Math212a1413
The Lebesgue integral.

Shlomo Sternberg

October 28, 2014

Simple functions.

In what follows, $(X, \mathcal{F}, m)$ is a space with a σ-field of sets, and m a measure on $\mathcal{F}$. The purpose of today's lecture is to develop the theory of the Lebesgue integral for functions defined on X. The theory starts with **simple** functions, that is functions which take on only finitely many non-zero values, say $\{a_1, \ldots, a_n\}$ and where

$$A_i := f^{-1}(a_i) \in \mathcal{F}.$$

In other words, we start with functions of the form

$$\phi(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \qquad A_i \in \mathcal{F}. \tag{1}$$

The integral of simple functions.

$$\phi(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \qquad A_i \in \mathcal{F}.$$

Then, for any $E \in \mathcal{F}$ we would like to define the integral of a simple function ϕ over E as

$$\int_E \phi dm = \sum_{i=1}^{n} a_i m(A_i \cap E) \tag{2}$$

and extend this definition by some sort of limiting process to a broader class of functions.

The range of the functions.

I haven't specified what the range of the functions should be. Even to get started, we have to allow our functions to take values in a vector space over $\mathbb{R}$, in order that the expression on the right of (2) make sense. I may eventually allow f to take values in a Banach space. However the theory is a bit simpler for real valued functions, where the linear order of the reals makes some arguments easier. Of course it would then be no problem to pass to any finite dimensional space over the reals. But we will on occasion need integrals in infinite dimensional Banach spaces, and that will require a little reworking of the theory.

Measurable maps.

If $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ are spaces with σ-fields, then

$$f : X \to Y$$

is called measurable if

$$f^{-1}(E) \in \mathcal{F} \quad \forall\ E \in \mathcal{G}. \tag{3}$$

Notice that the collection of subsets of Y for which (3) holds is a σ-field, and hence if it holds for some collection $\mathcal{C}$, it holds for the σ-field generated by $\mathcal{C}$.

Measurable real valued functions.

For the next few sections we will take $Y = \mathbb{R}$ and $\mathcal{G} = \mathcal{B}$, the Borel field. Since the collection of open intervals on the line generate the Borel field, a real valued function $f : X \to \mathbb{R}$ is measurable if and only if

$$f^{-1}(I) \in \mathcal{F} \quad \text{for all open intervals} \quad I.$$

Equally well, it is enough to check this for intervals of the form $(-\infty, a)$ for all real numbers a.

Proposition

If $F : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function and $f,\ g$ are two measurable real valued functions on X, then $F(f,g)$ is measurable.

Proof.

The set $F^{-1}(-\infty, a)$ is an open subset of the plane, and hence can be written as the countable union of products of open intervals $I \times J$. So if we set $h = F(f,g)$ then $h^{-1}((-\infty, a))$ is the countable union of the sets $f^{-1}(I) \cap g^{-1}(J)$ and hence belongs to $\mathcal{F}$. □

From this elementary proposition we conclude that if f and g are measurable real valued functions then:

Sums, products, etc.

- $f+g$ is measurable (since $(x,y) \mapsto x+y$ is continuous),
- fg is measurable (since $(x,y) \mapsto xy$ is continuous), hence
- $f\mathbf{1}_A$ is measurable for any $A \in \mathcal{F}$ hence
- f^+ is measurable since $f^{-1}([0,\infty]) \in \mathcal{F}$ and similarly for f^- so
- $|f|$ is measurable and so is $|f-g|$. Hence
- $f \wedge g$ and $f \vee g$ are measurable

and so on.

We are going to allow for the possibility that an integral might be infinite. We adopt the convention that

$$0 \cdot \infty = 0.$$

Recall that a non-negative function ϕ is **simple** if ϕ takes on a finite number of distinct non-negative values, $a_1, \ldots, a_n$, and that each of the sets

$$A_i = \phi^{-1}(a_i)$$

is measurable. These sets partition X:

$$X = A_1 \cup \cdots \cup A_n.$$

Of course since the values are distinct,

$$A_i \cap A_j = \emptyset \quad \text{for} \quad i \neq j.$$

With this definition, a simple function can be written as in (1):

$$\phi(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \qquad A_i \in \mathcal{F} \qquad (1)$$

and this expression is unique.

The integral of non-negative simple functions.

A *non-negative* simple function can be written as in (1):

$$\phi(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \qquad A_i \in \mathcal{F} \qquad (1)$$

with $a_i \geq 0$ and this expression is unique. So we may take

$$\int_E \phi dm = \sum_{i=1}^{n} a_i m(A_i \cap E) \qquad (2)$$

as the definition of the integral of a non-negative simple function.

Extending the definition to all non-negative functions.

We now extend the definition to an arbitrary ($[0,\infty]$ valued) measurable function f by

$$\int_E f dm = \sup I(E,f) \tag{4}$$

where

$$I(E,f) = \left\{ \int_E \phi dm : 0 \le \phi \le f,\ \phi \text{ simple} \right\}. \tag{5}$$

In other words, we take all integrals of expressions of simple functions ϕ such that $\phi(x) \le f(x)$ at all x. We then define the integral of f as the supremum of these values.

Notice that if $A := f^{-1}(\infty)$ has positive measure, then the simple functions $n\mathbf{1}_A$ are all $\leq f$ and so $\int_X f dm = \infty$.

Proposition

For (non-negative) simple functions, the definition (4) coincides with definition (2).

Proof.

Since ϕ is $\leq$ itself, the right hand side of (2) belongs to $I(E,\phi)$ and hence is $\leq \int_E \phi dm$ as given by (5). We must show the reverse inequality: Suppose that $\psi = \sum b_j \mathbf{1}_{B_j} \leq \phi$. We can write the right hand side of (2) as

$$\sum b_j m(E \cap B_j) = \sum_{i,j} b_j m(E \cap A_i \cap B_j)$$

since $E \cap B_j$ is the disjoint union of the sets $E \cap A_i \cap B_j$ because the A_i partition X, and m is additive on disjoint finite (even countable) unions. On each of the sets $A_i \cap B_j$ we must have $b_j \leq a_i$. Hence

$$\sum_{i,j} b_j m(E \cap A_i \cap B_j) \leq \sum_{i,j} a_i m(E \cap A_i \cap B_j) = \sum a_i m(E \cap A_i)$$

since the B_j partition X. $\square$

Monotonicity for non-negative simple functions.

In the course of the proof of the above proposition we have also established

$$\psi \leq \phi \text{ for non-negative simple functions implies } \int_E \psi dm \leq \int_E \phi dm. \tag{6}$$

A key tool in the approach we are taking

We will make repeated use of a proposition about measures that we proved in lecture 11:

$$A_n \nearrow A \Rightarrow m(A_n) \to m(A)$$

for measurable sets A_n.

Additivity over sets for non-negative simple functions.

Suppose that E and F are disjoint measurable sets. Then

$$m(A_i \cap (E \cup F)) = m(A_i \cap E) + m(A_i \cap F)$$

so each term on the right of (2) breaks up into a sum of two terms and we conclude that

If ϕ is a non-negative simple function and

$$E \cap F = \emptyset, \quad \text{then} \int_{E \cup F} \phi dm = \int_E \phi dm + \int_F \phi dm. \tag{7}$$

Multiplication by a positive constant for non-negative simple functions.

Also, it is immediate from (2) that if $a \geq 0$ then

If ϕ is a non-negative simple function then

$$\int_E a\phi dm = a \int_E \phi dm. \tag{8}$$

Extending these results to non-negative measurable functions.

It is now immediate that these results extend to all non-negative measurable functions. We list the results and then prove them. In what follows f and g are non-negative measurable functions, $a \geq 0$ is a real number and E and F are measurable sets:

$$f \leq g \quad \Rightarrow \quad \int_E f dm \leq \int_E g dm. \tag{9}$$

$$\int_E f dm \quad = \quad \int_X \mathbf{1}_E f dm \tag{10}$$

$$E \subset F \quad \Rightarrow \quad \int_E f dm \leq \int_F f dm. \tag{11}$$

$$\int_E a f dm \quad = \quad a \int_E f dm. \tag{12}$$

$$m(E) = 0 \quad \Rightarrow \quad \int_E f dm = 0. \tag{13}$$

$$E \cap F = \emptyset \quad \Rightarrow \quad \int_{E \cup F} f dm = \int_E f dm + \int_F f dm. \tag{14}$$

$$f = 0 \;\; \text{a.e.} \quad \Leftrightarrow \quad \int_X f dm = 0. \tag{15}$$

$$f \leq g \;\; \text{a.e.} \quad \Rightarrow \quad \int_X f dm \leq \int_X g dm. \tag{16}$$

Proofs.

$$f \leq g \;\Rightarrow\; \int_E f dm \leq \int_E g dm. \tag{9}$$

Proof.

$I(E, f) \subset I(E, g)$. □

$$\int_E f dm = \int_X \mathbf{1}_E f dm. \tag{10}$$

Proof.

If ϕ is a simple function with $\phi \leq f$, then multiplying ϕ by $\mathbf{1}_E$ gives a function which is still $\leq f$ and is still a simple function. The set $I(E, f)$ is unchanged by considering only simple functions of the form $\mathbf{1}_E \phi$ and these constitute all simple functions $\leq \mathbf{1}_E f$. □

$$E \subset F \ \Rightarrow \int_E fdm \leq \int_F fdm. \qquad (11)$$

Proof.

We have $\mathbf{1}_E f \leq \mathbf{1}_F f$ and we can apply (9) and (10). □

$$\int_E afdm = a\int_E fdm. \qquad (12)$$

Proof.

$I(E, af) = aI(E, f)$. □

$$m(E) = 0 \ \Rightarrow \ \int_E f dm = 0. \qquad (13)$$

Proof.

In the definition (2) all the terms on the right vanish since $m(E \cap A_i) = 0$. So $I(E, f)$ consists of the single element 0. □

$$E \cap F = \emptyset \;\Rightarrow\; \int_{E \cup F} f dm = \int_E f dm + \int_F f dm. \tag{14}$$

Proof. This is true for simple functions, so $I(E \cup F, f) = I(E, f) + I(F, f)$ meaning that every element of $I(E \cup F, f)$ is a sum of an element of $I(E, f)$ and an element of $I(F, f)$. Thus the sup on the left is $\leq$ the sum of the sups on the right, proving that the left hand side of (14) is $\leq$ its right hand side. To prove the reverse inequality, choose a simple function $\phi \leq \mathbf{1}_E f$ and a simple function $\psi \leq \mathbf{1}_F f$. Then $\phi + \psi \leq \mathbf{1}_{E \cup F} f$ since $E \cap F = \emptyset$. So $\phi + \psi$ is a simple function $\leq f$ and hence $\int_E \phi dm + \int_F \psi dm \leq \int_{E \cup F} f dm$. If we now maximize the two summands separately we get

$$\int_E f dm + \int_F f dm \leq \int_{E \cup F} f dm. \quad \square$$

$$f = 0 \text{ a.e. } \Leftrightarrow \int_X f dm = 0. \qquad (15):$$

Proof of $\Rightarrow$. If $f = 0$ almost everywhere, and $\phi \leq f$ then $\phi = 0$ a.e. since $\phi \geq 0$. This means that all sets which enter into the right hand side of (2) with $a_i \neq 0$ have measure zero, so the right hand side vanishes. So $I(X, f)$ consists of the single element 0. This proves $\Rightarrow$ in (15).

We wish to prove the reverse implication.

Proof.

Let $A = \{x|f(x) > 0\}$. We wish to show that $m(A) = 0$. Now

$$A = \bigcup A_n \quad \text{where} \quad A_n := \{x|f(x) > \frac{1}{n}\}.$$

The sets A_n are measurable and increasing, so we know that $m(A) = \lim_{n\to\infty} m(A_n)$. So it is enough to prove that $m(A_n) = 0$ for all n. But

$$\frac{1}{n}\mathbf{1}_{A_n} \leq f$$

and is a simple function. So

$$\frac{1}{n}\int_X \mathbf{1}_{A_n} dm = \frac{1}{n} m(A_n) \leq \int_X f dm = 0$$

implying that $m(A_n) = 0$. □

$$f \leq g \quad \text{a.e.} \quad \Rightarrow \int_X f dm \leq \int_X g dm. \qquad (16).$$

Proof.

Let $E = \{x | f(x) \leq g(x)\}$. Then E is measurable and E^c is of measure zero. By definition, $\mathbf{1}_E f \leq \mathbf{1}_E g$ everywhere, hence by (11)

$$\int_X \mathbf{1}_E f dm \leq \int_X \mathbf{1}_E g dm.$$

But

$$\int_X \mathbf{1}_E f dm + \int_X \mathbf{1}_{E^c} f dm = \int_E f dm + \int_{E^c} f dm = \int_X f dm$$

where we have used (14) and (15). Similarly for g. □

Fatou's lemma.

This says:

Theorem

If $\{f_n\}$ is a sequence of non-negative measurable functions, then

$$\lim_{n\to\infty} \inf_{k\geq n} \int f_k dm \geq \int \left(\lim_{n\to\infty} \inf_{k\geq n} f_k \right) dm. \tag{17}$$

Recall that the limit inferior of a sequence of numbers $\{a_n\}$ is defined as follows: Set

$$b_n := \inf_{k\geq n} a_k$$

so that the sequence $\{b_n\}$ is non-decreasing, and hence has a limit (possibly infinite) which is defined as the lim inf. For a sequence of functions, lim inf f_n is obtained by taking lim inf $f_n(x)$ for every x.

Consider the sequence of simple functions $\{\mathbf{1}_{[n,n+1]}\}$. At each point x the lim inf is 0, in fact $\mathbf{1}_{[n,n+1]}(x)$ becomes and stays 0 as soon as $n > x$. Thus the right hand side of (17) is zero. The numbers which enter into the left hand side are all 1, so the left hand side is 1.

Similarly, if we take $f_n = n\mathbf{1}_{(0,1/n]}$, the left hand side is 1 and the right hand side is 0. So without further assumptions, we generally expect to get strict inequality in Fatou's lemma.

Proof of Fatou's lemma, I.

Set

$$g_n := \inf_{k \geq n} f_k$$

so that

$$g_n \leq g_{n+1}$$

and set

$$f := \lim_{n \to \infty} \inf_{k \geq n} f_n = \lim_{n \to \infty} g_n.$$

Let

$$\phi \leq f$$

be a non-negative simple function. We must show that

$$\int \phi dm \leq \lim_{n \to \infty} \inf_{k \geq n} \int f_k dm. \tag{18}$$

There are two cases to consider:

Proof of Fatou's lemma, II.

a) $m(\{x : \phi(x) > 0\}) = \infty$. In this case $\int \phi dm = \infty$ and hence $\int f dm = \infty$ since $\phi \leq f$. We must show that $\liminf \int f_n dm = \infty$. Let

$$D := \{x : \phi(x) > 0\} \quad \text{so } m(D) = \infty.$$

Choose some positive number $b <$ all the positive values taken by ϕ. This is possible since there are only finitely many such values.

Let

$$D_n := \{x | g_n(x) > b\}.$$

Then $D_n \nearrow D$ since $b < \phi(x) \leq \lim_{n\to\infty} g_n(x)$ at each point of D. Hence $m(D_n) \to m(D) = \infty$. But

$$bm(D_n) \leq \int_{D_n} g_n dm \leq \int_{D_n} f_k dm \quad k \geq n$$

since $g_n \leq f_k$ for $k \geq n$. Now

$$\int f_k dm \geq \int_{D_n} f_k dm$$

since f_k is non-negative. Hence $\liminf \int f_n dm = \infty$.

Proof of Fatou's lemma, III.

b) $m(\{x : \phi(x) > 0\}) < \infty$. Choose $\epsilon > 0$ so that it is less than the minimum of the positive values taken on by ϕ and set

$$\phi_\epsilon(x) = \begin{cases} \phi(x) - \epsilon & \text{if} \quad \phi(x) > 0 \\ 0 & \text{if} \quad \phi(x) = 0. \end{cases}$$

Let

$$C_n := \{x | g_n(x) \geq \phi_\epsilon\}$$

and

$$C = \{x : f(x) \geq \phi_\epsilon\}.$$

Then $C_n \nearrow C$.

Proof of Fatou's lemma, IV.

We have

$$\begin{aligned}
\int_{C_n} \phi_\epsilon dm &\leq \int_{C_n} g_n dm \\
&\leq \int_{C_n} f_k dm \quad k \geq n \\
&\leq \int_C f_k dm \quad k \geq n \\
&\leq \int f_k dm \quad k \geq n.
\end{aligned}$$

So

$$\int_{C_n} \phi_\epsilon dm \leq \liminf \int f_k dm.$$

Proof of Fatou's lemma, V.

We will next let $n \to \infty$: Let c_i be the non-zero values of ϕ_ϵ so

$$\phi_\epsilon = \sum c_i \mathbf{1}_{B_i}$$

for some measurable sets $B_i \subset C$. Then

$$\int_{C_n} \phi_\epsilon dm = \sum c_i m(B_i \cap C_n) \to \sum c_i m(B_i) = \int \phi_\epsilon dm$$

since $(B_i \cap C_n) \nearrow B_i \cap C = B_i$. So

$$\int \phi_\epsilon dm \leq \liminf \int f_k dm.$$

Now

$$\int \phi_\epsilon dm = \int \phi dm - \epsilon m\left(\{x|\phi(x) > 0\}\right).$$

Since we are assuming that $m\left(\{x|\phi(x) > 0\}\right) < \infty$, we can let $\epsilon \to 0$ and conclude that $\int \phi dm \leq \liminf \int f_k dm$. $\square$

The monotone convergence theorem.

We assume that $\{f_n\}$ is a sequence of non-negative measurable functions, and that $f_n(x)$ is an increasing sequence for each x. Define $f(x)$ to be the limit (possibly $+\infty$) of this sequence. We describe this situation by $f_n \nearrow f$. The monotone convergence theorem asserts that:

$$f_n \geq 0, \quad f_n \nearrow f \quad \Rightarrow \quad \lim_{n\to\infty} \int f_n dm = \int f dm. \tag{19}$$

Proof. The f_n are increasing and all $\leq f$ so the $\int f_n dm$ are monotone increasing and all $\leq \int f dm$. So the limit exists and is $\leq \int f dm$. On the other hand, Fatou's lemma gives

$$\int f dm \leq \liminf \int f_n dm = \lim \int f_n dm. \quad \square$$

We only need convergence almost everywhere in the monotone convergence theorem.

In the monotone convergence theorem we need only know that

$$f_n \nearrow f \quad \text{a.e.}$$

Indeed, let C be the set where convergence holds, so $m(C^c) = 0$. Let $g_n = \mathbf{1}_C f_n$ and $g = \mathbf{1}_C f$. Then $g_n \nearrow g$ everywhere, so we may apply (19) to g_n and g. But $\int g_n dm = \int f_n dm$ and $\int g dm = \int f dm$ so the theorem holds for f_n and f as well.

Integrable functions.

We will say an $\mathbb{R}$ valued measurable function is **integrable** if both $\int f^+ dm < \infty$ and $\int f^- dm < \infty$. If this happens, we set

$$\int f dm := \int f^+ dm - \int f^- dm. \tag{20}$$

Since both numbers on the right are finite, this difference makes sense. Some authors prefer to allow one or the other numbers (but not both) to be infinite, in which case the right hand side of (20) might be $= \infty$ or $-\infty$. We will stick with the above convention.

The space $\mathcal{L}_1(X, \mathbb{R})$.

We will denote the set of all (real valued) integrable functions by $\mathcal{L}_1$ or $\mathcal{L}_1(X)$ or $\mathcal{L}_1(X, \mathbb{R})$ depending on how precise we want to be. Notice that if $f \leq g$ then $f^+ \leq g^+$ and $f^- \geq g^-$ all of these functions being non-negative. So

$$\int f^+ dm \leq \int g^+ dm, \quad \int f^- dm \geq \int g^- dm$$

hence

$$\int f^+ dm - \int f^- dm \leq \int g^+ dm - \int g^- dm$$

or

$$f \leq g \quad \Rightarrow \quad \int f dm \leq \int g dm. \tag{21}$$

Multiplication by a constant.

If a is a non-negative number, then $(af)^{\pm} = af^{\pm}$. If $a < 0$ then $(af)^{\pm} = (-a)f^{\mp}$ so in all cases we have

$$\int afdm = a \int fdm. \tag{22}$$

The integral of a sum.

We now wish to establish

$$f, g \in \mathcal{L}_1 \;\Rightarrow f + g \in \mathcal{L}_1 \;\text{ and }\; \int (f+g)dm = \int fdm + \int gdm. \tag{23}$$

We prove this in stages:

The sum of two non-negative simple functions.

Assume $f = \sum a_i \mathbf{1}_{A_i}, \; g = \sum b_i \mathbf{1}_{B_i}$ are non-negative simple functions, where the A_i partition X as do the B_j. Then we can decompose and recombine the sets to yield:

$$\begin{aligned}
\int (f+g)dm &= \sum_{i,j}(a_i+b_j)m(A_i\cap B_j)\\
&= \sum_i\sum_j a_i m(A_i\cap B_j) + \sum_j\sum_i b_j m(A_i\cap B_j)\\
&= \sum_i a_i m(A_i) + \sum_j b_j m(B_j)\\
&= \int f dm + \int g dm
\end{aligned}$$

where we have used the fact that m is additive and the $A_i \cap B_j$ are disjoint sets whose union over j is A_i and whose union over i is B_j.

Proof when f and g are non-negative measurable functions with finite integrals, 1.

Suppose that f and g are non-negative measurable functions with finite integrals. Set

$$f_n := \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbf{1}_{f^{-1}[\frac{k}{2^n},\frac{k+1}{2^n})}.$$

Each f_n is a simple function $\leq f$, and passing from f_n to f_{n+1} involves splitting each of the sets $f^{-1}([\frac{k}{2^n},\frac{k+1}{2^n}))$ in the sum into two, and choosing a larger value on the second portion, and also adding conributions from where $f > 2^n$.

Proof when f and g are non-negative measurable functions with finite integrals, 2.

So the f_n are increasing. Also, if $f(x) < \infty$, then $f(x) < 2^m$ for some m, and for any $n > m$ $f_n(x)$ differs from $f(x)$ by at most 2^{-n}. Hence $f_n \nearrow f$ a.e., since f is finite a.e because its integral is finite.

Proof when f and g are non-negative measurable functions with finite integrals, 3.

Similarly we can construct $g_n \nearrow g$. Also $(f_n + g_n) \nearrow f + g$ a.e. By the a.e. monotone convergence theorem

$$\int (f+g)dm = \lim \int (f_n + g_n)dm = \lim \int f_n dm + \lim \int g_n dm$$

$$= \int f dm + \int g dm,$$

where we have used (23) for simple functions. This argument shows that $\int (f+g)dm < \infty$ if both integrals $\int f dm$ and $\int g dm$ are finite.

Lebesgue's idea

This idea - that of "breaking up the y-axis", i.e. using the approximation

$$f_n := \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbf{1}_{f^{-1}[\frac{k}{2^n}, \frac{k+1}{2^n})}$$

to approximate the integral, instead of "breaking up the $x - axis$ " as in the Cauchy or Riemann definition of the integral, is one of the key ideas in Lebesgue's thesis.

Proof of the additivity of the integral in general

For any $f \in \mathcal{L}_1$ we conclude from the preceding that

$$\int |f| dm = \int (f^+ + f^-) dm < \infty.$$

Similarly for g. Since $|f + g| \leq |f| + |g|$ we conclude that both $(f+g)^+$ and $(f+g)^-$ have finite integrals. Now

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

or

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

All expressions are non-negative and integrable. So integrate both sides to get (23). $\square$

$\mathcal{L}_1$ is a vector space

We have thus established

Theorem

The space $\mathcal{L}_1(X, \mathbb{R})$ is a real vector space and $f \mapsto \int f dm$ is a linear function on $\mathcal{L}_1(X, \mathbb{R})$.

We also have

Proposition

If $h \in \mathcal{L}_1$ and $\int_A h dm \geq 0$ for all $A \in \mathcal{F}$ then $h \geq 0$ a.e.

Proof.

Let $A_n : \{x | h(x) \leq -\frac{1}{n}\}$. Then

$$\int_{A_n} h dm \leq \int_{A_n} \frac{-1}{n} dm = -\frac{1}{n} m(A_n)$$

so $m(A_n) = 0$. But if we let $A := \{x | h(x) < 0\}$ then $A_n \nearrow A$ and hence $m(A) = 0$. □

The $\mathcal{L}_1$ (semi-)norm

We have defined the integral of any function f as $\int f dm = \int f^+ dm - \int f^- dm$, and $\int |f| dm = \int f^+ dm + \int f^- dm$. Since for any two non-negative real numbers $a - b \leq a + b$ we conclude that

$$\left| \int f dm \right| \leq \int |f| dm. \tag{24}$$

The $\mathcal{L}_1$ (semi-)norm, continued

If we define

$$\|f\|_1 := \int |f| dm$$

we have verified that

$$\|f+g\|_1 \leq \|f\|_1 + \|g\|_1,$$

and have also verified that

$$\|cf\|_1 = |c|\|f\|_1.$$

In other words, $\|\cdot\|_1$ is a semi-norm on $\mathcal{L}_1$. From the preceding proposition we know that $\|f\|_1 = 0$ if and only if $f = 0$ a.e. The question of when we want to pass to the quotient and identify two functions which differ on a set of measure zero is a matter of taste.

The dominated convergence theorem.

This says that

Theorem

Let f_n be a sequence of measurable functions such that

$$|f_n| \leq g \text{ a.e.}, \quad g \in \mathcal{L}_1$$

$$f_n \to f \quad \text{a.e.} \;\Rightarrow\; f \in \mathcal{L}_1 \text{ and } \int f_n dm \to \int f dm.$$

The functions f_n are all integrable, since their positive and negative parts are dominated by g.

Proof in the non-negative case.

Assume for the moment that $f_n \geq 0$. Then Fatou's lemma says that

$$\int f dm \leq \liminf \int f_n dm.$$

Fatou's lemma applied to $g - f_n$ says that

$$\int (g - f) dm \leq \liminf \int (g - f_n) dm = \liminf \left(\int g dm - \int f_n dm \right)$$

$$= \int g dm - \limsup \int f_n dm.$$

Subtracting $\int g dm$ gives

$$\limsup \int f_n dm \leq \int f dm.$$

Proof of the dominated convergence theorem in the non-negative case, continued.

So

$$\limsup \int f_n dm \leq \int f dm \leq \liminf \int f_n dm$$

which can only happen if all three are equal. so we have proved the dominated convergence theorem for non-negative f_n.

Proof in general.

For general f_n we can write our hypothesis as

$$-g \leq f_n \leq g \quad \text{a.e..}$$

Adding g to both sides gives

$$0 \leq f_n + g \leq 2g \quad \text{a.e..}$$

We now apply the result for non-negative sequences to $g + f_n$ and then subtract off $\int g dm$. $\square$

Suppose that $X = [a, b]$ is an interval. What is the relation between the Lebesgue integral and the Riemann integral?

Let us suppose that $[a, b]$ is bounded and that f is a bounded function, say $|f| \leq M$.

Riemann lower and upper sums.

Each partition $P : a = a_0 < a_1 < \cdots < a_n = b$ into intervals $I_i = [a_{i-1}, a_i]$ with

$$m_i := m(I_i) = a_i - a_{i-1}, \quad i = 1, \ldots, n$$

defines a Riemann lower sum

$$L_P = \sum k_i m_i \qquad k_i = \inf_{x \in I_i} f(x)$$

and a Riemann upper sum

$$U_P = \sum M_i m_i \qquad M_i := \sup_{x \in I_i} f(x)$$

which are the Lebesgue integrals of the simple functions

$$\ell_P := \sum k_i \mathbf{1}_{I_i} \quad \text{and } u_P := \sum M_i \mathbf{1}_{I_i}$$

respectively.

Riemann's definition.

According to Riemann, we are to choose a sequence of partitions P_n which refine one another and whose maximal interval lengths go to zero. Write ℓ_i for ℓ_{P_i} and u_i for u_{P_i}. Then

$$\ell_1 \leq \ell_2 \leq \cdot \leq f \leq \cdots \leq u_2 \leq u_1.$$

Suppose that f is measurable. All the functions in the above inequality are Lebesgue integrable, so dominated convergence implies that

$$\lim U_n = \lim \int_a^b u_n dx = \int_a^b u dx$$

where $u := \lim u_n$ with a similar equation for the lower bounds. The Riemann integral is defined as the common value of $\lim L_n$ and $\lim U_n$ whenever these limits are equal.

Proposition

f is Riemann integrable if and only if f is continuous almost everywhere.

Proof.

Notice that if x is not an endpoint of any interval in the partitions, then f is continuous at x if and only if $u(x) = \ell(x)$. Riemann's condition for integrability says that $\int (u - \ell) dm = 0$ which implies that f is continuous almost everywhere.

Conversely, if f is continuous a.e. then $u = f = \ell$ a.e.. Since u is measurable so is f, and since we are assuming that f is bounded, we conclude that f Lebesgue integrable. As $\ell = f = u$ a.e. their Lebesgue integrals coincide. But the statement that the Lebesgue integral of u is the same as that of ℓ is precisely the statement of Riemann integrability. □

Notice that in the course of the proof we have also shown that the Lebesgue and Riemann integrals coincide when both exist.

The Beppo-Levi theorem.

We begin with a lemma:

Lemma

Beppo-Levi. *Let* $\{g_n\}$ *be a sequence of non-negative measurable functions. Then*

$$\int \sum_{n=1}^{\infty} g_n \, dm = \sum_{n=1}^{\infty} \int g_n dm.$$

Proof.

We have

$$\int \sum_{k=1}^{n} g_k dm = \sum_{k=1}^{n} \int g_k dm$$

for finite n by the linearity of the integral. Since the $g_k \geq 0$, the sums under the integral sign are increasing, and by definition converge to $\sum_{k=1}^{\infty} g_k$. The monotone convergence theorem implies the lemma. $\square$

But both sides of the equation in the lemma might be infinite.

Theorem

Beppo-Levi. *Let* $f_n \in \mathcal{L}_1$ *and suppose that*

$$\sum_{k=1}^{\infty} \int |f_k| dm < \infty.$$

Then $\sum f_k(x)$ *converges to a finite limit for almost all* x*, the sum is integrable, and*

$$\int \sum_{k=1}^{\infty} f_k \; dm = \sum_{k=1}^{\infty} \int f_k dm.$$

Proof.

Take $g_n := |f_n|$ in the lemma. If we set $g = \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} |f_n|$ then the lemma says that

$$\int g dm = \sum_{n=1}^{\infty} \int |f_n| dm,$$

and we are assuming that this sum is finite. So g is integrable, in particular the set of x for which $g(x) = \infty$ must have measure zero. In other words,

$$\sum_{n=1} |f_n(x)| < \infty \quad \text{a.e.} \ .$$

$$\sum_{n=1} |f_n(x)| < \infty \quad \text{a.e. }.$$

If a series is absolutely convergent, then it is convergent, so we can say that $\sum f_n(x)$ converges almost everywhere. Let

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

at all points where the series converges, and set $f(x) = 0$ at all other points. Now

$$\left| \sum_{n=0}^{\infty} f_n(x) \right| \le g(x)$$

at all points, and hence by the dominated convergence theorem, $f \in \mathcal{L}_1$ and

$$\int f dm = \int \lim_{n\to\infty} \sum^{n} f_k \, dm = \lim_{n\to\infty} \sum \int f_k dm = \sum^{\infty} \int f_k dm. \quad \square$$

$\mathcal{L}_1$ is complete.

This is an immediate corollary of the Beppo-Levi theorem and Fatou's lemma. Indeed, suppose that $\{h_n\}$ is a Cauchy sequence in $\mathcal{L}_1$. Choose n_1 so that

$$\|h_n - h_{n_1}\|_1 \leq \frac{1}{2} \quad \forall\ n \geq n_1.$$

Then choose $n_2 > n_1$ so that

$$\|h_n - h_{n_2}\|_1 \leq \frac{1}{2^2} \quad \forall\ n \geq n_2.$$

Continuing this way, we have produced a subsequence h_{n_j} such that

$$\|h_{n_{j+1}} - h_{n_j}\|_1 \leq \frac{1}{2^j}.$$

Proof that $\mathcal{L}_1$ is complete, using Beppo-Levy.

$$\|h_{n_{j+1}} - h_{n_j}\|_1 \leq \frac{1}{2^j}.$$

Let

$$f_j := h_{n_{j+1}} - h_{n_j}.$$

Then

$$\int |f_j| dm < \frac{1}{2^j}$$

so the hypotheses of the Beppo-Levy theorem are satisfied, and $\sum f_j$ converges almost everywhere to some limit $f \in \mathcal{L}_1$.

$\sum f_j$ converges almost everywhere to some limit $f \in \mathcal{L}_1$. But

$$h_{n_1} + \sum_{j=1}^{k} f_j = h_{n_{k+1}}.$$

So the subsequence h_{n_k} converges almost everywhere to some $h \in \mathcal{L}_1$.
We must show that this h is the limit of the h_n in the $\|\cdot\|_1$ norm.
For this we will use Fatou's lemma:

Using Fatou

For a given $\epsilon > 0$, choose N so that $\|h_n - h_m\|_1 < \epsilon$ for $m, n > N$. Since $h = \lim h_{n_j}$ we have, for $k > N$,

$$\|h - h_k\|_1 = \int |h - h_k| dm = \int \lim_{j \to \infty} |h_{n_j} - h_k| dm \leq \liminf \int |h_{n_j} - h_k| dm$$

$$= \liminf \|h_{n_j} - h_k\|_1 < \epsilon. \qquad \square$$

Summary of our logic so far.

Fatou's lemma

$\Downarrow$

The monotone convergence theorem

$\Downarrow$

$\mathcal{L}_1(X, \mathbb{R})$ is a real vector space and $f \mapsto \int f dm$ is a linear function on it

$\Downarrow$

The dominated convergence theorem

$\Downarrow$

The Beppo-Levi theorem

$\Downarrow$

$\mathcal{L}_1$ is complete.

Temporarily specializing to the case $X = \mathbb{R}$

Up until now we have been mostly studying integration on an arbitrary measure space $(X, \mathcal{F}, m)$. For the next few slides I will specialize and take $X = \mathbb{R}$, $\mathcal{F}$ to be the σ-field of Lebesgue measurable sets, and m to be Lebesgue measure, in order to simplify some of the formulations and arguments.

Approximation by simple functions

Suppose that f is a Lebesgue integrable non-negative function on $\mathbb{R}$. We know that for any $\epsilon > 0$ there is a non-negative simple function ϕ such that

$$\phi \leq f$$

and

$$\int f dm - \int \phi dm = \int (f - \phi) dm = \|f - \phi\|_1 < \epsilon.$$

To say that ϕ is simple implies that

$$\phi = \sum a_i \mathbf{1}_{A_i}$$

(finite sum) where each of the $a_i > 0$ and since $\int \phi dm < \infty$ each A_i has finite measure.

Approximation by simple functions, 2

$$\phi = \sum a_i \mathbf{1}_{A_i}$$

(finite sum) where each of the $a_i > 0$ and since $\int \phi dm < \infty$ each A_i has finite measure. Since $m(A_i \cap [-n, n]) \to m(A_i)$ as $n \to \infty$, we may choose n sufficiently large so that

$$\|f - \psi\|_1 < 2\epsilon \quad \text{where} \quad \psi = \sum a_i \mathbf{1}_{A_i \cap [-n,n]}.$$

For each of the sets $A_i \cap [-n, n]$ we can find a bounded open set U_i which contains it, and such that $m(U_i / A_i)$ is as small as we please. So we can find finitely many bounded open sets U_i such that

$$\|f - \sum a_i \mathbf{1}_{U_i}\|_1 < 3\epsilon.$$

Approximation by non-negative step functions

Each U_i is a countable union of disjoint open intervals, $U_i = \bigcup_j I_{i,j}$, and since $m(U_i) = \sum_j m(I_{i,j})$, we can find finitely many $I_{i,j}$, j ranging over a finite set of integers, J_i such that

$$m\left(\bigcup_{j \in J_i} I_{i,j}\right)$$

is as close as we like to $m(U_i)$. So let us call a **step function** a function of the form $\sum b_i \mathbf{1}_{I_i}$ where the I_i are bounded intervals. We have shown that we can find a step function with positive coefficients which is as close as we like in the $\|\cdot\|_1$ norm to f.

The step functions are dense

If f is not necessarily non-negative, we know (by definition!) that f^+ and f^- are in $\mathcal{L}_1$, and so we can approximate each by a step function. The triangle inequality then gives

Proposition

The step functions are dense in $\mathcal{L}_1(\mathbb{R}, \mathbb{R})$.

The continuous functions are dense

If $[a, b]$, $a < b$ is a finite interval, we can approximate $\mathbf{1}_{[a,b]}$ as closely as we like in the $\|\cdot\|_1$ norm by continuous functions: just choose n large enough so that $\frac{2}{n} < b - a$, and take the function which is 0 for $x < a$, rises linearly from 0 to 1 on $[a, a + \frac{1}{n}]$, is identically 1 on $[a + \frac{1}{n}, b - \frac{1}{n}]$, and goes down linearly from 1 to 0 from $b - \frac{1}{n}$ to b and stays 0 thereafter. As $n \to \infty$ this clearly tends to $\mathbf{1}_{[a,b}$ in the $\|\cdot\|$ norm. So

Proposition

The continuous functions of compact support are dense in $\mathcal{L}_1(\mathbb{R}, \mathbb{R})$.

The Riemann-Lebesgue lemma.

Let h be a bounded measurable function on $\mathbb{R}$. Recall a previous lecture that we say that h satisfies the **averaging condition** if

$$\lim_{|c|\to\infty} \frac{1}{|c|} \int_0^c h dm \to 0. \tag{25}$$

For example, if $h(t) = \cos \xi t,\ \xi \neq 0$, then the expression under the limit sign in the averaging condition is

$$\frac{1}{c\xi} \sin \xi c$$

which tends to zero as $c \to \infty$. Here the oscillations in h are what give rise to the averaging condition. As another example, let

$$h(t) = \begin{cases} 1 & |t| \leq 1 \\ 1/|t| & |t| \geq 1. \end{cases}$$

Then the left hand side of (25) is

$$\frac{1}{|c|}(1 + \log |c|), \qquad |c| \geq 1.$$

Here the averaging condition is satisfied because the integral in (25) grows more slowly than $|c|$.

Recall that we proved (using the density of step functions):

Theorem

Generalized Riemann-Lebesgue Lemma.
Let $f \in \mathcal{L}_1([c,d],\mathbb{R})$, $-\infty \leq c < d \leq \infty$. If h satisfies the averaging condition (25) then

$$\lim_{r\to\infty} \int_a^b f(t)h(rt)dt = 0. \tag{26}$$

The Cantor-Lebesgue theorem.

This says:

Theorem

If a trigonometric series

$$\frac{a_0}{2} + \sum_n d_n \cos(nt - \phi_n) \qquad d_n \geq 0$$

converges on a set E of positive Lebesgue measure then

$$d_n \to 0.$$

(I have written the general form of a real trigonometric series as a cosine series with phases since we are talking about only real valued functions at the present. Of course, applied to the real and imaginary parts, the theorem asserts that if $\sum a_n e^{inx}$ converges on a set of positive measure, then the $a_n \to 0$. Also, the notation suggests - and this is my intention - that the n's are integers. But in the proof below all that we will need is that the n's are any sequence of real numbers tending to ∞.)

Proof. The proof is a nice application of the dominated convergence theorem, which was invented by Lebesgue in part precisely to prove this theorem.
We may assume (by passing to a subset if necessary) that E is contained in some finite interval $[a, b]$. If $d_n \not\to 0$ then there is an $\epsilon > 0$ and a subsequence $d_{n_k} > \epsilon$ for all k. If the series converges, all its terms go to 0, so this means that

$$\cos(n_k t - \phi_k) \to 0 \quad \forall\ t \in E.$$

So

$$\cos^2(n_k t - \phi_k) \to 0 \quad \forall\ t \in E.$$

Now $m(E) < \infty$ and $\cos^2(n_k t - \phi_k) \leq 1$ and the constant 1 is integrable on $[a, b]$. So we may take the limit under the integral sign using the dominated convergence theorem to conclude that

$$\lim_{k\to\infty} \int_E \cos^2(n_k t - \phi_k) dt = \int_E \lim_{k\to\infty} \cos^2(n_k t - \phi_k) dt = 0.$$

But

$$\cos^2(n_k t - \phi_k) = \frac{1}{2}[1 + \cos 2(n_k t - \phi_k)]$$

so

$$\begin{aligned}\int_E \cos^2(n_k t - \phi_k)dt &= \frac{1}{2}\int_E [1 + \cos 2(n_k t - \phi_k)]dt \\ &= \frac{1}{2}\left[m(E) + \int_E \cos 2(n_k t - \phi_k)\right] \\ &= \frac{1}{2}m(E) + \frac{1}{2}\int_{\mathbb{R}} \mathbf{1}_E \cos 2(n_k t - \phi_k)dt.\end{aligned}$$

But $\mathbf{1}_E \in \mathcal{L}_1(\mathbb{R}, \mathbb{R})$ so the second term on the last line goes to 0 by the Riemann Lebesgue Lemma. So the limit is $\frac{1}{2}m(E)$ instead of 0, a contradiction. $\square$

Fubini's theorem.

This famous theorem asserts that under suitable conditions, a double integral is equal to an iterated integral. We will prove it for real (and hence finite dimensional) valued functions on arbitrary measure spaces. (The proof for Banach valued functions is a bit more tricky, and we shall omit it as we will not need it. This is one of the reasons why we have developed the real valued theory.) We begin with some facts about product σ-fields.

Product σ-fields.

Let $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ be spaces with σ-fields. On $X \times Y$ we can consider the collection $\mathcal{P}$ of all sets of the form

$$A \times B, \quad A \in \mathcal{F},\ B \in \mathcal{G}.$$

The σ-field generated by $\mathcal{P}$ will, by abuse of language, be denoted by

$$\mathcal{F} \times \mathcal{G}.$$

Sections.

If E is any subset of $X \times Y$, by an even more serious abuse of language we will let

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

and (contradictorially) we will let

$$E_y := \{x|(x,y) \in E\}.$$

The set $E_x \subset Y$ will be called the x-**section** of E and the set $E_y \subset X$ will be called the y-section of E.

Cylinder sets.

Finally we will let $\mathcal{C} \subset \mathcal{P}$ denote the collection of **cylinder sets**, that is sets of the form

$$A \times Y \qquad A \in \mathcal{F}$$

or

$$X \times B, \qquad B \in \mathcal{G}.$$

In other words, an element of $\mathcal{P}$ is a cylinder set when one of the factors is the whole space.

Theorem

.

- *$\mathcal{F} \times \mathcal{G}$ is generated by the collection of cylinder sets $\mathcal{C}$.*
- *$\mathcal{F} \times \mathcal{G}$ is the smallest σ-field on $X \times Y$ such that the projections*

$$\begin{array}{ll} \mathrm{pr}_X : X \times Y \to X & \mathrm{pr}_X(x,y) = x \\ \mathrm{pr}_Y : X \times Y \to Y & \mathrm{pr}_Y(x,y) = y \end{array}$$

  are measurable maps.
- *For each $E \in \mathcal{F} \times \mathcal{G}$ and all $x \in X$ the x-section E_x of E belongs to $\mathcal{G}$ and for all $y \in Y$ the y-section E_y of E belongs to $\mathcal{F}$.*

$\mathcal{F} \times \mathcal{G}$ is generated by the collection of cylinder sets $\mathcal{C}$.

Proof. $A \times B = (A \times Y) \cap (X \times B)$ so any σ-field containing $\mathcal{C}$ must also contain $\mathcal{P}$. $\square$

$\mathcal{F} \times \mathcal{G}$ is the smallest σ-field on $X \times Y$ such that the projections

$$\begin{aligned} \mathrm{pr}_X : X \times Y \to X \qquad \mathrm{pr}_X(x,y) = x \\ \mathrm{pr}_Y : X \times Y \to Y \qquad \mathrm{pr}_Y(x,y) = y \end{aligned}$$

are measurable maps.

Proof.

Since $\mathrm{pr}_X^{-1}(A) = A \times Y$, the map pr_X is measurable, and similarly for Y. But also, any σ-field containing all $A \times Y$ and $X \times B$ must contain $\mathcal{P}$ by what we just proved. $\square$

For each $E \in \mathcal{F} \times \mathcal{G}$ and $\forall x \in X$ the x-section E_x of E belongs to $\mathcal{G}$ and $\forall y \in Y$ the y-section E_y belongs to $\mathcal{F}$.

Proof.

Any set E of the form $A \times B$ has the desired section properties, since its x section is B if $x \in A$ or the empty set if $x \notin A$. Similarly for its y sections. So let $\mathcal{H}$ denote the collection of subsets E which have the property that all $E_x \in \mathcal{G}$ and all $E_y \in \mathcal{F}$. If we show that $\mathcal{H}$ is a σ-field we are done.
Now $E_x^c = (E_x)^c$ and similarly for y, so $\mathcal{H}$ is closed under taking complements. Similarly for countable unions:

$$\left(\bigcup_n E_n\right)_x = \bigcup_n (E_n)_x.$$

$\square$

Dynkin's π-systems.

Recall that the σ-field $\sigma(\mathcal{C})$ generated by a collection $\mathcal{C}$ of subsets of X is the intersection of all the σ-fields containing $\mathcal{C}$. Sometimes the collection $\mathcal{C}$ is closed under finite intersection. In that case, we call $\mathcal{C}$ a π-system. Examples:

- X is a topological space, and $\mathcal{C}$ is the collection of open sets in X.
- $X = \mathbb{R}$, and $\mathcal{C}$ consists of all half infinite intervals of the form $(-\infty, a]$. We will denote this π system by $\pi(\mathbb{R})$.

Dynkin's λ-systems.

A collection $\mathcal{H}$ of subsets of X will be called a λ-system if

1. $X \in \mathcal{H}$,
2. $A, B \in \mathcal{H}$ with $A \cap B = \emptyset \Rightarrow \; A \cup B \in \mathcal{H}$,
3. $A, B \in \mathcal{H}$ and $B \subset A \Rightarrow \; (A \setminus B) \in \mathcal{H}$, and
4. $\{A_n\}_1^\infty \subset \mathcal{H}$ and $A_n \nearrow A \; \Rightarrow A \in \mathcal{H}$.

From items 1) and 3) we see that a λ-system is closed under complementation, and since $\emptyset = X^c$ it contains the empty set. If $\mathcal{B}$ is both a π-system and a λ system, it is closed under any finite union, since $A \cup B = A \cup (B/(A \cap B)$ which is a disjoint union. Any countable union union can be written in the form $A = \nearrow A_n$ where the A_n are finite disjoint unions as we have already argued. So we have proved

Proposition

If $\mathcal{H}$ is both a π-system and a λ-system then it is a σ-field.

Also, we have

Proposition

[Dynkin's lemma.] *If $\mathcal{C}$ is a π-system, then the σ-field generated by $\mathcal{C}$ is the smallest λ-system containing $\mathcal{C}$.*

Let $\mathcal{M}$ be the σ-field generated by $\mathcal{C}$, and $\mathcal{H}$ the smallest λ-system containing $\mathcal{C}$. So $\mathcal{M} \supset \mathcal{H}$. By the preceding proposition, all we need to do is show that $\mathcal{H}$ is a π-system.

Proof.

Let

$$\mathcal{H}_1 := \{A|\ A \cap C \in \mathcal{H}\ \forall C \in \mathcal{C}\}.$$

Clearly $\mathcal{H}_1$ is a λ-system containing $\mathcal{C}$, so $\mathcal{H} \subset \mathcal{H}_1$ which means that $A \cap C \in \mathcal{C}$ for all $A \in \mathcal{H}$ and $C \in \mathcal{C}$.

Let

$$\mathcal{H}_2 := \{A|\ A \cap H \in \mathcal{H}\ \ \forall H \in \mathcal{H}\}.$$

$\mathcal{H}_2$ is again a λ-system, and it contains $\mathcal{C}$ by what we have just proved. So $\mathcal{H}_2 \supset \mathcal{H}$, which means that the intersection of two elements of $\mathcal{H}$ is again in $\mathcal{H}$, i.e. $\mathcal{H}$ is a π-system. □

The monotone class theorem.

Theorem

Let **B** *be a class of bounded real valued functions on a space* Z *satisfying*

1. **B** *is a vector space over* $\mathbb{R}$.
2. *The constant function* **1** *belongs to* **B**.
3. **B** *contains the indicator functions* $\mathbf{1}_A$ *for all* A *belonging to a* π*-system* $\mathcal{I}$.
4. *If* $\{f_n\}$ *is a sequence of non-negative functions in* **B** *and* $f_n \nearrow f$ *where* f *is a bounded function on* Z, *then* $f \in \mathbf{B}$.

Then **B** *contains every bounded* $\mathcal{M}$ *measurable function, where* $\mathcal{M}$ *is the* σ*-field generated by* $\mathcal{I}$.

Proof of the monotone class theorem, part I.

Let $\mathcal{H}$ denote the class of subsets of Z whose indicator functions belong to $\mathbf{B}$. Then $Z \in \mathcal{H}$ by item 2). If $B \subset A$ are both in $\mathcal{H}$, then $\mathbf{1}_{A \setminus B} = \mathbf{1}_A - \mathbf{1}_B$ and so $A \setminus B$ belongs to $\mathcal{H}$ by item 1). Similarly, if $A \cap B = \emptyset$ then $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B$ and so if A and B belong to $\mathcal{H}$ so does $A \cup B$ when $A \cap B = \emptyset$. Finally, condition 4) in the theorem implies condition 4) in the definition of a λ-system. So we have proved that that $\mathcal{H}$ is a λ-system containing $\mathcal{I}$. So by Dynkin's lemma, it contains $\mathcal{M}$.

Proof of the monotone class theorem, part II.

Now suppose that $0 \le f \le K$ is a bounded $\mathcal{M}$ measurable function, where we may take K to be an integer. For each integer $n \ge 0$ divide the interval $[0, K]$ up into subintervals of size 2^{-n}, and let

$$A(n,i) := \{z | i2^{-n} \le f(z) < (i+1)2^{-n}\}$$

where i ranges from 0 to $K2^n$. Let

$$s_n(z) := \sum_{i=0}^{K2^n} \frac{i}{2^n} \mathbf{1}_{A(n,i)}.$$

Since f is assumed to be $\mathcal{M}$-measurable, each $A(n,i) \in \mathcal{M}$, so by the preceding, and condition 1), $s_n \in \mathbf{B}$. But $0 \le s_n \nearrow f$, and hence by condition 4), $f \in \mathbf{B}$.

Proof of the monotone class theorem, conclusion.

For a general bounded $\mathcal{M}$ measurable f, both f^+ and f^- are bounded and $\mathcal{M}$ measurable, and hence by the preceding and condition 1), $f = f^+ - f^- \in \mathbf{B}$. $\square$

We now want to apply the monotone class theorem to our situation of a product space. So $Z = X \times Y$, where $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ are spaces with σ-fields, and where we take $\mathcal{I} = \mathcal{P}$ to be the π-system consisting of the product sets $A \times B, \quad A \in \mathcal{F}, \; B \in \mathcal{G}$.

Proposition

Let $\mathbf{B}$ *consist of all bounded real valued functions* f *on* $X \times Y$ *which are* $\mathcal{F} \times \mathcal{G}$*-measurable, and which have the property that*

- *for each* $x \in X$, *the function* $y \mapsto f(x, y)$ *is* $\mathcal{G}$*-measurable, and*
- *for each* $y \in Y$ *the function* $x \mapsto f(x, y)$ *is* $\mathcal{F}$*-measurable.*

Then $\mathbf{B}$ *consists of all bounded* $\mathcal{F} \times \mathcal{G}$ *measurable functions.*

Proof.

Indeed, $y \mapsto \mathbf{1}_{A\times B}(x,y) = \mathbf{1}_B(y)$ if $x \in A$ and $= 0$ otherwise; and similarly for $x \mapsto \mathbf{1}_{A\times B}(x,y)$. So condition 3) of the monotone class theorem is satisfied, and the other conditions are immediate. Since $\mathcal{F} \times \mathcal{G}$ was defined to be the σ-field generated by $\mathcal{P}$, the proposition is an immediate consequence of the monotone class theorem. $\square$

Fubini for finite measures and bounded functions.

Let $(X, \mathcal{F}, m)$ and $(Y, \mathcal{G}, n)$ be measure spaces with $m(X) < \infty$ and $n(Y) < \infty$. For every bounded $\mathcal{F} \times \mathcal{G}$-measurable function f, we know that the function

$$f(x, \cdot) : y \mapsto f(x, y)$$

is bounded and $\mathcal{G}$ measurable. Hence it has an integral with respect to the measure n, which we will denote by

$$\int_Y f(x, y) n(dy).$$

This is a bounded function of x (which we will prove to be $\mathcal{F}$ measurable in just a moment). Similarly we can form

$$\int_X f(x, y) m(dx)$$

which is a function of y.

Proposition

Let **B** *denote the space of bounded* $\mathcal{F} \times \mathcal{G}$ *measurable functions such that*

- $\int_Y f(x,y)n(dy)$ *is a* $\mathcal{F}$ *measurable function on* X*,*
- $\int_X f(x,y)m(dx)$ *is a* $\mathcal{G}$ *measurable function on* Y *and*
- $$\int_X \left(\int_Y f(x,y)n(dy) \right) m(dx) = \int_Y \left(\int_X f(x,y)m(dx) \right) n(dy). \tag{27}$$

Then **B** *consists of all bounded* $\mathcal{F} \times \mathcal{G}$ *measurable functions.*

Proof.

We have verified that the first two items hold for $\mathbf{1}_{A\times B}$. Both sides of (27) equal $m(A)n(B)$ as is clear from the proof of Proposition 9. So conditions 1-3 of the monotone class theorem are clearly satisfied, and condition 4) is a consequence of two double applications of the monotone convergence theorem. □

Now for any $C \in \mathcal{F} \times \mathcal{G}$ we define

$$(m \times n)(C)$$

$$:= \int_X \left(\int_Y \mathbf{1}_C(x,y) n(dy) \right) m(dx) = \int_Y \left(\int_X \mathbf{1}_C(x,y) m(dx) \right) n(dy), \tag{28}$$

both sides being equal on account of the preceding proposition. This measure assigns the value $m(A)n(B)$ to any set $A \times B \in \mathcal{P}$, and since $\mathcal{P}$ generates $\mathcal{F} \times \mathcal{G}$ as a sigma field, any two measures which agree on $\mathcal{P}$ must agree on $\mathcal{F} \times \mathcal{G}$. Hence $m \times n$ is the unique measure which assigns the value $m(A)n(B)$ to sets of $\mathcal{P}$.

Furthermore, we know that

$$\int_{X\times Y} f(x,y)(m\times n)$$

$$= \int_X \left(\int_Y f(x,y)n(dy)\right) m(dx) = \int_Y \left(\int_X f(x,y)m(dx)\right) n(dy) \tag{29}$$

is true for functions of the form $\mathbf{1}_{A\times B}$ and hence by the monotone class theorem it is true for all bounded functions which are measurable relative to $\mathcal{F}\times\mathcal{G}$.

The above assertions are the content of Fubini's theorem for bounded measures and functions. We summarize:

Statement of Fubini's theorem for the case of bounded measures and bounded measurable functions.

Theorem

Let $(X, \mathcal{F}, m)$ and $(Y, \mathcal{G}, n)$ be measure spaces with $m(X) < \infty$ and $n(Y) < \infty$. There exists a unique measure on $\mathcal{F} \times \mathcal{G}$ with the property that

$$(m \times n)(A \times B) = m(A)n(B) \quad \forall\ A \times B \in \mathcal{P}.$$

For any bounded $\mathcal{F} \times \mathcal{G}$ measurable function, the double integral is equal to the iterated integral in the sense that (29) holds.

Extensions to unbounded functions and to σ-finite measures.

Suppose that we temporarily keep the condition that $m(X) < \infty$ and $n(Y) < \infty$. Let f be any non-negative $\mathcal{F} \times \mathcal{G}$-measurable function. We know that (29) holds for all bounded measurable functions, in particular for all simple functions. We know that we can find a sequence of non-negative simple functions s_n such that $s_n \nearrow f$. Hence by several applications of the monotone convergence theorem, we know that (29) is true for all non-negative $\mathcal{F} \times \mathcal{G}$-measurable functions in the sense that all three terms are infinite together, or finite together and equal. Now we have agreed to call a $\mathcal{F} \times \mathcal{G}$-measurable function f integrable if and only if f^+ and f^- have finite integrals. In this case (29) holds.

A measure space $(X, \mathcal{F}, m)$ is called σ-**finite** if $X = \bigcup_n X_n$ where $m(X_n) < \infty$. In other words, X is σ-finite if it is a countable union of finite measure spaces. As usual, we can then write X as a countable union of disjoint finite measure spaces. So if X and Y are σ-finite, we can write the various integrals that occur in (29) as sums of integrals which occur over finite measure spaces. A bit of standard argumentation shows that Fubini continues to hold in this case.

If X or Y is not σ-finite, or, even in the finite case, if f is not non-negative or $m \times n$ integrable, then Fubini need not hold.

The Borel transform.

The following circle of ideas will be useful in one of our proofs of the spectral theorem, and is a nice application of the dominated convergence theorem and Fubini's theorem;
Suppose that μ is a finite non-negative Borel measure on $\mathbb{R}$. Its Borel transform $F = F_\mu$ is the function defined on $\mathbb{C} \setminus \mathbb{R}$ by

$$F(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z}.$$

F is clearly holomorphic on $\mathbb{C} \setminus \mathbb{R}$ and satisfies

$$F(z)^* = F(z^*) \quad \text{and} \quad |F(z)| \leq \frac{\mu(\mathbb{R})}{|\operatorname{Im} z|}.$$

there.

$$F(z)^* = F(z^*) \quad \text{and} \quad |F(z)| \leq \frac{\mu(\mathbb{R})}{|\operatorname{Im} z|}.$$

Also

$$\operatorname{Im}(F(z)) = \int \operatorname{Im}\left(\frac{1}{\lambda - z}\right) d\mu(\lambda) = \operatorname{Im}(z) \int_{\mathbb{R}} \frac{d\mu(\lambda)}{|\lambda - z|^2}$$

So F carries the upper and lower half-planes into themselves. A function with this property is called a **Herglotz function**.

The Stieltjes inversion formula.

This allows us to recover μ from its Borel transform. It says that

$$\frac{1}{2}\left(\mu((\lambda_1,\lambda_2)) + \mu([\lambda_1,\lambda_2])\right) = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{\lambda_1}^{\lambda_2} \mathrm{Im}(F(\lambda + i\epsilon)d\lambda. \quad (30)$$

To prove this formula set $z = x + i\epsilon$ in our formula

$$\mathrm{Im}(F(z)) = \mathrm{Im}(z)\int_{\mathbb{R}} \frac{d\mu(\lambda)}{|\lambda - z|^2}$$

to obtain

$$\mathrm{Im}\, F(x + i\epsilon) = \int_{\mathbb{R}} \frac{\epsilon}{(x-\lambda)^2 + \epsilon^2} d\mu(\lambda).$$

So the integral occurring on the right hand side of (30) can be written (by Fubini) as

$$\int_{\mathbb{R}} \int_{\lambda_1}^{\lambda_2} \frac{\epsilon}{(x-\lambda)^2+\epsilon^2} d\lambda d\mu(x).$$

We can do the inner integral to obtain

$$\frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \frac{\epsilon}{(x-\lambda)^2+\epsilon^2} d\lambda = \frac{1}{\pi} \left(\arctan\left(\frac{\lambda_2 - x}{\epsilon}\right) - \arctan\left(\frac{\lambda_1 - x}{\epsilon}\right)\right).$$

The expression on the right is uniformly bounded, and approaches zero if $x \notin [\lambda_1, \lambda_2]$, approaches $\frac{1}{2}$ if $x = \lambda_1$ or $x = \lambda_2$, and approaches 1 if $x \in (\lambda_1, \lambda_2)$. In short, it approaches

$$\frac{1}{2}\left(\mathbf{1}_{[\lambda_1,\lambda_2]} + \mathbf{1}_{(\lambda_1,\lambda_2)}\right).$$

We can apply the dominated convergence theorem to the outer integral to conclude the Stieltjes inversion formula.

Summary.