

$$\begin{aligned}
\langle \chi_n, \chi_m \rangle &= \int_0^1 \big(\cos(2\pi(n-m)x) + i \sin(2\pi(n-m)x)\big) \, \mathrm{d}x \\
&= \frac{\sin(2\pi(n-m)x)}{2\pi(n-m)} \bigg|_0^1 + i \, \frac{-\cos(2\pi(n-m)x)}{2\pi(n-m)} \bigg|_0^1 \\
&= 0.
\end{aligned}$$

We will show that $\{\chi_n : n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2([0,1])$. Once this is established, Theorem 13.9 will give, for every $f \in L^2([0,1])$,

$$f = \sum_{n \in \mathbb{Z}} \langle f, \chi_n \rangle \chi_n,$$

with convergence in $L^2([0,1])$. This means that $f = \lim_{N \to \infty} f_N$, where

$$f_N = \sum_{n=-N}^{N} \langle f, \chi_n \rangle \chi_n.$$

Then

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\langle f, \chi_n \rangle|^2 \tag{14.1}$$

(14.1) is called **Parseval's identity**.

Definition 14.6 For $f \in L^2([0,1])$, the nth **Fourier coefficient** of f is

$$\hat{f}(n) := \langle f, \chi_n \rangle = \int_0^1 f(x) e^{-2\pi i n x} \, \mathrm{d}x.$$

The series

$$\sum_{n \in \mathbb{Z}} \hat{f}(n) \chi_n$$

is called the **Fourier series** of f.

4. Fourier Analysis on the Circle

The Circle Group and Trigonometric Polynomials

Let $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ be the circle group. We identify $\mathbb{T}$ with $[0,1)$ (or $[0,1]$ with the endpoints 0 and 1 identified), via the map $\theta \mapsto e^{2\pi i \theta}$. With addition modulo 1, this identification makes $[0,1)$ into a compact group. Let $\mathrm{d}m$ be Lebesgue measure restricted to $[0,1)$; that is,

$$\int_{\mathbb{T}} f \, \mathrm{d}m = \int_0^1 f(x) \, \mathrm{d}x$$

for integrable functions on $\mathbb{T}$. We use $\mathbb{T}$ because functions on $\mathbb{T}$ correspond naturally to 1-periodic functions on $\mathbb{R}$.

Figure 8 makes the identification explicit: moving one unit along the parameter interval makes one full turn and returns to the same point of the circle.

Definition 14.7 For $1 \leq p < \infty$, define $L^p(\mathbb{T})$ to be the space of equivalence classes of measurable functions on $[0,1)$ such that

$$\|f\|_p = \left(\int_0^1 |f|^p \, \mathrm{d}x \right)^{1/p} < \infty.$$

Let $C(\mathbb{T})$ denote the space of continuous 1-periodic functions on $[0,1]$, and let $L^\infty(\mathbb{T})$ denote

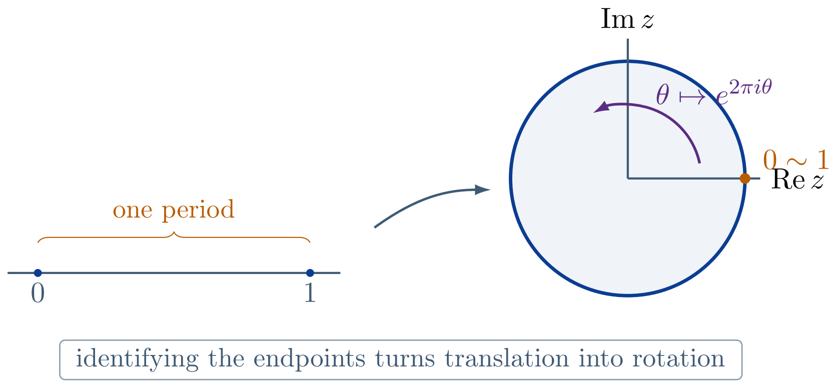


FIGURE 8

the space of equivalence classes of essentially bounded measurable functions on $[0, 1]$.

Theorem 14.8 Then $C(\mathbb{T}) \subsetneq L^\infty(\mathbb{T}) \subsetneq L^p(\mathbb{T})$ for $1 \leq p < \infty$.

Exercise 14.9 Prove that the space $C(\mathbb{T})$ is $\|\cdot\|_p$-dense in $L^p(\mathbb{T})$ for $1 \leq p < \infty$.

Recall that $L^q(\mathbb{T}) \subseteq L^p(\mathbb{T})$ when $p \leq q$.

Definition 14.10 Let $f \in L^1(\mathbb{T})$. Then define the nth **Fourier coefficient** of f to be

$$\hat{f}(n) := \int_0^1 f(x)e^{-2\pi inx}\,\mathrm{d}x.$$

The integral

$$\int_0^1 f(x)e^{-2\pi inx}\,\mathrm{d}x$$

is finite by Hölder's inequality, since $f \in L^1(\mathbb{T})$ and $\chi_{-n} = \overline{\chi_n} \in L^\infty(\mathbb{T})$. The formal series

$$S[f] \sim \sum_{n=-\infty}^{\infty} \hat{f}(n)\chi_n$$

is called the **Fourier series** of f.

Question 14.11 This raises natural questions. Does $S[f]$ converge to f in any sense? What if f is nice? What does "nice" even mean? What does anything mean? People living in the 1800s were deeply troubled by these questions.

Definition 14.12 A function of the form

$$\sum_{n=-N}^{N} a_n \chi_n$$

is called a **trigonometric polynomial**. $\text{Trig}(\mathbb{T}) = \{\text{all trigonometric polynomials on } \mathbb{T}\} = \text{span}(S)$, where $S = \{\chi_n : n \in \mathbb{Z}\}$.

Proposition 14.13 The trigonometric polynomials have the following properties:

1. $\text{Trig}(\mathbb{T}) \subseteq C(\mathbb{T})$.
2. $\text{Trig}(\mathbb{T})$ is an algebra: $\chi_n \cdot \chi_m = \chi_{n+m} \in \text{span}(S)$.
3. $\text{Trig}(\mathbb{T})$ is closed under conjugation: $\overline{\chi_n} = \chi_{-n}$.
4. $\text{Trig}(\mathbb{T})$ separates points.
5. $\text{Trig}(\mathbb{T})$ is $\|\cdot\|_\infty$-dense in $(C(\mathbb{T}), \|\cdot\|_\infty)$, and $\|\cdot\|_p$-dense in $L^p(\mathbb{T})$ for every $1 \le p < \infty$.

Proof. For (1), let $N \in \mathbb{N}$ and $a_{-N}, \ldots, a_N \in \mathbb{C}$. Then

$$\sum_{n=-N}^{N} a_n \chi_n(x) = \sum_{n=-N}^{N} a_n e^{2\pi i n x}$$

is continuous as a finite sum of continuous functions.

For (2), let $N, M \in \mathbb{N}$ and $a_{-N}, \ldots, a_N, b_{-M}, \ldots, b_M \in \mathbb{C}$. Then

$$\left(\sum_{n=-N}^{N} a_n \chi_n\right) \cdot \left(\sum_{m=-M}^{M} b_m \chi_m\right) = \sum_{n=-N}^{N} \sum_{m=-M}^{M} a_n b_m \chi_{n+m}$$

is a trigonometric polynomial.

For (3), let $N \in \mathbb{N}$ and $a_{-N}, \dots, a_N \in \mathbb{C}$. Then

$$\overline{\sum_{n=-N}^{N} a_n \chi_n} = \sum_{n=-N}^{N} \overline{a_n \chi_n} = \sum_{n=-N}^{N} \overline{a_n} \chi_{-n}$$

is a trigonometric polynomial.

For (4), let $x, y \in \mathbb{T}$ with $x \neq y$. Then $\chi_1(x) = e^{2\pi i x} \neq e^{2\pi i y} = \chi_1(y)$, so $\mathrm{Trig}(\mathbb{T})$ separates points.

For (5), the Stone–Weierstrass theorem implies that $\mathrm{Trig}(\mathbb{T})$ is $\|\cdot\|_\infty$-dense in $C(\mathbb{T})$. Since $\|\cdot\|_p \leq \|\cdot\|_\infty$ on $\mathbb{T}$, this also gives $\|\cdot\|_p$-density in $C(\mathbb{T})$. Combining that fact with the $\|\cdot\|_p$-density of $C(\mathbb{T})$ in $L^p(\mathbb{T})$ yields the claimed density in $L^p(\mathbb{T})$ for $1 \leq p < \infty$.

□

Corollary 14.14 S is an orthonormal basis for $L^2(\mathbb{T})$, where $S = \{\chi_n : n \in \mathbb{Z}\}$.

Proof. By Proposition 14.13, $\mathrm{Trig}(\mathbb{T}) = \mathrm{span}(S)$ is dense in $L^2(\mathbb{T})$. Since S is orthonormal, it follows that S is an orthonormal basis for $L^2(\mathbb{T})$. □

Corollary 14.15 (Parseval's identity) For all $f \in L^2(\mathbb{T})$,

$$f = \sum_{n=-\infty}^{\infty} \hat{f}(n)\chi_n$$

with $\|\cdot\|_2$-convergence. Consequently,

$$\|f\|_2 = \sqrt{\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2}$$

Proof. By Corollary 14.14, S is an orthonormal basis for $L^2(\mathbb{T})$. By Theorem 13.9, for every $f \in L^2(\mathbb{T})$,

$$f = \sum_{n=-\infty}^{\infty} \langle f, \chi_n\rangle \chi_n = \sum_{n=-\infty}^{\infty} \hat{f}(n)\chi_n$$

with $\|\cdot\|_2$-convergence. Moreover,

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\langle f, \chi_n \rangle|^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2.$$

□

Lecture 15 — Thursday, February 26, 2015

Definition 15.1 We define

$$\ell^2(\mathbb{Z}) = \{(a_n)_{n\in\mathbb{Z}} : \sum_{n\in\mathbb{Z}} |a_n|^2 < \infty\},$$

with the inner product

$$\langle (a_n), (b_n) \rangle = \sum_{n\in\mathbb{Z}} a_n \overline{b_n}$$

and the norm

$$\|(a_n)\|_2 = \sqrt{\sum_{n\in\mathbb{Z}} |a_n|^2}.$$

The **Fourier transform** $\mathcal{F} : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ is given by $f \mapsto (\hat{f}(n))_{n\in\mathbb{Z}}$.

Proposition 15.2 The Fourier transform is an isometric isomorphism of Hilbert spaces:

1. $\mathcal{F}(f+g) = \mathcal{F}(f) + \mathcal{F}(g)$.
2. $\mathcal{F}(cf) = c\mathcal{F}(f)$.
3. $\mathcal{F}$ is injective.
4. $\mathcal{F}$ is surjective.
5. $\mathcal{F}$ is an isometry.

Proof. For (1), let $n \in \mathbb{Z}$. Then

$$\begin{aligned}\widehat{f+g}(n) &= \int_0^1 (f(x)+g(x))e^{-2\pi inx}\,\mathrm{d}x \\ &= \int_0^1 f(x)e^{-2\pi inx}\,\mathrm{d}x + \int_0^1 g(x)e^{-2\pi inx}\,\mathrm{d}x \\ &= \hat{f}(n) + \hat{g}(n).\end{aligned}$$

For (2), the result follows from linearity of the integral.

For (3), suppose that $f \in L^2(\mathbb{T})$ and $\hat{f}(n) = 0$ for every n. Then

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2 = 0,$$

so $f = 0$ in $L^2(\mathbb{T})$.

For (4), take a sequence $(a_n)_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$. The general Hilbert space theory gives

$$f = \sum_{n=-\infty}^{\infty} a_n \chi_n$$

with convergence in $L^2(\mathbb{T})$. Continuity of the inner product gives

$$\hat{f}(j) = \langle f, \chi_j \rangle = a_j$$

for every $j \in \mathbb{Z}$, so $\mathcal{F}(f) = (a_n)_{n\in\mathbb{Z}}$. Finally, (5) follows from Parseval's identity:

$$\|\mathcal{F}(f)\|_2^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 = \|f\|_2^2.$$

□

Fourier Coefficients on $L^1(\mathbb{T})$

Recall that for $f \in L^1(\mathbb{T})$ and $n \in \mathbb{Z}$, the nth Fourier coefficient of f is

$$\hat{f}(n) = \int_0^1 f(x)e^{-2\pi inx}\,\mathrm{d}x,$$

and the Fourier series of f is

$$S[f] \sim \sum_{n=-\infty}^{\infty} \hat{f}(n)\chi_n.$$

Note that even when f is real-valued, the Fourier coefficients $\hat{f}(n)$ need not be real, since χ_n is complex-valued for $n \neq 0$.

Proposition 15.3 The Fourier coefficients of functions in $L^1(\mathbb{T})$ have the following properties:

1. $\widehat{f+g}(n) = \hat{f}(n) + \hat{g}(n)$.
2. $\widehat{cf}(n) = c\hat{f}(n)$.
3. $\widehat{\overline{f}}(n) = \overline{\hat{f}(-n)}$.
4. Let $f \in L^1(\mathbb{T})$, let $x \in [0,1]$, and define $f_x(y) := f(y-x) \pmod 1$. Then

$$\hat{f}_x(n) = \chi_{-n}(x)\hat{f}(n) = e^{-2\pi i n x}\hat{f}(n).$$

5. $|\hat{f}(n)| \leq \|f\|_1$.

Proof. For (1),

$$\widehat{f+g}(n) = \int_0^1 (f(t)+g(t))e^{-2\pi i n t}\,\mathrm{d}t = \int_0^1 f(t)e^{-2\pi i n t}\,\mathrm{d}t + \int_0^1 g(t)e^{-2\pi i n t}\,\mathrm{d}t = \hat{f}(n) + \hat{g}(n).$$

For (2),

$$\widehat{cf}(n) = \int_0^1 cf(t)e^{-2\pi i n t}\,\mathrm{d}t = c\int_0^1 f(t)e^{-2\pi i n t}\,\mathrm{d}t = c\hat{f}(n).$$

For (3), using conjugation through the integral,

$$\widehat{\overline{f}}(n) = \int_0^1 \overline{f(t)}e^{-2\pi i n t}\,\mathrm{d}t = \overline{\int_0^1 f(t)e^{2\pi i n t}\,\mathrm{d}t} = \overline{\hat{f}(-n)}.$$

For (4),

$$\begin{aligned}\hat{f}_x(n) &= \int_0^1 f(t-x)e^{-2\pi int}\,\mathrm{d}t \\ &= \int_0^1 f(t-x)e^{-2\pi in(t-x+x)}\,\mathrm{d}t \\ &= e^{-2\pi inx}\int_0^1 f(t-x)e^{-2\pi in(t-x)}\,\mathrm{d}t.\end{aligned}$$

Now use translation invariance of Lebesgue measure on $\mathbb{T}$ (equivalently, the change of variable $u = t - x \bmod 1$) to get

$$\int_0^1 f(t-x)e^{-2\pi in(t-x)}\,\mathrm{d}t = \int_0^1 f(u)e^{-2\pi inu}\,\mathrm{d}u = \hat{f}(n).$$

Hence

$$\hat{f}_x(n) = e^{-2\pi inx}\hat{f}(n) = \chi_{-n}(x)\hat{f}(n).$$

For (5), since $|e^{-2\pi int}| = 1$,

$$|\hat{f}(n)| = \left|\int_0^1 f(t)e^{-2\pi int}\,\mathrm{d}t\right| \le \int_0^1 |f(t)|\,\mathrm{d}t = \|f\|_1.$$

$\square$

Definition 15.4 The space $\ell^\infty(\mathbb{Z})$ is

$$\ell^\infty(\mathbb{Z}) = \{(a_n)_{n\in\mathbb{Z}} : a_n \in \mathbb{C}, \|a\|_\infty = \sup|a_n| < \infty\}.$$

Thus, if $f \in L^1(\mathbb{T})$, then $(\hat{f}(n))_{n\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ by Proposition 15.3(5).

Theorem 15.5 If $f_k \to f$ in $L^1(\mathbb{T})$, then $\hat{f}_k(n) \to \hat{f}(n)$ for all $n \in \mathbb{Z}$.

Proof. $|\hat{f}_k(n) - \hat{f}(n)| = |\widehat{f_k - f}(n)| \le \|f_k - f\|_1 \to 0$ as $k \to \infty$. $\square$

Theorem 15.6 (Riemann–Lebesgue lemma: circle version) If $f \in L^1(\mathbb{T})$, then $\hat{f}(n) \to 0$ as $n \to \pm\infty$.

Proof. Let $\varepsilon > 0$. Let $p \in \mathrm{Trig}(\mathbb{T})$ such that $\|f - p\|_1 < \varepsilon$. Write

$$p = \sum_{n=-N}^{N} a_n \chi_n.$$

For $|m| > N$, orthogonality gives

$$\hat{p}(m) = \langle p, \chi_m \rangle = \sum_{n=-N}^{N} a_n \langle \chi_n, \chi_m \rangle = 0,$$

and hence $\hat{p}(n) = 0$ for all $|n| > N$. Therefore, for $|n| > N$,

$$|\hat{f}(n)| = |\hat{f}(n) - \hat{p}(n)| \leq \|f - p\|_1 < \varepsilon.$$

Therefore $\hat{f}(n) \to 0$ as $n \to \pm\infty$. □

Remark 15.7 The Riemann–Lebesgue lemma has the following immediate trigonometric consequences.

1. If $f \in L^1(\mathbb{T})$, then
$$\lim_{n\to\infty} \int_0^1 f(x)\cos(2\pi nx)\,\mathrm{d}x = 0,$$
and
$$\lim_{n\to\infty} \int_0^1 f(x)\sin(2\pi nx)\,\mathrm{d}x = 0.$$

2. For $B \in \mathbb{R}$,
$$\begin{aligned}\lim_{n\to\infty} \int_0^1 f(x)\sin(2\pi nx + B)\,\mathrm{d}x &= \lim_{n\to\infty} \int_0^1 f(x)\sin(2\pi nx)\cos(B)\,\mathrm{d}x \\ &\quad + \lim_{n\to\infty} \int_0^1 f(x)\cos(2\pi nx)\sin(B)\,\mathrm{d}x \\ &= 0.\end{aligned}$$

So far, the Fourier transform sends

$$L^1(\mathbb{T}) \to c_0(\mathbb{Z}), \qquad f \mapsto (\hat{f}(n))_{n\in\mathbb{Z}},$$

where

$$c_0(\mathbb{Z}) = \{(a_n)_{n\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z}) : \lim_{n\to\pm\infty} a_n = 0\}.$$

Question 15.8 Does the Fourier series of f, $S[f]$, actually converge to f? If so, in what sense? Well? Say something!

Definition 15.9 For $n \in \mathbb{N}$, the **Dirichlet kernel** is

$$D_n := \sum_{k=-n}^{n} \chi_k.$$

The nth partial sum of the Fourier series of f is

$$S_n(f) := \sum_{k=-n}^{n} \hat{f}(k)\chi_k,$$

and we want to understand whether $S_n(f)$ converges to f for sufficiently regular functions.

Proposition 15.10 For $f \in L^1(\mathbb{T})$ and $x \in \mathbb{T}$,

$$S_n(f)(x) = f * D_n(x),$$

where $*$ is **convolution** in $L^1(\mathbb{T})$.

Proof. We have

$$\begin{aligned} S_n(f)(x) &= \sum_{k=-n}^{n} \hat{f}(k)\chi_k(x) \\ &= \sum_{k=-n}^{n} \left(\int_0^1 f(t)e^{-2\pi ikt}\,\mathrm{d}t\right) e^{2\pi ikx} \\ &= \int_0^1 f(t)\left(\sum_{k=-n}^{n} e^{-2\pi ikt}e^{2\pi ikx}\right)\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 f(t) \left(\sum_{k=-n}^{n} e^{2\pi i k(x-t)} \right) \mathrm{d}t \\
&= \int_0^1 f(t) D_n(x-t)\, \mathrm{d}t \\
&= f * D_n(x).
\end{aligned}
$$
□

The convolution $*$ is the operation that turns $L^1(\mathbb{T})$ into an algebra and makes partial sums accessible through kernels. If D_n behaved like an identity, then we would expect $S_n(f)(x) = f * D_n(x)$ to approach $f(x)$ for continuous f.

It remains to compute an explicit formula for D_n.

Proposition 15.11 For $n \in \mathbb{N}$ and $t \in \mathbb{T}$,

$$
D_n(t) = \begin{cases} \dfrac{\sin(2\pi(n+1/2)t)}{\sin(\pi t)} & \text{for } t \neq 0, \\ 2n+1, & t = 0. \end{cases}
$$

The profile in Figure 9 makes two important features visible: D_N has a tall central peak, but it also has oscillatory lobes of both signs away from the origin.

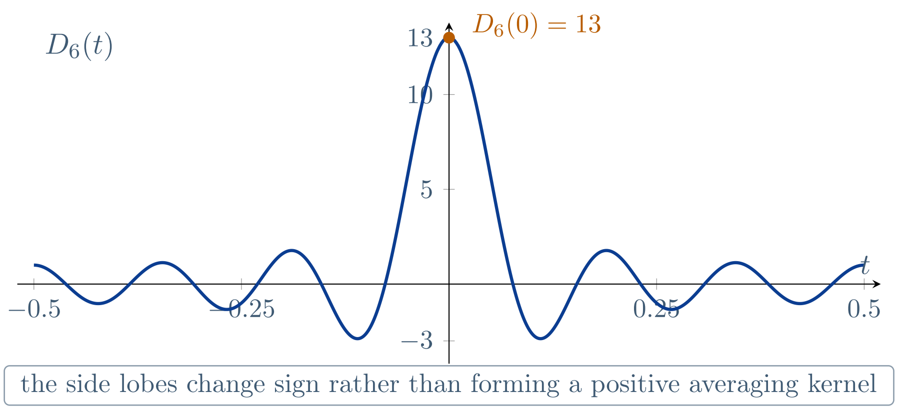


FIGURE 9

Proof. For $t \neq 0$, we have

$$
\sin(\pi t) D_n(t) = \left(\frac{e^{\pi i t} - e^{-\pi i t}}{2i} \right) \sum_{k=-n}^{n} e^{2\pi i k t}
$$

$$= \frac{\sum_{k=-n}^{n} e^{2\pi i(k+1/2)t} - \sum_{k=-n}^{n} e^{2\pi i(k-1/2)t}}{2i}$$
$$= \frac{e^{2\pi i(n+1/2)t} - e^{-2\pi i(n+1/2)t}}{2i}$$
$$= \sin(2\pi(n+1/2)t)$$

Thus,

$$D_n(t) = \frac{\sin(2\pi(n+1/2)t)}{\sin(\pi t)}, \quad t \neq 0.$$

The value $D_n(0) = 2n+1$ is the limit of this quotient as $t \to 0$. □

Lecture 16 — Tuesday, March 03, 2015

Dirichlet Kernels and Partial Sums

We now fix $x \in \mathbb{T}$ and pursue two goals:

1. Give a condition for convergence of $S_N(f)(x)$ when f is continuous.
2. Show that Fourier series of continuous functions do not necessarily converge at every $x \in \mathbb{T}$.

Proposition 16.1 Let $f \in L^1(\mathbb{T})$, let $x \in \mathbb{T} \cong [0,1)$, and define $f_x(t) := f(t-x) \pmod 1$. Then the following hold:

1. $$\lim_{n\to\infty} \int_0^1 f(t)\cos(2\pi nt)\,\mathrm{d}t = \lim_{n\to\infty} \int_0^1 f(t)\sin(2\pi nt)\,\mathrm{d}t = 0.$$

2. The Dirichlet kernel is
$$D_N = \sum_{k=-N}^{N} \chi_k,$$

and

$$D_N(t) = \begin{cases} \dfrac{\sin(2\pi(N+1/2)t)}{\sin(\pi t)}, & t \neq 0, \\ 2N+1, & t = 0. \end{cases}$$

3. $\|f_x\|_1 = \|f\|_1$.
4. $$S_N(f)(x) = \sum_{k=-N}^{N} \hat{f}(k)\chi_k(x) = \int_0^1 f(t)D_N(x-t)\,\mathrm{d}t = \int_0^1 f(x+t)D_N(t)\,\mathrm{d}t.$$

Since D_N is even, we may also use $D_N(t) = D_N(-t)$ in these representations.

Proof. For (1), apply the Riemann–Lebesgue lemma to f and use

$$\cos(2\pi nt) = \frac{e^{2\pi int} + e^{-2\pi int}}{2} \qquad \text{and} \qquad \sin(2\pi nt) = \frac{e^{2\pi int} - e^{-2\pi int}}{2i}.$$

Statement (2) was proved in Proposition 15.11.

For (3), translation invariance of Lebesgue measure on $\mathbb{T}$ gives

$$\|f_x\|_1 = \int_0^1 |f(t-x)|\,\mathrm{d}t = \int_0^1 |f(u)|\,\mathrm{d}u = \|f\|_1.$$

For (4),

$$S_N(f)(x) = \sum_{k=-N}^{N} \hat{f}(k)\chi_k(x)$$

by definition of the Nth partial Fourier sum. Using

$$D_N(x-t) = \sum_{k=-N}^{N} \chi_k(x)\chi_{-k}(t),$$

we obtain

$$\int_0^1 f(t)D_N(x-t)\,\mathrm{d}t = \sum_{k=-N}^{N} \chi_k(x) \int_0^1 f(t)\chi_{-k}(t)\,\mathrm{d}t$$

$$= \sum_{k=-N}^{N} \hat{f}(k)\chi_k(x) = S_N(f)(x).$$

Finally, substitute $u = x + t \pmod 1$ and use translation invariance:

$$\int_0^1 f(t)D_N(x-t)\,\mathrm{d}t = \int_0^1 f(x+t)D_N(t)\,\mathrm{d}t.$$

□

Fix $f \in C(\mathbb{T})$ and $x \in \mathbb{T}$. We ask when $S_N(f)(x)$ has a limit as $N \to \infty$. We calculate

$$\begin{aligned} S_N(f)(x) &= \int_0^1 f(x+t)D_N(t)\,\mathrm{d}t = \int_0^1 f(x+t)\frac{\sin(2\pi(N+1/2)t)}{\sin(\pi t)}\,\mathrm{d}t \\ &= \int_0^1 f(x+t)\frac{\sin(2\pi Nt)\cos(\pi t) + \sin(\pi t)\cos(2\pi Nt)}{\sin(\pi t)}\,\mathrm{d}t \\ &= \int_0^1 f(x+t)\cos(\pi t)\frac{\sin(2\pi Nt)}{\sin(\pi t)}\,\mathrm{d}t + \int_0^1 f(x+t)\cos(2\pi Nt)\,\mathrm{d}t \end{aligned}$$

Since $f \in L^1(\mathbb{T})$, the right-hand integral tends to 0 by the Riemann–Lebesgue lemma. Thus the convergence question is reduced to the left-hand integral

$$\int_0^1 f(x+t)\frac{\cos(\pi t)}{\sin(\pi t)}\sin(2\pi Nt)\,\mathrm{d}t = \int_0^1 f_{-x}(t)\frac{\cos(\pi t)}{\sin(\pi t)}\sin(2\pi Nt)\,\mathrm{d}t.$$

Proposition 16.2 (Riemann localization criterion) Let $f \in C(\mathbb{T})$, let $x \in \mathbb{T}$, and fix $\delta \in (0, 1/2)$. Then $S_N(f)(x)$ converges as $N \to \infty$ if and only if

$$\lim_{N\to\infty} 2\int_0^\delta \frac{f(x+t)+f(x-t)}{2}\,\frac{\sin(2\pi Nt)}{\pi t}\,\mathrm{d}t$$

exists.

Proof. Define

$$J_N := \int_0^1 f_{-x}(t)\frac{\cos(\pi t)}{\sin(\pi t)}\sin(2\pi Nt)\,\mathrm{d}t \qquad \text{and} \qquad C_N := \int_0^1 f(x+t)\cos(2\pi Nt)\,\mathrm{d}t.$$

By the preceding computation, $S_N(f)(x) = J_N + C_N$. Since $f \in L^1(\mathbb{T})$, Riemann–Lebesgue gives

$C_N \to 0$, so $S_N(f)(x)$ converges if and only if J_N converges.

Split

$$J_N = I_N^{(1)} + I_N^{(2)} + I_N^{(3)},$$

where the terms are the integrals over $[0,\delta]$, $[\delta, 1-\delta]$, and $[1-\delta, 1]$, respectively.

For the middle term, let

$$h_\delta(t) := f_{-x}(t)\frac{\cos(\pi t)}{\sin(\pi t)}\mathbb{1}_{[\delta,1-\delta]}(t).$$

Then $h_\delta \in L^1(\mathbb{T})$, and

$$I_N^{(2)} = \int_0^1 h_\delta(t)\sin(2\pi N t)\,\mathrm{d}t \to 0$$

by Riemann–Lebesgue.

For $I_N^{(3)}$, use $t = 1-s$ and periodicity on $\mathbb{T}$:

$$I_N^{(3)} = \int_0^\delta f_{-x}(1-s)\frac{\cos(\pi(1-s))}{\sin(\pi(1-s))}\sin(2\pi N(1-s))\,\mathrm{d}s = \int_0^\delta f_{-x}(-s)\frac{\cos(\pi s)}{\sin(\pi s)}\sin(2\pi N s)\,\mathrm{d}s.$$

Hence

$$I_N^{(1)} + I_N^{(3)} = 2\int_0^\delta A_x(t)\frac{\cos(\pi t)}{\sin(\pi t)}\sin(2\pi N t)\,\mathrm{d}t, \quad A_x(t) := \frac{f(x+t)+f(x-t)}{2}.$$

Define

$$g_\delta(t) := 2A_x(t)\left(\frac{1}{\pi t} - \frac{\cos(\pi t)}{\sin(\pi t)}\right)\mathbb{1}_{(0,\delta)}(t),$$

extended by 0 outside $(0,\delta)$. Since f is continuous, A_x is bounded. Also,

$$\frac{1}{\pi t} - \frac{\cos(\pi t)}{\sin(\pi t)} \to 0 \quad (t \downarrow 0),$$

so $g_\delta \in L^1(\mathbb{T})$. Therefore

$$\int_0^1 g_\delta(t)\sin(2\pi N t)\,\mathrm{d}t \to 0$$

by Riemann–Lebesgue. Consequently, the convergence of

$$2\int_0^\delta A_x(t)\frac{\cos(\pi t)}{\sin(\pi t)}\sin(2\pi N t)\,\mathrm{d}t$$

is equivalent to the convergence of

$$2\int_0^\delta A_x(t)\frac{\sin(2\pi Nt)}{\pi t}\,\mathrm{d}t.$$

Since $I_N^{(2)} \to 0$, this is also equivalent to the convergence of J_N, hence equivalent to the convergence of $S_N(f)(x)$. □

The Riemann localization theorem states that the convergence of $S_N(f)(x)$ when $N \to \infty$ depends only on the behavior of f near x.

Our updated goal is to show that there exists $f \in C(\mathbb{T})$ such that $S_N(f)(x)$ does not converge.

Divergence and the Uniform Boundedness Principle

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces over a field $\mathbb{K}$. Recall that a linear operator $T : X \to Y$ is a map satisfying $T(\alpha x + y) = \alpha T(x) + T(y)$ for all $\alpha \in \mathbb{K}$ and $x, y \in X$.

Definition 16.3 The **operator norm** of a linear operator $T : X \to Y$ is

$$\|T\|_{\text{op}} := \sup_{\|x\|_X = 1} \|Tx\|_Y.$$

An operator T with $\|T\|_{\text{op}} < \infty$ is called a **bounded linear operator**.

If T is bounded, then $\|Tx\|_Y \leq \|T\|_{\text{op}}\|x\|_X$. Also, every linear operator satisfies $T(0) = 0$.

Theorem 16.4 A linear operator $T : X \to Y$ is continuous if and only if it is bounded.

Proof. *Forward implication.* Suppose that T is bounded, so $\|T\|_{\text{op}} < \infty$. If $x_n \to x$ in X, then

$$\|T(x_n) - T(x)\|_Y = \|T(x_n - x)\|_Y \leq \|T\|_{\text{op}}\|x_n - x\|_X \to 0.$$

Thus T is continuous.

Reverse implication. Suppose that T is continuous, in particular at 0. Then there exists $\delta > 0$

such that $\|x\|_X < \delta$ implies $\|Tx\|_Y \leq 1$. Hence

$$\|T\|_{\text{op}} = \sup_{\|x\|_X=1} \frac{2}{\delta} \left\| T\left(\frac{\delta}{2}x\right) \right\|_Y \leq \frac{2}{\delta}.$$

□

Theorem 16.5 Fix $N \in \mathbb{N}$. Define $S_N : L^1(\mathbb{T}) \to L^1(\mathbb{T})$ by

$$f \mapsto S_N(f) = \sum_{k=-N}^{N} \hat{f}(k)\chi_k.$$

Then

$$\|S_N\|_{\text{op}} = \|D_N\|_1.$$

Proof. The operator S_N is a continuous (or equivalently, bounded, by Theorem 16.4), linear operator on $L^1(\mathbb{T})$. For $f \in L^1(\mathbb{T})$,

$$\begin{aligned}
\|S_N(f)\|_1 &= \int_0^1 |S_N(f)(x)| \, \mathrm{d}x \\
&= \int_0^1 \left| \int_0^1 f(t) D_N(x-t) \, \mathrm{d}t \right| \mathrm{d}x \\
&\leq \int_0^1 \int_0^1 |f(t)| \cdot |D_N(x-t)| \, \mathrm{d}t \, \mathrm{d}x \\
&= \int_0^1 |f(t)| \left(\int_0^1 |D_N(x-t)| \, \mathrm{d}x \right) \mathrm{d}t \\
&= \|f\|_1 \|D_N\|_1,
\end{aligned}$$

where the integrals were interchanged via Tonelli's theorem. Thus,

$$\|S_N\|_{\text{op}} \leq \|D_N\|_1.$$

Lecture 17 — Thursday, March 05, 2015

Proof continued. We now prove the reverse inequality $\|S_N\|_{\text{op}} \geq \|D_N\|_1$. It is enough to find functions $f_k \in L^1(\mathbb{T})$ such that $\|f_k\|_1 = 1$ and $S_N(f_k) \to D_N$ in the L^1-norm. Indeed, for every k,

$$\|S_N\|_{\text{op}} \geq \|S_N(f_k)\|_1.$$

Letting $k \to \infty$ gives $\|S_N\|_{\text{op}} \geq \|D_N\|_1$, which completes the proof of equality.

For the approximation, let $k \geq 2$, set $A_k = [0, 1/k] \cup [1 - 1/k, 1]$, and put $f_k = (k/2)\mathbb{1}_{A_k}$. Then

$$\|f_k\|_1 = \int_{[0,1/k]\cup[1-1/k,1]} \frac{k}{2} \,\mathrm{d}x = 1.$$

It remains to show convergence.

$$\begin{aligned}
\|S_N(f_k) - D_N\|_1 &= \int_0^1 |S_N(f_k)(x) - D_N(x)| \,\mathrm{d}x \\
&= \int_0^1 \left| \int_0^1 f_k(t) D_N(x-t) \,\mathrm{d}t - D_N(x) \right| \mathrm{d}x \\
&= \int_0^1 \left| \int_0^1 f_k(t)(D_N(x-t) - D_N(x)) \,\mathrm{d}t \right| \mathrm{d}x \\
&\leq \int_0^1 \int_0^1 |f_k(t)| \cdot |D_N(x-t) - D_N(x)| \,\mathrm{d}t \,\mathrm{d}x.
\end{aligned}$$

Since D_N is uniformly continuous and 1-periodic, for sufficiently large k we have $|D_N(x-t) - D_N(x)| < \varepsilon$ for all $x \in [0,1]$ and all $t \in A_k$. For such k,

$$\begin{aligned}
\|S_N(f_k) - D_N\|_1 &\leq \int_0^1 \int_{A_k} |f_k(t)| \cdot |D_N(x-t) - D_N(x)| \,\mathrm{d}t \,\mathrm{d}x \\
&\leq \int_0^1 \|f_k\|_1 \cdot \varepsilon \\
&= \varepsilon \cdot \|f_k\|_1 \\
&= \varepsilon.
\end{aligned}$$

Since this holds for every $\varepsilon > 0$, we have $\|S_N(f_k) - D_N\|_1 \to 0$. □

Next, recall the uniform boundedness principle from PMATH 351:

Theorem 17.1 (Uniform boundedness principle) Let X and Y be Banach spaces, and let $\mathcal{F}$ be a collection of bounded linear operators from X to Y. If

$$\sup_{T\in\mathcal{F}} \|Tx\|_Y < \infty$$

for every $x \in X$, then

$$\sup_{T\in\mathcal{F}} \|T\|_{\text{op}} < \infty.$$

Equivalently, if $\sup_{T\in\mathcal{F}} \|T\|_{\text{op}} = \infty$, then there exists $x \in X$ such that $\sup_{T\in\mathcal{F}} \|Tx\|_Y = \infty$.

Theorem 17.2 There exists $f \in L^1(\mathbb{T})$ for which the Fourier series is not L^1-convergent. Equivalently, the sequence $S_N(f)$ need not converge in $L^1(\mathbb{T})$.

Proof. Take $X = L^1(\mathbb{T})$, $Y = L^1(\mathbb{T})$, and

$$\mathcal{F} = \{S_N : L^1(\mathbb{T}) \to L^1(\mathbb{T}) : N \in \mathbb{N}\}.$$

We have $\|S_N\|_{\text{op}} = \|D_N\|_1$ by Theorem 16.5. We claim that there is a constant $C > 0$ such that $\|D_N\|_1 \geq C \log N$. Assuming this claim, we have

$$\sup_N \|S_N\|_{\text{op}} \geq \sup_N C \log N = \infty.$$

By the uniform boundedness principle, there exists $f \in L^1(\mathbb{T})$ such that $\sup_N \|S_N(f)\|_1 = \infty$. Hence $\{S_N(f)\}_{N\in\mathbb{N}}$ cannot be convergent in $L^1(\mathbb{T})$.

It remains to prove the claim. Fix $N \in \mathbb{N}$. Then

$$\begin{aligned}
\|D_N\|_1 &= \int_0^1 \left| \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)} \right| \mathrm{d}t \\
&\geq \sum_{j=0}^{4N+1} \frac{2(2N+1)}{\pi(j+1)} \int_{\frac{j}{2(2N+1)}}^{\frac{j+1}{2(2N+1)}} |\sin(\pi(2N+1)t)| \, \mathrm{d}t \\
&= \frac{2}{\pi^2} \sum_{j=0}^{4N+1} \frac{1}{j+1}.
\end{aligned}$$

The harmonic sum is bounded below by $\log(4N+2)$, so $\|D_N\|_1 \geq C \log N$ for some absolute constant $C > 0$ and every $N \geq 2$. $\square$

Corollary 17.3 There exists $C > 0$ such that

$$\|D_N\|_1 \geq C \log N$$

for every $N \geq 2$. In particular, $\|D_N\|_1 \to \infty$.

Proof. This estimate was proved at the end of the proof of Theorem 17.2. □

The following lemma will be used to transfer divergence from $L^1(\mathbb{T})$ to continuous functions.

Lemma 17.4 There exists $g_N \in C(\mathbb{T})$ such that $\|g_N\|_\infty = 1$ and $|S_N(g_N)(0)| \geq \|D_N\|_1/2$.

Proof. The function

$$D_N(t) = \frac{\sin(2\pi(N + 1/2)t)}{\sin(\pi t)}$$

is real-valued, and $\operatorname{sgn}(D_N)$ takes the values ± 1 except at finitely many points. Define a continuous g_N so that $|g_N| \leq 1$ everywhere and $\|g_N\|_\infty = 1$, and so that g_N agrees with $\operatorname{sgn}(D_N)$ outside small intervals around the discontinuities of $\operatorname{sgn}(D_N)$. Since $D_N \in L^1(\mathbb{T})$, absolute continuity of the integral lets us choose the union U of these intervals so that

$$\int_U |D_N(t)| \,\mathrm{d}t < \varepsilon.$$

Then

$$\begin{aligned} |S_N(g_N)(0)| &= |\int_0^1 g_N(t) D_N(t) \,\mathrm{d}t| \\ &= \left| \int_U g_N(t) D_N(t) \,\mathrm{d}t + \int_{[0,1]\setminus U} |D_N(t)| \,\mathrm{d}t \right|. \end{aligned}$$

Since $|g_N(t)D_N(t)| \leq |D_N(t)|$ for all t, we have

$$\begin{aligned} |S_N(g_N)(0)| &\geq \|D_N\|_1 - \int_U |D_N(t)| \,\mathrm{d}t - \left| \int_U g_N(t) D_N(t) \,\mathrm{d}t \right| \\ &\geq \|D_N\|_1 - 2\varepsilon. \end{aligned}$$

Taking $\varepsilon < \|D_N\|_1/4$ gives $|S_N(g_N)(0)| \geq \|D_N\|_1/2$. □

Lecture 18 — Tuesday, March 10, 2015

Consider the contrapositive form of the uniform boundedness principle: if X and Y are Banach spaces and $\mathcal{F}$ is a family of bounded linear operators from X to Y with $\sup_{T\in\mathcal{F}} \|T\|_{\text{op}} = \infty$, then there exists $x \in X$ such that $\sup_{T\in\mathcal{F}} \|Tx\|_Y = \infty$. Recall that we constructed functions $g_n \in C(\mathbb{T})$ with $\|g_n\|_\infty = 1$ and $|S_n(g_n)(0)| \geq c\log n$ for some $c > 0$.

Theorem 18.1 There exists $f \in C(\mathbb{T})$ such that $S_n(f)(0)$ does not converge.

Proof. Take $X = C(\mathbb{T})$ and $Y = \mathbb{C}$. Let $\mathcal{F} = \{T_n : C(\mathbb{T}) \to \mathbb{C}\}$ where

$$T_n(f) = S_n(f)(0) = \sum_{k=-n}^{n} \hat{f}(k)\chi_k(0).$$

For every $n \in \mathbb{N}$, T_n is bounded, because

$$\|T_n\|_{\text{op}} = \sup_{\|f\|_\infty=1} |T_n(f)| = \sup_{\|f\|_\infty=1} \left|\int_0^1 f(t)D_n(-t)\,\mathrm{d}t\right| \leq \sup_{\|f\|_\infty=1} \|f\|_\infty\|D_n\|_1 = \|D_n\|_1 < \infty$$

by Hölder's inequality. Using the functions $g_n \in C(\mathbb{T})$ constructed in Lemma 17.4, we have

$$\|T_n\|_{\text{op}} \geq |T_n(g_n)| = |S_n(g_n)(0)| \geq c\log(n).$$

Thus $\|T_n\|_{\text{op}} \to \infty$ as $n \to \infty$. By the uniform boundedness principle, there exists $f \in C(\mathbb{T})$ such that

$$\sup_n |T_n(f)| = \infty.$$

Since $T_n(f) = S_n(f)(0)$, the sequence $S_n(f)(0)$ is unbounded and therefore divergent. $\square$

Remark 18.2 We can, of course, replace 0 by any point of $\mathbb{T}$ in Theorem 18.1, since the Fourier partial sums are translation invariant.

Convolutions

We will use the following results on iterated integrals. **Tonelli's theorem** says that if $f : [0,1] \times [0,1] \to \mathbb{R}$ is nonnegative and measurable, then

$$\int_0^1 \int_0^1 f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \int_0^1 f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Fubini's theorem says that if $f : [0,1] \times [0,1] \to \mathbb{R}$ is integrable, then

$$\int_0^1 \int_0^1 f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^1 f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

Definition 18.3 For $f, g \in L^1(\mathbb{T})$, we define

$$(f * g)(x) := \int_0^1 f(t) g(x-t)\,\mathrm{d}t,$$

and call the function $f * g$ the **convolution** of f and g.

When g is a localized, smooth averaging kernel, the operation can be read as a moving weighted average of f, as in Figure 10. This is the intuition behind the summability kernels introduced later.

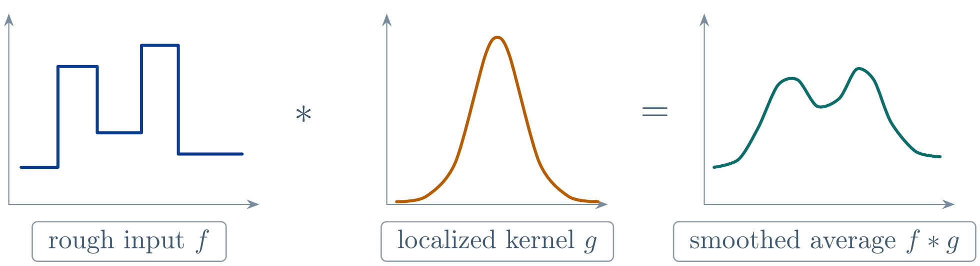


FIGURE 10

Proposition 18.4 Let $f, g \in L^1(\mathbb{T})$. The convolution has the following properties:

1. $f * g \in L^1(\mathbb{T})$ and
$$\|f * g\|_1 \le \|f\|_1 \, \|g\|_1.$$
2. $f * g = g * f$ almost everywhere.

3. If $g \in L^\infty(\mathbb{T})$, then $f * g \in L^\infty(\mathbb{T})$ and
$$\|f * g\|_\infty \leq \|f\|_1 \|g\|_\infty.$$
4. The partial sums satisfy
$$S_n(f) = f * D_n.$$
5. If $g = \chi_k$ for some $k \in \mathbb{Z}$, then
$$f * g(x) = \hat{f}(k)\chi_k(x).$$
6. For every $n \in \mathbb{Z}$,
$$\widehat{f * g}(n) = \hat{f}(n)\,\hat{g}(n).$$

Proof. For (1), by Tonelli's theorem,
$$\begin{aligned}
\|f * g\|_1 &= \int_0^1 \left| \int_0^1 f(t)g(x-t)\,\mathrm{d}t \right| \mathrm{d}x \\
&\leq \int_0^1 \int_0^1 |f(t)|\,|g(x-t)|\,\mathrm{d}t\,\mathrm{d}x \\
&= \int_0^1 |f(t)| \left(\int_0^1 |g(x-t)|\,\mathrm{d}x \right) \mathrm{d}t \\
&= \int_0^1 |f(t)|\,\|g\|_1\,\mathrm{d}t = \|f\|_1\,\|g\|_1 < \infty.
\end{aligned}$$
Hence $f * g \in L^1(\mathbb{T})$.

For (2), for almost every x,
$$(g * f)(x) = \int_0^1 g(t)f(x-t)\,\mathrm{d}t.$$
Set $u = x - t \pmod 1$. By translation invariance of Lebesgue measure on $\mathbb{T}$,
$$(g * f)(x) = \int_0^1 g(x-u)f(u)\,\mathrm{d}u = \int_0^1 f(u)g(x-u)\,\mathrm{d}u = (f * g)(x).$$

For (3), for almost every x,

$$|f * g(x)| = \left| \int_0^1 f(t)g(x-t)\,\mathrm{d}t \right| \leq \int_0^1 |f(t)|\,|g(x-t)|\,\mathrm{d}t \leq \|g\|_\infty \|f\|_1.$$

Taking the essential supremum in x gives the claimed norm bound.

For (4), from the kernel representation of partial sums,

$$S_n(f)(x) = \int_0^1 f(t)D_n(x-t)\,\mathrm{d}t = (f * D_n)(x).$$

For (5), if $g = \chi_k$, then

$$\begin{aligned}
(f * \chi_k)(x) &= \int_0^1 f(t)\chi_k(x-t)\,\mathrm{d}t \\
&= \int_0^1 f(t)e^{2\pi i k(x-t)}\,\mathrm{d}t \\
&= e^{2\pi i k x} \int_0^1 f(t)e^{-2\pi i k t}\,\mathrm{d}t \\
&= \chi_k(x)\hat{f}(k).
\end{aligned}$$

For (6), by Fubini's theorem,

$$\begin{aligned}
\widehat{f * g}(n) &= \int_0^1 (f * g)(x)\chi_{-n}(x)\,\mathrm{d}x \\
&= \int_0^1 \int_0^1 f(t)g(x-t)\chi_{-n}(x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \int_0^1 f(t) \left(\int_0^1 g(x-t)\chi_{-n}(x)\,\mathrm{d}x \right) \mathrm{d}t.
\end{aligned}$$

With $u = x - t \pmod 1$,

$$\int_0^1 g(x-t)\chi_{-n}(x)\,\mathrm{d}x = \int_0^1 g(u)\chi_{-n}(u+t)\,\mathrm{d}u = \chi_{-n}(t) \int_0^1 g(u)\chi_{-n}(u)\,\mathrm{d}u = \chi_{-n}(t)\hat{g}(n).$$

Therefore,

$$\widehat{f * g}(n) = \hat{g}(n) \int_0^1 f(t)\chi_{-n}(t)\,\mathrm{d}t = \hat{g}(n)\hat{f}(n).$$

□

Corollary 18.5 $L^1(\mathbb{T})$ does not have an identity for convolution; that is, there is no $g \in L^1(\mathbb{T})$ satisfying $f * g = f$ for all $f \in L^1(\mathbb{T})$.

Proof. Suppose there exists $g \in L^1(\mathbb{T})$ with $f * g = f$ for every $f \in L^1(\mathbb{T})$. Then

$$\widehat{f * g}(n) = \hat{f}(n)$$

for every n, and hence $\hat{f}(n)\hat{g}(n) = \hat{f}(n)$ for every n. Taking $f = \chi_n$ gives $\hat{f}(n) = 1$, so $\hat{g}(n) = 1$ for all $n \in \mathbb{Z}$. This contradicts the Riemann–Lebesgue lemma, since $\hat{g}(n) \not\to 0$. $\square$

Definition 18.6 The **Fourier transform** on the circle is the map $\mathcal{F} : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ given by

$$\mathcal{F}(f) = (\hat{f}(n))_{n \in \mathbb{Z}}.$$

On $L^1(\mathbb{T})$, addition, multiplication, and scalar multiplication are defined by

$$(f+g)(x) = f(x) + g(x), \qquad (f * g)(x) = \int_0^1 f(t)g(x-t)\,\mathrm{d}t, \qquad \text{and} \qquad (cf)(x) = cf(x),$$

and the norm is

$$\|f\|_1 = \int_0^1 |f(x)|\,\mathrm{d}x.$$

On $c_0(\mathbb{Z})$, addition, multiplication, and scalar multiplication are defined pointwise: for $a = (a_n)$ and $b = (b_n)$ in $c_0(\mathbb{Z})$,

$$a + b = (a_n + b_n), \qquad ab = (a_n b_n), \qquad \text{and} \qquad ca = (ca_n),$$

and the norm is

$$\|a\|_\infty = \sup_n |a_n|.$$

With these operations, the Fourier transform satisfies

$$\mathcal{F}(f+g) = \mathcal{F}(f) + \mathcal{F}(g) \qquad \text{and} \qquad \mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g).$$

Moreover, $\|\mathcal{F}(f)\|_\infty \le \|f\|_1$, so $\mathcal{F}$ is continuous. Thus $\mathcal{F}$ is a linear homomorphism of algebras.

Summability Kernels and the Fejér Kernel

Definition 18.7 A sequence $\{g_n\} \subset C(\mathbb{T})$ is called a **summability kernel** if it satisfies the following three properties:

1. $\displaystyle\int_0^1 g_n(t)\,\mathrm{d}t = 1$ for every n.
2. There exists $M > 0$ such that $\|g_n\|_1 \leq M$ for every n.
3. For every δ such that $0 < \delta < 1/2$,
$$\int_\delta^{1-\delta} |g_n(t)|\,\mathrm{d}t \longrightarrow 0 \text{ as } n \to \infty.$$

If $g_n(t) \geq 0$ for every $n \in \mathbb{N}$ and $t \in \mathbb{T}$, then $\{g_n\}$ is a **positive summability kernel**.

Figure 11 compares two summability kernels, g_{n_1} in blue and g_{n_2} in orange, on $\mathbb{T}$, where $n_1 < n_2$.

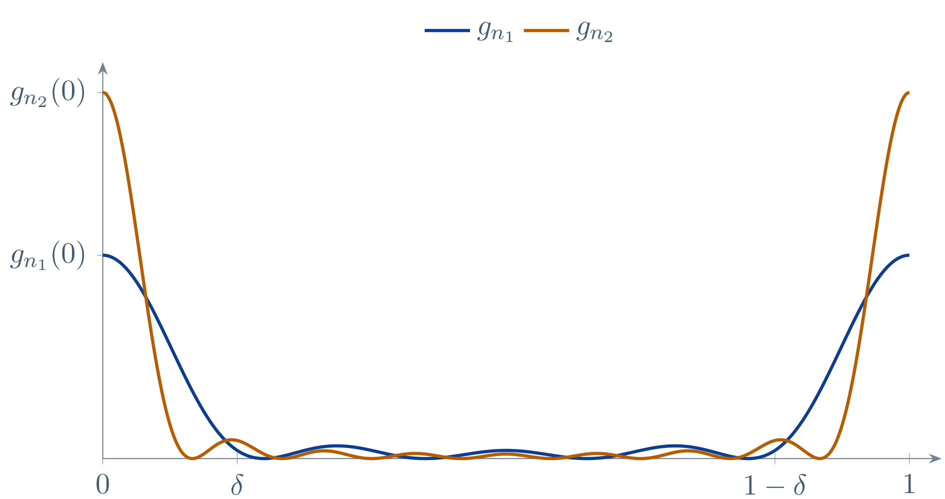


FIGURE 11

Example 18.8 For $n \geq 2$, the functions

$$K_n = \frac{n}{2}\mathbb{1}_{[0,1/n]\cup[1-1/n,1]}$$

satisfy the analytic approximate-identity conditions in the definition, except for continuity. In the proof of Theorem 16.5, the same family, indexed there by k, was used with N fixed to show that $S_N(K_k) \to D_N$ in the L^1 norm as $k \to \infty$.

Lecture 19 — Thursday, March 12, 2015

Recall that $\mathcal{F} : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ defined by $f \mapsto \hat{f} = (\hat{f}(n))_{n\in\mathbb{Z}}$ is the Fourier transform of f. Multiplication on $L^1(\mathbb{T})$ is handled through convolution:

$$f * g(x) = \int_0^1 f(t)g(x-t)\,\mathrm{d}t,$$

with $\|f * g\|_1 \le \|f\|_1\|g\|_1$. The map $\mathcal{F}$ is injective but not surjective.

Question 19.1 Can f be reconstructed from $\{\hat{f}(n)\}_{n\in\mathbb{Z}}$?

One strategy is to consider the partial sums

$$S_N(f) = \sum_{n=-N}^{N} \hat{f}(n)\chi_n$$

and ask whether $S_N(f)$ converges to f, and in what sense.

We can summarize what we already know. By Theorem 17.2, there exists $f \in L^1(\mathbb{T})$ such that $S_N(f)$ does not converge to f in $\|\cdot\|_1$. By Corollary 14.15, if $f \in L^2(\mathbb{T})$, then $S_N(f)$ does converge to f in $\|\cdot\|_2$. By Theorem 18.1, there exists $f \in C(\mathbb{T})$ such that $S_N(f)(0)$ does not converge to $f(0)$, and hence $S_N(f)$ does not converge to f in $\|\cdot\|_\infty$.

Theorem 19.2 Let $\{K_n\}$ be a summability kernel, and let $f \in C(\mathbb{T})$. Then $K_n * f \to f$ uniformly (that is, $\|K_n * f - f\|_\infty \to 0$).

Proof. Let $x \in [0,1]$. Then

$$|K_n * f(x) - f(x)| = \left|\int_0^1 K_n(t)f(x-t)\,\mathrm{d}t - f(x)\int_0^1 K_n(t)\,\mathrm{d}t\right|$$

$$
\begin{aligned}
&= \left| \int_0^1 K_n(t)(f(x-t) - f(x))\,\mathrm{d}t \right| \\
&\le \int_0^1 |K_n(t)| \cdot |f(x-t) - f(x)|\,\mathrm{d}t.
\end{aligned}
$$

Since $f \in C(\mathbb{T})$, it is uniformly continuous in the circle metric. Let $M = \sup_n \|K_n\|_1$, and choose $0 < \delta < 1/2$ such that

$$d_{\mathbb{T}}(x,y) < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \frac{\varepsilon}{2M}.$$

Then choose N such that

$$\int_\delta^{1-\delta} |K_n(t)|\,\mathrm{d}t < \frac{\varepsilon}{4(\|f\|_\infty + 1)}$$

whenever $n \ge N$. If $t \in [0,\delta] \cup [1-\delta, 1]$, then $d_{\mathbb{T}}(x-t, x) < \delta$; for t near 1, this follows from periodicity because $f(x-t) = f(x + (1-t))$. Hence, for $n \ge N$,

$$
\begin{aligned}
|K_n * f(x) - f(x)| &\le \frac{\varepsilon}{2M} \int_{[0,\delta]\cup[1-\delta,1]} |K_n(t)|\,\mathrm{d}t \\
&\quad + 2\|f\|_\infty \int_\delta^{1-\delta} |K_n(t)|\,\mathrm{d}t \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}
$$

The bound is independent of x, so $K_n * f \to f$ uniformly. □

Theorem 19.3 Let $\{K_n\}$ be a summability kernel, and let $f \in L^1(\mathbb{T})$. Then $K_n * f \to f$ in L^1 (that is, $\|K_n * f - f\|_1 \to 0$).

Proof. Let $\varepsilon > 0$. Since $C(\mathbb{T})$ is $\|\cdot\|_1$-dense in $L^1(\mathbb{T})$, there exists $g \in C(\mathbb{T})$ such that $\|f - g\|_1 < \varepsilon$. Then

$$
\begin{aligned}
\|K_n * f - f\|_1 &= \|K_n * f - K_n * g + K_n * g - g + g - f\|_1 \\
&\le \|K_n * (f-g)\|_1 + \|K_n * g - g\|_1 + \|g - f\|_1 \\
&\le \|K_n\|_1 \|f - g\|_1 + \|K_n * g - g\|_\infty + \|g - f\|_1 \\
&\le M\varepsilon + \|K_n * g - g\|_\infty + \varepsilon,
\end{aligned}
$$

where $M = \sup_{n \in \mathbb{N}} \|K_n\|_1$. By Theorem 19.2, $\|K_n * g - g\|_\infty \to 0$. Therefore

$$\limsup_{n\to\infty} \|K_n * f - f\|_1 \le (M+1)\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $K_n * f \to f$ in L^1. $\square$

Remark 19.4 The Dirichlet kernel

$$D_N = \sum_{n=-N}^{N} \chi_n.$$

satisfies

$$\int_0^1 D_N(t)\,\mathrm{d}t = \hat{D}_N(0) = 1,$$

but by Corollary 17.3, $\|D_N\|_1 \to \infty$, so D_N is not a summability kernel. This explains why $S_N(f)$ need not converge to f in L^1. We therefore look for a better kernel.

Definition 19.5 The **Fejér kernel** is defined by

$$F_N = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N+1}\right) \chi_n.$$

Lemma 19.6 Let $n \in \mathbb{N}$. Then

$$F_n(t) = \begin{cases} \dfrac{1}{n+1} \dfrac{\sin^2((n+1)\pi t)}{\sin^2(\pi t)}, & t \neq 0 \text{ in } \mathbb{T}, \\ n+1, & t = 0 \text{ in } \mathbb{T}. \end{cases}$$

Proof. We have

$$\sin^2(\pi t) = \left(\frac{e^{\pi i t} - e^{-\pi i t}}{2i}\right)^2 = \frac{e^{2\pi i t} + e^{-2\pi i t} - 2}{-4} = \frac{1}{2} - \frac{1}{4}(\chi_1 + \chi_{-1}).$$

Also,

$$\begin{aligned} \sin^2(\pi t) \cdot F_n(t) &= \left(\frac{1}{2} - \frac{1}{4}(\chi_1 + \chi_{-1})\right) \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \chi_j \\ &= \frac{1}{2} \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \chi_j \end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{4} \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \chi_{j+1} \\
&- \frac{1}{4} \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \chi_{j-1}.
\end{aligned}
$$

Expanding and collecting terms, the coefficient of χ_k is

$$\frac{1}{2}\left(1 - \frac{|k|}{n+1}\right) - \frac{1}{4}\left(1 - \frac{|k-1|}{n+1}\right) - \frac{1}{4}\left(1 - \frac{|k+1|}{n+1}\right).$$

For $k \neq 0, n+1, -(n+1)$, this simplifies to 0. Specifically, for $k = 0$, the coefficient is

$$\frac{1}{2} - \frac{1}{4}\left(1 - \frac{1}{n+1}\right) - \frac{1}{4}\left(1 - \frac{1}{n+1}\right) = \frac{1}{2} - \frac{1}{2} + \frac{1}{2(n+1)} = \frac{1}{2(n+1)}.$$

Thus,

$$
\begin{aligned}
\sin^2(\pi t) F_n(t) &= \frac{1}{2(n+1)}\chi_0 - \frac{1}{4(n+1)}\chi_{n+1} - \frac{1}{4(n+1)}\chi_{-(n+1)} \\
&= \frac{1}{4(n+1)}(2 - \chi_{n+1} - \chi_{-(n+1)}) \\
&= \frac{1}{4(n+1)}|1 - e^{2\pi i(n+1)t}|^2 \\
&= \frac{\sin^2((n+1)\pi t)}{n+1}.
\end{aligned}
$$

This proves the quotient formula away from 0. At 0, continuity gives

$$F_n(0) = \lim_{t \to 0} \frac{1}{n+1} \frac{\sin^2((n+1)\pi t)}{\sin^2(\pi t)} = n+1.$$

□

Remark 19.7 The Fejér kernel has the following properties:

1. $\hat{F}_N(0) = 1$.
2. $F_N = \dfrac{1}{N+1}(D_0 + D_1 + \cdots + D_N)$.
3. $F_N \geq 0$.

4. $\|F_N\|_1 = \int_0^1 F_N(t)\,\mathrm{d}t = \hat{F}_N(0) = 1$.
5. If $0 < \delta < 1/2$, then $\int_\delta^{1-\delta} F_N(t)\,\mathrm{d}t \longrightarrow 0$ as $N \to \infty$.

Thus, $\{F_N\}$ is a summability kernel.

Proof. For (1), by the definition

$$F_N = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N+1}\right) \chi_n,$$

the coefficient of χ_0 is 1. Hence $\widehat{F}_N(0) = 1$.

For (2), for each $j \in \{-N, \dots, N\}$, the term χ_j appears in D_k exactly when $|j| \le k \le N$, which occurs for $N - |j| + 1$ values of k. Therefore

$$\frac{1}{N+1} \sum_{k=0}^{N} D_k = \sum_{j=-N}^{N} \frac{N - |j| + 1}{N+1} \chi_j = \sum_{j=-N}^{N} \left(1 - \frac{|j|}{N+1}\right) \chi_j = F_N.$$

For (3), by Lemma 19.6,

$$F_N(t) = \frac{1}{N+1} \frac{\sin^2((N+1)\pi t)}{\sin^2(\pi t)}.$$

The right-hand side is nonnegative for every t (interpreting the quotient at zeros of $\sin(\pi t)$ by continuity), so $F_N \ge 0$.

For (4), we combine (3) with (1). Since $F_N \ge 0$,

$$\|F_N\|_1 = \int_0^1 |F_N(t)|\,\mathrm{d}t = \int_0^1 F_N(t)\,\mathrm{d}t = \widehat{F}_N(0) = 1.$$

For (5), fix $\delta \in (0, 1/2)$. By Lemma 19.6,

$$F_N(t) = \frac{1}{N+1} \frac{\sin^2((N+1)\pi t)}{\sin^2(\pi t)} \le \frac{1}{N+1} \frac{1}{\sin^2(\pi t)} \le \frac{1}{N+1} \frac{1}{\sin^2(\pi \delta)}$$

for all $t \in [\delta, 1-\delta]$, because $\sin(\pi t) \geq \sin(\pi\delta) > 0$ on that interval. Integrating gives

$$\int_{\delta}^{1-\delta} F_N(t)\,\mathrm{d}t \leq \frac{1-2\delta}{(N+1)\sin^2(\pi\delta)} \xrightarrow[N\to\infty]{} 0.$$

Thus all three summability kernel conditions hold for $\{F_N\}$, so $\{F_N\}$ is a summability kernel. □

Lecture 20 — Tuesday, March 17, 2015

Notation 20.1 We define the **Fejér means** of f by

$$\sigma_n(f) := F_n * f = \left(\sum_{j=-n}^{n}\left(1-\frac{|j|}{n+1}\right)\chi_j\right) * f = \sum_{j=-n}^{n}\left(1-\frac{|j|}{n+1}\right)\hat{f}(j)\chi_j.$$

Corollary 20.2 $\{F_n\}_{n\in\mathbb{N}}$ is a summability kernel. Consequently, we have the following:

1. If $f \in L^1(\mathbb{T})$, then $\sigma_n(f) \xrightarrow{\|\cdot\|_1} f$.
2. If $f \in C(\mathbb{T})$, then $\sigma_n(f) \xrightarrow{\|\cdot\|_\infty} f$.

Proof. (1) follows from Theorem 19.3, and (2) follows from Theorem 19.2. □

Remark 20.3 Corollary 20.2(2) proves that $\mathrm{Trig}(\mathbb{T})$ is dense in $C(\mathbb{T})$. We will show later that if $f \in C(\mathbb{T})$ and $S_n f(x_0)$ converges, then it converges to the same limit as $\sigma_n(f)(x_0)$ does, which is $f(x_0)$.

Theorem 20.4 The Fourier transform $\mathcal{F} : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ is injective.

Proof. Suppose there exists $f \in L^1(\mathbb{T})$ with $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$. Then

$$\sigma_n(f) = F_n * f = \sum_{j=-n}^{n} (1 - \frac{|j|}{n+1})\hat{f}(j)\chi_j = 0.$$

By Corollary 20.2(1), $\sigma_n(f) \to f$ in $\|\cdot\|_1$, which implies that $f = 0$ almost everywhere. This is the **uniqueness theorem**: if $f, g \in L^1(\mathbb{T})$ and $\hat{f}(n) = \hat{g}(n)$ for every n, then $\widehat{f-g}(n) = 0$ for every n, and hence $f = g$ almost everywhere. □

Recall that the Riemann–Lebesgue lemma says that if $f \in L^1(\mathbb{T})$, then $\hat{f} \in c_0(\mathbb{Z})$. Also, $f \in L^2(\mathbb{T})$ if and only if $\hat{f} \in \ell^2(\mathbb{Z})$.

Recall the functional-analytic open mapping theorem:

Theorem 20.5 (Open mapping theorem) If X and Y are Banach spaces and $T : X \to Y$ is bounded, linear, and surjective, then T is open.

Corollary 20.6 If T is continuous, linear, and bijective, then T^{-1} is continuous as well.

Proof. The inverse T^{-1} is linear because

$$T^{-1}(\alpha Tx + Ty) = T^{-1}(T(\alpha x + y)) = \alpha x + y = \alpha T^{-1}(Tx) + T^{-1}(Ty).$$

Moreover, T^{-1} is continuous because inverse images under T^{-1} are images under T, and T is open by the open mapping theorem. Since T^{-1} is linear and continuous, it is bounded. □

Theorem 20.7 The Fourier transform $\mathcal{F} : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$ is not surjective.

Proof. Suppose, toward a contradiction, that $\mathcal{F}$ is surjective. By Theorem 20.4, $\mathcal{F}$ is injective, so $\mathcal{F}$ is bijective.

Since $L^1(\mathbb{T})$ and $c_0(\mathbb{Z})$ are Banach spaces and $\mathcal{F}$ is bounded and bijective, the inverse map $\mathcal{F}^{-1} : c_0(\mathbb{Z}) \to L^1(\mathbb{T})$ is continuous by Corollary 20.6. Since $\mathcal{F}^{-1}$ is linear and continuous, it is bounded. In particular, there exists $M > 0$ such that

$$\|\mathcal{F}^{-1}(\mathcal{F}(f))\|_1 \leq M\|\mathcal{F}(f)\|_\infty$$

for all $f \in L^1(\mathbb{T})$. But taking $f = D_N$ gives

$$\|D_N\|_1 \leq M\|\mathcal{F}(D_N)\|_\infty = M$$

for every N, because

$$\mathcal{F}(D_N)(n) = \hat{D}_N(n) = \begin{cases} 1 & \text{if } |n| \leq N, \\ 0 & \text{if } |n| > N. \end{cases}$$

This is a contradiction, since $\|D_N\|_1 \to \infty$ by Corollary 17.3. Hence $\mathcal{F}$ is not surjective. $\square$

Pointwise Convergence of Fourier Series

What have we learned? If $f \in L^1(\mathbb{T})$, then $\sigma_n(f) \to f$ in $\|\cdot\|_1$. If $f \in C(\mathbb{T})$, then $\sigma_n(f) \to f$ in $\|\cdot\|_\infty$. However, there exists $f \in C(\mathbb{T})$ such that $S_n(f)(0)$ does not converge to $f(0)$, and hence $S_n(f)$ does not converge to f in $\|\cdot\|_\infty$.

Question 20.8 A natural question remains: what can we say about pointwise convergence when f is not continuous?

Theorem 20.9 (Fejér's theorem) Let $f \in L^1(\mathbb{T})$ and assume that

$$w_f(x_0) := \lim_{h\to 0} \frac{f(x_0+h) + f(x_0-h)}{2}$$

exists. Then $\sigma_n(f)(x_0) \to w_f(x_0)$.

Proof. Since $\int_0^1 F_n(t)\,\mathrm{d}t = 1$, we have

$$|\sigma_n(f)(x_0) - w_f(x_0)| = |F_n * f(x_0) - w_f(x_0)| = \left|\int_0^1 F_n(t)(f(x_0-t) - w_f(x_0))\,\mathrm{d}t\right|.$$

Split the integral into the two neighborhoods $[0,\delta]$ and $[1-\delta,1]$ of 0 in $\mathbb{T}$, together with the middle interval $[\delta, 1-\delta]$. By the symmetry of F_n,

$$\begin{aligned}&\left|\int_0^\delta F_n(t)(f(x_0-t) - w_f(x_0))\,\mathrm{d}t + \int_{1-\delta}^1 F_n(t)(f(x_0-t) - w_f(x_0))\,\mathrm{d}t\right| \\ &\quad\leq 2\int_0^\delta F_n(t)\left|\frac{f(x_0+t) + f(x_0-t)}{2} - w_f(x_0)\right|\mathrm{d}t.\end{aligned}$$

Let $\delta > 0$ be such that $|t| < \delta$ implies

$$\left| \frac{f(x_0 + t) + f(x_0 - t)}{2} - w_f(x_0) \right| < \varepsilon.$$

For this choice of δ, the contribution from the two neighborhoods of 0 is bounded by

$$2\varepsilon \int_0^\delta F_n(t)\,\mathrm{d}t \leq 2\varepsilon \|F_n\|_1 = 2\varepsilon.$$

For the middle interval, Lemma 19.6 gives

$$F_n(t) = \frac{1}{n+1} \frac{\sin^2((n+1)\pi t)}{\sin^2(\pi t)} \leq \frac{1}{(n+1)\sin^2(\pi\delta)}$$

whenever $\delta \leq t \leq 1 - \delta$. Hence

$$\begin{aligned}\int_\delta^{1-\delta} F_n(t)|f(x_0 - t) - w_f(x_0)|\,\mathrm{d}t &\leq \frac{1}{(n+1)\sin^2(\pi\delta)} \int_\delta^{1-\delta} \big(|f(x_0 - t)| + |w_f(x_0)|\big)\,\mathrm{d}t \\ &\leq \frac{\|f\|_1 + |w_f(x_0)|}{(n+1)\sin^2(\pi\delta)}.\end{aligned}$$

Choose N_0 such that

$$\frac{\|f\|_1 + |w_f(x_0)|}{(N_0 + 1)\sin^2(\pi\delta)} < \varepsilon.$$

Then the contribution from the middle interval is less than ε for every $n \geq N_0$. Combining the contributions from the two neighborhoods of 0 and the middle interval, we have

$$|\sigma_n(f)(x_0) - w_f(x_0)| \leq 2\varepsilon + \varepsilon = 3\varepsilon$$

for all $n \geq N_0$, and the result follows. □

Corollary 20.10

1. If $f \in L^1(\mathbb{T})$ is continuous at x_0, then $\sigma_n(f)(x_0) \to f(x_0)$.
2. If $\lim_{h\to 0^+} f(x_0 + h) = f(x_0^+)$ and $\lim_{h\to 0^+} f(x_0 - h) = f(x_0^-)$ exist, then

$$\sigma_n(f)(x_0) \to (f(x_0^+) + f(x_0^-))/2.$$

Proof. If f is continuous at x_0, then $w_f(x_0) = f(x_0)$ by Theorem 20.9. If $f(x_0^+)$ and $f(x_0^-)$ exist, then $w_f(x_0) = (f(x_0^+) + f(x_0^-))/2$. □

Lecture 21 — Thursday, March 19, 2015

Theorem 21.1 If $S_n f(x)$ converges, then it converges to the same limit as $\sigma_n(f)(x)$.

Proof. Suppose $S_n(f)(t_0) \to g(t_0)$. Since

$$\sigma_n(f)(t_0) = \frac{1}{n+1} \sum_{j=0}^{n} S_j(f)(t_0),$$

we estimate

$$\begin{aligned}
|\sigma_n(f)(t_0) - g(t_0)| &= \left| \frac{S_0(f)(t_0) + \cdots + S_n(f)(t_0)}{n+1} - g(t_0) \right| \\
&= \frac{1}{n+1} \left| \sum_{j=0}^{n} (S_j(f)(t_0) - g(t_0)) \right| \\
&\leq \frac{1}{n+1} \left[\sum_{j=0}^{N-1} |S_j(f)(t_0) - g(t_0)| + \sum_{j=N}^{n} |S_j(f)(t_0) - g(t_0)| \right]
\end{aligned}$$

Let N be such that $|S_j(f)(t_0) - g(t_0)| < \varepsilon$ for all $j \geq N$. Then

$$|\sigma_n(f)(t_0) - g(t_0)| \leq \frac{1}{n+1} \sum_{j=0}^{N-1} |S_j(f)(t_0) - g(t_0)| + \frac{n-N+1}{n+1} \varepsilon$$

and the first term tends to 0 as $n \to \infty$ while the second is at most ε. So for n large enough,

$$|\sigma_n(f)(t_0) - g(t_0)| \leq 2\varepsilon.$$

Hence $\sigma_n(f)(t_0) \to g(t_0)$. □

Corollary 21.2 Let $f \in L^1(\mathbb{T})$. If f is continuous at t_0, or if f has one-sided limits at t_0, and if $\{S_n(f)(t_0)\}$ converges, then

$$S_n(f)(t_0) \to \frac{f(t_0^+) + f(t_0^-)}{2}.$$

Proof. This follows from Theorems 20.9 and 21.1. □

Theorem 21.3 If $f \in L^1(\mathbb{T})$ and is differentiable at t_0, then $S_n(f)(t_0) \to f(t_0)$ as $n \to \infty$.

Proof. Since $\int_0^1 D_n(t)\,\mathrm{d}t = 1$, we have

$$\begin{aligned} |S_n(f)(t_0) - f(t_0)| &= \left| \int (f(t_0 - t) - f(t_0)) D_n(t)\,\mathrm{d}t \right| \\ &= \left| \int \frac{f(t_0 - t) - f(t_0)}{\sin(\pi t)} \sin(\pi(2n+1)t)\,\mathrm{d}t \right| \end{aligned}$$

It remains to prove that

$$g(t) = \frac{f(t_0 - t) - f(t_0)}{\sin(\pi t)}$$

belongs to $L^1(\mathbb{T})$. Since

$$\lim_{t \to 0} \frac{f(t_0 - t) - f(t_0)}{t} = -f'(t_0) \qquad \text{and} \qquad \lim_{t \to 0} \frac{t}{\sin(\pi t)} = \frac{1}{\pi},$$

the function g is bounded near 0. By periodicity, the same argument shows that g is bounded near 1. On every compact subinterval of $(0, 1)$, the denominator is bounded away from 0, so g is integrable there. Hence $g \in L^1(\mathbb{T})$.

$$\int_0^1 |g(t)|\,\mathrm{d}t = \int_0^\delta |g(t)|\,\mathrm{d}t + \int_\delta^{1-\delta} |g(t)|\,\mathrm{d}t + \int_{1-\delta}^1 |g(t)|\,\mathrm{d}t$$

Choose $\delta > 0$ such that $|g(t)| < |f'(t_0)|/\pi + 1$ whenever $0 < t < \delta$ or $1 - \delta < t < 1$. Then

$$\int_0^1 |g(t)|\,\mathrm{d}t < \infty.$$

Finally,

$$\sin(\pi(2n+1)t) = \sin(2\pi nt)\cos(\pi t) + \cos(2\pi nt)\sin(\pi t).$$

Both $g(t)\cos(\pi t)$ and $g(t)\sin(\pi t)$ belong to $L^1(\mathbb{T})$, so the Riemann–Lebesgue lemma, applied at the integer frequency n to these two functions, gives

$$\int_0^1 g(t)\sin(\pi(2n+1)t)\,\mathrm{d}t \to 0.$$

Hence $S_n(f)(t_0) \to f(t_0)$. $\square$

Remark 21.4 In the proof of Theorem 21.3, we only used differentiability to prove $g \in L^1(\mathbb{T})$.

Corollary 21.5 If f is constant on an interval, then $S_n(f)(t_0)$ converges to $f(t_0)$ for all t_0 in the interior of the interval.

Proof. If f is constant on an interval, then f is differentiable at every point in the interior of the interval, with derivative 0. Hence $S_n(f)(t_0) \to f(t_0)$ for every t_0 in the interior of the interval by Theorem 21.3. $\square$

Definition 21.6 A function f satisfies the **Lipschitz condition of order $\boldsymbol{\alpha}$** at a point t_0, where $\alpha \geq 0$, if there exist constants $c > 0$ and $\delta > 0$ such that

$$|f(t_0 - t) - f(t_0)| < c|t|^\alpha$$

whenever $|t| < \delta$.

Remark 21.7

1. It is harder to be Lipschitz for larger α: if f is Lipschitz of order α at t_0, then f is Lipschitz of order β at t_0 for every $0 \leq \beta \leq \alpha$.
2. f being Lipschitz of order α at t_0 with $\alpha = 0$ means that f is bounded close to t_0.
3. If f is Lipschitz of order α at t_0 for some $\alpha > 0$, then f is continuous at t_0. Indeed, if f is Lipschitz of order α at t_0, then there are constants $c > 0$ and $\delta > 0$ such that

$$|f(t_0 - t) - f(t_0)| < c|t|^\alpha$$

whenever $|t| < \delta$. Letting $t \to 0$ gives $f(t_0 - t) \to f(t_0)$, so f is continuous at t_0.

4. If f is differentiable at t_0, then f is Lipschitz of order 1. Indeed, differentiability gives a $\delta > 0$ such that
$$|f(t_0 - t) - f(t_0)| \leq (|f'(t_0)| + 1)|t|$$
whenever $|t| < \delta$. Thus differentiability at t_0 implies the Lipschitz condition of order 1 at t_0, which implies the Lipschitz condition of order α for every $0 \leq \alpha \leq 1$, and in particular implies continuity at t_0.

We can generalize Theorem 21.3 as follows.

Theorem 21.8 If $f \in L^1(\mathbb{T})$ and is Lipschitz of order α for some $\alpha > 0$ at t_0, then $S_n(f)(t_0) \to f(t_0)$.

Proof. As in the proof of Theorem 21.3, it is enough to show that
$$g(t) = \frac{f(t_0 - t) - f(t_0)}{\sin(\pi t)}$$
belongs to $L^1(\mathbb{T})$. Since f is Lipschitz of order α at t_0, there are constants $c > 0$ and $\delta > 0$ such that
$$|f(t_0 - t) - f(t_0)| \leq c|t|^\alpha$$
whenever $|t| < \delta$. Decreasing δ if necessary, we may assume that $0 < \delta < 1/2$ and that
$$|\sin(\pi t)| \geq 2|t|$$
whenever $|t| < \delta$. Hence, for $0 < t < \delta$,
$$|g(t)| \leq \frac{c}{2} t^{\alpha - 1}.$$
The same estimate holds near 1, with t replaced by $1 - t$, because $t = 1$ represents the same point as $t = 0$ on $\mathbb{T}$. On the interval $[\delta, 1 - \delta]$, the denominator is bounded away from 0, and $f(t_0 - t) - f(t_0)$ is integrable. Therefore
$$\int_0^1 |g(t)| \,\mathrm{d}t < \infty,$$

since $\displaystyle\int_0^\delta t^{\alpha-1}\,\mathrm{d}t < \infty$ when $\alpha > 0$. Using

$$\sin(\pi(2n+1)t) = \sin(2\pi nt)\cos(\pi t) + \cos(2\pi nt)\sin(\pi t)$$

and applying the Riemann–Lebesgue lemma to $g(t)\cos(\pi t)$ and $g(t)\sin(\pi t)$ now gives

$$S_n(f)(t_0) - f(t_0) = \int_0^1 g(t)\sin(\pi(2n+1)t)\,\mathrm{d}t \longrightarrow 0.$$

□

Definition 21.9 We say that f has a **derivative at t_0 from the right** if

$$\lim_{h\to 0^+} \frac{f(t_0+h) - f(t_0^+)}{h}$$

exists, where $f(t_0^+) = \lim_{h\to 0^+} f(t_0+h)$. Similarly, we say that f has a **derivative at t_0 from the left** if

$$\lim_{h\to 0^+} \frac{f(t_0-h) - f(t_0^-)}{-h}$$

exists, where $f(t_0^-) = \lim_{h\to 0^+} f(t_0-h)$.

Corollary 21.10 If $f \in L^1(\mathbb{T})$ is continuous at t_0 and has left and right derivatives at t_0, then $S_n(f)(t_0) \to f(t_0)$ as $n \to \infty$.

Proof. Since f is continuous at t_0, we have $f(t_0^+) = f(t_0^-) = f(t_0)$. Also, f is differentiable from the left and right at t_0.

The limits

$$\lim_{h\to 0^+} \frac{f(t_0+h) - f(t_0)}{h} \qquad \text{and} \qquad \lim_{h\to 0^+} \frac{f(t_0-h) - f(t_0)}{-h}$$

exist. Hence f is Lipschitz of order 1, and so Theorem 21.8 gives $S_n(f)(t_0) \to f(t_0)$. □

Example 21.11 Let $f(x) = x^2 - x$ on $[0,1]$. Since $f(0) = 0 = f(1)$, this defines a continuous periodic function on $\mathbb{T}$. At $t_0 = 0$, the right derivative is

$$(2x-1)|_{x=0} = -1,$$

while the left derivative, computed using the value at $x = 1$, is

$$(2x-1)|_{x=1} = 1.$$

Thus f is not differentiable at 0, but it has left and right derivatives there. By Theorem 21.8, $S_n(f)(0) \to f(0) = 0$. Therefore

$$\sum_{n\in\mathbb{Z}} \hat{f}(n) = 0 \tag{21.1}$$

by Corollary 21.2. For $n \neq 0$, we compute $\hat{f}(n)$ by integration by parts:

$$\begin{aligned}
\hat{f}(n) &= \int_0^1 (x^2-x)e^{-2\pi i n x}\,\mathrm{d}x \\
&= \left.\frac{(x^2-x)e^{-2\pi i n x}}{-2\pi i n}\right|_0^1 - \int_0^1 \frac{(2x-1)e^{-2\pi i n x}}{-2\pi i n}\,\mathrm{d}x \\
&= \frac{1}{2\pi i n}\int_0^1 (2x-1)e^{-2\pi i n x}\,\mathrm{d}x \\
&= \left.\frac{(2x-1)e^{-2\pi i n x}}{-4\pi^2 n^2}\right|_0^1 - \int_0^1 \frac{2e^{-2\pi i n x}}{-4\pi^2 n^2}\,\mathrm{d}x \\
&= \frac{1}{2\pi^2 n^2}.
\end{aligned}$$

On the other hand,

$$\hat{f}(0) = \int_0^1 (x^2-x)\,\mathrm{d}x = \left[\frac{x^3}{3} - \frac{x^2}{2}\right]_0^1 = \frac{1}{3} - \frac{1}{2} = -\frac{1}{6}.$$

Inserting these into (21.1) gives

$$\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{n^2} = \frac{2\pi^2}{6},$$

and therefore

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Surprise!

Remark 21.12 There is a sneakier way to arrive at the solution to the Basel problem. Take $f(x) = x$ on $[0, 1]$. Then

$$\|f\|_2^2 = \int_0^1 x^2 \, dx = \frac{1}{3}.$$

For all $n \neq 0$,

$$\hat{f}(n) = \int_0^1 xe^{-2\pi inx} \, dx = \frac{xe^{-2\pi inx}}{-2\pi in}\bigg|_0^1 - \int_0^1 \frac{e^{-2\pi inx}}{-2\pi in} \, dx = \frac{-1}{2\pi in}.$$

We also have

$$\hat{f}(0) = \int_0^1 x \, dx = \frac{1}{2}.$$

By Parseval's identity,

$$\sum_{n\in\mathbb{Z}\setminus\{0\}} \left|\frac{-1}{2\pi in}\right|^2 + \left|\frac{1}{2}\right|^2 = \frac{1}{3}.$$

Therefore

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Lecture 22 — Tuesday, March 24, 2015

5. Fourier Analysis on $\mathbb{R}$ and Applications

Fourier Transform on $\mathbb{R}$

The major difference between the circle and the real line is that $L^p(\mathbb{T}) \subseteq L^1(\mathbb{T})$, whereas $L^p(\mathbb{R}) \not\subseteq L^1(\mathbb{R})$ for $p \neq 1$. Recall that $L^1(\mathbb{R})$ is the space of Lebesgue measurable functions on $\mathbb{R}$ with

$$\|f\|_1 = \int_{\mathbb{R}} |f| \, dx.$$

For $f, g \in L^1(\mathbb{R})$, define their convolution by

$$(f * g)(x) = \int_{\mathbb{R}} f(y)g(x - y) \, dy.$$

As on $\mathbb{T}$, we can prove that the function $y \mapsto f(y)g(x-y)$ is integrable for almost every x. Moreover, $f * g \in L^1(\mathbb{R})$ and $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$.

Definition 22.1 For $\xi \in \mathbb{R}$, define $\chi_\xi : \mathbb{R} \to \mathbb{T}$ by $\chi_\xi(x) = e^{2\pi i \xi x}$.

We have

$$\chi_\xi(x+y) = \chi_\xi(x)\chi_\xi(y) \qquad \text{and} \qquad \chi_\xi(0) = 1.$$

Moreover,

$$|\chi_\xi(x) - \chi_\xi(y)| \leq 2\pi|\xi|\,|x-y|,$$

so χ_ξ is uniformly continuous.

Definition 22.2 A continuous function $\phi : \mathbb{R} \to \mathbb{T}$ is called a **character** if ϕ is a homomorphism: $\phi(x+y) = \phi(x)\phi(y)$.

Proposition 22.3 If ϕ is a character, then there exists $\xi \in \mathbb{R}$ such that $\phi = \chi_\xi$.

Proof. If ϕ is a character, then

$$\phi(0) = \phi(0+0) = \phi(0) \cdot \phi(0).$$

Thus $\phi(0) = 1$. By continuity, there exists $\varepsilon > 0$ such that

$$\phi((-\varepsilon, \varepsilon)) \subseteq \{w \in \mathbb{C} : \mathrm{Re}(w) > 0\}.$$

Write $\phi(\varepsilon) = e^{2\pi i \varepsilon y}$, where y is the unique element of $[-1/(4\varepsilon), 1/(4\varepsilon)]$ with this property. Since $\phi(\varepsilon/2)^2 = \phi(\varepsilon)$ and $\phi(\varepsilon/2)$ lies in the right half-plane, we have $\phi(\varepsilon/2) = e^{2\pi i \varepsilon y/2}$. Iterating gives

$$\phi(\varepsilon/2^n) = e^{2\pi i \varepsilon y/2^n}$$

for every $n \in \mathbb{N}$. Therefore, for $m \in \mathbb{Z}$,

$$\phi(m\varepsilon/2^n) = \phi(\varepsilon/2^n)^m = e^{2\pi i m \varepsilon y/2^n}.$$

The dyadic multiples $m\varepsilon/2^n$ are dense in $\mathbb{R}$, and continuity of ϕ extends this formula to every $x \in \mathbb{R}$. Hence $\phi(x) = e^{2\pi i y x}$ for every x, so $\phi = \chi_y$. $\square$

Remark 22.4 $\chi_\xi \notin L^p(\mathbb{R})$ for $1 \le p < \infty$.

Definition 22.5 Let $f \in L^1(\mathbb{R})$ and $\xi \in \mathbb{R}$. The **Fourier transform** of f at ξ is

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x)\overline{\chi_\xi(x)}\,\mathrm{d}x = \int_{\mathbb{R}} f(x)e^{-2\pi i\xi x}\,\mathrm{d}x.$$

Observe that

$$|\hat{f}(\xi)| \le \int_{\mathbb{R}} |f(x)| \cdot |e^{-2\pi i\xi x}|\,\mathrm{d}x = \int_{\mathbb{R}} |f(x)|\,\mathrm{d}x < \infty,$$

so $\hat{f}(\xi)$ is well-defined.

Proposition 22.6 If $f \in L^1(\mathbb{R})$, then $\hat{f}$ is uniformly continuous.

Proof. For every $\xi, h \in \mathbb{R}$,

$$|\hat{f}(\xi+h) - \hat{f}(\xi)| \le \int_{\mathbb{R}} |f(x)|\,|e^{-2\pi i h x} - 1|\,\mathrm{d}x.$$

The right-hand side is independent of ξ and tends to zero as $h \to 0$ by the dominated convergence theorem, since it is bounded by $2|f(x)|$. Hence $\hat{f}$ is uniformly continuous. □

Basic Properties and the Riemann–Lebesgue Lemma

Proposition 22.7 Let $f, g \in L^1(\mathbb{R})$ and let $\alpha \in \mathbb{C}$. The Fourier transform has the following properties:

1. $\widehat{f+g} = \hat{f} + \hat{g}$.
2. $\widehat{\alpha f} = \alpha\hat{f}$.
3. $\widehat{f * g} = \hat{f} \cdot \hat{g}$.
4. If $f_y(x) := f(x-y)$, then $\widehat{f_y}(\xi) = \chi_\xi(-y)\hat{f}(\xi)$.
5. If $\tilde{f}(x) := f(-x)$, then $\widehat{\tilde{f}}(\xi) = \hat{f}(-\xi)$.

Proof. For (1), linearity of the integral gives

$$\widehat{f+g}(\xi) = \int_{\mathbb{R}} (f(x)+g(x))e^{-2\pi i\xi x}\,\mathrm{d}x = \hat{f}(\xi) + \hat{g}(\xi).$$

For (2), again by linearity,

$$\widehat{\alpha f}(\xi) = \int_{\mathbb{R}} \alpha f(x)e^{-2\pi i\xi x}\,\mathrm{d}x = \alpha\hat{f}(\xi).$$

For (3), since $f, g \in L^1(\mathbb{R})$, Fubini's theorem applies:

$$\widehat{f*g}(\xi) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t)g(x-t)\,\mathrm{d}t\right) e^{-2\pi i\xi x}\,\mathrm{d}x = \int_{\mathbb{R}} f(t)\left(\int_{\mathbb{R}} g(x-t)e^{-2\pi i\xi x}\,\mathrm{d}x\right)\mathrm{d}t.$$

Set $u = x - t$. Then

$$\int_{\mathbb{R}} g(x-t)e^{-2\pi i\xi x}\,\mathrm{d}x = e^{-2\pi i\xi t}\int_{\mathbb{R}} g(u)e^{-2\pi i\xi u}\,\mathrm{d}u = e^{-2\pi i\xi t}\hat{g}(\xi).$$

Therefore,

$$\widehat{f*g}(\xi) = \hat{g}(\xi)\int_{\mathbb{R}} f(t)e^{-2\pi i\xi t}\,\mathrm{d}t = \hat{f}(\xi)\hat{g}(\xi).$$

For (4), with $u = x - y$,

$$\begin{aligned}
\widehat{f_y}(\xi) &= \int_{\mathbb{R}} f(x-y)e^{-2\pi i\xi x}\,\mathrm{d}x \\
&= \int_{\mathbb{R}} f(u)e^{-2\pi i\xi(u+y)}\,\mathrm{d}u \\
&= e^{-2\pi i\xi y}\int_{\mathbb{R}} f(u)e^{-2\pi i\xi u}\,\mathrm{d}u \\
&= \chi_\xi(-y)\hat{f}(\xi).
\end{aligned}$$

For (5), with $u = -x$,

$$\widehat{\tilde{f}}(\xi) = \int_{\mathbb{R}} f(-x)e^{-2\pi i\xi x}\,\mathrm{d}x = \int_{\mathbb{R}} f(u)e^{2\pi i\xi u}\,\mathrm{d}u = \hat{f}(-\xi).$$

$\square$

Lemma 22.8 Let $f \in L^1(\mathbb{R})$ be continuous and let

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

If $F \in L^1(\mathbb{R})$, then $\widehat{F}(\xi) = \hat{f}(\xi)/(2\pi i\xi)$ for every $\xi \neq 0$.

Proof. Integration by parts gives

$$(\widehat{F})(\xi) = F(x)\frac{e^{-2\pi i\xi x}}{-2\pi i\xi}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{e^{-2\pi i\xi x}}{2\pi i\xi} f(x)\,\mathrm{d}x.$$

The definition gives $F(x) \to 0$ as $x \to -\infty$ and $F(x) \to \int_{\mathbb{R}} f$ as $x \to \infty$. Since $F \in L^1(\mathbb{R})$ and the latter limit exists, it must also equal 0. Thus the boundary term vanishes, and

$$\widehat{F}(\xi) = \frac{\hat{f}(\xi)}{2\pi i\xi}.$$

□

Corollary 22.9 If f is continuously differentiable and $f, f' \in L^1(\mathbb{R})$, then

$$\widehat{f'}(\xi) = 2\pi i\xi\,\hat{f}(\xi).$$

Proof. Because $f' \in L^1(\mathbb{R})$, the function f has finite limits at both ends of the real line; because $f \in L^1(\mathbb{R})$, both limits are 0. Hence

$$f(x) = \int_{-\infty}^{x} f'(t)\,\mathrm{d}t.$$

Apply Lemma 22.8 with f' in place of the lemma's f and with this antiderivative f in place of F. □

Theorem 22.10 (Riemann–Lebesgue lemma: real version) Let $f \in L^1(\mathbb{R})$. Then $\lim_{|\xi|\to\infty} \hat{f}(\xi) = 0$.

Proof. First assume that f is compactly supported and continuously differentiable. By Corollary 22.9,

$$\widehat{f'}(\xi) = 2\pi i\xi\,\hat{f}(\xi).$$

Thus

$$|2\pi i\xi \hat{f}(\xi)| = |\widehat{f'}(\xi)| = \left|\int_{\mathbb{R}} f'(x)e^{-2\pi i\xi x}\,\mathrm{d}x\right| \le \int_{\mathbb{R}} |f'(x)|\,\mathrm{d}x = \|f'\|_1.$$

It follows that $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$ in this case.

Now let $f \in L^1(\mathbb{R})$ be arbitrary, and let $\varepsilon > 0$. Choose a compactly supported continuously differentiable function g such that $\|f - g\|_1 < \varepsilon$. Then

$$|\hat{f}(\xi) - \hat{g}(\xi)| = |\widehat{f-g}(\xi)| \le \|f - g\|_1 < \varepsilon,$$

and the first part gives $\hat{g}(\xi) \to 0$ as $|\xi| \to \infty$. Hence

$$\limsup_{|\xi|\to\infty} |\hat{f}(\xi)| \le \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$. □

Definition 22.11 Let $C_0(\mathbb{R})$ be the space of complex-valued continuous functions on $\mathbb{R}$ which vanish at infinity. The **Fourier transform** is the map

$$\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R}), \qquad f \mapsto \hat{f}.$$

	In $L^1(\mathbb{R})$	In $C_0(\mathbb{R})$
Addition	$(f+g)(x) = f(x) + g(x)$	$(f+g)(x) = f(x) + g(x)$
Multiplication	$f * g$ (convolution)	$(fg)(x) = f(x) \cdot g(x)$
Norm	$\|f\|_1$	$\|f\|_\infty$
Involution	f^*	$\overline{f}$

Summability Kernels and Fourier Inversion

Definition 22.12 A family of continuous functions $\{k_\lambda\}_{\lambda\in I} \subseteq L^1(\mathbb{R})$ is called a **summability kernel** if it satisfies the following properties:

1. $\displaystyle\int_{\mathbb{R}} k_\lambda(x)\,\mathrm{d}x = 1$ for every λ.
2. $\sup_{\lambda\in I}\|k_\lambda\|_1 < \infty$.
3. $\lim_{\lambda\to\infty}\displaystyle\int_{|x|>\delta} |k_\lambda(x)|\,\mathrm{d}x = 0$ for all $\delta > 0$.

Example 22.13 For a continuous function $k \in L^1(\mathbb{R})$ and $\lambda > 0$, define $k_\lambda(x) = \lambda k(\lambda x)$. If

$$\int_{\mathbb{R}} k(x)\,\mathrm{d}x = 1,$$

then $\{k_\lambda\}_{\lambda>0}$ is a summability kernel.

Proof. For $\lambda > 0$,

$$\int_{\mathbb{R}} k_\lambda(x)\,\mathrm{d}x = \int_{\mathbb{R}} \lambda k(\lambda x)\,\mathrm{d}x = \int_{\mathbb{R}} k(u)\,\mathrm{d}u = 1$$

by the change of variables $u = \lambda x$. Also

$$\|k_\lambda\|_1 = \int_{\mathbb{R}} \lambda|k(\lambda x)|\,\mathrm{d}x = \int_{\mathbb{R}} |k(u)|\,\mathrm{d}u = \|k\|_1,$$

so the L^1 norms are uniformly bounded. Finally, for any $\delta > 0$,

$$\int_{|x|>\delta} |k_\lambda(x)|\,\mathrm{d}x = \int_{|u|>\lambda\delta} |k(u)|\,\mathrm{d}u \xrightarrow[\lambda\to\infty]{} 0$$

by absolute integrability of k. These are exactly the summability kernel conditions. $\square$

Lecture 23 — Thursday, March 26, 2015

Plancherel and the Inverse Transform

Example 23.1 Two useful summability kernels are the following.

1. The **Fejér kernel** is
$$F_\omega(x) = \int_{-\omega}^{\omega} \left(1 - \frac{|\xi|}{\omega}\right) e^{2\pi i \xi x}\,\mathrm{d}\xi.$$
The family $\{F_\omega\}_{\omega>0}$ is a summability kernel, and
$$\begin{aligned} F_\omega * f(x) &= \int_{\mathbb{R}} f(y) F_\omega(x-y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}} f(y) \left(\int_{-\omega}^{\omega} \left(1 - \frac{|\xi|}{\omega}\right) e^{2\pi i \xi (x-y)}\,\mathrm{d}\xi\right)\mathrm{d}y \\ &= \int_{-\omega}^{\omega} \left(1 - \frac{|\xi|}{\omega}\right) \hat{f}(\xi) e^{2\pi i \xi x}\,\mathrm{d}\xi. \end{aligned}$$

2. The **Gauss kernel** is
$$G_\omega(x) = \int_{\mathbb{R}} e^{-\pi\xi^2/\omega^2} e^{2\pi i \xi x}\,\mathrm{d}\xi.$$
In this case, $G_\omega(x) = \omega e^{-\pi\omega^2 x^2}$. Then $\{G_\omega\}$ is a summability kernel, and
$$\begin{aligned} G_\omega * f(x) &= \int_{\mathbb{R}} f(y) G_\omega(x-y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}} f(y) \left(\int_{\mathbb{R}} e^{-\pi\xi^2/\omega^2} e^{2\pi i \xi (x-y)}\,\mathrm{d}\xi\right)\mathrm{d}y \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y) e^{-2\pi i \xi y}\,\mathrm{d}y\right) e^{-\pi\xi^2/\omega^2} e^{2\pi i \xi x}\,\mathrm{d}\xi \\ &= \int_{\mathbb{R}} e^{-\pi\xi^2/\omega^2} \hat{f}(\xi) e^{2\pi i \xi x}\,\mathrm{d}\xi. \end{aligned}$$

Theorem 23.2 Let $\{k_\alpha\}_{\alpha\in I}$ be a summability kernel. Then the following hold:

1. If $f \in L^1(\mathbb{R})$, then $k_\alpha * f \xrightarrow{\alpha\to\infty} f$ in $L^1(\mathbb{R})$.
2. If $f \in C_0(\mathbb{R})$, then $k_\alpha * f \to f$ in $C_0(\mathbb{R})$.
3. If $f \in L^p(\mathbb{R})$, $1 \le p < \infty$, then $k_\alpha * f \xrightarrow{\alpha\to\infty} f$ in $L^p(\mathbb{R})$.

Proof. The proofs of (1) and (2) are identical to the corresponding arguments on $\mathbb{T}$, with the middle interval replaced by the complement of $[-\delta, \delta]$.

For (3), first recall that translations are continuous in L^p: if $\tau_y f(x) := f(x-y)$, then

$$\|\tau_y f - f\|_p \longrightarrow 0 \qquad (y \to 0).$$

Indeed, this is immediate for $f \in C_c(\mathbb{R})$ by uniform continuity and compact support; the general case follows by approximating f in L^p by such a function and using $\|\tau_y h\|_p = \|h\|_p$.

Let $M := \sup_\alpha \|k_\alpha\|_1$. Since $\int k_\alpha = 1$, Minkowski's integral inequality gives

$$\begin{aligned}\|k_\alpha * f - f\|_p &= \left\| \int_{\mathbb{R}} k_\alpha(y)\big(\tau_y f - f\big)\,\mathrm{d}y \right\|_p \\ &\leq \int_{\mathbb{R}} |k_\alpha(y)|\, \|\tau_y f - f\|_p \,\mathrm{d}y.\end{aligned}$$

Given $\varepsilon > 0$, choose $\delta > 0$ so that $\|\tau_y f - f\|_p < \varepsilon/(2M)$ whenever $|y| < \delta$. Since $\|\tau_y f - f\|_p \leq 2\|f\|_p$, splitting the last integral gives

$$\|k_\alpha * f - f\|_p \leq \frac{\varepsilon}{2M}\|k_\alpha\|_1 + 2\|f\|_p \int_{|y|\geq\delta} |k_\alpha(y)|\,\mathrm{d}y.$$

The first term is at most $\varepsilon/2$, and the second tends to 0 by concentration of the kernel. Hence $k_\alpha * f \to f$ in $L^p(\mathbb{R})$. □

As a consequence, if $f \in L^1(\mathbb{R})$, then

$$F_\omega * f(x) = \int_{-\omega}^{\omega} \left(1 - \frac{|\xi|}{\omega}\right) \hat{f}(\xi) e^{2\pi i \xi x}\,\mathrm{d}\xi \longrightarrow f$$

in $L^1(\mathbb{R})$.

Corollary 23.3 The map $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is injective.

Proof. Suppose $f \in L^1(\mathbb{R})$ with $\hat{f}(\xi) = 0$ for all $\xi \in \mathbb{R}$. Then $F_\omega * f = 0$ for all $\omega > 0$. Then $f = \lim_{\omega\to\infty} F_\omega * f = 0$ in $L^1(\mathbb{R})$, so $f = 0$. □

We next restrict the Fourier transform to $L^2(\mathbb{R})$. Initially this gives

$$\mathcal{F} : L^2(\mathbb{R}) \cap L^1(\mathbb{R}) \to C_0(\mathbb{R}).$$

The space $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ contains all continuous compactly supported functions on $\mathbb{R}$, and is therefore $\|\cdot\|_2$-dense in $L^2(\mathbb{R})$.

Recall that $L^2(\mathbb{R})$ is a Hilbert space with inner product

$$\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}\,\mathrm{d}x$$

and norm $\|f\|_2 = \sqrt{\langle f, f\rangle}$.

Theorem 23.4 The map $\mathcal{F} : L^2(\mathbb{R}) \cap L^1(\mathbb{R}) \to C_0(\mathbb{R})$ extends to an isometric isomorphism $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$. In particular, we have

$$\|\mathcal{F}f\|_2 = \|f\|_2 \quad (\textbf{Plancherel's identity})$$

and

$$\langle f, g\rangle = \langle \mathcal{F}f, \mathcal{F}g\rangle \quad (\textbf{Parseval's identity}).$$

Proof. Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Recall that $G_\omega * f \to f$ in $L^2(\mathbb{R})$ since $\{G_\omega\}$ is a summability kernel. Then

$$\begin{aligned}
\|f\|_2^2 &= \langle f, f\rangle \\
&= \lim_{\omega\to\infty} \langle f, G_\omega * f\rangle \\
&= \lim_{\omega\to\infty} \int f(x)\overline{G_\omega * f(x)}\,\mathrm{d}x \\
&= \lim_{\omega\to\infty} \int \overline{\hat{f}(\xi)}\hat{f}(\xi)e^{-\pi\xi^2/\omega^2}\,\mathrm{d}\xi \\
&= \int |\hat{f}(\xi)|^2\,\mathrm{d}\xi \\
&= \|\mathcal{F}f\|_2^2,
\end{aligned}$$

which establishes Plancherel's identity for f. Thus $\mathcal{F}$ is an isometry on $L^1 \cap L^2(\mathbb{R})$.

The subspace $L^1 \cap L^2(\mathbb{R})$ is $\|\cdot\|_2$-dense in $L^2(\mathbb{R})$. Because $\mathcal{F} : L^1 \cap L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is an isometry, it extends uniquely to an isometry $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$. In particular, it is injective. Parseval's identity now follows from the **polarization identity** applied to the isometry:

$$\langle f, g\rangle = \frac{1}{4}\left(\|f+g\|_2^2 - \|f-g\|_2^2 + i\|f+ig\|_2^2 - i\|f-ig\|_2^2\right),$$

and the same formula with f, g replaced by $\mathcal{F}f, \mathcal{F}g$ gives

$$\langle f, g\rangle = \langle \mathcal{F}f, \mathcal{F}g\rangle.$$

Surjectivity will follow from the Fourier inversion formula, which we will prove next. $\square$

Definition 23.5 For $g \in L^1(\mathbb{R})$, define the **inverse Fourier transform**

$$(\mathcal{J}g)(x) = \int_{\mathbb{R}} g(\xi) e^{2\pi i \xi x} \, \mathrm{d}\xi.$$

Theorem 23.6 (Fourier inversion formula)

1. If $f, \hat{f} \in L^1(\mathbb{R})$, then f agrees almost everywhere with the continuous function $\mathcal{J}\hat{f}$.
2. If $g, \mathcal{J}g \in L^1(\mathbb{R})$, then g agrees almost everywhere with the continuous function $\mathcal{F}(\mathcal{J}g)$.

Thus, the Fourier transform $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is injective, and its inverse is given by $\mathcal{J}$ on the range of $\mathcal{F}$, so the following diagram commutes:

$$\begin{tikzcd}
L^1(\mathbb{R}) \arrow[r, "\mathcal{F}"] & C_0(\mathbb{R}) \\
\mathcal{F}(L^1(\mathbb{R})) \arrow[u, "\mathcal{F}^{-1}"] \arrow[r, "\mathcal{J}"] & \mathcal{J}(\mathcal{F}(L^1(\mathbb{R}))) \arrow[u, "\mathcal{F}"']
\end{tikzcd}$$

Proof. For (1), let $f, \hat{f} \in L^1(\mathbb{R})$. Recall that $F_\omega * f \to f$ in $L^1(\mathbb{R})$. Choose a sequence $\omega_k \to \infty$ along which $F_{\omega_k} * f \to f$ pointwise almost everywhere. For each $x \in \mathbb{R}$,

$$\begin{aligned}
\lim_{k\to\infty} (F_{\omega_k} * f)(x) &= \lim_{k\to\infty} \int_{-\omega_k}^{\omega_k} \left(1 - \frac{|\xi|}{\omega_k}\right) \hat{f}(\xi) e^{2\pi i \xi x} \, \mathrm{d}\xi \\
&= \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i \xi x} \, \mathrm{d}\xi
\end{aligned}$$

$$= (\mathcal{J}\hat{f})(x)$$

by the dominated convergence theorem. Along the chosen sequence, the left-hand side equals $f(x)$ almost everywhere. Thus $f = \mathcal{J}\hat{f}$ almost everywhere. Since the inverse transform of an L^1 function belongs to $C_0(\mathbb{R})$, this gives the claimed continuous representative.

The proof of (2) is the same argument with the signs in the exponent reversed.

We can now finish the surjectivity part of Theorem 23.4. If $g \in C_c^\infty(\mathbb{R})$, then $\mathcal{J}g$ is a Schwartz function, so $\mathcal{J}g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and (2) gives

$$\mathcal{F}(\mathcal{J}g) = g.$$

Thus the range of the L^2 Fourier transform contains $C_c^\infty(\mathbb{R})$, which is dense in $L^2(\mathbb{R})$. The range of an isometry from a complete space is closed, so the range is all of $L^2(\mathbb{R})$. □

Corollary 23.7 $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is a norm-decreasing (that is, $\|\hat{f}\|_\infty \le \|f\|_1$) injective algebra homomorphism (that is, $\widehat{f+g} = \hat{f} + \hat{g}$ and $\widehat{f * g} = \hat{f} \cdot \hat{g}$) with dense range.

Proof. The norm estimate is immediate:

$$\|\hat{f}\|_\infty = \sup_{\xi \in \mathbb{R}} \left| \int_{\mathbb{R}} f(x) e^{-2\pi i \xi x}\, \mathrm{d}x \right| \le \int_{\mathbb{R}} |f(x)|\, \mathrm{d}x = \|f\|_1.$$

Hence $\mathcal{F}$ is norm-decreasing. Linearity follows from linearity of the integral, and the convolution identity proved earlier gives

$$\widehat{f * g} = \hat{f}\,\hat{g},$$

so $\mathcal{F}$ is an algebra homomorphism from convolution on $L^1(\mathbb{R})$ to pointwise multiplication on $C_0(\mathbb{R})$. Injectivity is Corollary 23.3.

For density of the range, let $g \in C_c^\infty(\mathbb{R})$. Then $\mathcal{J}g$ is a Schwartz function and therefore belongs to $L^1(\mathbb{R})$. By Theorem 23.6(2),

$$\mathcal{F}(\mathcal{J}g) = g.$$

Hence the range contains $C_c^\infty(\mathbb{R})$, which is dense in $C_0(\mathbb{R})$ in the uniform norm. □

Lecture 24 — Thursday, April 02, 2015

The Heat Equation

Let us solve the heat equation on $\mathbb{R}$ using the Fourier transform. This is a partial differential equation which models the flow of heat in a medium.

For the following calculation, assume that $u(\cdot, t)$ is sufficiently regular and decays sufficiently rapidly in x to justify both integrations by parts without boundary terms. We also assume a locally integrable majorant for the time difference quotients, so that differentiation may pass through the Fourier integral by dominated convergence. These conditions hold, for example, for a classical solution whose relevant spatial derivatives are rapidly decreasing. We write $\hat{u} = \mathcal{F}u$.

Definition 24.1 Let $u(x,t)$ be the amount of heat at point x at time t. The **heat equation** is

$$u_t = \alpha u_{xx}$$

for $\alpha > 0$ and $t > 0$, with initial condition $u(x,0) = \phi(x)$.

We start with two observations:

1. For $k \in \mathbb{R}$, let
$$\mathcal{F}u(k,t) = \int_{\mathbb{R}} u(x,t)e^{-2\pi i x k}\,\mathrm{d}x,$$
Then the dominated convergence theorem gives
$$\begin{aligned}\mathcal{F}u_t(k,t) &= \int_{\mathbb{R}} u_t(x,t)e^{-2\pi i x k}\,\mathrm{d}x \\ &= \int_{\mathbb{R}} \lim_{h\to 0} \frac{u(x,t+h) - u(x,t)}{h} e^{-2\pi i x k}\,\mathrm{d}x \\ &= \lim_{h\to 0} \int_{\mathbb{R}} \frac{u(x,t+h) - u(x,t)}{h} e^{-2\pi i x k}\,\mathrm{d}x \\ &= \lim_{h\to 0} \frac{\hat{u}(k,t+h) - \hat{u}(k,t)}{h} = \partial_t \hat{u}(k,t),\end{aligned}$$

2. We have

$$\begin{aligned}\mathcal{F}u_{xx}(k,t) &= \int_{\mathbb{R}} u_{xx}(x,t)e^{-2\pi ixk}\,\mathrm{d}x \\ &= \left[u_x(x,t)e^{-2\pi ixk}\right]_{-\infty}^{\infty} - \int_{\mathbb{R}}(-2\pi ik)u_x(x,t)e^{-2\pi ixk}\,\mathrm{d}x \\ &= (2\pi ik)^2\int_{\mathbb{R}} u(x,t)e^{-2\pi ixk}\,\mathrm{d}x \\ &= -(2\pi)^2k^2\hat{u}(k,t).\end{aligned}$$

We now take the Fourier transform of the heat equation:

$$\partial_t\hat{u}(k,t) = -\alpha(2\pi)^2k^2\hat{u}(k,t),$$

or equivalently,

$$\partial_t\hat{u}(k,t) + \alpha(2\pi)^2k^2\hat{u}(k,t) = 0.$$

This is a linear differential equation, and those things are easy to solve! Using an integrating factor,

$$\partial_t\left(\hat{u}(k,t)e^{\alpha(2\pi)^2k^2t}\right) = 0.$$

Hence $\hat{u}(k,t)e^{\alpha(2\pi)^2k^2t}$ is constant in t. Using the initial condition at $t = 0$ gives

$$\hat{u}(k,0) = \hat{\phi}(k).$$

Therefore,

$$\hat{u}(k,t) = e^{-\alpha(2\pi)^2k^2t}\hat{\phi}(k).$$

We look for $S(x,t)$ such that $\hat{S}(k,t) = e^{-\alpha(2\pi)^2k^2t}$. By the Fourier inversion theorem,

$$S(x,t) = \int_{\mathbb{R}}\hat{S}(k,t)e^{2\pi ikx}\,\mathrm{d}k = \int_{\mathbb{R}} e^{-\alpha(2\pi)^2k^2t}e^{2\pi ikx}\,\mathrm{d}k = \frac{1}{\sqrt{4\pi\alpha t}}e^{-\frac{x^2}{4\alpha t}}$$

by a tedious but straightforward transformation of the standard Gaussian integral. Then

$$\hat{u}(k,t) = \hat{S}(k,t)\hat{\phi}(k),$$

so

$$u(x,t) = (S(\cdot,t)*\phi)(x) = \int_{-\infty}^{\infty}\phi(y)\frac{1}{\sqrt{4\pi\alpha t}}e^{-\frac{(x-y)^2}{4\alpha t}}\,\mathrm{d}y.$$

The kernels in Figure 12 show what the convolution does: as time increases, the same unit mass spreads over a wider region and its peak falls.

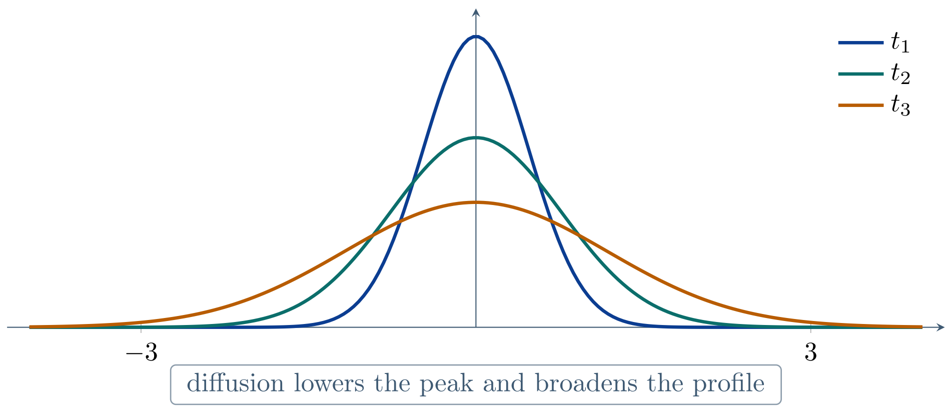


FIGURE 12

Remark 24.2 The solution above can be interpreted as $\mathbb{E}[\phi(X)]$ where $X \sim \mathcal{N}(x, 2\alpha t)$. This probabilistic interpretation connects the heat equation to Brownian motion and stochastic processes, where the diffusion of heat can be seen as the random movement of particles.

Translation-Invariant Subspaces of $L^2(\mathbb{R})$

For $f \in L^2(\mathbb{R})$ and $x \in \mathbb{R}$, we write $f_x(y) = f(y - x)$ in what follows.

Definition 24.3 We say that a subspace $\mathcal{U}$ of $L^2(\mathbb{R})$ is **translation-invariant** if it satisfies the following:

1. $\mathcal{U}$ is a closed subspace of $L^2(\mathbb{R})$ (that is, it is closed in $\|\cdot\|_2$ and for all $f, g \in \mathcal{U}$ and for all $c \in \mathbb{C}$, $f + cg \in \mathcal{U}$).
2. If $f \in \mathcal{U}$, then $f_x \in \mathcal{U}$ for all $x \in \mathbb{R}$.

The trivial examples of translation-invariant subspaces are $\{0\}$ and $L^2(\mathbb{R})$. To find less trivial examples, first observe that

$$\begin{aligned}
\hat{f}_x(k) &= \int_{\mathbb{R}} f(y - x)e^{-2\pi iyk}\,\mathrm{d}y \\
&= \int_{\mathbb{R}} f(y - x)e^{-2\pi i(y-x+x)k}\,\mathrm{d}y \\
&= \int_{\mathbb{R}} f(y - x)e^{-2\pi i(y-x)k}e^{-2\pi ixk}\,\mathrm{d}y
\end{aligned}$$

$$= e^{-2\pi ixk}\hat{f}(k).$$

Let $\chi_x(k) = e^{-2\pi ixk}$. Then we have the familiar identity $\hat{f}_x = \chi_x\hat{f}$.

Let $\mathcal{U}$ be a translation-invariant subspace of $L^2(\mathbb{R})$, and define $\hat{\mathcal{U}} = \{\hat{f} : f \in \mathcal{U}\}$. The translation-invariance of $\mathcal{U}$ implies that for all $g \in \hat{\mathcal{U}}$ and for all $x \in \mathbb{R}$, we have $\chi_x g \in \hat{\mathcal{U}}$. Recall that the Fourier transform $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is linear, surjective, and isometric by Theorem 23.4. Therefore $\hat{\mathcal{U}}$ is a closed subspace of $L^2(\mathbb{R})$ with the following invariance property:

$$g \in \hat{\mathcal{U}} \qquad \text{and} \qquad x \in \mathbb{R} \Longrightarrow \chi_x g \in \hat{\mathcal{U}}.$$

Now take a closed subset $F \subseteq \mathbb{R}$ and consider

$$V_F = \{f \in L^2(\mathbb{R}) : f = 0 \text{ almost everywhere on } \mathbb{R} \setminus F\}.$$

Then V_F is a closed subspace of $L^2(\mathbb{R})$ with the invariance property above. Hence

$$\mathcal{U} = \{f \in L^2(\mathbb{R}) : \hat{f} = 0 \text{ almost everywhere on } \mathbb{R} \setminus F\}$$

is a translation-invariant subspace of $L^2(\mathbb{R})$.

This matters for representation theory. A **representation** of a group G is a homomorphism $\pi : G \to GL_n$, so $\pi(xy) = \pi(x)\pi(y)$ and $\pi(e) = I_{n\times n}$. We often want to classify the **irreducible representations** of a group (i.e., representations that have no nontrivial invariant subspaces). If G is compact, then any translation-invariant subspace of $L^2(G)$ gives rise to representations that are closely related to irreducible representations of G (and vice versa). If G is noncompact, then the translation-invariant subspaces of $L^2(G)$ are more complicated, and this is one of the reasons why the representation theory of noncompact groups is more difficult than that of compact groups.

Appendix: Lecture and Tutorial Index

1. Lecture 1 — Tuesday, January 06, 2015
2. Lecture 2 — Thursday, January 08, 2015
3. Lecture 3 — Tuesday, January 13, 2015
4. Lecture 4 — Thursday, January 15, 2015
5. Lecture 5 — Tuesday, January 20, 2015
6. Lecture 6 — Thursday, January 22, 2015
7. Lecture 7 — Tuesday, January 27, 2015
8. Lecture 8 — Thursday, January 29, 2015
9. Lecture 9 — Tuesday, February 03, 2015
10. Lecture 10 — Thursday, February 05, 2015
11. Lecture 11 — Tuesday, February 10, 2015
12. Lecture 12 — Thursday, February 12, 2015
13. Lecture 13 — Thursday, February 19, 2015
14. Lecture 14 — Tuesday, February 24, 2015
15. Lecture 15 — Thursday, February 26, 2015
16. Lecture 16 — Tuesday, March 03, 2015
17. Lecture 17 — Thursday, March 05, 2015
18. Lecture 18 — Tuesday, March 10, 2015
19. Lecture 19 — Thursday, March 12, 2015
20. Lecture 20 — Tuesday, March 17, 2015
21. Lecture 21 — Thursday, March 19, 2015
22. Lecture 22 — Tuesday, March 24, 2015
23. Lecture 23 — Thursday, March 26, 2015
24. Lecture 24 — Thursday, April 02, 2015