Math 7350 1. Real Numbers Fall 2004

We assume the usual properties of the rationals $\mathbb{Q}$ and define a *real number* as a set of rationals $x \subseteq \mathbb{Q}$ with the following properties (p, q, r denote rationals, x, y, z denote reals):

D1. $x \neq \emptyset, \mathbb{Q}$.

D2. If $p \in x$ then $\forall q < p : q \in x$.

D3. If $p \in x$ then $\exists q > p : q \in x$.

Let the set of reals be denoted by $\mathbb{R}$. For every $p \in \mathbb{Q}$, the set $[p] = \{q \in \mathbb{Q} : q < p\}$ satisfies D1–D3, so we identify $p \in \mathbb{Q}$ with $[p] \in \mathbb{R}$. The sets x and $x^c = \mathbb{Q} \setminus x$ partition $\mathbb{Q}$, and one thinks of the real number as the dividing point between the two sets. We define an order $\leq$ on $\mathbb{R}$ by $x \leq y$ iff $x \subseteq y$.

Lemma 1 *$\leq$ is a total ordering on $\mathbb{R}$.*

Proof. Since $\subseteq$ is always a partial order, it is enough to show that if $x, y \in \mathbb{R}$ then either $x \subseteq y$ or $y \subseteq x$. Assume $x \not\subseteq y$, so that there exists a $p \in \mathbb{Q}$ with $p \in x$ but $p \notin y$. Assume $q \in y$. Then $q < p$ by D2. Thus $q \in x$ by D2. Hence $y \subseteq x$. □

Lemma (Least Upper Bound Axiom)
If a non-empty set of reals has an upper bound, then it has a least upper bound.

Proof. Assume S is a non-empty set of reals with upper bound $x_1 \in \mathbb{R}$. Let $x = \bigcup_{y \in S} y$. Clearly $x \neq \emptyset$, and $x \subseteq x_1 \subset \mathbb{Q}$, so D1 holds. If $p \in x$ then $p \in y$ for some $y \in S$. If $q < p$ then $q \in y \subseteq x$, so D2 holds. There is a $q > p$ with $q \in y$, so $q \in x$ and D3 also holds. Thus $x \in \mathbb{R}$. Clearly $y \leq x$ for all $y \in S$, and if $y \leq x'$ for all $y \in S$ then $x = \bigcup_{y \in S} y \subseteq x'$, so $x \leq x'$. Thus x is a least upper bound for S. □

If S is a non-empty set with an upper bound, we write the least upper bound as $\sup S$. Note that in general $\sup S$ may not lie in S. Define addition on $\mathbb{R}$ by $x+y = \{p+q : p \in x,\ q \in y\}$ and multiplication on $\mathbb{R}$ by $xy = \{r : \exists p \in x, p' \notin x, q \in y, q' \notin y : r < pq, p'q, pq', p'q'\}$.

Theorem 1 *These operations give elements of $\mathbb{R}$, and under these operations $\mathbb{R}$ forms an* ordered field, *i.e.*

1. *$+$ is associative, commutative, has identity 0, and inverses $-x$,*
2. *$\times$ is associative, commutative, has identity $1 \neq 0$, and inverses x^{-1} for $x \neq 0$,*
3. *$\times$ distributes over $+$: $x(y+z) = xy + xz$,*
4. *$\leq$ is a total order,*
5. *$\leq$ respects $+$: $x \leq y$ and $z \leq t$ imply $x + z \leq y + t$,*
6. *$\leq$ respects $\times$: $x \leq y$ and $0 \leq z$ imply $xz \leq yz$.*

Proof. Fairly easy, but tedious, check. □

From now on, we can forget the construction of $\mathbb{R}$, and just use the fact that it is an ordered field and satisfies the least upper bound axiom. These facts are enough to prove all the results about $\mathbb{R}$ that we shall need.

We can now define the usual notions of subtraction, division, and absolute value:

$$x - y = x + (-y), \qquad \tfrac{x}{y} = xy^{-1} \text{ for } y \neq 0, \qquad |x| = \max\{x, -x\}.$$

Exercise: Using just the ordered field and least upper bound axioms, prove

1. $0x = 0$,
2. $(-1)x = -x$,
3. $-(-x) = x$,
4. $0 < 1$,
5. $|x| \geq 0$,
6. $|x + y| \leq |x| + |y|$,
7. $x \leq y \Rightarrow -y \leq -x$
8. $z \leq 0$ and $x \leq y \Rightarrow xz \geq yz$,

Theorem 2 *If F is an ordered field satisfying the least upper bound axiom, then there exists a bijection $f : \mathbb{R} \to F$ such that for all $x, y \in \mathbb{R}$,*
(a) $x \leq y \Leftrightarrow f(x) \leq f(y)$, (b) $f(x + y) = f(x) + f(y)$, (c) $f(xy) = f(x)f(y)$.

Proof. (sketch)
1. Define $f_1 : \mathbb{N} \to F$ by $f_1(0) = 0$ and inductively $f_1(n + 1) = f_1(n) + 1$. One can show by induction that (a)–(c) hold for f_1.
2. Define $f_2 : \mathbb{Q} \to F$ by $f_2(p/q) = f_1(p)/f_1(q)$. One can show that this is well defined and satisfies (a)–(c).
3. Define $f : \mathbb{R} \to F$ by $f(x) = \sup_{p \in x} f_2(p)$. One can check that this is well defines and satisfies (a)–(c).
4. Finally, f is injective by (a), and if $z \in F$ one can check that $z = f(\{p \in \mathbb{Q} : f_2(p) < z\})$, so f is surjective. □

Lemma (Archimedean Property) *If $x, y \in \mathbb{R}$ and $x > 0$ then there exists a natural number n such that $nx > y$.*

Proof. Let $S = \{nx : n \in \mathbb{N}\}$. This set is non-empty and does not have a least upper bound: if $z = \sup S$ then $z - x$ is not an upper bound, so there exists an n with $z - x < nx$. But then $(n + 1)x > z$, a contradiction. Hence S has no upper bound, and so no y can exist such that $nx \leq y$ for all n. □

Corollary *Between any two real numbers there exists a rational number.*

Proof. Assume $x < y$. Then $\exists n \in \mathbb{N} : n(y - x) > 1$. Now $\exists s \in \mathbb{N} : s1 > -xn$ and $\exists t \in \mathbb{N} : t1 > xn + s$. Pick the smallest such t. Then $\frac{t-s}{n}$ will do. □

Math 7350 2. Limits Fall 2004

Definition A sequence (x_i) is *increasing* (*strictly increasing*) if $i < j$ implies $x_i \leq x_j$ ($x_i < x_j$). A sequence (x_i) is *decreasing* (*strictly decreasing*) if $i < j$ implies $x_i \geq x_j$ ($x_i > x_j$). A sequence is *(strictly) monotonic* if it is either (strictly) increasing or (strictly) decreasing. A *subsequence* of (x_i) is a sequence of the form (x_{n_i}) where (n_i) is strictly increasing.

Lemma *Every sequence has a monotonic subsequence.*

Proof. Assume (x_n) is a sequence and construct an *upper* subsequence (x_{u_i}) and a *lower* subsequence (x_{l_j}) as follows. Suppose we have constructed x_{u_i} for $i < i_0$ and x_{l_j} for $j < j_0$. Let the last elements of these subsequences be $x_u = x_{u_{i_0-1}}$ and $x_l = x_{l_{j_0-1}}$ (or $+\infty$ and $-\infty$ if $i_0 = 0$ or $j_0 = 0$ respectively). At each stage there will be infinitely many x_n such that $x_l \leq x_n \leq x_u$. Pick the smallest such $n > u_{i_0-1}, l_{j_0-1}$. If there are infinitely many x_m with $x_n \leq x_m \leq x_u$ then add x_n to the lower subsequence, $l_{j_0} = n$, otherwise there are infinitely many x_m with $x_l \leq x_m \leq x_n$ and we can add x_n to the upper subsequence, $u_{i_0} = n$. Repeat this process inductively. The subsequence (x_{u_n}) is decreasing and the subsequence (x_{l_n}) is increasing. At least one of these subsequences is infinite. □

We define the *extended real numbers* $\mathbb{R}^\star = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$, with the usual conventions, $\infty + x = \infty$, $\infty x = -\infty$ for $x < 0$, etc. We leave $\infty - \infty$ undefined, but let $\infty.0 = 0$. We now extend $\sup S$ to arbitrary sets of (extended) reals S by setting $\sup \emptyset = -\infty$, and $\sup S = +\infty$ if S has no upper bound. We define $\inf S = -\sup\{-x : x \in S\}$.

Assume $(x_n)_{n=0}^\infty$ is an infinite sequence of real numbers and $L \in \mathbb{R}$. We define

L is a *limit* of (x_n), $\lim_{n\to\infty} x_n = L$ $\Leftrightarrow$ $\forall \varepsilon > 0\colon \exists n_0\colon \forall n \geq n_0\colon |x_n - L| < \varepsilon$,
L is a *cluster* (or *accumulation*) point of (x_n) $\Leftrightarrow$ $\forall \varepsilon > 0\colon \forall n_0\colon \exists n \geq n_0\colon |x_n - L| < \varepsilon$.

Note that $\exists n_0\colon \forall n \geq n_0\colon$ can be translated "for all but finitely many (f.a.b.f.m.) n", whereas $\forall n_0\colon \exists n \geq n_0\colon$ can be translated "there exists infinitely many ($\exists\infty$-many) n". We also allow $L = \pm\infty$, by, replacing $\forall \varepsilon > 0\colon \cdots : |x_n - L| < \varepsilon$ by $\forall K > 0\colon \cdots : x_n > K$ when $L = +\infty$, and by $\forall K > 0\colon \cdots : x_n < -K$ when $L = -\infty$.

Definition We say $(x_n)_{n=0}^\infty$ *converges* or is *convergent* if $\lim x_n$ exists *and is finite.*

Definition We say $(x_n)_{n=0}^\infty$ is *bounded* if $\exists K > 0\colon \forall n\colon |x_n| < K$.

Lemma *Every bounded monotonic sequence converges.*

Proof. Assume w.l.o.g., (x_n) is increasing and let $L = \sup x_n$. Since (x_n) is bounded, $L \in \mathbb{R}$. Since $L - \varepsilon$ is not an upper bound, $\exists n_0 : x_{n_0} > L - \varepsilon$. But (x_n) is increasing, so $\forall n \geq n_0 : x_n > L - \varepsilon$. But $\forall n \geq n_0 : x_n \leq L$. Thus $|x_n - L| < \varepsilon$ and $\lim x_n = L$. □

Theorem (Bolzano-Weierstrass) *Every bounded sequence has a convergent subsequence.*

Proof. Pick a monotonic subsequence of (x_n). □

A sequence (x_n) is a *Cauchy sequence* iff $\forall \varepsilon > 0 \colon \exists n_0 \colon \forall n, m \geq n_0 \colon |x_n - x_m| < \varepsilon$.

Lemma *A sequence is a Cauchy sequence iff it converges.*

Proof. If (x_n) converges, then $\forall \varepsilon > 0 \colon \exists n_0 \colon \forall n \geq n_0 \colon |x_n - L| < \varepsilon/2$. But then $\forall n, m \geq n_0 \colon |x_n - x_m| \leq |x_n - L| + |L - x_m| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$, and (x_n) is Cauchy. Conversely, if (x_n) is Cauchy, pick a monotonic subsequence (x_{n_i}). Now $\forall \varepsilon > 0 : \exists n_0 \colon \forall n, m \geq n_0 \colon |x_n - x_m| < \varepsilon/2$. But then (x_{n_i}) is bounded (most terms by $x_{n_0} \pm \varepsilon/2$). If $\lim x_{n_i} = L$ then by taking $m = n_i$ for large enough i, $|x_m - L| < \varepsilon/2$, so $|x_n - L| < \varepsilon$. Thus (x_n) converges to L. □

We can define 'one-sided' limits:

$\overline{\lim}_{n\to\infty} a_n = L \quad \Leftrightarrow \quad \forall \varepsilon > 0 :$ f.a.b.f.m. $n : a_n \leq L + \varepsilon$ and $\exists\infty$-many $n : a_n \geq L - \varepsilon$,

$\underline{\lim}_{n\to\infty} a_n = L \quad \Leftrightarrow \quad \forall \varepsilon > 0 :$ f.a.b.f.m. $n : a_n \geq L - \varepsilon$ and $\exists\infty$-many $n : a_n \leq L + \varepsilon$.

And similarly for $L = \pm\infty$.

Lemma *Let (x_n) be a sequence of real numbers and let $L \in \mathbb{R}^\star$.*

1. *There is at least one cluster point $L \in \mathbb{R}^\star$ of (x_n),*
2. *(x_n) is bounded iff neither $+\infty$ nor $-\infty$ is a cluster point of (x_n),*
3. *$\underline{\lim}\, x_n$ is the smallest and $\overline{\lim}\, x_n$ is the largest cluster point of (x_n) (both exist $\in \mathbb{R}^\star$).*
4. *$\lim x_n = L$ exists $\Leftrightarrow$ there is exactly one cluster point of (x_n), namely L.*
5. *Any cluster point of a subsequence (x_{n_i}) is also a cluster point of (x_n).*
6. *L is a cluster point of (x_n) $\Leftrightarrow$ there exists a subsequence of (x_n) with limit L.*
7. *$\overline{\lim}_{n\to\infty} x_n = \lim_{n_0} \sup_{n \geq n_0} x_n, \quad \underline{\lim}_{n\to\infty} x_n = \lim_{n_0} \inf_{n \geq n_0} x_n$.*
8. *If (x_n) is increasing (resp. decreasing), then $\lim x_n = \sup_n x_n$ (resp. $\inf_n x_n$).*
9. *$\lim(x_n \pm y_n) = \lim x_n \pm \lim y_n$, $\lim x_n y_n = \lim x_n \lim y_n$, $\lim(x_n/y_n) = \lim x_n / \lim y_n$, provided these expressions are defined (and not $\infty.0$).*

We define the limit of a series $\sum_{i=0}^{\infty} x_i$ as $\lim_{n\to\infty} \sum_{i=0}^{n} x_i$.

Lemma

1. *If $\sum_{i=0}^{\infty} x_n$ converges then $\lim x_n = 0$.*
2. *If $\sum_{i=0}^{\infty} |x_n|$ converges then $\sum_{i=0}^{\infty} x_n$ converges.*

The examples $\sum \frac{1}{n}$ and $\sum \frac{(-1)^n}{n}$ show that the converses to these statements are false.

Warning: Don't interchange limits etc. In general, $\lim_n \lim_m x_{n,m} \neq \lim_m \lim_n x_{n,m}$.

Math 7350 3. Open and Closed Sets Fall 2004

Definition We say x is in the *interior* of a set S, if $\exists \varepsilon > 0$: $(x-\varepsilon, x+\varepsilon) \subseteq S$. The *interior* of S is the set $\mathring{S}$ of all interior points of S. A set U is *open* iff $U = \mathring{U}$.

Examples (a,b) is open, but $[a,b]$ and $[a,b)$ are not (take $x = a$). $\mathbb{Q}$ is not open.

Lemma *1. $\emptyset$ and $\mathbb{R}$ are open.*
2. The union of any collection of open sets is open.
3. The intersection of any finite *collection of open sets is open.*
4. $\mathring{S}$ is open, and is the largest open subset of S, $\mathring{S} = \bigcup_{open\ U \subseteq S} U$.

Definition We say x is a *point of closure* of S if $\forall \varepsilon > 0$: $(x-\varepsilon, x+\varepsilon) \cap S \neq \emptyset$. The *closure* $\bar{S}$ of S is the set of all points of closure of S. A set F is *closed* iff $F = \bar{F}$.

Lemma $(\bar{S})^c = (S^c)^\circ$. *In particular S is closed iff S^c is open.*

Corollary *1. $\emptyset$ and $\mathbb{R}$ are closed.*
2. The intersection of any collection of closed sets is closed.
3. The union of any finite *collection of closed sets is closed.*
4. $\bar{S}$ is closed, and is the smallest closed set containing S, $\bar{S} = \bigcap_{closed\ F \supseteq S} F$.

Note that *any* set is a union of closed sets (singletons) and an intersection of open sets (complements of singletons).

Exercise: Show that $\overline{S_1 \cup S_2} = \bar{S}_1 \cup \bar{S}_2$. Give an example where $\overline{S_1 \cap S_2} \neq \bar{S}_1 \cap \bar{S}_2$.

Lemma $L \in \bar{S}$ *iff there exists a sequence (x_n) with $x_n \in S$ such that* $\lim_{n\to\infty} x_n = L$.

Proof. $\Rightarrow$: $\forall n$: $\exists x \in S \cap (L - \frac{1}{n}, L + \frac{1}{n})$. Set x_n to be one such x.
$\Leftarrow$: $\forall \varepsilon > 0$: $\exists n_0$: $\forall n \geq n_0$: $|x_n - L| < \varepsilon$, hence $x_{n_0} \in S \cap (L-\varepsilon, L+\varepsilon)$. □

Example Define the *Cantor set* by $C = \bigcap_{i=0}^\infty C_i$ where $C_0 = [0,1]$, $C_{n+1} = \{x/3 : x \in C_n\} \cup \{(x+2)/3 : x \in C_n\}$. Then C_{n+1} is obtained from C_n by removing the middle third of each subinterval of C_n. C_n and hence C is closed. The set C can also be described as all $x \in [0,1]$ which can be written in base 3, $x = \sum_{i=1}^\infty a_n 3^{-n}$, with all $a_n \in \{0,2\}$. If $x, y \in C$ with $x < y$, then there exists $z \notin C$ with $x < z < y$. Thus C contains no non-trivial interval. Moreover, C is uncountable since each choice of (a_n) gives a distinct $x \in C$.

Definition An *interval* is a subset I of $\mathbb{R}$ such that if $x, y \in I$ then $\forall z : x \leq z \leq y \Rightarrow z \in I$. If $\inf I < z < \sup I$ then $z \in I$, so $I = [a,b]$, $(a,b]$, $[a,b)$, or (a,b) with $a, b \in \mathbb{R}^\star$.

Definition If $S \subseteq \mathbb{R}$, define an equivalence relation on S by $x \sim y$ iff $[x,y] \subseteq S$ if $x \leq y$ (or $[y,x] \subseteq S$ if $y \leq x$). We define a *component* of S to be an equivalence class of $\sim$. Note that components are non-empty intervals (possibly single points).

Lemma *Every open set is a disjoint union of countably many open intervals.*

Proof. Consider the component interval I_t of $t \in U$. If $x \in I_t$ then $\exists \varepsilon > 0\colon (x-\varepsilon, x+\varepsilon) \subseteq U$, so $\forall z \in (x-\varepsilon, x+\varepsilon)\colon z \sim x \sim t$ and so $(x-\varepsilon, x+\varepsilon) \subseteq I_t$. In particular I_t is open, so $U = \bigcup_{t\in U} I_t$ is a disjoint union of open intervals. There are only countably many components since each component contains a rational ($\exists$ a surjection from $\mathbb{Q}$ to components). $\square$

Note: Although any open set is countable union of open components, there may be uncountably many 'gaps' between the components (e.g., the complement of the Cantor set).

Definition If $S \subseteq \bigcup_{i\in I} U_i$ then $\{U_i : i \in I\}$ is said to be a *cover* of S. If the U_i are open, we call it an *open cover*. The cover is finite (resp. countable) if I is finite (resp. countable). We say S is *compact* if every open cover of S has a finite subcover, i.e., if $S \subseteq \bigcup_{i\in I} U_i$ then there is a finite $I_0 \subseteq I$ such that $S \subseteq \bigcup_{i\in I_0} U_i$. We say S is *connected* if whenever U_1 and U_2 are open, $S \subseteq U_1 \cup U_2$, and $U_1 \cap U_2 \cap S = \emptyset$, then either $S \subseteq U_1$ or $S \subseteq U_2$.

Theorem (Lindelöf) *If $S \subseteq \mathbb{R}$, then any open cover of S has a countable subcover.*

Proof. For all $x \in S \subseteq \bigcup_{i\in I} U_i$, $\exists i \in I$ such that $x \in U_i$ and $\exists p, q \in \mathbb{Q}$, $p < x < q$ with $x \in (p,q) \subseteq U_i$. Now for each pair of rationals $p < q$, pick (if it exists) $i \in I$ such that $(p,q) \subseteq U_i$. Let I_0 be the set of all i chosen. Then I_0 is countable and $x \in \bigcup_{i\in I_0} U_i$. $\square$

Theorem (Heine-Borel) *F is compact iff F is closed and bounded.*

Proof. Suppose first that $F = [a,b]$. Let $S = \{x \in [a,b] : \exists \text{ finite } I_0 \subseteq I\colon [a,x] \subseteq \bigcup_{i\in I_0} U_i\}$. Clearly $a \in S$ and b is an upper bound for S, so $c = \sup S$ exists and $c \in [a,b]$. Then $c \in U_{i_0}$ for some $i_0 \in I$, so $(c-\varepsilon, c+\varepsilon) \subseteq U_{i_0}$ for some $\varepsilon > 0$. Now $\exists d \in S : d > c - \varepsilon$, and a finite I_0 such that $[a,d] \subseteq \bigcup_{i\in I_0} U_i$. Hence $[a, c+\varepsilon) \subseteq \bigcup_{i\in I_0\cup\{i_0\}} U_i$. If $c < b$ then c is not an upper bound for S, and if $c = b$, $[a,b] \subseteq [a, c+\varepsilon)$ has a finite subcover.
In general, if F is bounded, then $F \subseteq [a,b]$ for some $a, b \in \mathbb{R}$. Now F^c is open and so $[a,b] \subseteq F^c \cup \bigcup U_i$ has a finite subcover. Removing F^c, this still gives a finite subcover of F.
If F is not bounded then $F \subseteq \bigcup_{n\in\mathbb{Z}}(n-1, n+1)$, but there is no finite subcover.
If F is not closed, pick a point of closure $x \notin F$. Then $F \subseteq \bigcup_{\varepsilon>0}[x-\varepsilon, x+\varepsilon]^c$, but there is no finite subcover. $\square$

Corollary *If F_i, $i \in I$, are closed, bounded, and $\bigcap_{i\in I_0} F_i \neq \emptyset$ for any finite I_0, $\bigcap_{i\in I} F_i \neq \emptyset$.*

Lemma *S is connected iff S is an interval.*

Proof. $\Rightarrow$: If $x, y \in S$ and $x < z < y$, consider $U_1 = (-\infty, z)$ and $U_2 = (z, \infty)$.
$\Leftarrow$: Assume $S \subseteq U_1 \cup U_2$, $S \not\subseteq U_i$. We may assume $x, y \in S$, $x \in U_1$, $y \notin U_1$, $x < y$. Let I_x be the component interval of $x \in U_1$, which we know is open, say (w,z), $x < z \leq y$. Now $z \in S$, so $z \in U_i$ for either $i = 1$ or $i = 2$. Let I_z be the component interval of $z \in U_i$. Then $(x,z) \cap I_z \cap I_x \neq \emptyset$. If $i = 1$ this implies $I_z = I_x$, so $z \in I_x = (w,z)$, a contradiction. If $i = 2$ then $(x,z) \cap I_z \cap I_x \subseteq S \cap U_2 \cap U_1$, so $S \cap U_1 \cap U_2 \neq \emptyset$. $\square$

Math 7350 4. Continuity Fall 2004

Definition Assume $S \subseteq \mathbb{R}$, $a \in \mathbb{R}$, and $f\colon S \setminus \{a\} \to \mathbb{R}$. Define limits by

$$\lim_{x \to a,\, x \in S} f(x) = L \quad \Leftrightarrow \quad \forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in S\colon 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon$$

One can also define cluster points, $\overline{\lim}$, and $\underline{\lim}$ similarly. One can also extend the definition to include $L = \pm\infty$, and/or $a = \pm\infty$. Results analogous to those of Section 2 hold, except that if $a \notin \overline{S \setminus \{a\}}$ then $\lim f(x) = L$ is vacuously true for all L but no L is a cluster point. (If $a \in \overline{S \setminus \{a\}}$ then a is called a *limit point* of S).

Definition Assume $S \subseteq \mathbb{R}$. We say $f\colon S \to \mathbb{R}$ is *continuous* at the point $x \in S$ if $\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall z \in S\colon |z - x| < \delta \Rightarrow |f(z) - f(x)| < \varepsilon$, i.e., $\lim_{z \to x,\, z \in S} f(z) = f(x)$. We say f is *continuous* if f is continuous at all $x \in S$.

Definition A subset T of a set S is called *relatively open* in S if $T = S \cap U$ for some open set U. Alternatively, $\forall x \in T\colon \exists \delta > 0\colon \forall z \in S\colon |z - x| < \delta \Rightarrow z \in T$.

Lemma *$f\colon S \to \mathbb{R}$ is continuous iff $f^{-1}[U]$ is relatively open in S for all open sets U.*

Theorem *If f is a continuous function on a (non-empty) closed and bounded set, then f is bounded and attains its bounds.*

Proof. Assume $f[F] \subseteq \bigcup_{i \in I} U_i$, where U_i are open. Then $f^{-1}[U_i] = F \cap V_i$ for some open V_i (e.g., $V_i = f^{-1}[U_i] \cup F^c$). $\bigcup f^{-1}[U_i] = f^{-1}[\bigcup U_i] \supseteq f^{-1}[f[F]] = F$, so $F \subseteq \bigcup_{i \in I} V_i$. Pick a finite subcover $F \subseteq \bigcup_{i \in I_0} V_i$. Now $f[F] = \bigcup_{i \in I_0} f[F \cap V_i] \subseteq \bigcup_{i \in I_0} U_i$. Hence $f[F]$ has a finite subcover, so is closed and bounded. Hence $f[F]$ contains $\sup f[F]$ and $\inf f[F]$. Thus f is bounded and attains its bounds. □

Theorem (Intermediate value theorem) *If f is continuous on an interval I then $f[I]$ is an interval.*

Proof. Pick U_1, U_2 open with $f[I] \subseteq U_1 \cup U_2$ and $U_1 \cap U_2 \cap f[I] = \emptyset$. Write $f^{-1}[U_i] = I \cap V_i$ where V_i are open. Then $I \subseteq V_1 \cup V_2$, $V_1 \cap V_2 \cap I = \emptyset$, thus $I \subseteq V_i$ for some i. But then $f[I] \subseteq f[V_i] \subseteq U_i$. Hence $f[I]$ is connected, and so is an interval. □

Uniformity

Definition $f\colon S \to \mathbb{R}$ is *uniformly continuous* if $\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x, y \in S\colon |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Note that the δ is independent of x.

Example $f(x) = x^2$ is continuous but not uniformly continuous on $\mathbb{R}$.

Theorem *If f is continuous on a closed bounded set F then f is uniformly continuous on F.*

Proof. Fix $\varepsilon > 0$. Then $\forall x \in F\colon \exists \delta_x > 0\colon \forall y \in F\colon |x-y| < \delta_x \Rightarrow |f(x) - f(y)| < \varepsilon/2$. Clearly $F \subseteq \bigcup_x (x - \delta_x/2, x + \delta_x/2)$. Pick a finite subcover corresponding to the points $x_1, \dots, x_n$ and assume $z \in (x_i - \delta_{x_i}/2, x_i + \delta_{x_i}/2)$. If $y \in F$ and $|z-y| < \delta_{x_i}/2$ then $|x_i - y| < \delta$, so $|f(z) - f(y)| \leq |f(z) - f(x_i)| + |f(x_i) - f(y)| < \varepsilon$. Thus if we set $\delta = \min\{\delta_{x_i}/2\} > 0$, $|z-y| < \delta \Rightarrow |f(z) - f(y)| < \varepsilon$ for all $z, y \in F$. □

Definition Suppose (f_n) is a sequence of functions $f_n\colon S \to \mathbb{R}$.
We say f_n *converges pointwise* to the function f if for all $x \in S$, $\lim_{n\to\infty} f_n(x) = f(x)$, i.e., $\forall x \in S\colon \forall \varepsilon > 0\colon \exists n_0\colon \forall n \geq n_0\colon |f_n(x) - f(x)| < \varepsilon$. We say f_n *converges uniformly* to f if $\forall \varepsilon > 0\colon \exists n_0\colon \forall n \geq n_0\colon \forall x \in S\colon |f_n(x) - f(x)| < \varepsilon$, i.e., the n_0 does not depend on x.

Theorem *If $f_n\colon S \to R$ are continuous and converge uniformly to $f\colon S \to \mathbb{R}$, then f is continuous.*

Proof. Fix $\varepsilon > 0$. Pick n_0 such that $|f_{n_0}(z) - f(z)| < \varepsilon/3$ for all $z \in S$. Fix $x \in S$. Now f_{n_0} is continuous, so $\exists \delta > 0\colon \forall y \in S\colon |y - x| < \delta \Rightarrow |f_{n_0}(y) - f_{n_0}(x)| < \varepsilon/3$. But then $|f(y) - f(x)| \leq |f(y) - f_{n_0}(y)| + |f_{n_0}(y) - f_{n_0}(x)| + |f_{n_0}(x) - f(x)| < \varepsilon$. □

Note: $f_n(x) = x^n$ converges pointwise but not uniformly on $[0,1]$ to the non-continuous function $f(x) = 0$, $x < 1$, $f(1) = 1$.

Borel sets

Definition A *Borel set* is an element of the σ-algebra, $\mathcal{B}$, generated by all open subsets.

Exercise: Show that $\mathcal{B}$ is also the σ-algebra generated by the intervals of the form (a, ∞), or by the intervals of the form by $[a, \infty)$, or by intervals of one of these forms with $a \in \mathbb{Q}$.

Definition A function $f\colon S \to \mathbb{R}$ is *Borel measurable* if for all $B \in \mathcal{B}$, $f^{-1}[B] = S \cap B'$ for some $B' \in \mathcal{B}$. Note that it is enough to assume $f^{-1}[(a, \infty)] = S \cap B'$ for all $a \in \mathbb{R}$.

Lemma *Any continuous function is Borel measurable.*

Definition An F-set is a closed set. A G-set is an open set. An F_σ-set is a countable union of closed sets. A G_δ-set is a countable intersection of open sets. An $F_{\sigma\delta}$-set is a countable intersection of F_σ-sets. A $G_{\delta\sigma}$-set is a countable union of G_δ-sets, etc.

Note that any G-set is also an F_σ-set, and any F-set is a G_δ-set. All $G_{\delta\sigma\delta\dots}$ or $F_{\sigma\delta\sigma\dots}$-sets are Borel sets. However, not all Borel sets are of one of these forms. On the other hand, not every subset of $\mathbb{R}$ is a Borel set. Indeed, $|\mathcal{B}| = |\mathbb{R}| < |\mathcal{P}(\mathbb{R})|$.

Exercises

1. Show that the set of points where $f\colon \mathbb{R} \to \mathbb{R}$ is continuous is a G_δ-set.
2. Show that the set of points where the sequence of continuous functions $f_n\colon \mathbb{R} \to \mathbb{R}$ converges is an $F_{\sigma\delta}$-set.

Math 7350 5. Lebesgue Measure Fall 2004

Our aim is to construct a notion of the "length" $\lambda(S)$ of a subset $S \subseteq \mathbb{R}$. We would like the following conditions:

1. $\lambda(S)$ is defined and $\lambda(S) \in [0, \infty]$ for all $S \subseteq \mathbb{R}$,
2. $\lambda([a, b]) = \lambda((a, b)) = b - a$,
3. If $S_1, S_2, \dots$ are disjoint then $\lambda(\bigcup S_i) = \sum \lambda(S_i)$
4. $\lambda(a + S) = \lambda(S)$ where $a + S = \{a + x : x \in S\}$ is a translate of S.

Unfortunately, no such λ exists. However, if we drop Condition 1 and only require λ to be defined on a large σ-algebra of subsets, then one can construct such a λ.

Definition A *measure* μ on a σ-algebra $\mathcal{A}$ is a $[0, \infty]$-valued *countably additive* set function, i.e., if $S_i \in \mathcal{A}$, $i \in I$, are disjoint and I is countable, then $\mu(\bigcup_{i \in I} S_i) = \sum_{i \in I} \mu(S_i)$. We allow I to be empty (so $\mu(\emptyset) = 0$), finite, or countably infinite.

Note: We do not allow uncountable unions: Condition 2 implies $\lambda(\{x\}) = 0$, $\lambda([0, 1]) = 1$, but $\sum_{x \in [0,1]} \lambda(\{x\}) = \sum 0 = 0$.

Lemma 1. *Any measure μ is* monotonic*: if $A, B \in \mathcal{A}$, $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*

Proof. If $A, B \in \mathcal{A}$, then $B \setminus A \in \mathcal{A}$ and $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$. □

As our first attempt, we define the Lebesgue *outer measure* of $S \subseteq \mathbb{R}$ to be $\lambda^*(S) = \inf_{\cup I_i \supseteq S} \sum_i l(I_i)$, where the infimum is over all countable unions of open intervals $I_i = (a_i, b_i)$ that contain S, and $l(I_i) = b_i - a_i$ is the length of I_i.

Lemma 2. *λ^* is monotonic, and* countably subadditive*: if S_i, $i \in I$, is a countable collection of sets then $\lambda^*(\bigcup_i S_i) \leq \sum_i \lambda^*(S_i)$.*

Proof. Monotonicity is clear. For subadditivity, choose open intervals I_{ij} so that $S_i \subseteq \bigcup_j I_{ij}$ and $\sum_j l(I_{ij}) \leq \lambda^*(S_i) + \varepsilon/2^i$, $i = 1, 2, \dots$. Then $\lambda^*(\bigcup S_i) \leq \sum_{ij} l(I_{ij}) \leq \sum_i \lambda^*(S_i) + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, $\lambda^*(\bigcup S_i) \leq \sum_i \lambda^*(S_i)$. □

Lemma 3. *If I is an interval then $\lambda^*(I) = l(I)$.*

Proof. First assume $I = [a, b]$ is closed and bounded. We need to show that if $I \subseteq \bigcup I_i$ then $l(I) \leq \sum l(I_i)$. By compactness of I, we can assume there are only finitely many I_i. W.l.o.g., we may assume that no I_i is contained in another I_j (otherwise remove it, and use induction on the number of I_i). Now order the $I_i = (a_i, b_i)$ so that $a_1 < a_2 < \cdots < a_n$ (the a_i are distinct, since if $a_i = a_j$ then either $I_i \subseteq I_j$ or $I_j \subseteq I_i$). Now it is clear that $b_1 > a_2$, $b_2 > a_3$, ..., $b_n > b$. Hence $(b_1 - a_1) + (b_2 - a_2) + \cdots + (b_n - a_n) \geq b_n - a_1 \geq b - a$ as required. Conversely, $I \subseteq (a - \varepsilon, b + \varepsilon)$, so $\lambda^*(I) \leq b - a + 2\varepsilon$ for any $\varepsilon > 0$. Thus $\lambda^*(I) = b - a$. For general intervals, use monotonicity, e.g., $\lambda^*([a + \varepsilon, b - \varepsilon]) \leq \lambda^*((a, b)) \leq \lambda^*([a, b])$ and let $\varepsilon \to 0$, or $\lambda^*((-\infty, a]) \geq \lambda^*([a - n, a]) = n$ and let $n \to \infty$. □

It is clear now that λ^* satisfies all the conditions except countable additivity. However, it is possible to gain additivity at the expense of being defined on a smaller class of sets.

Definition A set X is *Lebesgue measurable* if for all $A \subseteq \mathbb{R}$, $\lambda^*(A) = \lambda^*(A \cap X) + \lambda^*(A \setminus X)$.

Note we have $\leq$ by subadditivity, so we only need to check $\geq$.

Lemma 4. *The collection of Lebesgue measurable sets is a σ-algebra.*

Proof. Clearly $\emptyset$ is measurable. Also if X is measurable, then so is X^c. Thus it is enough to show that if $X_1, X_2, \ldots$ are measurable, then so is $\bigcup X_i$. Define inductively $A_0 = A$ and $A_{i+1} = A_i \setminus X_i$. By measurability of X_i, $\lambda^*(A_i) = \lambda^*(A_i \cap X_i) + \lambda^*(A_{i+1})$. Hence

$$\lambda^*(A) = \sum_{i=1}^n \lambda^*(A_i \cap X_i) + \lambda^*(A_{n+1}).$$

However $A \setminus X \subseteq A_{n+1}$, so $\lambda^*(A) \geq \sum_{i=1}^n \lambda^*(A_i \cap X_i) + \lambda^*(A \setminus X)$. Since this is true for all n, we have

$$\lambda^*(A) \geq \sum_{i=1}^\infty \lambda^*(A_i \cap X_i) + \lambda^*(A \setminus X).$$

However, $\bigcup(A_i \cap X_i) = A \cap X$, so by subadditivity $\lambda^*(A \cap X) \leq \sum_{i=1}^\infty \lambda^*(A_i \cap X_i)$. Thus

$$\lambda^*(A) \geq \lambda^*(A \cap X) + \lambda^*(A \setminus X),$$

as required. (If there are only finitely many X_i, set the other $X_i = \emptyset$.) □

Lemma 5. *The interval (a, ∞) is measurable.*

Proof. Assume $A \subseteq \bigcup I_i$ and $\lambda^*(A) \geq \sum l(I_i) - \varepsilon$. Now $\lambda^*(A \cap (a, \infty)) \leq \sum l(I_i \cap (a, \infty))$, and $\lambda^*(A \setminus (a, \infty)) \leq \sum l(I_i \cap (-\infty, a)) + l((a - \varepsilon, a + \varepsilon))$. But $l(I_i) = l(I_i \cap (a, \infty)) + l(I_i \cap (-\infty, a))$. Thus $\lambda^*(A) \geq \lambda(A \cap (a, \infty)) + \lambda(A \setminus (a, \infty)) - 3\varepsilon$. Since this is true for all $\varepsilon > 0$, (a, ∞) is measurable. □

Theorem *All Borel sets are Lebesgue measurable*

Proof. Any σ-algebra that contains all (a, ∞) contains all Borel sets. □

Definition The *Lebesgue measure* $\lambda(S)$ of a Lebesgue measurable set S is $\lambda(S) = \lambda^*(S)$.

Lemma 6. *Lebesgue measure is a measure on the σ-algebra of Lebesgue measurable sets.*

Proof. Since Lebesgue measurable sets form a σ-algebra and $\lambda = \lambda^*$ is countably subadditive, it only remains to show that if $X_1, \ldots$ are *disjoint* measurable sets, then $\lambda(\bigcup_{i=1}^\infty X_i) \geq \sum_{i=1}^\infty \lambda(X_i)$. Let $A_n = \bigcup_{i=n}^\infty X_i$. Then $\lambda(A_n) = \lambda(A_n \cap X_n) + \lambda(A_n \setminus X_n) = \lambda(X_n) + \lambda(A_{n+1})$. Thus $\lambda(\bigcup_{i=1}^\infty X_i) = \lambda(A_1) = \sum_{i=1}^n \lambda(X_i) + \lambda(A_{n+1}) \geq \sum_{i=1}^n \lambda(X_i)$. Now let $n \to \infty$. □

It is now clear that $\lambda(S)$ satisfies Conditions 2 to 4 at the beginning of this section.

Math 7350 6. Measurability Fall 2004

Lemma 1. *If S is Lebesgue measurable then for all $\varepsilon > 0$, there exists an open set U and a closed set F with $F \subseteq S \subseteq U$ and $\lambda(U \setminus F) < \varepsilon$.*

Proof. Assume first that $S \subseteq [0,1]$. Then $\lambda(S) \geq \sum l(I_i) - \varepsilon \geq \lambda(U) - \varepsilon$ where $S \subseteq U = \bigcup I_i$, U open. Similarly $\lambda([0,1] \setminus S) \geq \sum l(I'_i) - \varepsilon \geq \lambda(U') - \varepsilon$. Setting $F = [0,1] \setminus U'$ we have F closed, $F \subseteq S$ and $\lambda(U \setminus F) < 2\varepsilon$. For general S, apply this to $S_n = S \cap [n, n+1]$ with ε replaced by $\varepsilon/2^{|n|}$ giving $\lambda(U_n \setminus F_n) < \varepsilon/2^{|n|}$. Set $U = \bigcup U_n$, $F = \bigcup F_n$. $\square$

Lemma 2. *If $X \subseteq \mathbb{R}$ and for all $\varepsilon > 0$ there exist measurable F and G with $F \subseteq X \subseteq G$ and $\lambda(G \setminus F) < \varepsilon$, then X is measurable.*

Proof. Subadditivity and monotonicity gives $\lambda^*(A \setminus X) \leq \lambda^*(A \setminus G) + \lambda^*(G \setminus F)$. Hence $\lambda^*(A) = \lambda^*(A \cap G) + \lambda(A \setminus G) \geq \lambda^*(A \cap X) + \lambda^*(A \setminus X) - \varepsilon$. Now let $\varepsilon \to 0$. $\square$

Note: $\lambda^*(X) = 0 \Rightarrow X$ is measurable, since $\emptyset \subseteq X \subseteq \bigcup I_i$ with $\lambda(\bigcup I_i \setminus \emptyset) \leq \sum l(I_i) < \varepsilon$.

Lemma 3. *If X is measurable, there exists a G_δ-set G and an F_σ-set F with $F \subseteq X \subseteq G$ and $\lambda(F \setminus G) = 0$.*

Proof. Pick $F_n \subseteq X \subseteq G_n$, F_n closed, G_n open, $\lambda(G_n \setminus F_n) < \frac{1}{n}$. Let $F = \bigcup F_n$, $G = \bigcap G_n$. $\square$

Lemma 4. *If $X_1 \subseteq X_2 \subseteq \dots$, with X_i measurable, then $\lambda(\bigcup_{i=1}^\infty X_i) = \lim_{n\to\infty} \lambda(X_n)$.*
If $X_1 \supseteq X_2 \supseteq \dots$, with X_i measurable and $\lambda(X_1) < \infty$, then $\lambda(\bigcap_{i=1}^\infty X_i) = \lim_{n\to\infty} \lambda(X_n)$.

Proof. Write $Y_1 = X_1$ and $Y_n = X_n \setminus X_{n-1}$ for $n > 1$, so Y_n are disjoint and $\lambda(\bigcup_{i=1}^\infty X_i) = \lambda(\bigcup_{i=1}^\infty Y_i) = \sum_{i=1}^\infty \lambda(Y_i) = \lim_n \sum_{i=1}^n \lambda(Y_i) = \lim_n \lambda(X_n)$. For the second part, apply the first part to $Y_n = X_1 \setminus X_n$. $\square$

Note: If $X_n = [n, \infty)$, then $\bigcap_{i=1}^\infty X_i = \emptyset$, but $\lim_n \lambda(X_n) = +\infty$.

Exercise: Show that the Cantor set C has measure 0.

Theorem 1. *There exists a non-measurable set.*

Proof. Define an equivalence relation on $[0,1]$ by letting $x \sim y$ iff $x - y \in \mathbb{Q}$. Construct a set X consisting of precisely one element in $[0,1]$ from each equivalence class of $\sim$ (using Axiom of Choice). Every $x \in [0,1]$ is of the form $x_0 + q$, $x_0 \in X$, $q \in \mathbb{Q} \cap [-1,1]$. Let $Y = \bigcup_{q \in \mathbb{Q} \cap [-1,1]} X + q$, where $X + q = \{x_0 + q : x_0 \in X\}$. Then $[0,1] \subseteq Y \subseteq [-1,2]$. In particular $1 \leq \lambda^*(Y) \leq 3$. Assume X were measurable and $\lambda(X) = c$. Then $X + q$ is measurable, $\lambda(X + q) = c$, and the $X + q$ are disjoint for distinct q's. Hence $1 \leq \lambda(Y) = \sum_q c \leq 3$. If $c = 0$ then $\sum_q c = 0$ and if $c > 0$ then $\sum_q c = +\infty$, a contradiction. $\square$

Definition Define the *Cantor function* $f\colon [0,1] \to [0,1]$ by writing x in base 3, $x = \sum_{n=1}^\infty a_n 3^{-n}$, $a_n \in \{0,1,2\}$, and setting $f(x) = \sum_{n=1}^{n_0} b_n 2^{-n}$ where n_0 is the first n for which $a_n = 1$, or ∞ if all $a_n \in \{0,2\}$, and $b_n = 0$ if $a_n = 0$ and $b_n = 1$ otherwise.

Exercise: Show that f is continuous, f maps the cantor set C *onto* $[0,1]$, and maps the set $[0,1]\setminus C$ to dyadic rationals (rationals of the form $\frac{a}{2^n}$, $a\in\mathbb{Z}$, $n\geq 0$).

Theorem 2. *There exists a Lebesgue measurable set that is not Borel.*

Proof. Take the non-measurable set X from Theorem 1 and remove the (unique) rational number from X. Then $A=f^{-1}[X]\subseteq C$, so $\lambda^*(A)\leq\lambda(C)=0$. Hence A is measurable. The inverse function $g(\sum b_n 2^{-n})=\sum 2b_n 3^{-n}$ is continuous except at dyadic rationals. Hence if A were Borel, $g^{-1}[A]=X$ would be Borel, and so measurable. □

Definition Suppose S is measurable. A function $f\colon S\to\mathbb{R}^*$ is *Lebesgue measurable* iff $f^{-1}[(a,\infty]]=\{x\in S: f(x)>a\}$ is measurable for all $a\in\mathbb{R}$.

Extend the Borel sets to $\mathbb{R}^*$ by setting $\mathcal{B}^*=\{B\cup I: B\in\mathcal{B},\ I\subseteq\{-\infty,+\infty\}\}$. Since the measurable sets form a σ-algebra, this implies that $f^{-1}[B]$ is measurable for all $B\in\mathcal{B}^*$. However, the inverse image of a Lebesgue measurable set may not be measurable (e.g., the function g above). To show a function is measurable, it is enough to show $f^{-1}[B]$ is measurable for any collection of Bs that generate $\mathcal{B}^*$, for example $\{[a,\infty]: a\in\mathbb{Q}\}$, or $\{[-\infty,a): a \text{ dyadic rational}\}$.

Definition The *characteristic function* χ_X of a set X is defined by $\chi_X(x)=1$ if $x\in X$ and $\chi_X(x)=0$ if $x\notin X$.

Lemma 4. *Suppose S is measurable and $f,g,f_n\colon S\to\mathbb{R}^*$.*

1. *Every continuous real valued function is measurable.*
2. *If f and g are measurable then so are $f\pm g$, fg, f/g, if defined.*
3. *If f_n are measurable, then so are* $\inf f_n$, $\sup f_n$, $\underline{\lim}\, f_n$, *and* $\overline{\lim}\, f_n$
4. *If X is measurable then χ_X is measurable.*

Proof. If $f\colon S\to\mathbb{R}$ is continuous then $f^{-1}[(a,\infty]]=S\cap U$ where U is open, and so $S\cap U$ is measurable. If f is measurable and $h(x)=-f(x)$ then $h^{-1}[(a,\infty]]=f^{-1}[[-\infty,a)]$ is measurable. If $h(x)=1/f(x)$ (defined to be $+\infty$, say, if $f(x)=0$), then $h^{-1}[(a,\infty]]=f^{-1}[[0,1/a)]$ is measurable if $a\geq 0$, and $h^{-1}[(a,\infty]]=f^{-1}[[1/a,0)^c]$ is measurable if $a<0$. If $h(x)=f(x)^2/2$ then $h^{-1}[(a,\infty]]=f^{-1}[[-\infty,-\sqrt{2a})\cup(\sqrt{2a},\infty]]$ (or $f^{-1}[\mathbb{R}^*]$ if $a<0$) is measurable. Now if f and g are measurable and $h(x)=f(x)+g(x)$ is never of the form $\infty-\infty$, then $h(x)>a$ iff $\exists q\in\mathbb{Q}\colon f(x)>q$ and $g(x)>a-q$. Hence

$$h^{-1}[(a,\infty]]=\textstyle\bigcup_{q\in\mathbb{Q}}(f^{-1}[(q,\infty)]\cap g^{-1}[(a-q,\infty]])$$

is measurable. Now $f-g=f+(-g)$, $fg=(f+g)^2/2-f^2/2-g^2/2$ and $f/g=(f)(1/g)$ are measurable. If $f=\sup_n f_n$, then $f(x)>a$ iff $\exists n\colon f_n(x)>a$. Hence $f^{-1}[(a,\infty]]=\bigcup_n f_n^{-1}[(a,\infty]]$ is measurable. Now $\inf f_n=-\sup(-f_n)$, $\underline{\lim}\, f_n=\sup_{n_0}\inf_{n\geq n_0} f_n$, and $\overline{\lim}\, f_n=\inf_{n_0}\sup_{n\geq n_0} f_n$ are measurable. Finally $\chi_X^{-1}[(a,\infty]]=S$, $S\cap X$ or $\emptyset$, all of which are measurable. □

Math 7350 7. Lebesgue Integration Fall 2004

Definition A *simple function* is a real-valued measurable function $f\colon \mathbb{R} \to \mathbb{R}$ that takes only finitely many values, i.e., $f[\mathbb{R}]$ is finite.

Definition If $f, g\colon \mathbb{R} \to \mathbb{R}^*$ are two functions, write $f \geq g$ if $f(x) \geq g(x)$ for all $x \in \mathbb{R}$. Note that $\geq$ is a partial order on functions. In particular, we say f is non-negative, $f \geq 0$, if $f(x) \geq 0$ for all $x \in \mathbb{R}$.

Lemma 1. *Any simple function can be written in the form $f = \sum_{i=1}^{n} a_i \chi_{S_i}$ where the S_i are measurable and form a partition of $\mathbb{R}$ (so in particular are disjoint).*

Proof. Let $f[\mathbb{R}] = \{a_1, \dots, a_n\}$ and set $S_i = f^{-1}[\{a_i\}]$. □

Definition If $\phi = \sum_{i=1}^{n} a_i \chi_{S_i}$ is a non-negative simple function with S_i disjoint, define $\int \phi = \int_{\mathbb{R}} \phi(x)\, dx = \sum_{i=1}^{n} a_i \lambda(S_i)$.

Note: We could drop the non-negative condition, but we would have to leave $\int \phi$ undefined if the sum was of the form $\infty - \infty$.

Lemma 2. *Assume ϕ and ψ are non-negative simple functions.*

(a) $\int \phi$ *is well defined and lies in* $[0, \infty]$

(b) If $\phi \geq \psi$ *then* $\int \phi \geq \int \psi$.

(c) If $c \geq 0$ *then* $\int c\phi = c \int \phi$.

(d) $\int (\phi + \psi) = \int \phi + \int \psi$

Proof. (b) Let $\phi = \sum_{i=1}^{n} a_i \chi_{S_i}$ and $\psi = \sum_{j=1}^{m} b_j \chi_{T_j}$. By including extra terms $a_0 = b_0 = 0$, $S_0 = (\bigcup S_i)^c$, $T_0 = (\bigcup T_i)^c$, we may assume $\bigcup S_i = \bigcup T_i = \mathbb{R}$. Let $E_{ij} = S_i \cap T_j$. Then the E_{ij} are disjoint and so $\sum_i \lambda(E_{ij}) = \lambda(\bigcup_i E_{ij}) = \lambda(T_j)$ and $\sum_j \lambda(E_{ij}) = \lambda(\bigcup_j E_{ij}) = \lambda(S_i)$. Thus $\sum_i a_i \lambda(S_i) = \sum_{ij} a_i \lambda(E_{ij})$ and $\sum_j b_j \lambda(T_j) = \sum_{ij} b_j \lambda(E_{ij})$. But if $\phi \geq \psi$ then either $a_i \geq b_j$ or $E_{ij} = \emptyset$. Hence $\int \phi \geq \int \psi$.
(a) Applying (b) to $\psi = \phi$ we see that $\int \phi$ is well defined. Since $\phi \geq 0$, $\int \phi \geq \int 0 = 0$.
(c) Clear.
(d) Defining E_{ij} as above, $\phi = \sum_{ij} a_i \chi_{E_{ij}}$, $\psi = \sum_{ij} b_j \chi_{E_{ij}}$ and $\phi + \psi = \sum_{ij} (a_i + b_j) \chi_{E_{ij}}$, and $\sum_{ij} (a_i + b_j) \lambda(E_{ij}) = \sum_{ij} a_i \lambda(E_{ij}) + \sum_{ij} b_j \lambda(E_{ij})$. The result follows. □

Note that $\int \sum a_i \chi_{S_i} = \sum a_i \lambda(S_i)$ holds for any $a_i \geq 0$, and measurable S_i. In particular, we did not need the S_i to be disjoint in the definition of $\int \phi$.

Definition Suppose $f\colon \mathbb{R} \to \mathbb{R}^*$ is measurable and $f \geq 0$. Define $\int f = \int_{\mathbb{R}} f(x)\, dx = \sup_{\phi \leq f} \int \phi$ where the supremum is over simple non-negative functions ϕ with $\phi \leq f$.

Example Suppose $f(x) = 0$ if x is irrational and $f(p/q) = q$ if p/q is a rational number in lowest terms. Show that $\int f = 0$.

Lemma 3. *Assume f and g are non-negative measurable functions.*

(a) If $f \geq g$ then $\int f \geq \int g$.

(b) If $c \geq 0$ then $\int cf = c \int f$.

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

Proof. (a) and (b) are clear. For (c), if $\phi \leq f$ and $\psi \leq g$ then $\phi + \psi$ is a simple function and $\phi + \psi \leq f + g$. Thus $\int (f+g) \geq \int (\phi+\psi) = \int \phi + \int \psi$. Taking supremums over ϕ and ψ gives $\int (f+g) \geq \int f + \int g$. Now suppose $\upsilon = \sum a_i \chi_{S_i} \leq f+g$. Let N be large and set $E_j = \{x : \frac{j}{N} \leq \frac{f(x)}{f(x)+g(x)} < \frac{j+1}{N}\}$. (If $g(x) = 0$ or $f(x) = \infty$, put x in E_{N-1}). Then $E_0, \ldots, E_{N-1}$ is a measurable partition of $\mathbb{R}$. Define $\phi = \sum_{ij} \frac{j}{N} a_i \chi_{S_i \cap E_j}$, and $\psi = \sum_{ij} \frac{N-1-j}{N} a_i \chi_{S_i \cap E_j}$. Then $\phi \leq f$, $\psi \leq g$, and $\phi + \psi = \frac{N-1}{N} \upsilon$. Thus $\int \upsilon \leq \frac{N}{N-1}(\int \phi + \int \psi) \leq \frac{N}{N-1}(\int f + \int g)$. Letting $N \to \infty$ gives $\int \upsilon \leq \int f + \int g$. Hence $\int (f+g) \leq \int f + \int g$. □

Definition Suppose f is a measurable function. Set $f_+(x) = \max\{f(x), 0\}$ and $f_-(x) = \max\{-f(x), 0\}$ so that $f = f_+ - f_-$, $|f| = f_+ + f_-$, and $f_+, f_- \geq 0$. Define $\int f = \int f_+ - \int f_-$ provided this is not of the form $\infty - \infty$.

Definition A function f is *Lebesgue integrable*, if f is measurable and $\int f_+, \int f_- < \infty$ (equivalently $\int |f| < \infty$).

Lemma 4. *Assume f and g are measurable functions.*

(a) If $f \geq g$ then $\int f \geq \int g$ (if both defined).

(b) If $c \in \mathbb{R}$ then $\int cf = c \int f$ (if RHS defined).

(c) $\int (f+g) = \int f + \int g$ (if RHS defined).

Proof. For (a) note that $f \geq g$ implies $f_+ \geq g_+$ and $f_- \leq g_-$. For (b), write $f = f_+ - f_-$ and treat the cases $c = 0$, $c > 0$, and $c < 0$ separately. For (c), let $h = f + g$ and note that $h_+ + f_- + g_- = h_- + f_+ + g_+$. Use Lemma 3(c). □

Definition If S and $f \colon S \to \mathbb{R}^*$ are measurable, define $\int_S f = \int_{\mathbb{R}} \chi_S(x) f(x)\, dx$.

Exercises

1. Show that $\frac{1}{\sqrt{x}}$ is Lebesgue integrable on $S = (0,1]$ (i.e., $\chi_{(0,1]}(x)\frac{1}{\sqrt{x}}$ is integrable).
2. Show that $\frac{\sin(x)}{x}$ is not Lebesgue integrable (on $S = \mathbb{R}$).
3. Show that if $f = g$ a.e., then $\int f = \int g$.

Math 7350 8. Convergence Theorems Fall 2004

Lemma 1. *If f is measurable then $\int |f| = 0$ iff $f = 0$ a.e.*

Proof. Let $E_k = \{x : |f(x)| \geq 1/k\}$. Then $\int |f| \geq \int (1/k)\chi_{E_k} = \lambda(E_k)/k$. Hence if $\int |f| = 0$ then $\lambda(E_k) = 0$, so $\lambda(\bigcup_k E_k) = 0$. But $\bigcup_k E_k = \{x : f(x) \neq 0\}$. Conversely, if $f = 0$ a.e., and $\phi = \sum a_i \chi_{S_i} \leq |f|$ with $a_i \geq 0$ and S_i disjoint, then for all i, either $a_i = 0$ or $\lambda(S_i) = 0$. This implies $\int \phi = 0$ so $\int |f| = \sup_{\phi \leq |f|} \int \phi = 0$. □

Theorem (Monotone Convergence Theorem, MCT) *If $0 \leq f_1 \leq f_2 \leq \dots$ is an increasing sequence of non-negative measurable $\mathbb{R}^*$-valued functions and $f(x) = \lim_{n\to\infty} f_n(x)$, then $\int f = \lim_{n\to\infty} \int f_n$.*

Note: The limits and/or the $\int f_n$ may be $+\infty$.

Proof. Suppose $\phi = \sum_{i=1}^m a_i \chi_{S_i}$ is a simple function with $0 \leq \phi \leq f$. Fix $\varepsilon > 0$ and let $E_n = \{x : f_n(x) > (1-\varepsilon)\phi(x)\}$. Since $f_1 \leq f_2 \leq \dots$, clearly $E_1 \subseteq E_2 \subseteq \dots$. Also, for any x, $f_n(x) \to f(x) \geq \phi(x)$. But $\phi(x)$ is finite, so for some n, $f_n(x) > (1-\varepsilon)\phi(x)$. Thus $\bigcup_n E_n = \mathbb{R}$. Let $\phi_n = (1-\varepsilon)\phi\chi_{E_n} = (1-\varepsilon)\sum_{i=1}^m a_i \chi_{S_i \cap E_n}$. Now $\phi_n \leq f_n$ and $\lim_n \lambda(S_i \cap E_n) = \lambda(S_i)$. Hence $\lim_n \int f_n \geq \lim_n \int \phi_n = \lim_n (1-\varepsilon)\sum_{i=1}^m a_i \lambda(S_i \cap E_n) = (1-\varepsilon)\sum_{i=1}^m a_i \lim_n \lambda(S_i \cap E_n) = (1-\varepsilon)\sum_{i=1}^m a_i \lambda(S_i) = (1-\varepsilon)\int \phi$. Since this holds for all ε and ϕ, $\lim_n \int f_n \geq \int f$. However, $f_n \leq f$, so $\lim_n \int f_n \leq \int f$ and so $\int f = \lim_n \int f_n$. □

Theorem (Fatou's Lemma) *If $f_n \geq 0$ are non-negative measurable $\mathbb{R}^*$-valued functions, then $\int \underline{\lim}_{n\to\infty} f_n \leq \underline{\lim}_{n\to\infty} \int f_n$.*

Note: We can have $<$, e.g., $f_n = \chi_{[0,1]}$ if n even and $f_n = \chi_{[1,2]}$ if n odd.

Proof. Let $g_{n_0}(x) = \inf_{n \geq n_0} f_n(x)$. Then g_{n_0} is an increasing sequence of non-negative measurable $\mathbb{R}^*$-valued functions. Also $\int g_{n_0} \leq \int f_n$ for all $n \geq n_0$, so $\int g_{n_0} \leq \inf_{n \geq n_0} \int f_n$. MCT implies $\int \underline{\lim}\, f_n = \int \lim_{n_0} g_{n_0} = \lim_{n_0} \int g_{n_0} \leq \lim_{n_0} \inf_{n \geq n_0} \int f_n = \underline{\lim} \int f_n$. □

Theorem (Lebesgue Dominated Convergence Theorem, DCT) *If f_n are measurable functions which converge a.e. to $f(x)$, and $|f_n(x)| \leq g(x)$ where $\int g < \infty$, then $\int f = \lim \int f_n$.*

Note: Dominating f_n by g is important, e.g., theorem fails for $f_n = \chi_{[n,n+1]}$.

Proof. Applying Fatou to $g + f_n \geq 0$ gives $\underline{\lim} \int (g + f_n) \geq \int \underline{\lim}(g + f_n) = \int (g + f)$. Hence $\int f \leq \underline{\lim} \int f_n$. Applying Fatou to $g - f_n \geq 0$ gives $\underline{\lim} \int (g - f_n) \geq \int \underline{\lim}(g - f_n) = \int (g - f)$. Hence $\overline{\lim} \int f_n \leq \int f$. Thus $\int f \leq \underline{\lim} \int f_n \leq \overline{\lim} \int f_n \leq \int f$, so $\lim \int f_n = \int f$. □

Definition Assume f_n and f are measurable functions. Recall that $f_n \to f$ a.e. if $\lambda(\{x : f_n(x) \not\to f(x)\}) = 0$. We say $f_n \to f$ *in mean* if $\int |f_n - f| \to 0$ as $n \to \infty$. We say $f_n \to f$ *in measure* if $\forall \varepsilon > 0\colon \exists n_0\colon \forall n \geq n_0\colon \lambda(\{x : |f_n(x) - f(x)| > \varepsilon\}) < \varepsilon$.

Examples

1. Convergence in mean always implies convergence in measure. (If $\lambda(\{x : |f_n - f| > \varepsilon\}) \geq \varepsilon$ then $\int |f_n - f| \geq \varepsilon^2$.)
2. If $f_n = nx^n$ on $[0,1]$ then $f_n \to 0$ in measure and a.e., but not in mean.
3. If $f_n = \chi_{[n,n+1]}$ then $f_n \to 0$ a.e., but not in mean nor in measure.
4. If $f_n = \chi_{[a/2^k,(a+1)/2^k]}$ where $2^k \leq n = 2^k + a < 2^{k+1}$, then $f_n \to 0$ in mean and in measure but not a.e.
5. If $f_n = 2^k \chi_{[a/2^k,(a+1)/2^k]}$ where $2^k \leq n = 2^k + a < 2^{k+1}$, then $f_n \to 0$ in measure but not in mean nor a.e.

Lemma 2. *If $f_n \to f$ in measure, then there exists a subsequence such that $f_{n_k} \to f$ a.e.*

Proof. Choose an increasing sequence n_k so that $\lambda(\{x : |f_{n_k}(x) - f(x)| > 2^{-k}\}) < 2^{-k}$. Let $E_{jk} = \{x : |f_{n_k}(x) - f(x)| \geq 2^{-j}\}$. Then when $k \geq j$, $\lambda(E_{jk}) \leq 2^{-k}$, so for $k_0 \geq j$, $\lambda(\bigcup_{k \geq k_0} E_{jk}) \leq 2^{1-k_0}$, and thus $\lambda(\bigcap_{k_0} \bigcup_{k \geq k_0} E_{jk}) = 0$. But then $E = \bigcup_j \bigcap_{k_0} \bigcup_{k \geq k_0} E_{jk}$ has measure zero. But this is just the set of points x where $f_{n_k}(x) \not\to f(x)$. $\square$

Corollary *DCT also holds if we assume $f_n \to f$ in measure.*

Lemma 3. *Assume $|f_n| \leq g$, $\int \min\{g, 1\} < \infty$. If $f_n \to f$ a.e., then $f_n \to f$ in measure.*

Note: Hypothesis is satisfied (by $g = \infty.\chi_S$) if f_n is zero outside S and $\lambda(S) < \infty$.

Proof. Fix $0 < \varepsilon < 1$. Let $E_n = \{x : |f_n(x) - f(x)| > \varepsilon\}$. Now $\bigcup_{n \geq n_0} E_n$ is decreasing in n_0 and since $f_n \to f$ a.e., $\lambda(\bigcap_{n_0} \bigcup_{n \geq n_0} E_n) = 0$. But if $x \in E_n$ then $g(x) > \varepsilon/2$. Hence $\int \min\{g, 1\} \geq (\varepsilon/2)\lambda(\bigcup E_n)$, so $\lambda(\bigcup_{n \geq 1} E_n) < \infty$. Thus $\lambda(\bigcup_{n \geq n_0} E_n) \to 0$. Hence there is an n_0 with $\lambda(\bigcup_{n \geq n_0} E_n) < \varepsilon$. So $\forall n \geq n_0$: $\lambda(\{x : |f_n(x) - f(x)| > \varepsilon\}) = \lambda(E_n) < \varepsilon$. $\square$

Summary: 'a.e.'⇒'in measure' if the f_n are bounded horizontally, 'in measure'⇒'in mean' if the f_n are bounded in the plane, and 'in mean'⇒'in measure'⇒'a.e. on a subsequence'.

Exercises

1. Show that if $f_n \geq 0$ are measurable, then $\sum_{n=1}^{\infty} \int f_n = \int \sum_{n=1}^{\infty} f_n$.
2. Show that if $f \geq 0$ is measurable then $\int_{-\infty}^{\infty} f(x)\, dx = \lim_{n \to \infty} \int_{-n}^{n} f(x)\, dx$.
3. Show that if $f \geq 0$ is integrable then $F(z) = \int_{-\infty}^{z} f(x)\, dx$ is continuous.
4. Suppose $f_n \geq 0$, $f_n \to f$ a.e., and $\int f_n \to \int f$. If S is a measurable set, show that $\int_S f_n \to \int_S f$.
5. Suppose that f_n and f are integrable. Show that if $f_n \to f$ in mean then $\int f_n \to \int f$. Is it true that if $\int f_n \to \int f$ and $f_n \to f$ a.e., then $f_n \to f$ in mean?

Math 7350 9. Riemann Integration Fall 2004

Definition A *step function* on $[a,b]$ is a simple function of the form $\phi = \sum_{i=1}^{n} a_i \chi_{S_i}$ where the S_i are intervals. Equivalently, there exist a finite partition $a = a_0 < a_1 < \dots < a_n = b$ with ϕ constant on each of the intervals (a_i, a_{i+1}).

Definition Suppose $f\colon [a,b] \to [-K,K]$ is a bounded function on the finite interval $[a,b]$. The *lower Riemann integral* of f is ${}_R\underline{\int} f = \sup_{\phi\le f} \int \phi$ where the supremum is over all step functions ϕ with $\phi \le f$. The *upper Riemann integral* of f is ${}_R\overline{\int} f = \inf_{\phi\ge f} \int \phi$ where the infimum is over all step functions ϕ with $\phi \ge f$. If ${}_R\underline{\int} f = {}_R\overline{\int} f$ then we call f *(properly) Riemann integrable* and define the (proper) Riemann integral ${}_R\int f = {}_R\underline{\int} f = {}_R\overline{\int} f$.

Example $\chi_{\mathbb{Q}}$ is not Riemann integrable on $[0,1]$ since if $\phi \ge \chi_{\mathbb{Q}}$ and ϕ is a step function then $\phi \ge 1$ except at a finite number of points. Thus ${}_R\overline{\int}_0^1 \chi_{\mathbb{Q}} = 1$. Similarly ${}_R\underline{\int}_0^1 \chi_{\mathbb{Q}} = 0$.

Strictly speaking, the above definition is a minor variant of the Darboux integral, which is normally defined as the common limit of $L_{(a_i)} = \sum m_i(a_{i+1}-a_i)$ and $U_{(a_i)} = \sum M_i(a_{i+1}-a_i)$ where $M_i = \sup_{x\in[a_i,a_{i+1}]} f(x)$ and $m_i = \inf_{x\in[a_i,a_{i+1}]} f(x)$. The limit is taken over finer and finer partitions of $[a,b]$ (i.e., over any sequence of partitions where $\max|a_{i+1} - a_i| \to 0$). We call f integrable if both limits exist and are equal. In fact, both limits always exist and are independent of the partitions provided $\max|a_{i+1} - a_i| \to 0$. In particular, one can assume the partitions are regular, $a_i = a + (b-a)\frac{i}{n}$. To see this, note that if $a = a'_0 < \dots < a'_{n'} = b$ is another partition with $\max|a'_{i+1} - a'_i| < \varepsilon \min|a_{i+1} - a_i|$, then $L_{(a_i)} - 2\varepsilon(b-a)K \le L_{(a'_i)} \le U_{(a'_i)} \le U_{(a_i)} + 2\varepsilon(b-a)K$. Hence by choosing a $L_{(a_i)}$ close to $\sup_{(a_i)} L_{(a_i)} < \infty$ we see that *every* sufficiently fine partition gives a value of $L_{(a'_i)}$ close to $\sup_{(a_i)} L_{(a_i)} = \sup_{\phi\le f} \int \phi$. Similarly for $U_{(a'_i)}$.

The Riemann integral is defined as the limit of $\sum f(x_i)(a_{i+1} - a_i)$ where $x_i \in [a_i, a_{i+1}]$. The limit is taken over finer and finer partitions ($\max|a_{i+1} - a_i| \to 0$)), and f is integrable if the limit exists and is independent of both the sequence of partitions and the choices of the x_i for each partition. By choosing x_i so that $f(x_i)$ is close to $m_i = \inf_{x\in[a_i,a_{i+1}]} f(x)$ or $M_i = \sup_{x\in[a_i,a_{i+1}]} f(x)$, it is clear that this is equivalent to the Darboux integral. In particular, we may fix the sequence of partitions. It is also possible to fix x_i, e.g., as $x_i = a_i$ (left Riemann sum) or $x_i = a_{i+1}$ (right Riemann sum), but we *cannot* then fix the sequence of partitions, e.g., $2 = \int_0^2 \chi_{\mathbb{Q}} \ne \int_0^{\sqrt{2}} \chi_{\mathbb{Q}} + \int_{\sqrt{2}}^2 \chi_{\mathbb{Q}} = 0$ if we use $x_i = a_i$ and a regular partition.

The Riemann, Darboux, and the variant of the Darboux integral defined above are all equivalent, however the version I am using is easier to work with.

Note that the MCT and DCT do not hold for Riemann integrals: Enumerate the rationals in $[0,1]$ as $q_1, q_2, \dots$, then ${}_R\int \chi_{\{q_1,\dots,q_n\}} = 0$ but ${}_R\int \chi_{\mathbb{Q}\cap[0,1]}$ is undefined.

Theorem *Assume $f\colon [a,b] \to [-K,K]$ is a bounded function. Then f is properly Riemann integrable iff f is measurable function which is continuous a.e.. In this case $_R\!\int_a^b f = \int f$.*

Proof. $\Rightarrow$: If f is not continuous at $x \in (a,b)$ then $x \in E_k$ for some k, where $E_k = \{x : \forall$ open intervals I with $x \in I \subseteq [a,b]\colon \sup_I f - \inf_I f > 1/k\}$. Let $\varepsilon > 0$ and pick step functions $\phi \le f \le \psi$ with $\int \psi - \int \phi < \varepsilon/k$. If $x \in E_k$ is not one of the finite set of points at which either ϕ or ψ jumps, then $\psi(x) - \phi(x) > 1/k$. Thus $\lambda(E_k)/k \le \int(\psi - \phi) < \varepsilon/k$. Hence $\lambda(E_k) < \varepsilon$ for all $\varepsilon > 0$, and so $\lambda(E_k) = 0$. Thus $\lambda(\bigcup E_k) = \lim \lambda(E_k) = 0$ so f is continuous a.e.. Since continuous functions are measurable, f is measurable on a set of the form $[a,b] \setminus E$, where $\lambda(E) = 0$. Thus f is measurable on $[a,b]$.
$\Leftarrow$: Let E be the set of discontinuities of f. Then $\lambda(E) = 0$ and so $E \subseteq U$ where U is open and $\lambda(U) < \varepsilon$. Set $F = [a,b] \setminus U$. Then F is closed and bounded and f is continuous on F. Thus $\exists \delta > 0\colon \forall x \in F, y \in [a,b]\colon |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. (Note: this is slightly stronger than uniform continuity since we do not require $y \in F$. However the proof of this result is the same as the proof that continuous functions on closed bounded sets are uniformly continuous.) Choose a partition $a = a_0 < a_1 < \dots < a_n = b$ with $|a_{i+1} - a_i| < \delta$. If $[a_i, a_{i+1}) \cap F = \emptyset$, set $\phi = -K$ and $\psi = +K$ on $[a_i, a_{i+1})$. Otherwise, pick $z \in [a_i, a_{i+1}) \cap F$ and set $\phi = f(z) - \varepsilon$ and $\psi = f(z) + \varepsilon$ on $[a_i, a_{i+1})$. Then $\phi \le f \le \psi$ and $\int(\psi - \phi) \le 2K\varepsilon + 2\varepsilon(b - a)$. This can be made arbitrarily small by suitable choice of ε. Thus f is Riemann integrable. Finally $\int \phi \le \int f, {}_R\!\int f \le \int \psi$, so $\int f = {}_R\!\int f$. ∎

One can extend the Riemann integral to more general functions by introducing the *improper* Riemann integral. For example, if f is bounded on each interval $[a+\varepsilon, b]$ but not on $[a,b]$, we define $_R\!\int_a^b f = \lim_{\varepsilon \to 0^+} {}_R\!\int_{a+\varepsilon}^b f$ if this limit exists.

Example The function $f(x) = \frac{1}{x} \cos \frac{1}{x}$ is improperly Riemann integrable on $[0,1]$. since $\int_\varepsilon^1 f = [-x \sin \frac{1}{x}]_\varepsilon^1 + \int_\varepsilon^1 \sin \frac{1}{x}\, dx$ which tends to a limit as $\varepsilon \to 0$. However, f is not Lebesgue integrable since $\int_0^1 |f| = \infty$.

It is true however that if the (improper) Riemann integral and Lebesgue integral both exist, then they are equal.

One can modify the definition of the proper Riemann integral in such a way that it includes both improperly Riemann integrable functions and also Lebesgue integrable functions. If $f\colon [a,b] \to \mathbb{R}$ is any function (not bounded in general), we define the *Gauge Integral* (a.k.a. Generalized Riemann Integral, a.k.a. Henstock-Kurzweil integral, a.k.a. Denjoy-Perron integral) of f to be L if for all $\varepsilon > 0$, there is a positive function $\delta\colon [a,b] \to (0,\infty)$ such that whenever we have a partition $a = a_0 < a_1 < \dots < a_n = b$, and points $x_i \in [a_i, a_{i+1}]$ with $|a_{i+1} - a_i| < \delta(x_i)$, then $|\sum f(x_i)(a_{i+1} - a_i) - L| < \varepsilon$. Note that this differs from the Riemann integral only in the assumption that δ may depend on the x_i. The Gauge integral does satisfy some nice properties, (a version of DCT holds), but not as many as the Lebesgue integral. Also, both the Gauge and Riemann integrals to not generalize well to integration over more general spaces, whereas the Lebesgue integral is defined on any space with a measure.

Math 7350 10. Absolute Continuity Fall 2004

Definition A function $f\colon [a,b] \to \mathbb{R}$ is *of bounded variation* if there exists a K such that for all $a = a_0 < a_1 < \cdots < a_n = b$, $\sum_{i=1}^n |f(a_i) - f(a_{i-1})| < K$.

Lemma 1. *A function $f\colon [a,b] \to \mathbb{R}$ is of bounded variation iff $f = g - h$ for some increasing functions $g, h\colon [a,b] \to \mathbb{R}$.*

Proof. Define $h(x) = \sup \sum_{i=1}^n |f(a_i) - f(a_{i-1})|$ where the supremum is over all n and all partitions $a = a_0 < a_1 < \cdots < a_n = x$ of $[a,x]$. If $y > x$ then $h(y) \geq h(x) + |f(y) - f(x)|$, since any partition $a = a_0 < a_1 < \cdots < a_n = x$, gives rise to a partition $a = a_0 < a_1 < \cdots < a_n = x < a_{n+1} = y$. If f is of bounded variation, then $h(x) \leq h(b) < \infty$ for all $x \in [a,b]$, and both h and $g = f + h$ are increasing. Conversely, if $f = g - h$ and g and h are increasing and $y > x$ then $|f(y) - f(x)| \leq g(y) - g(x) - h(y) + h(x)$, so $\sum_{i=1}^n |f(a_i) - f(a_{i-1})| \leq K = g(b) - g(a) - h(b) + h(a) < \infty$. □

Definition A function $F\colon [a,b] \to \mathbb{R}$ is *absolutely continuous* if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever (a_i, b_i), $i = 1, \ldots, n$, are disjoint intervals with $\sum_{i=1}^n (b_i - a_i) < \delta$ then $\sum_{i=1}^n |F(b_i) - F(a_i)| < \varepsilon$.

Lemma 2. *If $f\colon [a,b] \to \mathbb{R}$ is absolutely continuous then f is continuous and of bounded variation.*

Proof. Pick δ corresponding to $\varepsilon = 1$ in the definition of absolute continuity. Fix $K > (b-a)/\delta$. For any $a = a_0 < \cdots < a_n = b$ we may add division points $a_{n_i} = a + i(b-a)/K$. Now $\sum_{j=n_i+1}^{n_{i+1}} (a_j - a_{j-1}) < \delta$, so $\sum_{j=n_i+1}^{n_{i+1}} |f(a_j) - f(a_{j-1})| < 1$ and $\sum_{i=1}^n |f(a_i) - f(a_{i-1})| < K$. Hence f is bounded variation. (Uniform) continuity follows by taking $n = 1$ in the definition of absolute continuity. □

Example The Cantor function is continuous and of bounded variation, but not absolutely continuous. (Take the (a_i, b_i) to be the 2^n intervals of length 3^{-n} defining the nth stage of the construction of the Cantor set. Then $\sum |F(b_i) - F(a_i)| = 1$ but $\sum (b_i - a_i) = (2/3)^n$).

Exercise: Assume f and g are absolutely continuous and h is absolutely continuous and monotonic. Show that $f(h(x))$, $f + g$, $f - g$, fg, f/g (if $g \neq 0$), $|f|$, $f_\pm$, and $f(x)^\alpha$ ($\alpha \in \mathbb{R}$, $f > 0$) are all absolutely continuous.

Lemma 3. *If f is integrable, then $\forall \varepsilon > 0\colon \exists \delta > 0\colon\ \lambda(S) < \delta \ \Rightarrow\ \int_S |f| < \varepsilon$.*

Proof. Pick simple $\phi = \sum_{i=1}^n a_i \chi_{S_i} \leq |f|$ with $\int \phi \geq \int |f| - \varepsilon/2$. Let $\delta = \varepsilon/(2 \max a_i)$. □

Corollary *If $f\colon [a,b] \to \mathbb{R}^*$ is integrable, then $F(x) = \int_a^x f(t)\,dt$ is absolutely continuous.*

Proof. $\sum_{i=1}^n |F(b_i) - F(a_i)| = \sum |\int_{a_i}^{b_i} f| \leq \sum \int_{a_i}^{b_i} |f| = \int_S |f|$ where $S = \bigcup (a_i, b_i)$. Now apply Lemma 3. □

Math 7350 11. Derivatives Fall 2004

Lemma (Vitali) *Suppose S is set of finite measure and $\mathcal{I}$ is a collection of intervals with the property that for all $x \in S$, and all $\varepsilon > 0$, there exists a non-trivial interval $I \in \mathcal{I}$ with $x \in I$ and $\lambda(I) < \varepsilon$. Then for all $\varepsilon > 0$, there exists disjoint $I_i \in \mathcal{I}$ with $\lambda(S \setminus \bigcup_{i=1}^n I_i) < \varepsilon$.*

Proof. We can assume w.l.o.g., that the intervals are closed. Pick an open set $U \supseteq S$ of finite measure. Inductively define $I_n \in \mathcal{I}$ with $I_n \subseteq U$ and I_n disjoint as follows. If $S \subseteq \bigcup_{i=1}^n I_n$, we are done. If $S \not\subseteq \bigcup_{i=1}^n I_n$, then if $x \in S \setminus \bigcup_{i=1}^n I_i$ there exists an $I \in \mathcal{I}$ with $I \subseteq U \setminus \bigcup_{i=1}^n I_i$. Pick I_{n+1} so that $\lambda(I_{n+1}) > \frac{1}{2} \sup_I \lambda(I)$ where the supremum is over $I \in \mathcal{I}$ with $I \subseteq U \setminus \bigcup_{i=1}^n I_i$. Now $\sum_{i=1}^\infty \lambda(I_i) \leq \lambda(U) < \infty$, so we can pick N such that $\sum_{i=N+1}^\infty \lambda(I_i) < \varepsilon/5$. Suppose $x \in S \setminus \bigcup_{i=1}^N I_i$. Pick $I \in \mathcal{I}$ with $x \in I \subseteq U \setminus \bigcup_{i=1}^N I_i$. Since $\sum \lambda(I_i) < \infty$, there must be an n with $\lambda(I_n) < \lambda(I)/2$. But then either I would have been chosen as a I_m for some m, or it intersects some I_m with $N < m \leq n$ and $\lambda(I_m) \geq \lambda(I)/2$. In either case $I \subseteq J_m$ where J_m is the interval of length $5\lambda(I_m)$ centered on I_m. Hence $\lambda(S \setminus \bigcup_{i=1}^N) \leq \lambda(\bigcup_{m>N} J_m) \leq 5 \sum \lambda(I_m) = \varepsilon$. □

Lemma 1. *If $F \colon [a, b] \to \mathbb{R}$ is increasing then F' exists a.e., and $\int_a^b F' \leq F(b) - F(a)$.*

Proof. Let $E_{u,v} = \{x : \underline{\lim}_{h\to 0} \frac{F(x+h)-F(x)}{h} < u < v < \overline{\lim}_{h\to 0} \frac{F(x+h)-F(x)}{h}\}$. If $F'(x)$ does not exist then for some rationals u and v, $x \in E_{u,v}$. Thus it is enough to show that $\lambda(E_{u,v}) = 0$. Assume $\lambda(E_{u,v}) = s > 0$ Pick an open $U \supseteq E_{u,v}$ with $\lambda(U) < s + \varepsilon$. For each $x \in E_{u,v}$ there are arbitrarily small intervals $[a, b] = [x, x + h]$ or $[x - h, x]$ with $x \in [a, b] \subseteq U$ and $F(b) - F(a) \leq u(b - a)$. Let $V = \bigcup_{i=1}^N (a_i, b_i)$ where the $I_i = [a_i, b_i]$ are a disjoint family of such intervals with $\lambda(E_{u,v} \setminus V) < \varepsilon$, (so $s - \varepsilon \leq \lambda(V) \leq s + \varepsilon$). Now for each $x \in E_{u,v} \cap V$, there are arbitrarily small intervals $[c, d] = [x, x + h]$ or $[x - h, x]$ with $x \in [c, d] \subseteq V$ and $F(d) - F(c) \geq v(d - c)$. Let $W = \bigcup_{j=1}^M (c_j, d_j)$ where the $J_j = [c_j, d_j]$ are a disjoint family of such intervals with $\lambda((E_{u,v} \cap V) \setminus W) < \varepsilon$ (so $\lambda(W) \geq s - 2\varepsilon$). Now since the J_j are disjoint and each is contained in some I_i, $\sum_{j=1}^M F(d_j) - F(c_j) \leq \sum_{i=1}^N F(b_i) - F(a_i)$. but then $v\lambda(W) \leq u\lambda(V)$, so $v(s - 2\varepsilon) \leq u(s + \varepsilon)$. Since this holds for all $\varepsilon > 0$, $vs \leq us$. But $v > u$, so $s = 0$. Thus F' exists a.e..

Let $f_n(x) = n(F(x + 1/n) - F(x))$ (extend F by setting $F(x) = F(b)$ for $x > b$). Then $f_n \to F'$ a.e., and $f_n \geq 0$ since F is increasing. Now by Fatou, $\int_a^b F' = \int_a^b \underline{\lim}\, f_n \leq \underline{\lim} \int_a^b f_n = \underline{\lim}\, n(\int_{a+1/n}^{b+1/n} F - \int_a^b F) = \underline{\lim}(n \int_b^{b+1/n} F - n \int_a^{a+1/n} F) \leq F(b) - F(a)$. □

Corollary *If $F \colon [a, b] \to \mathbb{R}$ is absolutely continuous then F' exists a.e..*

Proof. F absolutely continuous $\Rightarrow$ F is of bounded variation $\Rightarrow$ F is the difference of two increasing functions $\Rightarrow$ F' exists a.e.. □

Example If F is the Cantor function, then $F' = 0$ a.e., but $F(1) - F(0) = 1$.

Lemma 2. *If $F\colon [a,b] \to \mathbb{R}$ is absolutely continuous and $F' = 0$ a.e., then F is constant.*

Proof. Pick $c \in (a,b]$. We need to show that $F(c) = F(a)$. Pick $\varepsilon > 0$, and let $\delta > 0$ be as in the definition of absolute continuity. Let $S = \{x \in (a,c) : F' = 0\}$. Then for all $x \in S$ there exists arbitrarily small intervals $[p,q] = [x, x+h]$ or $[x-h, x]$ such that $|F(q) - F(p)| < \varepsilon(q-p)$. Hence, by Vitali, there is a partition $a = q_0 < p_1 < q_1 < p_2 < q_2 < \cdots < p_n < q_n < c = p_n$ such that $|F(q_i) - F(p_i)| < \varepsilon(q_i - p_i)$ and $\sum(q_i - p_i) > \lambda(S) - \delta = (c-a) - \delta$. But then $\sum(p_{i+1} - q_i) < \delta$. But then $\sum |F(p_{i+1}) - F(q_i)| < \varepsilon$. Thus $|F(c) - F(a)| < \varepsilon + \varepsilon(c-a)$. Since this holds for all $\varepsilon > 0$, $F(c) = F(a)$. $\square$

Lemma 3. *If $f\colon [a,b] \to \mathbb{R}^*$ is integrable and $\int_a^x f = 0$ for all $x \in [a,b]$, then $f = 0$ a.e.*

Proof. Assume that $S_k = \{x : f(x) > 1/k\}$ has positive measure. Then there is a closed set $F \subseteq S$ with $\lambda(F) > 0$, and so $\int_F f \geq \lambda(F)/k > 0$. Write $U = (a,b) \setminus F$. Then U is open and $\int_U f = -\int_F f \neq 0$. Now U is a countable disjoint union of intervals (a_n, b_n) and by DCT $\int_U f = \sum \int_{a_n}^{b_n} f \neq 0$, so $\int_{a_n}^{b_n} f \neq 0$ for some n. But then either $\int_a^{a_n} f \neq 0$ or $\int_a^{b_n} f \neq 0$, a contradiction. Similarly $S'_k = \{x : f(x) < -1/k\}$ has measure zero, so $\bigcup_k (S_k \cup S'_k) = \{x : f(x) \neq 0\}$ has measure zero. $\square$

Theorem (1st Fundamental Theorem of Calculus) *If $f\colon [a,b] \to \mathbb{R}^*$ is integrable and $F(x) = \int_a^x f(t)\,dt$ then F is absolutely continuous, differentiable a.e., and $F' = f$ a.e..*

Proof. W.l.o.g., we may assume $f \geq 0$. We know F is absolutely continuous, so F is continuous and F' exists a.e.. First assume that f is bounded, $f \leq K$. Let $f_n(x) = n(F(x+1/n) - F(x))$. Then $|f_n| = |n \int_x^{x+1/n} f| \leq K$ and $f_n \to F'$ a.e., so by DCT, $\int_a^x F' = \lim \int_a^x f_n = \lim(n \int_x^{x+1/n} F - n \int_a^{a+1/n} F) = F(x) - F(a) = F(x)$, since F is continuous. Now drop the condition that f is bounded. The function $f_n(x) = \min\{f(x), n\}$ is bounded, and if $F_n(x) = \int_a^x f_n$ then $F(x) - F_n(x) = \int_a^x (f - f_n)$ is increasing in x. Hence $(F - F_n)' \geq 0$ a.e., so $\int_a^x F' \geq \int_a^x F'_n = F_n(x)$ for all n. Thus $\int_a^x F' \geq F(x)$. But F is increasing so $\int_a^x F' \leq F(x) - F(a) = F(x)$ by Lemma 1. Hence $\int_a^x F' = F(x)$, so $\int_a^x (F' - f) = 0$ for all $x \in [a,b]$ and thus $F' = f$ a.e. by Lemma 3. $\square$

Theorem (2nd Fundamental Theorem of Calculus) *If $F\colon [a,b] \to \mathbb{R}$ is absolutely continuous, then F' exists a.e., is integrable, and $\int_a^b F'(t)\,dt = F(b) - F(a)$.*

Proof. F is of bounded variation, so $F = G - H$ where G and H are increasing. Thus, by Lemma 1, $F' = G' - H'$ exists a.e., and $\int |F'| \leq \int (G' + H') \leq G(b) - G(a) + H(b) - H(a) < \infty$, so F' is integrable. Let $F_0(x) = \int_a^x F'$. Then $F_0 - F$ is absolutely continuous and $F'_0 - F' = 0$ a.e., thus $F_0 - F$ is constant by Lemma 2. Hence $\int_a^b F' = F_0(b) = F_0(b) - F_0(a) = F(b) - F(a)$. $\square$

Math 7350 12. Convex Functions Fall 2004

Definition A function $\phi\colon (a,b) \to \mathbb{R}$ is *convex* if for all $x, y \in (a,b)$ and $\mu \in [0,1]$, $\phi(\mu x + (1-\mu)y) \leq \mu\phi(x) + (1-\mu)\phi(y)$.

Lemma 1. *If $\phi\colon (a,b) \to \mathbb{R}$ is convex, $a < a_i < b_i < b$ for $i = 1, 2$ and $a_1 \leq a_2$, $b_1 \leq b_2$, then $\frac{\phi(b_1)-\phi(a_1)}{b_1-a_1} \leq \frac{\phi(b_2)-\phi(a_2)}{b_2-a_2}$.*

Proof. Set $\mu = \frac{b_1-a_1}{b_2-a_1}$. Then $\phi(b_1) = \phi(\mu b_2 + (1-\mu)a_1) \leq \mu\phi(b_2) + (1-\mu)\phi(a_1)$. Rearranging gives $\frac{\phi(b_1)-\phi(a_1)}{b_1-a_1} \leq \frac{\phi(b_2)-\phi(a_1)}{b_2-a_1}$. Similarly $\frac{\phi(b_2)-\phi(a_1)}{b_2-a_1} \leq \frac{\phi(b_2)-\phi(a_2)}{b_2-a_2}$. Combining these inequalities gives the result. □

Lemma 2. *$\phi\colon (a,b) \to \mathbb{R}$ is convex iff for each $c \in (a,b)$ there is a linear function $\psi_c(x) = m_c(x-c) + \phi(c)$ with $\psi_c(c) = \phi(c)$ and $\psi_c \leq \phi$ on (a,b). In this case, m_c is an increasing function of c, and $\phi'(c) = m_c$ except at a countable number of points.*

Proof. $\Leftarrow$: Let $c = \mu x + (1-\mu)y$ and consider $\psi_c(x) = m(x-c) + \phi(c)$. Then $\mu\phi(x) + (1-\mu)\phi(y) \geq \mu\psi_c(x) + (1-\mu)\psi_c(y) = \psi_c(c) = \phi(c)$.
$\Rightarrow$: Let $m_-(c) = \sup_{x<c} \frac{\phi(x)-\phi(c)}{x-c}$ and $m_+(c) = \inf_{x>c} \frac{\phi(x)-\phi(c)}{x-c}$. Then by Lemma 1 $m_-(c) \leq m_+(c)$. Any $m_c \in [m_-(c), m_+(c)]$ will do to define ψ_c. Also, by Lemma 1 $\frac{\phi(x)-\phi(c)}{x-c}$ is increasing in x, so $m_\pm(c) = \lim_{x\to c^\pm} \frac{\phi(x)-\phi(c)}{x-c}$. If $d > c$ then by Lemma 1, $m_+(c) \leq m_-(d)$, so the intervals $(m_-(c), m_+(c))$ are disjoint for distinct values of c. Since any non-trivial open interval contains a rational, $m_- = m_+ = \phi'$ for all but a countable number of values of c. □

Lemma 3. *If $\phi\colon (a,b) \to \mathbb{R}$ is convex, then it is absolutely continuous on every closed interval $[c,d] \subseteq (a,b)$.*

Proof. Pick $x < y$, $x, y \in [c,d]$. Then $m_c(y-x) \leq m_x(y-x) \leq \psi_x(y) - \psi_x(x) \leq \phi(y) - \phi(x) \leq \psi_y(y) - \psi_y(y) \leq m_y(y-x) \leq m_d(y-x)$. Hence $|\phi(y)-\phi(x)| \leq M|y-x|$ where $M = \max\{|m_c|, |m_d|\}$. Hence if $\sum |b_i - a_i| < \delta$ then $\sum |\phi(b_i) - \phi(a_i)| < M\delta$. Taking $\delta = \varepsilon/M$ gives absolute continuity. □

Lemma 4. *If $\phi\colon (a,b) \to \mathbb{R}$ is differentiable and ϕ' is increasing, then ϕ is convex.*

Proof. Let $\psi(x) = \phi(x) - \phi'(c)(x-c)$. Then $\psi' \geq 0$ for $x \geq c$ and $\psi' \leq 0$ for $x \leq c$. But then $\psi(x) \geq 0$ for all $x \in (a,b)$ (Mean Value Theorem). Result follows from Lemma 2. □

Theorem (Jensen's Inequality) *If $f\colon [a,b] \to I$ is integrable and $\phi\colon I \to \mathbb{R}$ is convex then*

$$\frac{1}{b-a}\int_a^b \phi(f(t))\,dt \geq \phi\left(\frac{1}{b-a}\int_a^b f(t)\,dt\right).$$

Proof. Write $\phi(x) \geq \phi(c) + m(x-c)$ where $c = \frac{1}{b-a}\int_a^b f(t)\,dt$, and integrate. □

Math 7350 13. L^p Spaces Fall 2004

Definition If f is measurable, the *essential supremum* $\operatorname{ess\,sup} f = \inf\{c : f(x) \le c \text{ a.e.}\}$.

Definition Assume f is measurable. For $1 \le p < \infty$, define $\|f\|_p = \left(\int |f|^p\right)^{1/p}$. If $p = \infty$, define $\|f\|_\infty = \operatorname{ess\,sup} |f|$.

Note: For $1 \le p \le \infty$,

(0) if $f = g$ a.e., then $\|f\|_p = \|g\|_p$.

(1) $\|f\|_p \ge 0$ and if $\|f\|_p = 0$ then $f = 0$ a.e..

(2) If $\mu \in \mathbb{R}$ then $\|\mu f\|_p = |\mu| \|f\|_p$.

Theorem (Minkowski) *If $1 \le p \le \infty$, then $\|f + g\|_p \le \|f\|_p + \|g\|_p$. For $1 < p < \infty$, equality (if finite) occurs iff f and g are proportional a.e..*

Proof. The case $p = \infty$ is clear, so assume $p < \infty$ and set $\alpha = \|f\|_p$, $\beta = \|g\|_p$. W.l.o.g., $0 < \alpha, \beta < \infty$. Let $\mu = \alpha/(\alpha + \beta)$, so that $1 - \mu = \beta/(\alpha + \beta)$. The function $x \to |x|^p$ is convex on $(0, \infty)$ for $p \ge 1$, so

$$\left|f + g\right|^p = \left|\mu \tfrac{f}{\mu} + (1 - \mu) \tfrac{g}{1-\mu}\right|^p \le \mu \left|\tfrac{f}{\mu}\right|^p + (1 - \mu) \left|\tfrac{g}{1-\mu}\right|^p.$$

Integrating gives

$$\|f + g\|_p^p \le \mu \|\tfrac{f}{\mu}\|_p^p + (1 - \mu) \|\tfrac{g}{1-\mu}\|_p^p = \mu(\alpha + \beta)^p + (1 - \mu)(\alpha + \beta)^p = (\alpha + \beta)^p.$$

Now take pth roots to get $\|f + g\|_p \le \alpha + \beta = \|f\|_p + \|g\|_p$. If $p > 1$ then $x \to |x|^p$ is strictly convex, so equality implies $\frac{f}{\mu} = \frac{g}{1-\mu}$ a.e.. $\square$

Definition A *normed space* is a real vector space V with a *norm* $\|\cdot\|$ with the following properties:

(1) $\|v\| \in [0, \infty)$ and $\|v\| = 0$ iff $v = 0$.

(2) $\|\mu v\| = |\mu| \|v\|$ where $\mu \in \mathbb{R}$.

(3) $\|u + v\| \le \|u\| + \|v\|$.

The functions $\|\cdot\|_p$ are not quite norms since they may by ∞ and are zero if $f = 0$ a.e., rather than if $f = 0$. We can construct a vector space on which they are norms as follows.

Definition Let S be a measurable subset of $\mathbb{R}$. Define $L^p(S) = V/\sim$, where $V = \{f \colon S \to \mathbb{R}^* : f \text{ measurable with } \|f\|_p < \infty\}$ and $f \sim g$ iff $f = g$ a.e.. V is a vector space under pointwise addition and scalar multiplication. If we set $Z = \{f \in V : f = 0 \text{ a.e.}\}$, then Z is a vector subspace of V and $V/\sim = V/Z$ is the quotient vector space. Since $\|f\|_p$ only depends on f up to $\sim$, $\|\cdot\|_p$ defines a norm on $L^p(S)$.

Note: $L^1(\mathbb{R})$ is the set of integrable functions and $L^\infty(\mathbb{R})$ is the set of bounded functions (up to $=$ a.e.).

Definition Let l^p be the vector space of infinite sequences $(x_n)_{n=1}^\infty$ with $\|(x_n)\|_p = (\sum x_n^p)^{1/p}$ (or $\sup|x_n|$ if $p = \infty$) finite. Addition and scalar multiplication are componentwise.

Exercise: Show that Minkowski's Theorem holds for l^p, making it a normed spaces with the norm $\|\cdot\|_p$.

Examples

1. Suppose $\lambda(S) < \infty$ and $p < q$, then $f \in L^q(S)$ implies $f \in L^p(S)$.
2. Suppose $p < q$, then $(x_n) \in l^p$ implies $(x_n) \in l^q$.
3. $f(x) = x^{-\alpha} \in L^p([0,1])$ iff $\alpha < 1/p$. In particular, if $p < q$ then $x^{-1/q} \in L^p([0,1])$ but $x^{-1/q} \notin L^q([0,1])$.
4. $f(x) = x^{-\alpha} \in L^p([1,\infty))$ iff $\alpha > 1/p$. In particular, if $p < q$ then $x^{-1/p} \in L^q([1,\infty))$ but $x^{-1/p} \notin L^p([1,\infty))$.
5. Similarly, $(x_n) = (n^{-\alpha}) \in l^p$ iff $\alpha > 1/p$. In particular, if $p < q$ then $(n^{-1/p}) \in l^q$ but $(n^{-1/p}) \notin l^p$.

Lemma (Young's Inequality) *If $p, q > 1$, $a, b \geq 0$ and $\frac{1}{p} + \frac{1}{q} = 1$ then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. Equality holds iff $a^p = b^q$.*

Proof. The function $f(t) = a^{p(1-t)}b^{qt} = a^p(b^q/a^p)^t$ is convex in t (exponential function). Expand $f(\mu 0 + (1-\mu)1) \geq \mu f(0) + (1-\mu)f(1)$ when $\mu = 1/p$. For equality, f must be linear (constant). □

Theorem (Hölder's Inequality) *If $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, and if $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$ then*

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

Equality holds iff f^p is proportional to g^q a.e..

Proof. For $1 < p, q < \infty$ integrate Young's inequality with $a = f/\|f\|_p$ and $b = g/\|g\|_q$. The case $p = 1$, $q = \infty$, is clear. □

Exercises

1. Derive an analogue of Hölder's inequality for sequences.
2. Show that $\int_0^1 \sqrt{x^4 + 4x^2 + 3}\, dx \leq \frac{2}{3}\sqrt{10}$ using Hölder's inequality.
3. Show that if $f(x) > 0$ then $\left(\int_0^1 f(x)\, dx\right)\left(\int_0^1 f(x)^{-1}\, dx\right) \geq 1$

Math 7350 14. Completeness Fall 2004

Definition A sequence of vectors v_n in a normed space converges to v iff $\|v_n - v\| \to 0$ as $n \to \infty$. A sequence (v_n) is a *Cauchy sequence* if $\forall \varepsilon > 0\colon \exists n_0\colon \forall n, m \geq n_0\colon \|v_n - v_m\| < \varepsilon$. Note that any convergent sequence is a Cauchy sequence.

Definition A normed space is *complete* if every Cauchy sequence converges. A complete normed space is also called a *Banach space*.

Definition If $f_n, f \in L^p$ then $f_n \to f$ *in mean of order* p if $f_n \to f$ in L^p, i.e., $\|f_n - f\|_p \to 0$. Note, 'convergence in mean of order 1' is just 'convergence in mean', and 'convergence in mean of order ∞' is a.e. uniform convergence.

Examples Assume $1 \leq p < r < q \leq \infty$.

(a) If $f_n(x) = n^{1/r}\chi_{[0,1/n]}$ then $\|f_n\|_p \to 0$, but $\|f_n\|_q \to \infty$.

(b) If $g_n(x) = n^{-1/r}\chi_{[0,n]}$ then $\|g_n\|_q \to 0$, but $\|g_n\|_p \to \infty$.

(c) If $(x_n) \in l^p$ then $\|(x_n)\|_q \leq \|(x_n)\|_p$. [Scale so that $\|(x_n)\|_p = 1$.]

(d) If $f \in L^q(S)$, $\lambda(S) < \infty$, then $\|f\|_p \leq \|f\|_q \lambda(S)^{1/p-1/q}$. [Use Hölder.]

Lemma 1. *A normed space is complete iff every absolutely convergent sum is convergent, i.e.,* $\sum_{n=1}^{\infty} \|v_n\| < \infty \Rightarrow \sum_{n=1}^{N} v_n$ *converges to some* v *as* $N \to \infty$.

Proof. $\Rightarrow$: If $u_N = \sum_{n=1}^{N} v_n$, then $\|u_n - u_m\| \leq \sum_{i=m+1}^{n} \|v_i\| \to 0$ as $\min\{m, n\} \to \infty$. Thus (u_n) is a Cauchy sequence, so converges.
$\Leftarrow$: If (v_n) is Cauchy, then there is a subsequence (v_{n_i}) with $\|v_n - v_{n_i}\| < 2^{-i}$ for all $n > n_i$. Let $u_0 = v_{n_1}$ and $u_i = v_{n_{i+1}} - v_{n_i}$ for $i > 0$. Then $\sum \|u_i\| < \infty$, so $\sum_{i=0}^{N-1} = v_{n_N}$ converges to v, say, as $N \to \infty$. But for all $n > n_i$, $\|v_n - v\| \leq 2^{-i} + \|v_{n_i} - v\|$, so $v_n \to v$. $\square$

Theorem (Riesz-Fischer) $L^p(S)$ *is complete for* $1 \leq p \leq \infty$.

Proof. Suppose $\sum_{n=1}^{\infty} \|f_n\|_p = L < \infty$. Let $g_N(x) = \sum_{n=1}^{N} |f_n(x)|$ and $g(x) = \sum_{n=1}^{\infty} |f_n(x)|$. Since $\||f|\|_p = \|f\|_p$, $\|g_N\|_p \leq L$ for all N. If $p < \infty$, then $\int |g_N|^p \leq L^p$ and $|g_N|^p$ is increasing, so by MCT $\int |g|^p \leq L^p < \infty$. But then $g(x) < \infty$ a.e.. Similarly, if $p = \infty$, $g_N \leq L$ a.e., so $g = \lim g_N \leq L$ a.e.. But when $g(x) < \infty$, $\sum |f_n(x)|$ converges. But then $f(x) = \sum_{i=1}^{\infty} f_n(x)$ converges a.e., and $|f| \leq |g|$. If $p < \infty$ then $|f - \sum_{n=1}^{N} f_n|^p \leq |g|^p$, so by DCT $\int |f - \sum_{n=1}^{N} f_n|^p \to 0$ and $\sum_{n=1}^{\infty} f_n$ converges to f in L^p. Similarly if $p = \infty$, $|f - \sum_{n=1}^{N} f_n| \leq \sum_{N+1}^{\infty} \|f_n\|_\infty$ a.e., so $\sum_{n=1}^{\infty} f_n$ converges to f in L^∞. $\square$

Approximating functions in L_p

Just as any real is a limit of rationals, it is sometimes convenient to express an element of L^p as a limit of functions of a special form.

Theorem *If $1 \le p < \infty$, $f \in L^p(\mathbb{R})$, and $\varepsilon > 0$ then*

(a) there is a simple function $\phi \in L^p$ with $\|\phi - f\|_p < \varepsilon$,

(b) there is a step function $\psi \in L^p$ with $\|\psi - f\|_p < \varepsilon$,

(c) there is a continuous function $g \in L^p$ with $\|g - f\|_p < \varepsilon$.

Proof. (a) Let $\phi_n(x) = \frac{a}{2^n}$, $-n2^n \le a \le n2^n$ be the closest value to $f(x)$ with $|\phi(x)| \le |f(x)|$. Then $|\phi_n - f|^p \le |f|^p$ and $\phi_n \to f$ a.e.. Thus by DCT $\|\phi_n - f\|_p \to 0$.
(b) By (a) and linearity it is enough to approximate χ_S with a step function. Since $\chi_S \in L^p$, $\lambda(S) < \infty$. Thus there is a finite disjoint union of intervals U with $\lambda(S \triangle U) < \varepsilon^p$. But then $\|\chi_S - \chi_U\|_p^p = \int \chi_{S\triangle U}^p < \varepsilon^p$ and χ_U is a step function.
(c) By (b) and linearity it is enough to approximate $\chi_{[a,b]}$ with a continuous function. Let $g(x)$ be a piecewise linear function with $g(a - \varepsilon^p/4) = 0$, $g(a + \varepsilon^p/4) = g(b - \varepsilon^p/4) = 1$, $g(b + \varepsilon^p/4) = 0$. Then $\|g - \chi_{[a,b]}\|_p^p \le \varepsilon^p$. □

Part (a) also holds for $p = \infty$, but (b) and (c) fail. However, we do have:

Theorem (Lusin's Theorem) *Let $f\colon \mathbb{R} \to \mathbb{R}$ be measurable. Then for any $\varepsilon > 0$ there is a continuous function $g\colon \mathbb{R} \to \mathbb{R}$ such that $\lambda(\{x : f(x) \ne g(x)\}) < \varepsilon$.*

Note: This is *not* the same as saying f is continuous on a large set (e.g., consider $\chi_{\mathbb{Q}}$).

Proof. Enumerate the rational numbers as $\mathbb{Q} = \{q_i : i = 1, 2, \dots\}$. For each q_i, find an open U_i and a closed F_i with $F_i \subseteq f^{-1}[(q_i, \infty)] \subseteq U_i$ and $\lambda(U_i \setminus F_i) < \varepsilon/2^i$. Let $U = \bigcup_i (U_i \setminus F_i)$. Then $\lambda(U) < \varepsilon$. Now U is open, so a union of disjoint open intervals (a_j, b_j). Define $g(x) = f(x)$ if $x \notin U$, and let g be linear on $[a_j, b_j]$. Clearly g is continuous at all $x \in U$. Now assume $x \notin U$ and $\eta > 0$. Pick rationals q_i, q_j, with $g(x) - \eta < q_i < f(x) = g(x) < q_j < g(x) + \eta$. Then $x \in U_i \setminus F_j$, which is open. Thus $y \in U_i \setminus F_j$ for $|y - x| < \delta$, say. If $y \notin U$ then $y \in F_i$ and $y \notin U_j$, so $q_i < f(y) = g(y) \le q_j$. If $y \in U$ then $y \in (a_i, b_i)$ for some i. If this interval is entirely within $(x - \delta, x + \delta)$ then $q_i < g(y) \le q_j$ since this holds at both endpoints a_i and b_i and g is linear on $[a_i, b_i]$. There can be at most two such intervals (a_i, b_i) that do not lie in $(x - \delta, x + \delta)$, and since g is linear on these, one can ensure, by reducing δ if necessary, that $|g(y) - g(x)| < \varepsilon$ for these values of y as well. □

Similarly, although $f_n \to f$ a.e. does not imply $f_n \to f$ in L^p, we do have:

Theorem (Egorov's Theorem) *If $f_n\colon S \to \mathbb{R}$ is a sequence of measurable functions that converge a.e., to $f\colon S \to \mathbb{R}$, and $\lambda(S) < \infty$, then $\exists E\colon\ \lambda(E) < \varepsilon$ and $f_n \to f$ uniformly on $S \setminus E$ (and hence $f_n \to f$ in $L^p(S \setminus E)$ for all $1 \le p \le \infty$).*

Proof. Since $f_n \to f$ a.e., the set $E_0 = \{x : f_n(x) \not\to f(x)\}$ has measure zero. Let $E_{kn} = \{x \in S : |f_n(x) - f(x)| \ge 1/k\}$. If $\forall n_0\colon \exists n \ge n_0\colon |f_n(x) - f(x)| \ge 1/k$ then $f_n(x) \not\to f(x)$. Thus for all k, $\bigcap_{n_0} \bigcup_{n \ge n_0} E_{kn} \subseteq E_0$. Since $\lambda(\bigcup_n E_{kn}) \le \lambda(S) < \infty$, $\lim_{n_0} \lambda(\bigcup_{n\ge n_0} E_{kn}) = 0$. For each k, pick n_k such that $\lambda(\bigcup_{n \ge n_k} E_{kn}) < \delta/2^k$. Set $E = \bigcup_k \bigcup_{n \ge n_k} E_{kn}$. Then $\lambda(E) \le \sum \delta/2^i = \delta$. Moreover, if $n \ge n_k$ then $|f_n(x) - f(x)| < 1/k$ for all $x \in S \setminus E$. Hence $f_n \to f$ uniformly on $S \setminus E$. □

Math 7350 15. Linear Functionals Fall 2004

Definition A *linear functional* on a normed space $(V, \|\cdot\|)$ is a linear function $F\colon V \to \mathbb{R}$, i.e., $\forall \alpha, \beta \in \mathbb{R}$, $u, v \in V$: $F(\alpha u + \beta v) = \alpha F(u) + \beta F(v)$. A linear functional F is *bounded* iff there is some constant M such that $|F(v)| \leq M\|v\|$ for all $v \in V$. We define the *norm* of F to be the smallest such M:

$$\|F\| = \sup_{v \in V \setminus 0} \tfrac{|F(v)|}{\|v\|}.$$

A linear functional is bounded iff it is 'continuous', i.e.,

$$\forall v\colon \forall \varepsilon > 0\colon \exists \delta > 0\colon \forall u\colon \|u - v\| < \delta \Rightarrow \|F(u) - F(v)\| < \varepsilon.$$

To see this, set $w = \frac{1}{\delta}(u - v)$, then F is continuous iff

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall w\colon \|w\| < 1 \Rightarrow \|F(w)\| < \varepsilon/\delta.$$

But this is equivalent to F being bounded ($\Rightarrow$: take $\varepsilon = 1$. $\Leftarrow$: take $\delta = \varepsilon/M$).

Theorem (Riesz Representation Theorem) *Let F be a bounded linear functional on $L^p(S)$, $1 \leq p < \infty$. Then there is a function $g \in L^q(S)$, $\frac{1}{p} + \frac{1}{q} = 1$, such that*

$$F(f) = \int fg \qquad \text{for all } f \in L^p(S).$$

Moreover, for all $g \in L^q(S)$, the above formula defines a linear functional with $\|F\| = \|g\|_q$.

Proof. Extending F to $L^p(\mathbb{R})$ by setting $F(f) = F(\chi_S f)$, we can assume F is a bounded linear functional on $L^p(\mathbb{R})$ with the same norm ($|F(f)| = |F(\chi_S f)| \leq M\|\chi_S f\|_p \leq M\|f\|_p$). Thus we may assume $S = \mathbb{R}$.

Define $\Phi(x) = F(\chi_{[0,x]})$ ($= -F(\chi_{[x,0]})$ if $x < 0$). Then Φ is absolutely continuous on any finite interval: If $I_i = (a_i, b_i)$ are disjoint and $\sum \lambda(I_i) < (\varepsilon/M)^p$ then $\sum |\Phi(b_i) - \Phi(a_i)| = F(f)$ where $f = \sum \pm\chi_{I_i}$. But $\|f\| = (\sum \lambda(I_i))^{1/p} < \varepsilon/M$, so $|F(f)| < \varepsilon$.

Since Φ is absolutely continuous, $\Phi(x) = \int_0^x g$ ($= -\int_x^0 g$ if $x < 0$), where $g = \Phi'$. For any finite interval $I = [a, b]$, $F(\chi_I) = \Phi(b) - \Phi(a) = \int \chi_I g$. Thus by linearity $F(\chi_U) = \int \chi_U g$ for all finite unions U of finite intervals. If $E \subseteq [-K, K]$ is measurable and $\lambda(E \triangle U_n) < 2^{-n}$, then $\chi_{U_n} \to \chi_E$ in L^p and a.e., so $F(\chi_E) = \lim F(\chi_{U_n}) = \lim \int \chi_{U_n} g = \int \chi_E g$ by continuity of F and DCT ($|g|$ is integrable on $[-K, K]$). Now by linearity, $F(f) = \int fg$ for every simple function f supported on $[-K, K]$.

Suppose $\|g\|_q > M = \|F\|$. Then one can find a K and a simple function g_0 supported on $[-K, K]$ with $0 \leq |g_0| \leq |g|$, $\operatorname{sgn} g_0 = \operatorname{sgn} g$, and $\|g_0\|_q > M$. Now $f = (\operatorname{sgn} g_0)|g_0|^{q/p}$ is simple and supported on $[-K, K]$. By equality in Hölder, $F(f) = \int fg \geq \int |fg_0| = \|g_0\|_q \|f\|_p > M\|f\|_p$, a contradiction. (For $q = \infty$, take $f = (\operatorname{sgn} g)\chi_{\{|g|>M\} \cap [-K,K]}$.)

Finally, suppose $f \in L^p$. We can find a simple function supported on some $[-K, K]$ with $\|f - \phi\|_p < \varepsilon$. Now $|F(f) - F(\phi)| = |F(f - \phi)| \leq M\varepsilon$ and $|\int fg - \int \phi g| = |\int (f - \phi)g| \leq \|f - \phi\|_p \|g\|_q \leq M\varepsilon$. Thus letting $\varepsilon \to 0$, $F(f) = \int fg$ for all $f \in L^p$. Finally, $|F(f)| \leq \|f\|_p \|g\|_q$, so $M = \|F\| \leq \|g\|_q$. Thus $\|F\| = \|g\|_q$. $\square$