Monday, 1-10-2005

Exam dates:
Final - Thursday, May 5, 1 - 3 PM Midterms - Wednesday, February 16 and March 30. 5:30 - 7:30 PM

Review
We say \( \nu \) (signed measure) is absolutely continuous to \( \mu \) ("regular" measure) (denoted \( \nu \ll \mu \)) if
\[
\mu(E) = 0 \implies |\nu|(E) = 0.
\]

1. Radon-Nikodym and Lebesgue Decomposition Theorems

Theorem 1. (Radon-Nikodym): Let \( \nu \) be a \( \sigma \)-finite signed measure on \( (X, S) \) and let \( \mu \) be a \( \sigma \)-finite positive measure on \( (X, S) \). If \( \nu \ll \mu \), then there exists an extended-real valued \( \mu \)-integrable function \( f : X \rightarrow \mathbb{R} \) such that
\[
\nu(E) = \int_E f \, d\mu.
\]
If \( f_1 : X \rightarrow \mathbb{R} \) is \( \mu \)-integrable and satisfies
\[
\nu(E) = \int_E f_1 \, d\mu,
\]
then \( f = f_1 \) \( \mu \)-almost everywhere.

Theorem 2. (Lebesgue-Decomposition): Let \( \nu \) be a \( \sigma \)-finite signed measure on \( (X, S) \) and \( \mu \) be a \( \sigma \)-finite positive measure on \( (X, S) \). Then there exists unique \( \sigma \)-finite signed measures \( \nu_0 \) and \( \nu_1 \) on \( (X, S) \) such that
\[
\nu = \nu_0 + \nu_1, \quad \nu_0 \perp \mu \text{ and } \nu_1 \ll \mu.
\]

First, we need a lemma.

Lemma 1.1. Suppose \( \nu \) and \( \mu \) are positive measures on \( (X, S) \) such that neither takes the a value in \( \{\infty, -\infty\} \). Then either \( \nu \perp \mu \) or there exists a \( \epsilon > 0 \) and \( E \in S \) such that \( \mu(E) > 0 \) and \( \nu(F) \geq \epsilon \mu(F) \) for all \( F \in S, F \subset E \).

Remark: Note that the conclusion is that \( E \) is a positive set for \( \nu - \epsilon \mu \).

Proof of Lemma:
For each \( n \in \mathbb{N} \), let \( (P_n, N_n) \) be the Hahn Decomposition for \( \nu - \frac{1}{n} \mu \). Put
\[
P = \bigcup_{n} P_n \text{ and } N = \bigcap_{n} N_n.
\]
Notice that \( N \) is a negative set for \( \nu - \frac{1}{n} \mu, \forall n \in \mathbb{N} \). Hence, \( 0 \leq \nu(N) \leq \frac{\mu(N)}{n} \) for all \( n \). Hence, \( \nu(N) = 0 \).

(a) If \( \mu(P) = 0 \), then since \( P \cup N = X \) and \( P \cap N = \emptyset \), then we see that \( \mu \perp \nu \).
(b) If \( \mu(P) > 0 \), then \( \mu(P_n) > 0 \) for some \( n \in \mathbb{N} \). Since \( P_n \) is a positive set for \( \nu - \frac{1}{n} \mu \), we find that \( \nu(F) \geq \frac{1}{n} \mu(F) \) for all \( F \in P_n \). Take \( P_n = E \), and \( \epsilon = \frac{1}{n} \), and the lemma is proved.
Proof. First, suppose \( \mu \) and \( \nu \) are finite positive measures.

Let \( \mathcal{F} = \{ f : X \to [0, \infty] : f \) is measurable and \( \int_E f d\mu \leq \nu(E), \forall E \in \mathcal{S} \}. \)

Note that \( \mathcal{F} \) is finite because \( \mathcal{S} \) is finite subalgebra. Let \( f, g \in \mathcal{F} \), then we claim \( h = \max f, g \in \mathcal{F} \).

If \( A = \{ x \in X : f(x) > g(x) \} \), then given any \( E \in \mathcal{S} \),
\[
hd\mu = \int_{E \cap A} hd\mu + \int_{E \setminus A} hd\mu = \int_{E \cap A} f d\mu + \int_{E \setminus A} g d\mu \\
\leq \nu(E \cap A) + \nu(E \setminus A) = \nu(E)
\]

Let \( a = \sup_f \{ \int_X f d\mu : f \in \mathcal{F} \} \). Then \( a \leq \nu(V) < \infty \). Choose \( f_n \in \mathcal{F} \) such that \( \int_X f_n d\mu \to a \). Put
\[
g_n = \max \{ f_1, \ldots, f_n \}
\]

Then, \( g_n \leq g_{n+1} \) and as \( g_n \in \mathcal{F} \), we have
\[
a \geq \int_X g_n d\mu \geq \int_X f_n d\mu,
\]
so
\[
\int_X g_n d\mu \to a.
\]

Put \( f(x) = \lim_{n \to \infty} g_n(x) \). By the Monotone Convergence Theorem, we get
\[
\int_X f d\mu = a < \infty
\]

Thus \( f \) is integrable and \( f \) is finite almost everywhere. WLOG, we may take \( f \) to be finite everywhere. Put \( \nu_0(E) = \nu(E) - \int_E f d\mu \) for \( E \in \mathcal{S} \). Note that \( \nu_0 \) is a positive measure because \( f \in \mathcal{F} \).

We claim that \( \mu \perp \nu_0 \). If not, then by Lemma 1.1 there exists a \( \epsilon > 0 \) and \( E_0 \in \mathcal{S} \) which is positive for \( \nu_0 - \epsilon \mu \) and such that \( \mu(E_0) > 0 \). Then,
\[
\int_E (f + \epsilon \chi_{E_0}) d\mu = \int_E f d\mu + \epsilon \mu(E \cap E_0) \\
\leq \int_E f d\mu + \nu_0(E \cap E_0) \leq \int_E f d\mu + \nu_0(E) = \nu(E).
\]

Thus, \( f + \epsilon \chi_{E_0} \in \mathcal{F} \). But,
\[
\int_X f + \epsilon \chi_{E_0} d\mu = a + \epsilon \mu(E_0) > a,
\]
which contradicts the fact that \( \int_X f = a \) is a supremum in \( \mathcal{F} \). Thus, \( \mu \perp \nu_0 \).

We define \( \nu_1(E) = \int_E f d\mu \).

Wednesday, 1-12-2005

Uniqueness: Suppose \( \nu = \tau_0 + \tau_1 \), where \( \tau_i \) are signed measures, \( \tau_0 \perp \mu \), and \( \tau_1 \ll \mu \). We have \( \nu = \tau_1 + \tau_0 = \nu_0 + \nu_1 \). So, \( \tau_0 - \nu_0 = \nu_1 - \tau_1 \).

Claim: \( \tau_0 - \nu_0 \perp \mu \) and \( \nu_1 - \tau_1 \ll \mu \).

Note that the claim gives uniqueness - last semester we proved that if \( \sigma \perp \mu \) and \( \sigma \ll \mu \), then \( \sigma = 0 \).

Proof of claim:
Find measurable sets \( A, B \) such that \( A \cap B = \emptyset \) and \( A \cup B = X \), \( C, D \) such that \( C \cap D = \emptyset \) and \( C \cup D = X \), and \( \nu_0(B) = \mu(A) = 0 = |\tau_0|(D) = \mu(C) \). Then \( A \cup C \)
and $B \cap D$ are disjoint and have union $X$.
Also, $|\tau_0 - \nu_0| \leq |\tau_0| + |\nu_0|$ (proof of this inequality is left as an exercise), so $|\tau_0 - \nu_0|(B \cap D) = 0$ and $\mu(A \cup C) = 0$. Thus, $\tau_0 - \nu_0 \perp \mu$.
Next, if $\mu(E) = 0$ and $F \subset E$, then
$$\nu_1(F) = \tau_1(F) = 0.$$ 
So,
$$|\tau_0 - \nu_0|_0 = (B \cap D) = 0$$
Hence, $E$ is a null set for $\nu_1 - \tau_1$, so $\tau_1 - \nu_1 \ll \mu$.
This proves Theorems 1 and 2 for finite cases.

For the Radon-Nikodyn Theorem 
??, since $\nu \ll \mu$ and we have $\nu = 0 + \nu$, the decomposition breaks $\nu$ into singular and absolutely continuous parts. By uniqueness of the Lebesgue Decomposition theorem, $\nu = \int_X f\,d\mu$.

In order to complete the proofs for the $\sigma$-finite case, we first need this lemma:

**Lemma 1.2.** Let $\lambda$ be a $\sigma$-finite positive measure on $(X, S)$ and suppose $\lambda(X) = \infty$.
Then there exists an integrable function $w : X \to (0, 1)$ such that if $\lambda_w(E) = \int_E w\,d\lambda$
we have the following:

- (a) $\lambda_w$ is a finite measure.
- (b) $\forall f \geq 0$ measurable, $\int_X f\,d\lambda_w = \int_X f\,w\,d\lambda$.
- (c) $\lambda(E) = \int_E \frac{1}{w}\,d\lambda_w$.
- (d) $\forall f \geq 0$ measurable, $\int_X f\,d\lambda = \int_X f\,\frac{1}{w}\,d\lambda_w$.
- (e) For $E \in S$, $\lambda(E) = 0 \iff \lambda_w(E) = 0$.

**Proof.** Let $A_i$ be disjoint measurable sets such that $\cup_{i=1}^{\infty} A_i = X$ and $1 < \lambda(A_1) < \infty$.

Put $w = \sum_{i=1}^{\infty} \frac{1}{2^j \lambda(A_j)} \chi_{A_j}$. Note that $w(x) \in (0, 1)$ for all $x \in X$. Also, by the MCT
$$\int_X w\,d\lambda = \sum_{j=1}^{\infty} \frac{1}{2^j \lambda(A_j)} \int_X \chi_{A_j} \,d\lambda$$
Hence, $w$ is integrable, and $\lambda_w$ is a finite measure.
We can easily prove (b) by noting that it is true for simple functions, passing to approximations, then repeat as usual.
In order to prove (c), simply apply (b) with $f = \frac{1}{w}$.
Part (d) follows from (c) as in part (b).
Finally, (e) follows from the definition of $\lambda_w$ and part (c).

Now, we return to the proof of Theorems 1 and 2 for $\sigma$-finite positive measures.

**Proof.** Let $\nu, \mu$ be $\sigma$-finite positive measures. Apply the lemma to obtain measurable functions $v, w : X \to (0, 1)$ with $v$ $\nu$-integrable and $w$ $\mu$-integrable.

Put
$$\nu_v(E) = \int_E v\,d\nu$$
and
$$\mu_w(E) = \int_E w\,d\mu.$$
Find unique measures $\nu^s_v$ and $\nu^a_v$ such that $\nu^s_v \perp \mu_w$, $\nu^a_v \ll \mu_w$, and $\nu^s_v + \nu^a_v = \nu_v$.

Define

$$\nu_0(E) = \int_E \frac{1}{v} d\nu^s_v$$

and

$$\nu_1(E) = \int_E \frac{1}{v} d\nu^a_v.$$ 

Then,

$$\nu_0(E) + \nu_1(E) = \int_E \frac{1}{v} d(\nu^s_v + \nu^a_v) = \nu(E)$$

Since $\nu^s_v \perp \mu_w$, we have $\nu_0 \perp \mu$ and since $\nu^a_v \ll \mu_w$ we have $\nu_1 \ll \mu$.

Since $\nu^a_v \ll \mu_w$, there exists an $h \mu_w$-integrable such that

$$\nu^a_v(E) = \int_E h d\mu_w$$

Define,

$$\nu_1(E) = \int_E \frac{1}{v} d\nu^a_v = \int_E \frac{1}{v} h d\mu_w = \int_E \frac{1}{v} h w d\mu$$

Finally, to establish the theorem for a signed measure $\nu$, apply the above to $\nu^+$ and $\nu^-$ and take the difference $\nu = \nu^+ - \nu^-$. □

**Notation:** If $\nu \ll \mu$, we write the $\nu(E) = \int_E f d\mu$. The function $f$ is called the Radon-Nikodym derivative of $\nu$ with respect to $\mu$. We denote $f$ by $\left[\frac{d\nu}{d\mu}\right]$. Often when $\nu \ll \mu$, we write $d\nu = f d\mu$. This is motivated by the fact

$$\int h d\nu = \int h f d\mu.$$ 

**Proposition 1.1.** Let $\mu$ be a $\sigma$-finite signed measure and suppose $\lambda, \mu$ are positive $\sigma$-finite measures such that $\nu \ll \mu$ and $\mu \ll \lambda$.

(a) If $g \in L^1(\nu)$, then $\int g d\nu = \int_X g \left[\frac{d\nu}{d\mu}\right] d\mu$

(b) If $\nu \ll \lambda$, then $\left[\frac{d\nu}{d\lambda}\right] = \left[\frac{d\nu}{d\mu}\right] \cdot \left[\frac{d\mu}{d\lambda}\right]$

**Proof.** To prove (a), prove it for $\nu^+$ and $\nu^-$ (through standard methods) and subtract.

For (b), we have

$$\nu(E) = \int_E \left[\frac{d\nu}{d\mu}\right] d\mu = \int_E \left[\frac{d\nu}{d\mu}\right] \cdot \left[\frac{d\mu}{d\lambda}\right] d\lambda$$

Therefore, $\nu \ll \mu$ and $\left[\frac{d\nu}{d\mu}\right] = \nu(E)$.

□

2. **Complex Measures**

**Definition 2.1.** A complex measure on a measurable space $(X, S)$ is a function $\nu : S \to \mathbb{C}$ such that

(a) $\nu(\emptyset) = 0$, and
(b) If \( \{E_i\}_{i=1}^{\infty} \) are pairwise disjoint, \( E_i \in \mathcal{S} \), then

\[
\nu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \nu(E_i),
\]

where the sum converges absolutely.

**Remark:** A positive measure is a complex measure when it is finite \((\mu(X) < \infty)\).

For a complex measure, \( \nu \), write \( \nu_{re}, \nu_{im} \) for its real and imaginary parts, respectively. Thus \( \nu_{im} \) and \( \nu_{re} \) are signed measures:

\[
\nu = \nu_{re} + \nu_{im} = (\nu_{re}^+ - \nu_{re}^-) + i(\nu_{im}^+ - \nu_{im}^-)
\]

Therefore, the range of \( \nu \) is a bounded set.

For a complex measure, \( \nu \),

\[
L^1(\nu) := L^1(\nu_{re}) \cap L^1(\nu_{im}).
\]

For \( f \in L^1(\nu) \), define

\[
\int_X f \, d\nu = \int_X f \, d\nu_{re} + i \int_X f \, d\nu_{im}.
\]

For \( \nu \) a complex measure and \( \mu \) a positive measure, \( \nu \perp \mu \) means \( \nu_{re} \perp \mu \) and \( \nu_{im} \perp \mu \).

Similarly, \( \nu \ll \mu \) means \( \nu_{re} \ll \mu \) and \( \nu_{im} \ll \mu \).

**Theorem 3. Lebesgue-Radon-Nikodym**

Let \((X, \mathcal{S})\) be a measurable space, \( \mu \) a \( \sigma \)-finite positive measure, and \( \nu \) a complex measure. Then, there exists unique complex measures \( \nu_0 \) and \( \nu_1 \) such that \( \nu = \nu_0 + \nu_1 \), \( \nu_0 \perp \mu \), and \( \nu_1 \ll \mu \). Moreover, there exists a \( \mu \)-integrable function \( f : X \to \mathbb{C} \) such that \( \nu_1(E) = \int_E f \, d\mu \). Finally, \( f \) is unique \( \mu \)-almost everywhere.

Wednesday, 1-19-2005

**Definition 2.2. Extended Integrable:**

Let \( \nu \) be a signed measure. Then, we say that \( f : X \to \mathbb{R} \) is extended integrable if \( f \) is integrable of at least one of \( \nu^+ \) or \( \nu^- \). For such \( f \),

\[
\int f \, d\nu = \int f \, d\nu^+ - \int f \, d\nu^-
\]

(Thus, we don’t have a problem such as \( \infty - \infty \))

Because of this, we must slightly change the statement of the Radon-Nikodym Theorem:

**Theorem 4. Radon-Nikodym**

Let \( \nu \) be a \( \sigma \)-finite signed measure on \((X, \mathcal{S})\) and let \( \mu \) be a \( \sigma \)-finite positive measure on \((X, \mathcal{S})\). If \( \nu \ll \mu \), then there exists an extended integrable function \( f \) (i.e., either \( f^+ \) or \( f^- \) is integrable) such that

\[
\nu(E) = \int_E f \, d\mu
\]

for all \( E \in \mathcal{S} \).
Let $\nu$ be a complex measure. We write
\[
\mu = \nu_{Re}^+ + \nu_{Re}^- + \nu_{Im}^+ + \nu_{Im}^- 
\]
where $\mu$ is the total variation of $\nu$. Note that $\mu$ is a positive measure and $\nu \ll \mu$.

**Lemma 2.1.** Suppose $\lambda_1$, $\lambda_2$ are positive measures such that $\nu \ll \lambda_1$ and $\nu \ll \lambda_2$. Write $d\nu = f_1d\lambda_1$. Then,
\[
|f_1|d\lambda_1 = |f_2|d\lambda_2 
\]

*Proof.* Let $\rho = \lambda_1 + \lambda_2$, so $\rho$ is a positive measure with $\lambda_1 \ll \rho$ and $\lambda_2 \ll \rho$. By the Theorem 1, there exists non-negative functions $h_1$, $h_2$ such that $d\lambda_1 = h_1d\rho$ and $d\lambda_2 = h_2d\rho$.

So, $d\nu = f_1h_1d\rho$ and similarly $d\nu = f_2h_2d\rho$. By the uniqueness statement, we have $f_1h_1 = f_2h_2 \rho$-almost everywhere. So, $|f_1|h_1 = |f_2|h_2 \rho$-almost everywhere. In other words, $|f_1|d\lambda_1 = |f_2|d\lambda_2$. $\square$

**Definition 2.3. Total Variation:**
The total variation $|\nu|$ of the complex measure $\nu$ is defined by $d|\nu| = |f|d\lambda$ where $\lambda$ is any positive measure with $\nu \ll \lambda$ and $d\nu = f d\lambda$.

Properties of $|\nu|$:
(a) $\forall E$ measurable, $|\nu(E)| \leq |\nu|(E)$
(b) $\nu \ll |\nu|$ and $\left|\frac{d\nu}{d|\nu|}\right| = 1$ almost everywhere.
(c) If $\nu_1, \nu_2$ are complex measures, then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$.

*Proof.* Let $\lambda$ be a positive measure such that $\nu \ll \mu$ and write $d\nu = f d\lambda$. Then,
(a) $|\nu(E)| = \int_E f d\lambda \leq \int_E |f|d\lambda = |\nu|(E)$.
(b) Part (a) give $\nu \ll |\nu|$. Write
\[
\nu(E) = \int_E \left|\frac{d\nu}{d|\nu|}\right| d|\nu|
\]

Then,
\[
\int_E 1 d|\nu| = |\nu|(E) = \left|\frac{d\nu}{d|\nu|}\right| =
\]

Therefore, by uniqueness of Theorem 1,
\[
\left|\frac{d\nu}{d|\nu|}\right| = 1, \ |\nu|-\text{almost everywhere}
\]

(c) Let $\lambda_1, \lambda_2$ be positive measures such that $\nu_1 \ll \lambda_1$ and $\nu_2 \ll \lambda_2$; write $d\nu_1 = f_1d\lambda_1$. Put $\rho = \lambda_1 + \lambda_2$. Then,
\[
(\nu_1 + \nu_2)(E) = \int_E f_1d\lambda_1 + \int_E f_2d\lambda_2 =
\]
\[
= \int_E f_1 \left[ \frac{d\lambda_1}{d\rho} \right] d\rho + \int_E f_2 \left[ \frac{d\lambda_2}{d\rho} \right] d\rho =
\]
\[
= \int_E \left( f_1 \left[ \frac{d\lambda_1}{d\rho} \right] + f_2 \left[ \frac{d\lambda_2}{d\rho} \right] \right) d\rho.
\]

So,
\[
|\nu_1 + \nu_2|(E) = \int_E \left| f_1 \left[ \frac{d\lambda_1}{d\rho} \right] + f_2 \left[ \frac{d\lambda_2}{d\rho} \right] \right| d\rho
\]
\[
\leq \int_E |f_1| \left| \frac{d\lambda_1}{d\rho} \right| d\lambda_1 + \int_E |f_2| \left| \frac{d\lambda_2}{d\rho} \right| d\lambda_2 = \int_E |f_1| d\lambda_1 + \int_E |f_2| d\lambda_2 = |\nu_1|(E) + |\nu_2|(E)
\]

**Theorem 5.** Let \( \nu \) be a complex measure. Then for any measurable \( E \),
\[
|\nu|(E) = \sup \left\{ \sum_{i=1}^{\infty} |\nu(E_i)| : E = \bigcup_{i=1}^{\infty} E_i, E_i \cap E_j = \emptyset (i \neq j), \text{ and } E_i \ \text{measurable} \right\}
\]

Proof. Found in Rudin, *Real and Complex Analysis.*

An application to Harmonic Analysis:
Let \( X = \mathbb{R} \), \( \mathcal{L} \) be Lebesgue-measurable sets. For complex measures \( \nu_1, \nu_2 \), define
\[
(\nu_1 * \nu_2)(E) = \int_\mathbb{R} \left[ \int_\mathbb{R} \chi_E(x+y) d\nu_1(x) \right] d\nu_2(y)
\]
This is the convolution of \( \nu_1 \) and \( \nu_2 \). Note that by using the Radon-Nikodym derivatives, we can convert this to the product of two positive measures.

This makes the set of complex measures into an algebra over \( \mathbb{C} \), with vector space under + and usual associative multiplication.

The measures which are absolutely continuous with respect to the Lebesgue measure \( m \) form an ideal.

Let \( \nu_t \) be the point mass at \( t \). Then, \( \nu_t * \nu_s = \nu(t+s) \). Hence, the real line is embedded in this algebra in a natural sort of way.

3. Differentiating Measures on \( \mathbb{R}^n \)
Let \( m \) be Lebesgue measurable on \( \mathbb{R}^k \). An integrable \( f \) is with respect to \( m \).

**Definition 3.1. Shrinks Nicely:**
Fix \( x \in \mathbb{R}^k \). A sequence of Borel sets \( \{E_i\}_{i=1}^{\infty} \) is said to shrink nicely to \( x \) if \( \exists \alpha > 0 \) and \( \exists r_i > 0 \) such that
(a) \( E_i \subset B(x, r_i) \)
(b) \( m(E_i) > \alpha m(B(x, r_i)) \)
(c) \( r_i \to 0 \)

Remarks:
(1) Notice that we do not require \( x \in E_i \).
(2) Some authors use an index \( r \in (0, \infty) \) and define \( E_r \) shrinking nicely to \( x \).

**Definition 3.2.** Let \( \mu \) be a complex Borel measure on \( \mathbb{R}^k \) and let \( x \in \mathbb{R}^k \). If \( \lambda \in \mathbb{C} \) and
\[
\lim_{i \to \infty} \frac{\mu(E_i)}{m(E_i)} = \lambda
\]
for every sequence \( \{E_i\}_{i=1}^{\infty} \) of Borel sets shrinking nicely to \( x \), we say \( \lambda \) is the derivative of \( \mu \) at \( x \) and write \( (D\mu)(x) = \lambda \).
Lemma 3.1. **Covering Lemma:**
Let $\mathcal{C}$ be a nonempty collection of open balls in $\mathbb{R}^k$ and let $U = \bigcup_{B \in \mathcal{C}} B$. If $c < m(U)$, there exists disjoint balls $B_1, \ldots, B_n \in \mathcal{C}$ such that $\sum_{i=1}^n m(B_i) > 3^{-k}c$.

**Proof.** First recall that the Lebesgue measure is regular in the sense that given a measurable set $A$, 
\[ m(A) = \sup\{m(K) : K \text{ is compact, } K \subset A\}. \]

Let $K \subset U$ be a compact set such that $m(K) > c$. Finitely many of the balls in $\mathcal{C}$ cover $K$, say $A_1, \ldots, A_p \in \mathcal{C}$ and $\bigcup_{i=1}^p A_i \supseteq K$. We may assume that $\text{radius}(A_1) \geq \cdots \geq \text{radius}(A_p)$.

Let $B_1 = A_1$. Discard all the $A_i$ with $i > 1$ such that $A_i \cap B_1 \neq \emptyset$. The first remaining $A_{i_0}$ is $B_2$. Define $B_1, \ldots, B_n$ in this manner. It is obvious from this construction that all the $B_i$’s are disjoint.

Let $A_j$ be one of the discarded balls. Then $A_j$ meets at least one of the $B_1, \ldots, B_n$. Let $i$ be the smallest integer such that $B_i \cap A_j \neq \emptyset$. We know $\text{radius}(B_i) \geq \text{radius}(A_j)$, so $A_j \subset B_i$ where $B_i$ is concentric with $B_i$ and $\text{radius}(B_i) = 3\text{radius}(B_i)$.

Hence, $K \subset \bigcup_{i=1}^n A_j \subset \bigcup_{j=1}^n B_j$, and 
\[ c \leq \sum_{j=1}^n m(B_j) = 3^k \sum_{j=1}^n m(B_j) \]

Rearranging, we get 
\[ \sum_{i=1}^n m(B_i) = 3^{-k}c. \]

\[ \square \]

**Definition 3.3. Locally Integrable:** If $f : \mathbb{R}^k \to \mathbb{C}$ is measurable, we say $f$ is **locally integrable** if for all bounded measurable sets $E \subset \mathbb{R}^k$, $\int_E |f(x)|dm < \infty$.

Write $L^1_{\text{loc}}(\mathbb{R}^k)$ for the collection of locally integrable functions.

**Definition 3.4.** For $f \in L^1_{\text{loc}}(\mathbb{R}^k)$ and $x \in \mathbb{R}^k$ let 
\[ (A_r f)(x) := \frac{1}{m(B(x,r))} \int_{B(x,r)} f dm. \]

**Lemma 3.2.** If $f \in L^1_{\text{loc}}(\mathbb{R}^k)$ and $x \in \mathbb{R}^k$, then the map $G(x,r) = (A_r f)(x)$ is jointly continuous in $x$ and $r$. (i.e., $G : \mathbb{R}^k \times (0,\infty) \to \mathbb{C}$ is continuous).

**Proof.** Let $S(x,r)$ be the sphere of radius $r$ centered at $x$ (the edge of the ball $B(x,r)$). Now, $m(B(x,r)) = cm(B(0,1))$ for some $c$ and $m(S(x,r)) = 0$. Fix $r_0 \in (0,\infty)$ and $x_0 \in \mathbb{R}^k$. As $r \to r_0$ and $x \to x_0$, 
\[ \chi_{B(x,r)} \to \chi_{B(x_0,r_0)} \]

pointwise on $\mathbb{R}^k \setminus S(x_0,r_0)$. So, 
\[ \chi_{B(x,r)} \to \chi_{B(x_0,r_0)} \] almost everywhere 
and 
\[ |\chi_{B(x,r)}| \leq \chi_{B(x_0,r_0+1)} \]
if \( r < r_0 + \frac{1}{2} \) and \( |x - x_0| < \frac{1}{2} \). If \( r_n \to r_0 \) and \( x_n \to x_0 \), the Dominated Convergence Theorem says

\[
\int_{B(x_n, r_n)} f \, dm \to \int_{B(x_0, r_0)} f \, dm
\]

Hence the map

\[
(x, r) \to \int_{B(x, r)} f \, dm
\]

is continuous. Then so is

\[
(A_r f)(x) = \frac{1}{c(x, r)m(B(0, 1))} \int_{B(x, r)} f \, dm
\]

Therefore, \((A_r f)(x) = G(x, r)\) is continuous.

\[\square\]

**Definition 3.5. Hardy-Littlewood maximal function:**
For \( f \in L_{loc}(\mathbb{R}^k) \), the Hardy-Littlewood maximal function for \( f \) is

\[
(Hf)(x) = \sup_{r > 0} (A_r |f|)(x)
\]

Claim: \( Hf \) is measurable.

By the above lemma, for \( r > 0 \),

\[
(A_r |f|)^{-1}(a, \infty)
\]

is an open set for \( a \in \mathbb{R}^k \), as \( A \) is continuous. Therefore,

\[
\bigcup_{r > 0} (A_r |f|)^{-1}(a, \infty)
\]

is open (an arbitrary union of open sets is open). But,

\[
(Hf)^{-1}(a, \infty) = \bigcup_{r > 0} (A_r |f|)^{-1}(a, \infty)
\]

If \( x \in \bigcup_{r > 0} (A_r |f|)^{-1}(a, \infty) \), then for some \( r > 0, \)

\[
(A_r |f|)(x) \in (a, \infty)
\]

So, \( \sup_{r > 0} (A_r |f|)(x) > a \). Thus, \( x \in (Hf)^{-1}(a, \infty) \). If \( (Hf)(x) \in (a, \infty) \), then for some \( r > 0, A_r |f|(x) > a \), so \( x \in \bigcup_{r > 0} (A_r |f|)^{-1}(a, \infty) \). We conclude \((Hf)^{-1}(a, \infty)\) is open for all \( a \in \mathbb{R}^k \), and \( Hf \) is measurable.

**Theorem 6. Hardy-Littlewood Maximal Theorem:**
There exists a constant \( c > 0 \) such that for all \( f \in L^1_{loc}(\mathbb{R}^k) \) and all \( \alpha > 0 \),

\[
m(\{x : (Hf)(x) > \alpha\}) \leq \frac{C}{\alpha} \int_{\mathbb{R}^k} |f| \, dm
\]

In fact, \( C \leq 3^k \).

Proof. Given \( \alpha > 0 \) and \( f \in L^1_{loc} \), let \( E_\alpha = \{x \in \mathbb{R}^k : (Hf)(x) > \alpha\} \). For each \( x \in E_\alpha \), let \( r_x > 0 \) such that \( A_{r_x} |f|(x) > \alpha \). Then,

\[
\bigcup_{x \in E_\alpha} B(x, r_x) \supset E_\alpha
\]
so by the covering lemma, if \( c < m(E_\alpha) \), then there exists \( x_1, \ldots, x_n \in E_\alpha \) such that \( B_j := B(x_j, r_{x_j}) \) are a disjoint family and

\[
\sum_{j=1}^{n} m(B_j) > 3^{-k}c
\]

So,

\[
c < 3^k \sum_{j=1}^{n} m(B_j) \leq 3^k \sum_{j=1}^{n} m(B_n) \frac{A_{x_j}|f|(x_j)}{\alpha} \leq \frac{3^k}{\alpha} \sum_{j=1}^{n} \int_{B_j} |f(y)| dy
\]

\[
\leq \frac{3^k}{\alpha} \int_{\mathbb{R}^k} |f(y)| dy.
\]

By taking supremaums over \( c \), we get

\[
m(E_\alpha) \leq \frac{3^k}{\alpha} \int_{\mathbb{R}^k} |f| dm.
\]

\[\square\]

Recall: If \( f : \mathbb{R}^k \to \mathbb{R} \), then

\[
\limsup_{y \to x} f(y) = \inf_{\delta > 0} \sup_{0 < |x-y| < \delta} f(y)
\]

Fact: \( \lim_{y \to x} f(y) = c \equiv \limsup_{y \to x} |f(y) - c| = 0. \)

**Theorem 7.** For \( f \in L^1_{\text{loc}}(\mathbb{R}^k) \),

\[
\lim_{r \to 0} (A_r f)(x) = f(x)
\]

for almost every \( x \in \mathbb{R}^k \).

**Proof.** For any \( N \in \mathbb{N} \), we’ll show that

\[
\lim_{r \to 0} (A_r f)(x) = f(x)
\]

for almost every \( x \in \mathbb{R}^k \) with \( |x| \in N \).

Note that for \( |x| < N \) and \( r \leq 1 \), the values of \( (A_r f)(x) \) only depend on the values of \( f \) in \( B(0, N + 1) \). So, we can replace \( f \) with \( f \chi_{B(0, N+1)} \) as necessary. Hence, we may assume that \( f \in L^1(\mathbb{R}^k) \).

So, WLOG, take \( f \in L^1(\mathbb{R}^k) \). Let \( \epsilon > 0 \). Then, there exists a continuous function \( g \in L^1(\mathbb{R}^k) \) such that

\[
\int_{\mathbb{R}^k} |f(y) - g(y)| dm(y) < \epsilon
\]

Then,

\[
(A_r g)(x) - g(x) = \frac{1}{m(B(x, r))} \int_{B(x, r)} g(y) - g(x) dm(y)
\]

Since \( g \) is continuous, given \( \delta > 0 \), there exists some \( r > 0 \) such that

\[
|g(y) - g(x)| < \delta
\]

if \( y \in B(x, r) \). So,

\[
|(A_r g)(x) - g(x)| < \delta
\]

Thus,

\[
\lim_{r \to 0} (A_r g)(x) = g(x)
\]
We have
\[ \limsup_{r \to 0} |(A_r f)(x) - f(x)| = \limsup_{r \to 0} |(A_r (f - g))(x) + ((A_r g)(x) - g(x)) + (g(x) - f(x))| \]
\[ \leq \limsup_{r \to 0} |A_r (f - g)(x)| + \limsup_{r \to 0} |(A_r g)(x) - g(x)| + \limsup_{r \to 0} |g(x) - f(x)| \]
\[ \leq (H(f - g))(x) + 0 + \limsup_{r \to 0} |g(x) - f(x)| \]

Thus, we have the following inequality
\[ (1) \quad \limsup_{r \to 0} |(A_r f)(x) - f(x)| \leq H(f - g)(x) + \limsup_{r \to 0} |g(x) - f(x)| \]

For \( \alpha > 0 \), let
\[ E_\alpha := \{ x : \limsup_{r \to 0} |(A_r f)(x) - f(x)| > \alpha \} \]
\[ F_\alpha := \{ x : |f(x) - g(x)| > \alpha \} \]

Then, we claim \( E_\alpha = F_{\frac{\alpha}{2}} \cup \{ x : H(f - g)(x) > \frac{\alpha}{2} \} \). This follows directly from the above inequality (1). So,
\[ \epsilon > \int_{\mathbb{R}^k} |f - g| dm \geq \int_{F_{\frac{\alpha}{2}}} |f - g| dm \geq \frac{\alpha}{2} m(F_{\frac{\alpha}{2}}) \]

By the Hardy-Littlewood Maximal Theorem,
\[ m(\{ x : H(f - g)(x) > \frac{\alpha}{2} \}) < \frac{2C}{\alpha} \int_{\mathbb{R}^k} |f - g| dm \]

So,
\[ m(E_\alpha) < \frac{2\epsilon}{\alpha} + \frac{2C}{\alpha} \int_{\mathbb{R}^k} |f - g| dm = \frac{2\epsilon}{\alpha} + \frac{2C\epsilon}{\alpha} \]

As \( \epsilon \) is arbitrary, \( m(E_\alpha) = 0 \). This is true for all \( \alpha \), so
\[ m(\bigcup_{n=1}^{\infty} E_{\frac{\alpha}{n}}) = 0 \]

Hence, if \( x \notin \bigcup_{n=1}^{\infty} E_{\frac{\alpha}{n}} \), then
\[ \limsup_{r \to 0} |(A_r f)(x) - f(x)| = 0 \]

Hence,
\[ \lim_{r \to 0} (A_r f)(x) = f(x) \]

for almost all \( x \in \mathbb{R}^k \). \( \square \)

**Remark:** The conclusion of the theorem is equivalent to
\[ \lim_{r \to 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} f(y) - f(x) dy = 0. \]

In fact, we can do better than this; the conclusion is still valid if we use absolute values under the integrand.
**Definition 3.6. Lebesgue Set:**
Given \( f \in L_{1\text{loc}}(\mathbb{R}^k) \), the Lebesgue set for \( f \) is the set
\[
L_f := \{ x \in \mathbb{R}^k : \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| \, dy = 0 \}
\]

Note that for \( x \in L_f \), \( \lim_{r \to 0} (A_r f)(x) = f(x) \).

**Theorem 8.** For \( f \in L_{1\text{loc}}(\mathbb{R}^k) \), \( m(\mathbb{R}^k \setminus L_f) = 0 \).

**Proof.** For \( \lambda \in \mathbb{C} \), put \( g_\lambda(x) = |f(x) - \lambda| \). Notice that \( g_\lambda \in L_{1\text{loc}} \), so the previous theorem applies to \( g_\lambda \), i.e., the set
\[
E_\lambda := \{ x \in \mathbb{R}^k : \lim_{r \to 0} (A_r g_\lambda)(x) \neq g_\lambda(x) \}
\]
has measure 0. Let \( \{ \lambda_i \}_{i=1}^\infty \) be a countable dense set in \( \mathbb{C} \) and put
\[
E = \bigcup_{i=1}^\infty E_{\lambda_i}
\]
It is clear that \( m(E) = 0 \).

**Friday, 1-28-2005**

Let \( \epsilon > 0 \), and \( x \notin E \). Pick \( i \in \mathbb{N} \) such that \( |\lambda_i - f(x)| < \epsilon \). For any \( y \in \mathbb{R}^k \),
\[
|f(y) - f(x)| < |f(y) - \lambda_i| + \epsilon.
\]
Then,
\[
\limsup_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y) - f(x)| \, dy < \limsup_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y) - \lambda_i| \, dy + \epsilon
\]
As \( x \notin E \),
\[
(A_r g_{\lambda_i})(x) \to g_{\lambda_i}(x)
\]
I.e.,
\[
\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y) - \lambda_i| \, dy = |f(x) - \lambda_i|
\]
Therefore,
\[
\limsup_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y) - f(x)| \, dy < \epsilon + \epsilon
\]
Thus, for \( x \in E \),
\[
\lim_{r \to 0} \frac{1}{B_r(x)} \int_{B_r(x)} |f(y) - f(x)| \, dy = 0
\]
Hence, \( x \notin E \) implies \( x \notin L_f \). Turning this expression around,
\[
m(\mathbb{R}^k \setminus L_f) \leq m(E)
\]

**Theorem 9. Lebesgue Differentiation Theorem:**
Suppose \( f \in L_{1\text{loc}} \). Then for every \( x \in L_f \), we have
\[
\lim_{r \to 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)| \, dy = 0
\]
for every family \( \{ E_r \}_{r \geq 0} \) which shrinks nicely to \( x \).
Proof. For some $\alpha > 0$, we know that $E_r \subset B_r(x)$ and $m(E_r) > \alpha m(B_r(x))$. So,
$$\frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|dy \leq \frac{1}{m(E_r)} \int_{B_r(x)} |f(y) - f(x)|dy$$
$$\leq \frac{1}{\alpha m(B_r(x))} \int_{B_r(x)} |f(y) - f(x)|dy.$$ 

Now, apply the previous theorem. So, if $x \in L_f$, the limit is 0.

\[\square\]

**Theorem 10.** Suppose $\mu$ is a complex Borel measure on $\mathbb{R}^k$ and $\mu \ll m$. Then, $(D\mu)(x)$ exists for almost all $x \in \mathbb{R}^k$ and if $f = \left[ \frac{d\mu}{dm} \right]$, then $(D\mu)(x) = f(x)$ almost everywhere.

**Proof.** We know that $f \in L^1_{loc}(m)$, so by the Lebesgue Differentiation Theorem, if $x \in L_f$,
$$\lim_{r \to 0} \frac{1}{m(E_r)} \int_{E_r} f(y)dm(y) = f(x)$$

i.e.,
$$\frac{\mu(E_r)}{m(E_r)} \to f(x) \text{ for almost all } x$$

\[\square\]

**Pseudo-Example**

Suppose that for $f \in L^1(\mathbb{R})$, we can write
$$F(x) = \int_{-\infty}^{x} f(t)dt.$$ 
By the theorem, $F'(x) = f(x)$ for almost all $x$. Note that we have no need for continuity of $f$. Also, $F$ is continuous almost everywhere.

For $f \in L^1_{loc}$, define
$$F(x) = \int_{0}^{x} f(t)dt$$
and we can get
$$F'(x) = f(x), \text{ almost everywhere}$$

From a previous homework, we’ve seen the function
$$f(x) = \int_{1}^{\infty} \frac{g(x - r_i)}{2r_i}$$
where $\{r_i\}$ is a countable dense subset of the real line. Then, we get that the integral of $f$ is continuous almost everywhere.

Suppose we have
$$d\nu = d\lambda + fdm$$
where $\lambda \perp m$ (we have used the Lebesgue Decomposition) - how would we start with a general measure and find these? We’ll show that $(D\mu)$ exists almost everywhere and is equal to $f$. 


Definition 3.7. A positive Borel measure $\mu$ on $\mathbb{R}^k$ is regular if

(a) For all compact sets $K \subset \mathbb{R}^k$, $m(K) < \infty$.

(b) For all $E \subset \mathbb{R}^k$, $E$ Borel,

$$\mu(E) = \inf \{ \mu(U) : U \text{ is open }, U \supset E \}.$$ 

A complex or signed measure is regular if its total variation is.

Observation: If $f \geq 0$ is measurable, then $d\mu = fdm$ is regular if and only if $f \in L^{1}_{\text{loc}}(\mathbb{R}^k)$.

Proof. If $\mu$ is regular, then for any bounded set $E$,

$$\int_E fdm \leq \int_{\bar{E}} fdm = \mu(\bar{E}) < \infty$$

Thus, $f \in L^{1}_{\text{loc}}(\mathbb{R}^k)$.

If $f \in L^{1}_{\text{loc}}$, then (a) in the definition of regular is obvious. We leave part (b) as an exercise.

Monday, 1-31-2005

Let $E \subset \mathbb{R}$ be Borel and bounded. Let $\epsilon > 0$, let $B \subset R$ be a bounded open set. Since $f \in L^{1}_{\text{loc}}$, there exists a $\delta > 0$ such that whenever $F \subset B$ and $m(F) < \delta$ then

$$\int_F fdm < \epsilon.$$ 

Find an open set $U$ such that $U \supset E$ and $m(U \setminus E) < \delta$. We can do this because last semester we showed that the Lebesgue measure was regular. Then,

$$m(B \cap U) = \int_{U \cap B} fdm = \int_E fdm + \int_{(U \cap B) \setminus E} fdm$$

$$< \mu(E) + \epsilon$$

Therefore, (b) holds for $E$ bounded.

If $E$ isn’t bounded ($E$ Borel), cover it with bounded sets, and approximate each of these bounded sets by $2^{-n}\epsilon$ (use standard techniques).

Example: Suppose $\mu$ is a regular (signed / complex) Borel measure on a $\sigma$-finite measure space. (Note that $\sigma$-finite is also implied by the regularity conditions.)

Let

$$d\mu = d\lambda + fdm$$

be the Lebesgue Decomposition of $\mu$ with respect to $m$, where $\lambda \perp m$.

Claim: $d\lambda$ and $fdm$ are regular.

To see this, we first observe

$$d|\mu| = d|\lambda| + |f|dm$$
Proof. \( \lambda \perp m \), so \(|\lambda| \perp m\). Find Borel sets \( A, B \subset \mathbb{R}^k \) such that \( A \cup B = \mathbb{R}^k \), \(|\lambda|(B) = m(A) = 0\), and \( A \cap B = \emptyset \). Let \( g \) be such that \( d\lambda = gd|\lambda| \), where \(|g| = 1\).

For \( E \subset \mathbb{R}^k \), \( \mu \)-Borel,

\[
\mu(E) = \int_E gd|\lambda| + \int f dm = \int_E g\chi_A |\lambda| + \int_E f\chi_B dm = \int_E (g\chi_A + f\chi_B)|\lambda| + \int_E (g\chi_A + f\chi_B) dm = \int_E (g\chi_A + f\chi_B) (|\lambda| + m) = \mu(E)
\]

Also, \( \mu \ll (|\lambda| + m) \). Hence,

\[
|\mu|(E) = \int_E |g\chi_A + f\chi_B| d(|\lambda| + m)
\]

As \( A \cap B = \emptyset \),

\[
= \chi_E\chi_A + |f|\chi_B d(|\lambda| + m)
\]

so, \( d|\mu| = d|\lambda| + |f| dm \).

Since \( |\mu| \) is regular, \( \mu \) is regular. So, if \( K \) is compact, then

\[
\int_K |f| dm \leq |\mu|(K) < \infty
\]

So, \( f \in L^1_{loc} \). Hence, \( f dm \) is regular by our previous work.

To see \( |\lambda| \) is regular, let \( E \) be Borel and find \( U \) such that \( U \supset E \) is open and \( |\mu|(U \setminus E) < \epsilon \). Then, \( |\lambda|(U \setminus E) < \epsilon \), so

\[
|\lambda|(U) = |\lambda|(E) + |\lambda|(U \setminus E) < |\lambda|(E) + \epsilon
\]

\( \square \)

Theorem 11. Let \( \lambda \) be a regular complex or signed Borel measure such that \( \lambda \perp m \). Then for \( m \)-almost all \( x \in \mathbb{R}^k \),

\[
\lim_{r \to 0} \frac{\lambda(E_r)}{m(E_r)} = 0
\]

for every family \( \{E_r\}_{r \geq 0} \) of Borel sets which shrink nicely to \( x \).

Proof. Without loss of generality, we may assume that \( \lambda \) is a positive regular Borel measure and that \( E_r = B(r, x) \).

Let \( \lambda \perp m \). Then, there exists a Borel set \( A \subset \mathbb{R}^k \) such that \( \lambda(A) = m(A^c) = 0 \). Put

\[
F_n = \{ x \in A : \limsup_{r \to 0} \frac{\lambda(B(r, x))}{m(B(r, x))} > \frac{1}{n} \}
\]

We’ll prove that \( m(F_n) = 0 \) for each \( n \). Hence the theorem follows: for if

\[
x \notin \bigcup_{n=1}^{\infty} F_n \cup A^c
\]
then \( \limsup_{r \to 0} \frac{\lambda(B(r, x))}{m(B(r, x))} = 0 \), as desired.

Fix \( n \in \mathbb{N} \). By the regularity of \( \lambda \), given \( \epsilon > 0 \), there exists an open set \( U_\epsilon \) such that \( U_\epsilon \supset A \) and \( \lambda(U_\epsilon) < \epsilon \). By the definition of \( F_n \), given \( x \in F_n \), there exists a \( r_x > 0 \) such that \( B(r_x, x) \subset U_\epsilon \) and

\[
\frac{\lambda(B(r, x))}{m(B(r, x))} > \frac{1}{n}
\]

Let \( V_\epsilon = \bigcup_{x \in F_n} B(r_x, x) \). Let \( 0 < c < m(V_\epsilon) \). By the covering lemma, there exists \( x_1, \ldots, x_p \in F_n \) such that \( B(r_{x_1}, x_1), \ldots, B(r_{x_n}, x_n) \) are disjoint and

\[
c < 3^k \sum_{j=1}^{p} m(B(r_{x_j}, x_j)) < 3^k n \sum_{j=1}^{p} \lambda(B(r_{x_j}, x_j)).
\]

Because the balls are disjoint and \( \lambda \) is a measure,

\[
= 3^k n \lambda \left( \bigcup_{j=1}^{p} B(r_{x_j}, x_j) \right) \leq 3^k n \lambda(V_\epsilon)
\]

By our clever construction, we get

\[
\leq 3^k n \lambda(U_\epsilon) < 3^k n \epsilon.
\]

Hence,

\[
m(V_\epsilon) \leq 3^k n \epsilon.
\]

\[
m(F_n) \leq 3^k n \epsilon
\]

As \( \epsilon \) is arbitrary, we get \( m(F_n) = 0 \).

Combining this result with the Lebesgue Differentiation Theorem, we immediately see the following:

**Theorem 12.** Let \( \mu \) be a regular signed or complex Borel measure on \( \mathbb{R}^k \), and let \( d\mu = d\lambda + fdm \) be its Lebesgue decomposition with respect to \( m \). Then, for \( m \)-almost every \( x \in \mathbb{R}^k \),

\[
\lim_{r \to 0} \frac{\mu(E_r)}{m(E_r)} = f(x) \quad \text{for every family } \{E_r\}_{r \geq 0} \text{ shrinking nicely to } x.
\]

**Example:**

Let \( F(x) \) be the Cantor function \( F : [0, 1] \to [0, 1] \). We can extend \( F \) to a function \( F : \mathbb{R} \to [0, 1] \).

Define \( \mu([a, b]) = F(b) - F(a) \). Then this determines a unique Borel measure on \( \mathbb{R} \) such that \( F(x) = \int_{0}^{x} d\mu \). Since \( F'(x) = 0 \) for almost every \( x \), \( \mu \perp m \).

**Application:**

**Theorem 13.** Let \( F : \mathbb{R} \to \mathbb{R} \) be non-decreasing and let

\[
G(x) = F(x^+)
\]

Then, \( F \) is continuous, except at countably many points. Further, for almost all \( x \in \mathbb{R} \), \( F \) and \( G \) are differentiable and their derivatives are equal almost everywhere.
Proof. Since $F$ is nondecreasing, for all $x \in \mathbb{R}$

$$F(x^-) \leq F(x) \leq F(x^+)$$

Hence the open intervals, $(F(x^-), F(x^+))$ are disjoint and if $|x| < N \in \mathbb{N}$, we have $(F(x^-), F(x^+)) \subset (F(-N), F(N))$. So, the sum

$$\sum_{|x|<N} (F(x^+)-F(x^-)) \leq F(N) - F(-N).$$

Hence except for at most countably many $x$, $F(x^+) = F(x^-)$; i.e., $F$ is continuous on $(-N, N)$ except at countably many points. This gives the first statement.

We have $G(x) = F(x)$ whenever $F$ is continuous at $x$. Observe that there exists a regular Borel measure $\mu_G$ such that

$$\mu_G((a,b]) = G(b) - G(a).$$

Intervals such as $(x-r,x]$ and $(x,x+r]$ shrink nicely to $x$. Since $\mu_G$ is regular, $D\mu_G$ exists for almost every $x \in \mathbb{R}$. Hence

$$\lim_{h \to 0^+} \frac{\mu_G(x,x+h]}{h}$$

and

$$\lim_{h \to 0^-} \frac{\mu_G(x+h,x]}{h}$$

exist for almost every $x$. Hence, $G$ is differentiable for almost all $x$. Next, we need to prove that $F'(x) = G'(x)$ almost everywhere. Define

$$H(x) = G(x) - F(x);$$

from the definition of $G$, $H$ is nonnegative. We want to show $H$ is diff. almost everywhere and $H'(x) = 0$ almost everywhere. We have

$$F(x^-) \leq F(x) \leq F(x^+) = G(x)$$

Therefore,

$$H(x) \leq F(x^+) - F(x^-).$$

We get, for $N \in \mathbb{N}$,

$$\sum_{|x|<N} H(x) \leq \sum_{|x|<N} F(x^+)-F(x^-) < F(N) - F(-N) < \infty$$

So, $H(x)$ has to be zero except on countably may points. Let $\{x_j\}_{j=1}^\infty$ be an enumeration of $B = \{x : H(x) \neq 0\}$.

Let $\delta_{x_j}$ be the point mass at $x_j$ and put

$$\mu = \sum_{j=1}^\infty H(x_j)\delta_{x_j},$$

Then, $\mu(K) < \infty$ if $K \subset \mathbb{R}$ is compact (by the previous argument, as $K$ has to be bounded). Also if $E$ is Borel, it isn’t hard to prove that $\mu(E) = \inf\{\mu(U) : U \text{ open}\}$. Hence, $\mu$ is regular.

Note that

$$\mu(B^c) = 0 = m(B)$$
and \( \mu \perp m \). Hence, \((D\mu)(x) = 0\) for almost all \( x \). Note,

\[
\frac{|H(x + h) - H(x)|}{h} \leq \frac{H(x + h) + H(x)}{|h|} \leq \frac{\mu(x - 2h, x + 2h)}{|h|} = 4\frac{\mu(x - 2h, x + 2h)}{4|h|}
\]

Hence, since \((D\mu)(x) = 0\) for almost all \( x \), we get

\[
H'(x) = \lim_{h \to 0} 4\frac{\mu(x - 2h, x + 2h)}{4|h|} = 0
\]

for almost all \( x \).

\[\square\]

**Problem:** Which functions are of the form

\[F(x) = \int_{-\infty}^{x} d\mu\]

for a Borel complex measure?

We already know this answer, in part. For every increasing real valued function that is continuous on the right, we can get a Baire measure (look at last semester’s notes).

If \( F \) is continuously differentiable (and was nice enough) and

\[F(x) = \int_{-\infty}^{x} F'(t)dt,\]

we’d expect

\[\int_{-\infty}^{x} |F'(t)|dt < \infty\]

for all \( x \).

Given a function \( F : \mathbb{R} \to \mathbb{C} \), we define the total variation of \( F \) over \( \mathbb{R} \) by

\[
\sup\left\{ \sum_{j=0}^{n+1} |F(x_{j+1}) - F(x_j)| : x_0 \leq x_1 \leq \cdots \leq n \text{ is a finite partition of } \mathbb{R} \right\}
\]

More generally, if \([a, b]\) is an interval, then

\[T_a^b(f) = \sup\left\{ \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| : \{x_j\} \text{ partition } [a, b] \right\}\]

is the total variation of \( F \) from \( a \) to \( b \).

If \( a = -\infty \), write \( V_f(x) = T_{-\infty}^x(f) \). \( V_f(x) \) is the total variation function of \( f \).

**Definition 3.8. Bounded Variation:**

If \( V_f(+\infty) < \infty \), then we say \( f \) is of bounded variation on \( \mathbb{R} \). If \( T_a^b(f) < \infty \), we say \( f \) is of bounded variation on \([a, b]\).
Definition 3.9. $BV[a, b]$: 
$BV[a, b]$ is the set of all functions of bounded variation on $[a, b]$.

Some Examples:
- A bounded increasing function has finite total variation.
- The set of all functions of bounded variation is a linear space.
- A function with bounded derivative is of bounded variation on a finite interval.
- $\sin x$ is an example of a function with bounded variation on a finite interval, but not of bounded variation on the real line.
- 
  $$f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0 \\ x = 0, & x = 0. \end{cases}$$

If $0 \notin [a, b]$, then $f \in BV[a, b]$. If $0 \in [a, b]$, $f \notin BV[a, b]$.

Remark:
Notice that if $f$ has bounded variation, then (for $x > y$)

$$V_f(x) - V_f(y) = T_y^x(f).$$

Proposition 3.1. Suppose $F : \mathbb{R} \to \mathbb{R}$ and $F \in BV$. Then $V_F + F$ and $V_F - F$ are both increasing functions. Thus in particular,

$$F = \frac{(V_F + F)}{2} - \frac{(V_F - F)}{2}.$$

In other words, all functions of bounded variation is the difference of two increasing functions.

Remark: The proposition is valid if we use $F : [a, b] \to \mathbb{R}$ instead of $F : \mathbb{R} \to \mathbb{R}$.

Proof. Let $x < y$ and $\epsilon > 0$. Find a partition $\{x_0, \ldots, x_n\}$ such that

$$\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| \geq V_F(x) - \epsilon$$

Then,

$$\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| + |F(y) - F(x)|$$

approximates $V_F(y)$ and $F(y) = (F(y) - F(x)) + F(x)$.

So,

$$V_F(y) + F(y) \geq \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| + |F(y) - F(x)| + [F(y) - F(x)] + F(x).$$

$$\geq V_F(x) - \epsilon + F(x)$$

So,

$$V_F(y) + F(y) \geq V_F(x) + F(x),$$
and $V_F + F$ is an increasing function. Changing the appropriate plus signs to minus signs, we get that $V_F - F$ is also increasing. □

Combining this with our work on increasing functions leads to this theorem:

**Theorem 14.** The following are true:

1. $F \in BV \iff Re(F) \in BV$ and $Im(F) \in BV$
2. If $F : \mathbb{R} \to \mathbb{R}$, then $F \in BV \iff F$ is a difference of two bounded increasing functions.
3. If $F \in BV$, then the following limits exist for every $x \in \mathbb{R}$:
   - $\lim_{y \to x^+} F(y)$
   - $\lim_{y \to x^-} F(y)$
   - $\lim_{y \to \pm \infty} F(y)$
4. If $F \in BV$, the set of discontinuities of $F$ is countable.
5. For $F \in BV$, let $G(x) = F(x^+)$. Then $F', G'$ exists almost everywhere and $F' = G'$ almost everywhere.

**Proof.**

1. Use

   $Re(F) = \frac{F + \bar{F}}{2}, \quad Im(F) = \frac{F - \bar{F}}{2i}$

2. We’ve done $\implies$, and need to go the other way. However, this is obvious from the definition of bounded variation and the triangle inequality.

3. - (5) have been proved previously □

**Definition 3.10.** If $F : \mathbb{R} \to \mathbb{R}$ is of bounded variation, we call the decomposition

   $F = \frac{V_F + F}{2} - \frac{V_F - F}{2}$

the Jordan decomposition of $F$.

**Definition 3.11.** A function $F \in NBV$ (normalized bounded variation) if $F \in BV$, $F(-\infty) = 0$, and $F$ is right continuous.

We can always make a function of $BV$ into a function of $NBV$. If $F \in BV$, then $G(x) := F(x^+) - F(-\infty) \in NBV$. Moreover, $F'(x) = G'(x)$ for almost all $x$.

**Lemma 3.3.** If $F \in BV$, then $V_F(-\infty) = 0$ and if $F$ is right continuous, then so is $V_F$. 
Proof. Wednesday, 2-9-2005

Let $\epsilon > 0$, $x \in \mathbb{R}$. Let $x_0 < x_1 < \cdots < x_n = x$ satisfy

$$\sum_{i=1}^{n} |F(x_j) - F(x_{j-1})| \geq V_F(x) - \epsilon$$

We have then,

$$T_{x_0}^x(F) = V_F(x) - V_F(x_0) = \sup \left\{ \sum_{i=1}^{n} |F(y_i) - F(y_{i-1})| : x_0 = y_0 < \cdots < y_n = x \right\}$$

Therefore,

$$V_F(x) - V_F(x_0) \geq \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| \geq V_F(x) - \epsilon$$

and $V_F(x-0) < \epsilon$. Hence, if $y < x_0$, we get $V_F(y) < \epsilon$, so $V_F(-\infty) = 0$.

Suppose that $F$ is right continuous. Fix $x \in \mathbb{R}$, $\epsilon > 0$, and put $\alpha = V_F(x^+) - V_F(x)$. Find $\delta > 0$ such that $|F(x + h) - F(x)| < \epsilon$ and $V_F(x + h) - V_F(x^+) < \epsilon$ if $0 < h < \delta$.

For $0 < h < \delta$, find a partition $x = x_0 < \cdots < x_n = x + h$ such that

$$\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| \geq \frac{3}{4}(V_F(x + h) - V_F(x)).$$

$$\geq \frac{3}{4}\alpha, \text{ as } V_F(x + h) > V_F(x^+)$$

Now, partition the interval $x_0 = t_0 < \cdots < t_m = x_1$. such that

$$\sum_{j=1}^{m} |F(t_j) - F(t_{j-1})| \geq \frac{3}{4}\alpha.$$

We have

$$\alpha + \epsilon > V_F(x + h) - V_F(x) \geq \sum_{j=1}^{m} |F(t_j) - F(t_{j-1})| + \sum_{i=2}^{n} |F(x_i) - F(x_{i-1})|$$

$$\geq \frac{3}{4}\alpha + \frac{3}{4}\alpha - |F(x_1) - F(x)| \geq \frac{3}{2} - \epsilon$$

Manipulating this, we get

$$\alpha < 4\epsilon$$

for all $\epsilon$ - and hence, $\alpha = 0$. \hfill \Box

**Theorem 15.** If $\mu$ is a complex, regular Borel measure on $\mathbb{R}$ and if $F(x) = \mu(-\infty, x]$, then $F \in \text{NBV}$. Conversely, if $F \in \text{NBV}$, then there exists a unique regular complex Borel measure $\mu_F$ such that $F(x) = \mu(-\infty, x]$. 
Proof. If $\mu$ is a regular Borel complex measure, write

$$\mu = \mu_1^+ - \mu_1^- + i(\mu_2^+ - \mu_2^-)$$

where $\mu_1 = \text{Re}\mu$, $\mu_2 = \text{Im}\mu$, and $\mu_i^\pm$ are positive (finite) measures. For $j = 1, 2$, let $F_j^\pm(x) = \mu_j^\pm(-\infty, x]$. Then $F_j^\pm$ are increasing functions and the continuity theorem shows they are continuous on the right and $F_j^\pm(-\infty) = 0$. We conclude that $F_j^\pm \in \text{NBV}$. Hence, $F = F_1^+ - F_1^- + i(F_2^+ - F_2^-) \in \text{NBV}$.

Conversely, write

$$F = \text{Re}F + i\text{Im}F$$

Since $\text{Re}F$ and $\text{Im}F$ are real-valued functions of bounded variation, we can write the Jordan Decomposition:

$$\begin{align*}
V_{F_1} + F_1 &= \frac{V_{F_1} - F_1}{2} + \frac{V_{F_2} + F_2}{2} + i\left(\frac{V_{F_2} + F_2}{2} - \frac{V_{F_2} - F_2}{2}\right)
\end{align*}$$

By the Lemma,

$$\begin{align*}
V_{F_1} \pm F_1, \frac{V_{F_2} \pm F_2}{2} \in \text{NBV}.
\end{align*}$$

By last semester’s work, there are unique regular and Borel measures $\mu_i^\pm$ such that

$$\mu_i^\pm(-\infty, x] = \frac{V_{F_i} \pm F_i}{2}(x).$$

So, if $\mu = (\mu_1^+ - \mu_1^-) + i(\mu_2^+ - \mu_2^-)$, we have $\mu$ is regular, Borel, and $\mu(-\infty, x] = F(x)$. 

Remark: It can be shown that for $F \in \text{NBV}$, $|\mu_F| = \mu_{V_F}$, and if $F$ is real, then the Jordan Decomposition of $\mu_F$ corresponds to $\mu_{\frac{1}{2}(V_F \pm F)}$. This naturally leads to the question: Which $F$ satisfy $\mu_F \ll m$?

**Proposition 3.2.** If $F \in \text{NBV}$, then $F' \in L^1(m)$. Moreover, $\mu_F \perp m \Leftrightarrow F' = 0$ (Lebesgue) almost-everywhere. Finally, $\mu_F \ll m$ if and only if

$$F(x) = \int_{(-\infty, x]} F'(t) dt.$$

Proof. Observe that

$$F'(x) = \lim_{r \to 0} \frac{\mu_F(E_r)}{m(E_r)}$$

where $E_r = (x, x+r]$ or $E_r = (x-r, x]$. Write $\mu = \lambda + \nu$ where $\lambda \perp m$ and $\nu \ll m$; then, for almost every $x$,

$$D_{\mu_F}(x) = F'(x) = \left[\frac{d\nu}{dm}\right](x).$$

Hence, $\mu_F \perp m \Leftrightarrow \nu = 0$ almost everywhere $\Leftrightarrow F'(x) = 0$ almost everywhere. On the other hand,

$$\mu_F \ll m \Leftrightarrow F(x) = \mu_F(-\infty, x] = \int_{-\infty}^{x} d\mu_F$$

$$= \int_{(-\infty, x]} F'(x) dm(x).$$
Finally, $F' \in L^1(m)$ because the Radon-Nikodym derivative is, and $F' = \frac{d\nu}{dm}$ almost everywhere.

□

This proposition isn’t very satisfactory because we must compute both the derivative and the integral of the derivative.

**Definition 3.12.** We say that $F : \mathbb{R} \to \mathbb{C}$ is an absolutely continuous function if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for every finite family $(a_1, b_1), \ldots, (a_n, b_n)$ of disjoint intervals with

$$\sum_{j=1}^{n} (b_j - a_j) < \delta$$

we have

$$\sum_{j=1}^{n} |F(b_j) - F(a_j)| < \epsilon.$$

We say that $F$ is absolutely continuous on the interval $[a, b]$ if the same condition holds for all disjoint intervals contained in $[a, b]$.

**Remarks:**

1. If $F$ is absolutely continuous, then $F$ is uniformly continuous.
2. If $F$ is everywhere differentiable and $F'$ is bounded, then $F$ is absolutely continuous.

**Proposition 3.3.** If $F \in NBV$, then $F$ is absolutely continuous if and only if $\mu_F \ll m$.

**Proof.** Suppose $\mu_F \ll m$. Then $F' \in L^1$ and

$$F(x) = \int_{-\infty}^{x} F'(t)dt$$

by the previous proposition. Given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\int_E |F'(t)|dm(t) < \epsilon$$

whenever $m(E) < \delta$. Taking $E$ to be a collection of disjoint intervals, we find that $F$ is absolutely continuous.

Conversely, suppose that $F$ is absolutely continuous. Let $E$ be a Borel set such that $m(E) = 0$. Given $\epsilon > 0$, we may find $\delta > 0$ which satisfies the absolutely continuous condition for $F$.

Let $\{U_j\}_{j=1}^{\infty}$ be a sequence of open sets such that $U_1 \supset U_2 \supset \cdots \supset E$, $m(U_j) < \delta$, and

$$\lim_{j \to \infty} \mu_F(U_j) = \mu_F(E).$$

Every $U_j$ is a disjoint union of open intervals, so write

$$U_j = \bigcup_{k=1}^{\infty} (a_j^k, b_j^k).$$
Then for any \( N \in \mathbb{N} \),
\[
\sum_{k=1}^{N} (b_j^k - a_j^k) < \delta.
\]
Hence,
\[
\sum_{k=1}^{N} |F(b_j^k) - F(a_j^k)| < \epsilon.
\]
In particular,
\[
\sum_{k=1}^{\infty} |F(b_j^k) - F(a_j^k)| \leq \epsilon.
\]
Therefore,
\[
\left| \sum_{k=1}^{\infty} F(b_j^k) - F(a_j^k) \right| = |\mu_F(U_j)| \leq \epsilon.
\]
Taking \( j \to \infty \), we get \( \mu_F(E) = 0 \) so that \( \mu_F \ll m \).
\( \square \)

Remark: As an immediate corollary, we get the following.

**Theorem 16.** If \( f \in L^1(m) \), then \( F(x) := \int_{-\infty}^{x} f(t)dt \) is in NBV and \( F'(x) = f(x) \) for almost all \( x \). Conversely, if \( F \in NBV \) and is absolutely continuous, then \( F' \in L^1(m) \) and
\[
F(x) = \int_{-\infty}^{x} F'(t)dt.
\]

**Lemma 3.4.** If \( F \in AC \) on \([a, b]\), then \( F \in BV[a, b] \).

**Proof.** For you. \( \square \)

Monday, 2-14-2005

**Theorem 17.** If \(-\infty < a < b < \infty \) and \( F : [a, b] \to C \), then the following are equivalent:

(a) \( F \) is absolutely continuous on \([a, b]\).

(b) \( F(x) - F(a) = \int_{a}^{x} f(t)dt \) for some \( f \in L^1([a, b], m) \)

(c) \( F \) is differentiable for \( m \)-almost every \( x \in [a, b] \), \( F' \in L^1([a, b], m) \), and
\[
F(x) - F(a) = \int_{a}^{x} F'(t)dt.
\]

**Proof.** (c) \( \implies \) (b) is trivial

(b) \( \implies \) (a) Take \( f = 0 \) outside of \([a, b]\). Then \( f \in L^1(\mathbb{R}) \), and apply previous results.

(a) \( \implies \) (c). Put \( G(x) = F(x) - F(a) \). Then, \( G(a) = 0 \) and we extend \( G \) to all of \( \mathbb{R} \) by \( G(t) = 0 \) for \( t < a \), \( G(t) = G(b) \) if \( x > b \). Then, \( G \in NBV \) and is absolutely continuous so \( G' \in L^1(m) \) and
\[
G(x) = \int_{-\infty}^{x} G'(t)dm(t)
\]
i.e.,

$$F(x) - F(a) = \int_a^x F'(t)dt.$$ 

\[\square\]

**Final Remarks:**

Let $\mu$ be a Borel measure on $\mathbb{R}^k$. $\mu$ is called discrete if there exists a countable set $\{x_j\}_{j=1}^\infty \subset \mathbb{R}^k$ and $c_j \in \mathbb{C}$ such that

$$(1) \sum_{j=1}^\infty |c_j| < \infty \text{ and }$$

$$(2) \mu = \sum_{j=1}^\infty c_j \delta_{x_j}$$

We say that $\mu$ is continuous if $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}^k$.

**Observe:**

If $\mu$ is a complex Borel measure, then $\mu = \mu_c + \mu_d$ where $\mu_d$ is discrete and $\mu_c$ is continuous. Write $E = \{x : \mu(\{x\}) \neq 0\}$.

If $F \subset E$ is countable, the series

$$\sum_{x \in F} \mu(\{x\})$$

converge absolutely. Therefore,

$$\{x \in E : |\mu(\{x\})| > \frac{1}{k}\}$$

is finite and as

$$E = \bigcup\{x : |\mu(\{x\})| \leq \frac{1}{k}\},$$

$E$ is countable. Now put $\mu_d(A) = \mu(A \cap E)$ and $\mu_c(A) = \mu(A \cap E^c)$. Notice that $\mu_d \perp m$. We can further decompose to get $\mu_c = \mu_{ac} + \mu_{sc}$ where $\mu_{ac} \perp m$ and $\mu_{ac} \ll m$. Thus,

$$\mu = \mu_d + \mu_{sc} + \mu_{ac}$$

This decomposition is important in function theory (branch of complex analysis). For an example of a singular (with respect to $m$) and continuous function, look at the measure generated by the Cantor function.

If $F \in NBV$, write $\mu_F$ for the associated complex measure. Then, we often write $\int gd\mu_F$ as $\int gdF$.

### 4. Baire Category

**Definition 4.1.** Let $X$ be a metric space. A set $E \subset X$ is nowhere dense if $\overline{E}$ (the closure of $E$) has empty interior.

Notice that if $E$ is nowhere dense, then $E^c$ is a dense set.

**Example:** $\mathbb{Q}$ is not nowhere dense, but the (regular) Cantor set is nowhere dense.
Definition 4.2. \( E \subset X \) is of first category if \( E \) is a countable union of nowhere dense sets. A set \( E \subset X \) is of second category if it is not first category.

**Synonyms:** A meager set is a set of first category. A non-meager set is of second category.

**Examples:** \( \mathbb{Q} \) is first category, as is the Cantor set (hence we can have uncountable sets which are first category).

Some key theorems:

**Theorem 18. Baire’s Theorem:**

Let \( X \) be a complete metric space. Suppose for each \( k \in \mathbb{N} \), \( G_k \) is a dense open set. Then, \( \cap_{k=1}^{\infty} G_k \) is dense.

**Proof.** Fix \( x \in X \) and \( \epsilon > 0 \). \( G_1 \) is dense so there exists a \( y_1 \in G_1 \) and \( 0 < r_1 < \frac{1}{2} \epsilon \) such that \( B(y_1, r_1) \subset B(x) \cap G_1 \). Now, \( B(r_1, y_1) \cap G_2 \) is open and nonempty so there exists a \( y_2 \) and \( 0 < r_2 < \frac{1}{2} r_1 \) such that \( B(r_2, y_2) \subset B(r_1, y_1) \cap G_2 \).

Continue this process inductively to obtain the sequence \( \{y_k\}_{k=1}^{\infty} \) in \( X \) such that \( B(r_{k+1}, y_{k+1}) \subset B(r_k, y_k) \cap G_{k+1} \).

**Claim:** \( y_k \) is Cauchy. For \( n, m \in \mathbb{N} \), \( n, m > N \), \( y_n, y_m \in B(r_N, y_N) \) so

\[
d(y_n, y_m) \leq d(y_n, y_N) + d(y_N, y_m) < 2 r_N < 2 \left( \frac{1}{2} \right)^N \epsilon = \left( \frac{1}{2} \right)^{N-1} \epsilon.
\]

Hence, there exists \( y \in X \) such that \( y = \lim_{k \to \infty} y_k \).

For each \( N \in \mathbb{N} \),

\[
y \in B(r_N, y_N) \subset B(r_{N-1}, y_{N-1}) \cap G_N
\]

Therefore, \( y \in G_n \) for all \( N \). Further,

\[
B(r_N, y_N) \subset B(r_1, y_1) \subset B(x)
\]

so \( \bigcap_{N=1}^{\infty} G_N \) is dense in \( X \).

**Observation:** Suppose that, for all \( n \), \( H_n \) is a dense \( G_\delta \)-set in the complete metric space \( X \). Then, \( \bigcap_{n=1}^{\infty} H_n \) is dense in \( X \).

**Proof.** Let

\[
H_n = \bigcap_{m=1}^{\infty} G_{n,m}
\]

where \( G_{n,m} \) is open. Since \( H_n \) is dense, so is \( G_{n,m} \) for all \( n, m \). Thus,

\[
\bigcap_{n=1}^{\infty} H_n = \bigcap_{n,m} G_{n,m}
\]
is dense by Baire’s Theorem.

\[ \square \]

**Theorem 19. Baire Category Theorem:**
Let \( X \) be a complete metric space. Then every open subset of \( X \) is second category in \( X \).

**Proof.** Let \( G \) be a non-empty open set. Let \( E_n \) be a countable collection of nowhere dense sets. Then \( X \setminus \bar{E}_n \) is dense and open, so \( \bigcap_{n=1}^{\infty} (X \setminus \bar{E}_n) \) is dense. Then,

\[
G \cap \left( \bigcap_{n=1}^{\infty} X \setminus \bar{E}_n \right) \neq \emptyset
\]

Hence \( G \) is not a subset of \( \bigcup_{n=1}^{\infty} \bar{E}_n \), so in particular \( G \neq \bigcup_{n=1}^{\infty} E_n \). Thus, \( G \) is of second category.

\[ \square \]

**Examples:**
Let \( C \) be the set of real valued continuous functions on \([0, 1]\). Make \( C \) a metric space by defining

\[
d(f, g) = \sup_{x \in [0, 1]} |f(x) - g(x)|.
\]

Then if \( \epsilon > 0 \) and \( f \in C \), then the sequence \( d(f_n, f) \to 0 \) if and only if \( f_n \to f \) uniformly on \([0, 1]\).

**Recall:** The derivates of \( f \):

\[
(D^+ f)(x_0) = \limsup_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}
\]

\[
(D^- f)(x_0) = \limsup_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}
\]

\[
(D_{\pm} f)(x_0) = \liminf_{x \to x_0^\pm} \frac{f(x) - f(x_0)}{x - x_0}
\]

Not that \( f \) is differentiable at \( x_0 \in (0, 1) \) if and only if the derivates all agree.

**Define:**

\[ \mathcal{F} = \{ f \in C : (D^+ f)(x_0) \text{ is finite for some } x_0 \in [0, 1) \} \]

**Proposition 4.1.** \( \mathcal{F} \) is 1st category in \( C \).

**Proof.** Let

\[ \mathcal{F}_n = \{ f \in C : \exists x_0 \in [0, 1 - \frac{1}{n}] \text{ s.t. } |f(x) - f(x_0)| \leq n(x - x_0) \forall x \in [x_0, 1) \} \]

Observe that \( \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F} \). Actually,

\[
\bigcup_{n=1}^{\infty} \mathcal{F}_n = \mathcal{F}.
\]

To see this, let \( f \in \mathcal{F} \). Then \( |D^+ f(x_0)| < \infty \) for some \( x_0 \in [0, 1) \). Hence,

\[
|D^+ f(x_0)| < N_0 \in \mathbb{N}
\]
Also,
\[ x \in [0, 1 - \frac{1}{N_1}] \]
for some \( N_1 \). Hence, \( f \in F_n \) if \( n \geq \max \{N_0, N_1\} \).

**Claim:** \( F_n \) is nowhere dense in \( C \). Once we show that \( F_n \) is closed, given any \( f \in C \), we can add a bunch of “zig-zags” that all have slope greater than \( n \) so we remain within \( \epsilon \) of \( f \). Hence, there is no nearby points of \( F_n \), i.e.,

If \( F_n \) is not nowhere dense, then \( F_n = F_n \) contains an open ball, say \( B(\epsilon, f) \subset F_n \).

We now show that \( F_n \) is closed. Suppose \( f_k \in F_n \) and \( f_k \to f \in C \). Then there exists \( x_k \in [0, 1 - \frac{1}{N}] \) such that
\[ |f_k(x) - f_k(x_k)| \leq n(x - x_k), \forall \in [x_k1] \]

By passing to a subsequence if necessary, we may assume that \( x_k \) converges; put \( x_0 = \lim_{k \to \infty} x_k \). Then, \( x_0 \in [0, 1 - \frac{1}{N}] \). For any \( x \in [x_0, 1] \),
\[ |f(x) - f(x_0)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(x_k)| + |f_k(x_k) - f(x_k)| + |f(x_k) - f(x_0)| \]

Given \( x \in [x_0, 1] \), if \( x = x_0 \) then \( |f(x) - f(x_0)| \leq n(x - x_0) \) is trivial. Assume \( x_0 < x < 1 \). Let \( \epsilon > 0 \). Since \( x_k \to x_0 \) and \( f \) is continuous, find \( N \in \mathbb{N} \) such that if \( k \geq N \),
\[ |f(x_k) - f(x_0)| < \frac{\epsilon}{4} \]
and
\[ |x_k - x_0| < \frac{\epsilon}{4n} \]
and \( x_k < x \). Choose \( k > N \) such that \( |f_k(t) - f(t)| < \frac{\epsilon}{4} \) for all \( t \in [0, 1] \). We can do this as convergence is uniform in \( C \). Apply our inequality to get
\[ |f(x) - f(x_0)| < \frac{\epsilon}{4} + n(x - x_k) + \frac{\epsilon}{4} + \frac{\epsilon}{4} \leq \frac{3\epsilon}{4} + n(x - x_0) + n|x_k - x| \]
\[ \leq \epsilon + n(x - x_0). \]

\( \Box \)

We again would like to use the Baire Category theorem to show the existence of a differentiable function \( h : \mathbb{R} \to \mathbb{R} \) such that for all \( (a, b) \subset \mathbb{R} \), there exists \( x, y \in (a, b) \) such that \( h'(x) > 0 > h'(y) \).

Let
\[ D = \{ f : \mathbb{R} \to \mathbb{R} : f \text{ is bounded and } \exists F : \mathbb{R} \to \mathbb{R} \text{ with } F' = f \} \]

For \( f, g \in D \), define \( d(f, g) = \sup_{x \in \mathbb{R}} |f(x) - g(x)| \).

**Claim:** \( D \) is a complete metric space. If \( f_n \in D \) is Cauchy, then \( \{f_n(0)\} \) converges. Let \( F_n : \mathbb{R} \to \mathbb{R} \) satisfy \( F_n(0) = 0 \) and \( F_n' = f_n \). By a theorem from 826, \( F_n \) converges uniformly to a differentiable function \( F \) and \( F' \) is the uniform limit of \( f_n' \). (Rudin, 7.17).

Let
\[ D_0 = \{ f \in D : f^{-1}([0]) \text{ is dense in } \mathbb{R} \} \]
Given \( f \in D_0 \), let \( Z_f = f^{-1}([0]) \).
Claim: $Z_f$ is a dense $G_\delta$ set. Proof is an exercise.

Thus, if $f_n \in D_0$ and $f_n \to f \in D$ then

$$\bigcup_{n=1}^{\infty} Z_{f_n} \subset Z_f$$

hence by Baire, $Z_f$ is dense and so is $f \in D_0$. Therefore, $D_0$ is a complete metric space. If $f, g \in D_0$ then $f + g \in D_0$. Thus,

$$Z_{f+g} \supset Z_f \cap Z_g$$

and $f + g \in D_0$. It is obvious that if $f \in D_0$, then $\lambda f \in D_0$ for all $\lambda \in \mathbb{R}$. Hence, $D_0$ is a vector space.

Claim: $D_0 \neq \{0\}$.

Let $\{d_n\}_{n=1}^{\infty}$ be a countable dense subset of $(-1, 1)$. For $x \in [-1, 1]$, let

$$F(x) = \sum_{n=1}^{\infty} \frac{(x - d_n)^{\frac{1}{3}}}{2^n}.$$

Notice the series converges uniformly, and so the limit is a continuous function. The function is increasing, so we know that it has a derivative almost everywhere. Then,

$$\frac{F(x) - F(t)}{x - t} = \sum_{n=1}^{\infty} \frac{1}{[\frac{1}{3}(t - d_n)^{\frac{1}{3}} + \frac{1}{3}(x - d_n)^{\frac{1}{3}} + \frac{1}{3}(x - d_n)^{\frac{1}{3}}]2^n}$$

For $a, b \in \mathbb{R}$,

$$\frac{1}{2}(a^2 + b^2) \leq a^2 + ab + b^2 \leq \frac{3}{2}(a^2 + b^2)$$

Let

$$K = \{x \in (-1, 1) : \sum_{n=1}^{\infty} \frac{1}{3(x - d_n)^{\frac{1}{3}}2^n} \text{ converges} \}$$

Using the DCT, one can show that if $x \in K$, then $F'(x)$ exists and

$$F'(x) = \sum_{n=1}^{\infty} \frac{1}{3(x - d_n)^{\frac{1}{3}}2^n}$$

Also if $x \in K$,

$$\lim_{t \to x} \frac{F(x) - F(t)}{x - t} = \infty$$

Finally, for all $x \in (-1, 1)$

$$\lim_{t \to x} \frac{F(t) - F(x)}{t - x} > \frac{\sqrt{2}}{3}.$$

Let $G : (F(-1), F(1)) \to (-1, 1)$ be the inverse of $F$. Then, $G$ is differentiable on $(F(-1), F(1))$ and $G'(F(x)) = \frac{1}{F'(x)}$. Notice that $G'$ is bounded and $G'(x) = 0$ if and only if $x \in F(K)$. So $G' = 0$ on a dense subset of $(F(-1), F(1))$.

*Monday, 2-21-2005*
Let
\[
g(x) = \begin{cases} 
(x + 1)^2(x - 1)^2, & x \in (F(-1), F(1)) \\
0 & \text{otherwise}
\end{cases}
\]
Then, \(g'\) exists for all \(x \in \mathbb{R}\); consider
\[
H(x) = \begin{cases} 
g(G(x)), & x \in (F(-1), F(1)) \\
0, & \text{otherwise}
\end{cases}
\]
Then, \(H'(x) = g'(G(x))G'(x)\) for \(x \in (F(-1), F(1))\) and 0 otherwise. Thus \(H\) is a nonzero function in \(D_0\).

Now, let \(E = \{f \in D_0 : \exists \text{ an interval } (a,b) \subset \mathbb{R} \text{ with } f(a,b) \subset [0, +\infty) \text{ or } f(a,b) \subset (-\infty, 0] \} \)

**Claim:** \(E\) is first category in \(D_0\). Once we show this, there must be a function \(D_0\) whose antiderivative is monotone on no subinterval on \( \mathbb{R} \). By Baire then, \(D_0 \setminus E \neq \emptyset\).

To show this, let \(\{I_n\}_{n=1}^{\infty}\) be an enumeration of the open intervals with rational endpoints, and put

\[
A_n = \{f \in E : f(I_n) \subset [0, +\infty)\}; \\
B_n = \{f \in E : f(I_n) \subset (-\infty, 0]\}.
\]

Notice that
\[
E = \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n=1}^{\infty} B_n,
\]
so we just need to show that the \(A_n\) and \(B_n\) are nowhere dense.

Notice that both \(A_n\) and \(B_n\) are closed, so we really just need to show that the interior is empty. Let \(\epsilon > 0\) and \(f \in A_n\).

**Claim:** \(B(\epsilon, f)\) is not a subset of \(A_n\).

Since \(f \in A_n\) and \(Z_f\) is dense, there exists \(x_0 \in I_n\) with \(f(x_0) = 0\). By translating a nonzero element of \(H\) of \(D_0\), we find an element \(g \in D_0\) such that \(g(x_0) < 0\) (multiply by \(-1\) if necessary). Let
\[
M > \sup x \in \mathbb{R} |g(x)|;
\]
then,
\[
d(f, \frac{\epsilon g}{m} + f) < \epsilon.
\]
As both \(g\) and \(f\) are in \(D_0\),
\[
\frac{\epsilon g}{m} + f \in D_0.
\]
But,
\[
\left(\frac{\epsilon g}{M} + f\right)(x_0) < 0
\]
is not in \(A_n\). Finally, we get that \(A_n\) is nowhere dense. Similarly, \(B_n\) is nowhere dense, so \(E\) is first category.

Putting everything together, we finish the proof.

**Example:** A complete metric space with no isolated points is uncountable (\(t \in X\) is isolated if \(\{t\}\) is open).

If \(X = \{x_n\}_{n=1}^{\infty}\) is countable, then since each \(\{x_n\}\) is nowhere dense, we’d have \(X\)
first category.

**Theorem 20. Uniform Boundedness:**

Let $\mathcal{F}$ be a family of real valued continuous functions on a complete metric space $X$. Suppose $\forall x \in X$,

$$\sup\{|f(x)| : f \in \mathcal{F}\} := M_x < \infty$$

Then, there exists an open set $G \subset X$ and a constant $C > 0$ such that

$$|f(x)| < C, \forall x \in G$$

and for all $f \in \mathcal{F}$.

**Proof.** Let

$$E_{m,f} = \{x \in X : |f(x)| \leq m\}.$$ 

Then $E_{m,f}$ is closed and

$$\bigcap_{f \in \mathcal{F}} E_{m,f} =: E_m$$

is also closed. By hypothesis,

$$\bigcup_{f \in \mathcal{F}} E_m = X.$$ 

By Baire, at least one of the $E_m$’s has non-empty interior. Then if $E_m^\circ =: G \neq \emptyset$, the theorem holds for this open set. □

Suppose $X, Y$ are metric spaces, $\mathcal{F}$ is a family of functions from $X$ into $Y$. We say that $\mathcal{F}$ is equicontinuous if for all $\epsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ whenever $f \in \mathcal{F}$ and $y \in X$ with $d_X(x, y) < \delta$.

The family is equicontinuous on $X$ if it is equicontinuous at each $x \in X$.

**Theorem 21. Ascoli-Arzela**

Let $X$ be a separable metric space and $Y$ be a complete metric space. Let $\mathcal{F}$ be an equicontinuous family of functions from $X$ into $Y$. Suppose that $(f_n)_{n=1}^\infty$ is a sequence in $\mathcal{F}$ such that $\{f_n(x) : 1 \leq n \leq \infty\}$ is compact. Then there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $f_{n_k}$ converges pointwise to a continuous function $f : X \to Y$ and if $K \subset X$ is a compact set, $f_{n_k} \to f$ uniformly on $K$.

**Remark:** Notice that if $X$ is itself compact and $Y = \mathbb{R}$. Take $\mathcal{C} \subset \mathcal{F} \subset \mathcal{C}(X)$ such that $\mathcal{F}$ is equicontinuous and pointwise bounded. Then, $\mathcal{C}$ is a compact set.

**Proof.** Let $(f_n)$ be a sequence in $\mathcal{C}$. By the Ascoli-Arzela theorem, there exists a subsequence which converges uniformly on $X$. So as then $\mathcal{C}$ is a metric space such that every sequence has a convergent subsequence, we conclude that $\mathcal{C}$ is a compact subset of $\mathcal{C}(X)$.

Proof of Ascoli-Arzela is same as it has always been.
4.1. Applications of Arzela-Ascoli. Example 1: For \( f \in C[0,1] \), let
\[
(Tf)(x) = \int_0^x f(t)dt
\]
Then, \( Tf \in C[0,1] \), so \( T \) is a linear map from \( C[0,1] \) to itself. Let
\[
\mathcal{F} := \{Tf : f \in C[0,1], \|f\|_\infty \leq 1\}.
\]
We would like to see whether \( \mathcal{F} \) is equicontinuous -
\[
|(Tf)(x) - (Tf)(y)| = |\int_x^y f(t)dt| \leq |x - y|.
\]
Hence, \( \mathcal{F} \) is an equicontinuous family. Also,
\[
|(Tf)(x)| \leq x \leq 1.
\]
Hence, \( \bar{\mathcal{F}} \) is compact.

Example 2: (Sketch) -

**Theorem 22. (Peano)** Suppose \( D \in \mathbb{R}^2 \) is an open set, \( f : D \rightarrow \mathbb{R} \) is continuous, and \( (x_0, y_0) \in D \). Then the differential equation \( y' = f(x, y) \) has a local solution passing through \( (x_0, y_0) \).

Proof idea - construct a set of approximations that are equicontinuous, then use the Ascoli-Arzela to show that there is a convergent subsequence. The point where that subsequence converges to is the solution function.

5. \( L^p \) Spaces

Let \((X, S, \mu)\) be a complete measure space. Let \( 0 < p, \infty \). Then,
\[
L^p(X, S, \mu) = \{f : X \rightarrow \mathbb{F}, mble : \int_X |f|^p d\mu < \infty\}.
\]
We also want to define \( L^\infty(X) \).

Given \( f : X \rightarrow \mathbb{F} \),
\[
esssup |f| = \inf\{t > 0 : m\{x : |f(x)| > t\} = 0\}.
\]
Then,
\[
L^\infty(X) = \{f : X \rightarrow \mathbb{F}, mble : esssup |f| < \infty\}
\]

For \( 1 \leq p < \infty \), define
\[
\|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}
\]
For \( p = \infty \), define
\[
\|f\|_\infty = esssup |f|
\]
Think of \( \mathbb{R}^2 \) as the set of functions on a two point space. Then, draw
\[
\{(x, y) : \|(x, y)\|_p = 1, 1 \leq p \leq \infty\}.
Theorem 23. Minkowski’s Inequality:
Let \((X, S, \mu)\) be a complete metric space. Let \(1 \leq p \leq \infty\). If \(f, g \in L^p\), then \(f + g \in L^p\) and
\[
\|f + g\|_p \leq \|f\|_p + \|g\|_p
\]

Proof. If \(f, g \in L^\infty\), let \(a > \|f\|_\infty\) and \(b > \|g\|_\infty\). Then,
\[
\mu\{x : |f(x)| > a\} = \mu\{x : |g(x)| > b\} = 0.
\]

By the triangle inequality,
\[
\{x : |f(x) + g(x)| > a + b\} \subset \{x : |f(x)| > a\} \cup \{x : |g(x)| > b\},
\]
so
\[
\mu\{x : |f(x) + g(x)| > a + b\} = 0.
\]
Therefore \(f + g \in L^\infty\) and \(\|f + g\|_\infty < a + b\). Taking the infimum over \(a\) and \(b\), we establish Minkowski’s Inequality for \(p = \infty\).

We must use the following:
Fact: For any real numbers \(a, b\) and \(p > 0\), we have
\[
|a + b|^p \leq 2^p(|a|^p + |b|^p)
\]

Proof. For \(s \in [0, 1]\), we know
\[
\frac{1 + s}{2} \in \left[\frac{1}{2}, 1\right]
\]
So,
\[
\left(\frac{1 + s}{2}\right)^p \in (0, 1]
\]
Thus,
\[
\left(\frac{1 + s}{2}\right)^p \leq 1 + s^p.
\]
Rearranging,
\[
(1 + s)^p \leq 2^p(1 + s^p)
\]
WLOG, assume \(|a|, |b| > 0\). Then,
\[
|a + b|^p = |a + \left(\frac{b}{a}\right)|^p = |a|^p|1 + \left(\frac{|b|}{|a|}\right)|^p
\]
\[
\leq |a|^p2^p\left(1 + \left(\frac{|b|}{|a|}\right)^p\right) = 2^p(|a|^p + |b|^p)
\]
\[
\square
\]

Recall that a function \(\phi : \mathbb{R} \to \mathbb{R}\) is convex if
\[
\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y)
\]
whenever \(x < y\) and \(\lambda \in [0, 1]\).
If \(\|f\|_p = 0\) or \(\|g\|_p = 0\), then there is nothing to do, so suppose
\[
\alpha = \|f\|_p
\]
\[
\beta = \|g\|_p
\]
are both nonzero. Put $f_0 = \frac{f}{\alpha}$ and $g_0 = \frac{g}{\beta}$. Then $\|f_0\|_p = \|g_0\|_p = 1$. Put

$$\lambda = \frac{\alpha}{\alpha + \beta}$$

and

$$(1 - \lambda) = \frac{\beta}{\alpha + \beta}$$

Now,

$$|f(x) + g(x)|^p \leq (|f(x)| + |g(x)|)^p = (\alpha|f_0(x)| + \beta|g_0(x)|)^p$$

$$(\alpha + \beta)^p(\lambda|f_0(x)| + (1 - \lambda)|g_0(x)|)^p \leq (\alpha + \beta)^p\lambda|f_0(x)|^p + (1 - \lambda)|g_0(x)|^p$$

Hence,

$$\int_X |f(x) + g(x)|^p d\mu \leq (\alpha + \beta)^p \left[ \lambda \int_X |f_0(x)|^p d\mu + (1 - \lambda) \int_X |g_0(x)|^p d\mu \right]$$

$$= (\alpha + \beta)^p (\|f\|_p + \|g\|_p)^p$$

Taking $p^{th}$ roots, $f + g \in L^p$ and the inequality is established. \hfill \Box

Wednesday, 3-2-2005: Dr. Rammaha teaches

**Definition 5.1.** A normed space $(X, \| \cdot \|)$ is a linear space $X$ (closed under addition and multiplication by scalars over a field $F$) and a function $\| \cdot \| : X \to \mathbb{R}$ with the properties:

1. $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$.
2. $\|cx\| = |c|\|x\|$, for all $c \in F$ and all $x \in X$.
3. Triangle Inequality:

$$\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in X$$

Let us agree that $(X, S, \mu)$ be any measure space. Then, for $1 \leq p < \infty$,

$$L^p(\mu) = L^p(X, S, \mu) := \{ f : X \to \mathbb{F} : f \text{ is measurable}, \|f\|_p < \infty \}$$

where

$$\|f\|_p = \left( \int_X |f|^p d\mu \right)^{\frac{1}{p}}.$$

Because of the Minkowski inequality, it is now trivial to verify that $(L^p(\mu), \| \cdot \|_p)$ is a normed space for $1 \leq p \leq \infty$.

**Definition 5.2.** Let $(X, \| \cdot \|)$ be a normed space. Let $x_n, x \in X$ for $n \in \mathbb{N}$. Then, we say that $x_n \to x$ in $X$ if and only if

$$\|x_n - x\|_p \to 0$$

as $n \to \infty$. We say $\{x_n\}$ is Cauchy in $X$ if

$$\|x_n - x_m\| \to 0$$

as $n, m \to \infty$. We say that $X$ is complete if all Cauchy sequences converge to a point in $X$. In this case, $X$ is called a Banach space ($B$-space).
Definition 5.3. Let $(X, \| \cdot \|)$ be a normed space and take a sequence $\{u_n\}_{n=1}^{\infty} \subset X$. We say the series
\[ \sum_{n=1}^{\infty} u_n \]
converges absolutely if
\[ \sum_{n=1}^{\infty} \| u_n \| \]
converges in $\mathbb{R}$.

Theorem 24. A normed space $(X, \| \cdot \|)$ is complete if and only if every absolutely convergent series converges in $X$.

Proof. Suppose $X$ is very complete - every Cauchy sequence converges. Let $\{u_n\}_{n=1}^{\infty} \subset X$ such that
\[ \sum_{n=1}^{\infty} \| u_n \| \]
is convergent - say to $M \in [0, \infty)$. Let $\epsilon > 0$ be given. Then, there exists $N \in \mathbb{N}$ such that
\[ \sum_{n=N}^{\infty} \| u_n \| < \epsilon \]
Set
\[ S_n = \sum_{k=1}^{n} u_k \]
Then, for all $n > m \geq N$, we have
\[ \| S_n - S_m \| = \| \sum_{k=m+1}^{n} u_k \| \leq \sum_{k=m+1}^{n} \| u_k \| \]
(where the last step is by the triangle inequality).
\[ \leq \sum_{k=N}^{\infty} \| u_k \| < \epsilon \]
So, $S_n$ is Cauchy in $X$, which is complete. Hence, $S_n$ converges to some point $u \in X$. We write:
\[ \lim_{n \to \infty} S_n =: \sum_{k=1}^{\infty} u_k = u \in X \]

The converse is a bit more tricky. Let $\{u_n\}_{n=1}^{\infty} \subset X$ such that $\{u_n\}$ is Cauchy. For every $k \in \mathbb{N}$, there exists a strictly increasing sequence $\{n_k\}_{k=1}^{\infty} \subset \mathbb{N}$ such that $\| u_n - u_m \| < 2^{-k}$ for all $n, m \geq n_k$ (we can do this by the definition of a Cauchy sequence).
So, we have a subsequence
\[ \{u_{n_k}\}_{k=1}^{\infty} \subset \{u_n\} \]
Set
\[ g_1 = u_{n_1} \]
and
\[ g_k = u_{n_k} - u_{n_{k-1}} \]
Set
\[ S_j = \sum_{k=1}^{j} g_k = u_{n_j} \]
Also,
\[ \sum_{k=1}^{j} \|g_k\| = \sum_{k=1}^{j} \|u_{n_k} - u_{n_{k-1}}\| \leq \sum_{k=1}^{j} 2^{-k} \]
Hence,
\[ \sum_{k=1}^{j} \|g_k\| \leq \sum_{k=1}^{\infty} 2^{-k} = 1 < \infty. \]
Hence, by hypothesis, \( S_k \) converges to some point, \( u \in X \). So,
\[ \|u_{n_j} - u\| = \|S_j - u\| \to 0 \]
Then, given any \( \epsilon > 0 \), there exists a \( N_1 > 0 \) such that \( \|u_{n_j} - u\| \leq \frac{\epsilon}{2} \) for \( j \geq N_1 \) and a \( N_2 > 0 \) such that \( \|u_m - u_{n_j}\| \leq \frac{\epsilon}{2} \) for \( m \geq N_2 \) and \( j \geq N_1 \). Hence, for any \( m > N := \max\{N_1, N_2\} \),
\[ \|u_m - u\| \leq \|u_{m} - u_{n_j}\| + \|u_{n_j} - u\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \]
Hence, the Cauchy sequence converges to \( u \in X \), and \( X \) is complete. \( \Box \)

**Theorem 25. Riesz-Fisher Theorem:**
\((L^p(\mu), \| \cdot \|)\) is a Banach space for all \( 1 \leq p \leq \infty \).

**Proof.** First, assume \( 1 \leq p < \infty \). By the last theorem, we need to show every sequence \( \{f_n\}_{n=1}^{\infty} \subset L^p(\infty) \) such that \( \sum_{k=1}^{\infty} \|f_k\| \) is absolutely convergent, then \( \sum_{k=1}^{\infty} f_k \) is convergent in \( L^p(\mu) \). So, let \( \{f_k\} \subset L^p(\mu) \) such that
\[ \sum_{k=1}^{\infty} \|f_k\|_p = B < \infty \]
Define
\[ g_n(x) = \sum_{k=1}^{n} |f_k(x)|. \]
Then, \( g_n \) is an increasing sequence on \( X \) with values in \([0, \infty]\). And so,
\[ g(x) := \sum_{k=1}^{\infty} |f_k(x)| \]
is measurable and \( g(x) \in [0, \infty] \). Now, by the triangle (Minkowski) inequality,
\[ \|g_n\|_p = \| \sum_{k=1}^{n} g_k \|_p \leq \sum_{k=1}^{n} \|g_k\|_p \]
I.e.,
\[ \int_X |g_n|^p d\mu \leq B^p \]
So, by the Monotone Convergence Theorem, we have
\[ \int_X |g(x)|^p d\mu = \lim_{n \to \infty} \int_X |g_n(x)|^p d\mu \leq B^p \]
Hence,
\[ \|g\|_p \leq B. \]
So, \( g \in L^p(\mu) \). By an old theorem, \( g < \infty \) almost everywhere and so the series
\[ \sum_{k=1}^{\infty} f_k(x) \]
is absolutely convergent for almost all \( x \). Let \( E \) be the set on which \( f_k(x) \) converges absolutely. Then, define
\[ F(x) = \begin{cases} \sum_{k=1}^{\infty} f_k(x), & x \in E \\ 0, & x \in E^c \end{cases} \]
Then,
\[ |F(x)| \leq g(x) \]
for all \( x \in X \) and \( F \) is \( \mu \)-measurable. Note
\[ \int_X |F(x)|^p d\mu \leq \int_X g(x)^p d\mu \leq B^p < \infty \]
So, \( F \in L^p(\mu) \).

Finally, we note the following:
\[ 0 \text{ almost everywhere} = \lim_{n \to \infty} |F - \sum_{k=1}^{n} f_k|^p \]
and, for all \( n \),
\[ \leq \left( |F| + \sum_{k=1}^{n} |f_k| \right)^p \leq (g + g)^p = 2^p g^p \in L^1 \]
So, by the Dominated Convergence Theorem,
\[ \int_X |F - \sum_{k=1}^{\infty} f_k|^p = 0 \]
almost everywhere. Thus,
\[ \lim_{n \to \infty} \|F - \sum_{k=1}^{n} f_k\|_p = 0. \]
I.e.,
\[ \sum_{k=1}^{\infty} f_k = F. \]
in \( L^p(\mu) \). This proves the completeness of \( L^p(\mu) \) for \( p \) finite.

The case \( p = \infty \). Let \( \{f_n\}_1^{\infty} \subset L^\infty(\mu) \) be a Cauchy sequence. By the definition of \( L^\infty(\mu) \), for all \( n \in \mathbb{N} \), there exists a set \( E_n \) such that \( |f_n(x)| \leq \|f_n\|_\infty \) on \( E_n^c \) and
where \( \mu(E_n) = 0 \). Similarly, \( f_n - f_m \in L^\infty(\mu) \) and so for all \( m, n \in \mathbb{N} \), there exists a \( F_{n,m} \) such that

\[
|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty
\]
on \( F_{n,m}^c \) and where \( \mu(F_{n,m}) = 0 \). Now, for all \( n, m \in \mathbb{N} \), set

\[
A = \left( \bigcup_{n=1}^{\infty} E_n \right) \cup \left( \bigcup_{n,m=1}^{\infty} F_{n,m} \right)
\]

Then, \( \mu(A) = 0 \). Also, we still have

\[
|f_n(x)| \leq \|f_n\|_\infty
\]
and

\[
|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty
\]
on \( A^c \). Now, \( \|f_n - f_m\|_\infty \to 0 \) as \( n, m \to \infty \).

Therefore, \( \{f_n(x)\}_{n=1}^{\infty} \) is uniformly Cauchy as \( n, m \to \infty \) on \( A^c \). Thus, there exists a function \( f \) such that \( f(x) = \lim_{n \to \infty} f_n(x) \) for all \( x \in A^c \). Then, set

\[
F(x) = \begin{cases} f(x) = \lim_{n \to \infty} f_n(x), & x \in A^c \\ 0, & x \in A. \end{cases}
\]

Now, there exists \( N \in \mathbb{N} \) (using \( f_n(x) \to f(x) \) uniformly on \( A^c \)) such that

\[
|f(x) - f_N(x)| < 1, \forall x \in A^c.
\]

So, for all \( x \in A^c \), we have

\[
|F(x)| \leq |f(x) - f_N(x)| + |f_N(x)| < 1 + \|f_N\|_\infty < \infty.
\]

So, \( F \in L^\infty(\mu) \). Now,

\[
\|F - f_n\|_\infty = \sup_{x \in A^c} |F(x) - f_n(x)| = \sup_{x \in A^c} |f(x) - f_n(x)| \to 0
\]

So, \( f_n \to F \) in \( L^\infty(\mu) \). Hence, \( L^\infty(\mu) \) is complete. \( \square \)

**Monday, 3-7-2005:**

**Lemma 5.1.** “Young’s Inequality”:

If \( a, b \geq 0 \) and \( 0 \leq \lambda \leq 1 \), then

\[
a^{\lambda}b^{1-\lambda} \leq \lambda a + (1 - \lambda)b
\]

and equality holds if and only if \( a = b \).

**History** Young’s original statement: Let \( a : [0, \infty) \to [0, \infty) \) be continuous and strictly increasing. Let \( \tilde{a} := a^{-1} \). So, \( \tilde{a} \) is continuous and strictly increasing on \([0, \infty)\). Assume \( \lim_{x \to \infty} a(x) = +\infty \). For \( x \geq 0 \), set

\[
A(x) = \int_0^x a(t)dt \text{ and } \tilde{A}(x) = \int_0^x \tilde{a}(t)dt
\]

Then, Young’s Original Inequality says:

\[
xy \leq A(x) + \tilde{A}(y), \forall x, y \geq 0
\]
A special case of this inequality is with \( a(t) = t^{p-1} \) where \( 1 < p < \infty \). Here, \( \tilde{a}(t) = t^{1/p} \). So, the original inequality gives us that

\[
x y \leq \frac{1}{p} x^p + \frac{p-1}{p} y^{p/\lambda}
\]

**Definition 5.4.** Let \( 1 \leq p, q \leq \infty \). We say \( p \) and \( q \) are Holder conjugates if \( \frac{1}{p} + \frac{1}{q} = 1 \). Here, if \( 1 < p < \infty \), then \( q = \frac{p}{p-1} \) is the Holder conjugate to \( p \).

So, we get that

\[
x y \leq \frac{x^p}{p} + \frac{y^q}{q}, \forall x, y \geq 0 \text{ and } p, q \text{ are Holder conjugates}
\]

This is the statement we’d like to prove.

**Proof.** (Young’s Inequality)

If \( b = 0 \), then eq. 2 is trivial. Suppose \( b > 0 \) - then 2 is equivalent to

\[
\left( \frac{a}{b} \right)^\lambda \leq \lambda \left( \frac{a}{b} \right) + 1 - \lambda
\]

Setting \( t = \frac{a}{b} \). Then \( t \geq 0 \) and 2 is equivalent to

\[
t^\lambda \leq \lambda t + 1 - \lambda, t \geq 0
\]

and equality holds if and only if \( t = 1 \).

Set

\[
f(t) = t^\lambda - \lambda t, t \geq 0
\]

Note,

\[
f'(t) = \lambda t^{\lambda-1} - \lambda = \lambda(t^{\lambda-1} - 1)
\]

\[
= \begin{cases} 
< 0, & t > 1 \\
> 0, & 0 < t < 1 \\
0, & t = 1
\end{cases}
\]

i.e., \( f \) has a global maximum at \( t = 1 \) and so

\[
f(t) = t^\lambda - \lambda t \leq f(1) = 1 - \lambda
\]

and equality holds if and only if \( t = 1 \).

**Remark:** Let \( a, b \geq 0 \) and \( p > 1 \) and \( \frac{1}{q} = 1 - \frac{1}{p} \) (i.e., \( p \) and \( q \) are Holder conjugates.) Take \( \tilde{a} = a^p \) and \( \tilde{b} = b^q \) and \( \lambda = \frac{1}{p} \). Note that \( \frac{1}{q} = 1 - \lambda \). From eq. 2, we have

\[
\tilde{a}^{1-\lambda} \tilde{b}^{1-\lambda} \leq \lambda \tilde{a} + (1 - \lambda) \tilde{b}
\]

Rewriting, this is

\[
ab \leq \frac{a^p}{p} + \frac{b^q}{q}
\]

**Example:** Take \( a, b, \epsilon > 0 \).

\[
ab = a^\epsilon \epsilon^{-\frac{1}{p}} b \leq \frac{1}{p} (a^\epsilon)^p + \frac{1}{q} (\epsilon^{-\frac{1}{q}} b)^q
\]

\[
= \frac{ea^p}{p} + \frac{1}{q} \epsilon^{-\frac{p}{q}} b^q
\]
**Theorem 26. “Holder’s Inequality”:**

Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $f, g : X \to \mathbb{C}$ be $\mu$-measurable functions. Then,

$$
\|fg\|_1 = \int_X |fg|d\mu \leq \|f\|_p \|g\|_q = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} \left(\int_X |g|^q d\mu\right)^{\frac{1}{q}}
$$

In particular, if $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$, and in this case, equality holds in 3 if and only if

$$
\|g\|_q^p |f| = \|f\|_p^p |g|.
$$

**Proof.** If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $f = 0$ $\mu$-almost everywhere or $g = 0$ $\mu$-almost everywhere, and eq. 3 is trivial. Also, if $\|f\|_p = \infty$ or $\|g\|_q = \infty$, then eq. 3 is again trivial. So, we assume that $0 < \|f\|_p, \|g\|_q < \infty$.

So, eq. 3 is equivalent to

$$
\frac{1}{\|f\|_p \|g\|_q} \int_X |fg|d\mu \leq 1
$$

Let

$$
a(x) = \frac{|f(x)|}{\|f\|_p}, b(x) = \frac{|g(x)|}{\|g\|_q}
$$

Then, by Young’s Inequality,

$$
\frac{|fg|}{\|f\|_p \|g\|_q} = ab \leq \frac{a^p}{p} + \frac{b^q}{q} = \frac{1}{p} \|f(x)|^p \|g\|_q^q + \frac{1}{q} |g(x)|^q \|g\|_q^q.
$$

This holds $\mu$-almost everywhere, with equality if and only if

$$
\|g\|_q^p |f| = \|f\|_p^p |g|.
$$

By integrating, we have:

$$
\frac{1}{\|f\|_p \|g\|_q} \int_X |fg|d\mu \leq \frac{1}{p} \|f\|_p^p + \frac{1}{q} \|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1
$$

So, we get

$$
\|fg\|_1 \leq \|f\|_p \|g\|_q \quad \square
$$

Wednesday, 3-9-2005:

**Proposition 5.1.** For $1 \leq p < \infty$, the set

$$
\mathcal{S} := \{f = \sum_{j=1}^n a_j \chi_{E_j} : \mu(E_j) < \infty\}
$$

is dense in $L^p$.

**Proof.** As $\mu(E_j) < \infty$, each element in the sum is in $L^p$, so $\mathcal{S} \subset L^p$. Let $f \in L^p$.

Find a sequence of simple functions $f_n$ such that $|f_n| \leq |f|$ and $f_n \to f$ pointwise.
Writing these functions in standard form, \( f_n = \sum_{k=1}^{r} a_k \chi_{E_k} \), the condition that \( |f_n| \leq f \) shows that \( \mu(E_k) < \infty \), so \( f_n \in \mathcal{S} \). Also,

\[
|f_n(x) - f(x)|^p \leq 2^p(|f_n(x)|^p + |f(x)|^p) \leq 2^p+1|f(x)|^p
\]

Hence, by the DCT, \( \int |f_n - f|^p \, d\mu \to 0 \), and \( \|f_n - f\|_p \to 0 \). □

**Remark:** A similar result is true for \( p = \infty \).

**Definition 5.5.** Let \( X, Y \) be normed spaces over \( \mathbb{F} \) and let \( T : X \to Y \) be a linear map. We say that \( T \) is a bounded linear map if

\[
\sup\{\|Tx\|_Y : \|x\|_X \leq 1\} < \infty
\]

If \( T \) is bounded, we define

\[
\|T\| := \sup\{\|Tx\|_Y : \|x\|_X \leq 1\}
\]

**Example:** Let \( X = \{f : (0, 1) \to \mathbb{R} : f \text{ is bounded and has derivatives of all orders}\} \). For \( f \in X \), let \( \|f\| = \sup_{t \in (0, 1)} |f(t)| \). Then \( X \) is a normed space. Define \( D : X \to X \) by \( Df = f' \). Then \( D \) isn’t bounded. Reason: \( f_n(x) = \sin nx \).

So, \( \|f_n\| = 1 \) but \( \|Df_n\| = n \).

**Proposition 5.2.** Let \( X, Y \) be normed spaces and \( T : X \to Y \) a linear map. The following are equivalent.

(a) \( T \) is continuous at 0.
(b) \( T \) is continuous.
(c) \( T \) is bounded.

**Proof.** (b) \( \Rightarrow \) (a) is trivial. (a) \( \Rightarrow \) (b) Suppose \( x_n \in X \) and \( x_n \to x \in X \). Then, \( x_n - x \to 0 \), so by hypotheses, \( T(x_n - x) \to 0 \), i.e., \( Tx_n \to Tx \).

(b) \( \Rightarrow \) (c) (Prove the contrapositive). Suppose \( T \) isn’t bounded. Then \( \exists x_n \in X \) such that \( \|x_n\| \leq 1 \), yet \( \|Tx_n\| > n \). Then, \( \|\frac{x_n}{n}\| \to 0 \) so \( \frac{x_n}{n} \to 0 \) in \( X \). But, \( \|T\frac{x_n}{n}\| = \frac{\|Tx_n\|}{n} > 1 \), so \( T(\frac{x_n}{n}) \) does not converge to zero. Hence, \( T \) isn’t continuous at \( 0 \) so \( T \) isn’t continuous.

Finally, suppose that (c) holds. Let \( x_n \to 0 \). Then,

\[
\|Tx_n\| = \begin{cases} 0, & x_n = 0 \\ \|x_n\|\|T(\frac{x_n}{\|x_n\|})\|, & x_n \neq 0 \end{cases}
\]

Hence, \( \|Tx_n\| \leq \|T\|\|x_n\| \to 0 \). As \( Tx_n \to 0 \) so \( T \) is continuous at 0. □

**Remark:**

If \( T \) is a bounded linear transformation, then \( \forall x \in X, \|Tx\| \leq \|T\|\|x\| \).

**Definition 5.6.** Linear Functional:

Let \( X \) be a normed space. A linear functional is a linear map \( f : X \to \mathbb{F} \) and is bounded if it is bounded as a linear transformation.

**Definition 5.7.** Given a normed space, \( X \), the dual space of \( X \) is the collection of all bounded linear functionals on \( X \), as is denoted \( X^* \).
Example:
Let $1 < p < \infty$ and suppose $p, q$ are Hölder conjugates. Let $X = L^p$ and fixed $g \in L^q$. Define

$$\phi_g(f) = \int_X f g d\mu.$$ 

By Hölder’s inequality, $\phi_g$ is a linear functional on $L^p$ and

$$|\phi_g(f)| \leq \int_X |f g| d\mu \leq \|f\|_p \|g\|_q.$$

Dividing by $\|f\|_p$, we get that $|\phi_g(f)| \leq \|g\|_q$. Taking the sup over $\|f\|_p \leq 1$, we get that $\phi_g$ is bounded and $\|\phi_g\| \leq \|g\|_q$.

Friday, 3-11-2005:
(Aside on notation:)
$$\ell^p := L^p(\mathbb{N}, \text{counting measure})$$
That is,

$$= \{(x_n)_{n=1}^\infty : \sum_{n=1}^\infty |x_n|^p < \}$$

Actually, when $1 < p \leq \infty$, we get equality for any measure and equality also holds when $p = 1$ and $\mu$ is semifinite.

Proof. For $1 < p < \infty$:
We have

$$\|g\|_q^p = \int_X |g|^q d\mu = \mathbb{E}_X |g|^{q-1} |g| d\mu$$

Put $f_0 = |g|^{q-1}$. Then,

$$|f_0|^p = |g|^{pq-p} = |g|^q$$

Therefore, $|f_0| \in L^p$. Moreover,

$$\|f_0\|_p = \left( \int_X |f_0|^p \right)^{\frac{1}{p}} = \left( \int_X |g|^q \right)^{\frac{1}{p}} = \|g\|_q^p = \|g\|_q^{q-1}$$

Put

$$f = \frac{|g|^{q-1}}{\|g\|_q} \text{sign}(g)$$

where $\text{sign}(g) : X \to \mathbb{F}$ satisfies

$$\text{sign}(g)|g| = g$$

Notice that $\text{sign}(g)$ is measurable. Then $\|f\|_p = 1$ and

$$\phi_g(f) = \int_X \frac{|g|^{q-1}}{\|g\|_q^q} \text{sign}(g) g d\mu$$

$$= \int_X \frac{|g|^q}{\|g\|_q^q} d\mu = \left\| \frac{|g|^q}{\|g\|_q^q} \right\|_1 = \|g\|_q.$$ 

Hence, $\|\phi_g\| = \|g\|_q$. We leave the other cases for the interested reader. When $p = \infty$, $q = 1$ and take $f = \text{sign}(g)$. Then $f \in L^\infty$ and

$$\int_X fg = \int |g| = \|g\|_1.$$ 

\[\square\]
Theorem 27. **Reisz Representation Theorem** Let $1 \leq p < \infty$ and let $q$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Let $\phi : L^p \rightarrow \mathbb{F}$ be a bounded linear functional. Then,

1. If $1 < p < \infty$, then there exists a unique $g \in L^q$ such that for all $f \in L^p$,
   \[ \phi(f) = \int_X fg \, d\mu. \]
   and $\| \phi \| = \| g \|_q$.

2. If $X$ is $\sigma$-finite and $g : L^1 \rightarrow \mathbb{F}$ is a bounded linear functional, then there exists a unique $g \in L^\infty$ such that
   \[ \phi(f) = \int_X fg \, d\mu \]
   and $\| \phi \| = \| g \|_\infty$.

**Proof.** First assume that $(X, \mathcal{S}, \mu)$ is a finite measure. Define $\nu(E) = \phi(\chi_E)$ for $E \in \mathcal{S}$. Then, $\| \chi_E \|_p = \mu(E)^{\frac{1}{p}}$. Let $E = \bigcup_{j=1}^{\infty} E_j$ be a disjoint union and let $\alpha_i \in \mathbb{F}$ satisfy $|\alpha_i| = 1$ and $\alpha_i \nu(E_j) \geq 0$. Let

\[ f = \sum_{i=1}^{\infty} \alpha_i \chi_{E_i} \]

Note that the series converges in $L^p$:

\[ \| f - \sum_{j=1}^{n} \alpha_j \chi_{E_j} \|_p = \| \sum_{j=n+1}^{\infty} \alpha_j \chi_{E_j} \|_p = \mu(\bigcup_{j=n+1}^{\infty} E_j)^{\frac{1}{p}} \]

Now, we are in a finite measure so the right hand side always exists. By the continuity theorem, we get that this term goes to zero. Hence, the partial sums converge in $L^p$, so the whole series converge in $L^p$. Then

\[ \phi(f) = \lim_{n \to \infty} \phi(\sum_{j=1}^{n} \alpha_j \chi_{E_j}) \]

\[ = \lim_{n \to \infty} \sum_{j=1}^{n} \alpha_j \phi(\chi_{E_j}) = \lim_{n \to \infty} \sum_{j=1}^{n} |\nu(E_j)| = \sum_{j=1}^{\infty} |\nu(E_j)| \leq \| \phi \| \| f \|_p \]

Hence, $\sum_{j=1}^{\infty} |\nu(E_j)|$ is absolutely convergent and $\nu(E) = \sum_{j=1}^{\infty} \nu(E_j)$. Hence, $\nu$ is a complex measure.

If $\mu(E) = 0$, then $\chi_E = 0 \in L^p$,

\[ \nu(E) = \phi(\chi_E) = \phi(0) = 0 \]

Hence, $\nu \ll \mu$. By the Radon-Nikodym theorem, there exists a function $g \in L^1(\mu)$ such that

\[ \phi(\chi_E) = \nu(E) = \int_X g \chi_E \, d\mu \]

By linearity, if $f = \sum_{j=1}^{n} c_j \chi_{E_j}$ is a simple function, $\phi(f) = \int_X g \chi_E \, d\mu$. As simple functions are dense in $L^p$ and $\phi$ is continuous, we see this holds for all $f \in L^p$. 

\[ \text{ANALYSIS NOTES 43} \]
We next show that \( g \in L^q \). Write \( g = \text{sign}(g)|g| \), where
\[
\text{sign}(g)(x) = \begin{cases} 
\frac{g(x)}{|g(x)|}, & g(x) \neq 0 \\
1, & g(x) = 0 
\end{cases}
\]
Let \( \psi_n \) be a sequence of simple functions with \( 0 \leq \psi_n \leq |g|^q \) and \( \lim_{n \to \infty} \psi_n(x) = |g(x)|^q \). Because the measure is finite, we conclude that each \( \psi_n \) is integrable.
Then, \( \psi_n^{\frac{1}{q}} \text{sign}(g) \in L^p \) and when \( p \in (1, \infty) \),
\[
\psi_n = \psi_n^{\frac{1}{q}} \psi_n^{\frac{1}{p}} \leq \psi_n^{\frac{1}{q}} |g| = \psi_n^{\frac{1}{q}} \text{sign}(g) |g|
\]
so
\[
\int \psi_n d\mu \leq \int \psi_n^{\frac{1}{q}} \text{sign}(g) |g| d\mu
\]
\[
= \phi(\psi_n^{\frac{1}{q}} \text{sign}(g)) \leq \|\phi\| \|\psi_n^{\frac{1}{q}} \text{sign}(g)\|_p
\]
\[
= \|\phi\| \left( \int_X \psi_n d\mu \right)^{\frac{1}{p}}
\]
Therefore,
\[
\left( \int_X \psi_n d\mu \right)^{\frac{1}{q}} \leq \|\phi\|
\]
By the Monotone Convergence Theorem, we can take the limit to get that \( g \in L^q \) and \( \|g\|_q \leq \|\phi\| \). By our previous work, \( \|\phi\| \leq \|g\|_q \), so \( \|\phi\| = \|g\|_q \).
We now establish uniqueness:
Let \( g_1 \in L^q \) and
\[
\int_X fg d\mu = \int_X f g_1 d\mu
\]
for all \( f \in L^p \). Then,
\[
\int_X (f - f g_1) d\mu = 0
\]
for all \( f \in L^p \). Take \( f = \text{sign}(g - g_1) \in L^p \) (\( \mu \) is finite), so
\[
\int |g - g_1| = 0
\]
Hence, \( g = g_1 \) almost everywhere.
We now finish the proof for (b). When \( p = 1 \), argue as before to obtain \( g \in L^1 \) such that
\[
\phi(f) = \int_X fg d\mu
\]
Let \( 0 < M < \text{esssup}|g| \) (if \( \text{esssup}|g| = 0 \), then we really have nothing to do). Let
\[
E := \{ x \in X : |g(x)| > M \}
\]
Note that \( \mu(E) > 0 \). Put
\[
f = \frac{\chi_E}{\mu(E)} \text{sign}(g)
\]
Then,
\[
\|f\|_1 = 1
\]
and
\[
\phi(f) = \int_X \frac{\chi_E}{\mu(E)} \text{sign}(g) g d\mu = \int_X \frac{\chi_E}{\mu(E)} |g| d\mu
\]
≥ \frac{M}{\mu(E)} \mu(E) = M

Therefore,

M < \phi(f) \leq \|\phi\| \|f\|_1 = \|\phi\|

Taking the sup over possible values of M, we get that

\text{esssup}|g| \leq \|\phi\|,

i.e., g \in L^\infty. Uniqueness follows similarly. Hence, this completes the theorem for the case of a finite measure.

Wednesday, 3-23-2005:
We now proceed from the finite case to the \( \sigma \)-finite case.

Write \( X = \bigcup_{n=1}^{\infty} X_n \) where \( \mu(X_n) < \infty \) and \( X_n \subset X_{n+1} \). Consider the linear functional \( \phi_n \) on \( L^p(X, \mu) \) defined by

\[ \phi_n(f) = \phi(f \chi_{X_n}) \]

Note that \( \phi_n \) is bounded since

\[ \|\phi_n(f)\| = \|\phi(f \chi_{X_n})\| \leq \|\phi\| \|f\|_p \leq \|\phi\| \|f\|_p \]

This also shows that \( \|\phi_n\| \leq \|\phi\| \).

By our previous work, there exists a unique \( g_n \in L^q \) such that

\[ \phi_n(f) = \int_X f g_n d\mu \]

and \( \|g_n\|_q = \|\phi_n\| \leq \|\phi\| \). But,

\[ g_{n+1}|_{X_n} = g_n \]

by the uniqueness assumption. Define \( g : X \to F \) by

\[ g(x) = g_n(x) \]

if \( x \in X_n \). We have \( |g \chi_{X_n}| = |g_n \chi_{X_n}| \) and thus \( |g \chi_{X_n}| \to |g| \) pointwise. Further,

\[ \int_X |g|^q d\mu = \lim_{n \to \infty} \int_X |g \chi_{X_n}|^q d\mu \]

\[ = \lim_{n \to \infty} \|g_n\|_q^q \leq \phi^q \]

Therefore, \( g \in L^q \) and \( \|g\|_q \leq \|\phi\| \). Finally, for \( f \in L^p \), put \( f_n = f \chi_{X_n} \). Then \( f_n \to f \) pointwise and

\[ |f_n g| \leq |fg| \in L^1 \]

by Holder’s Inequality. By the DCT:

\[ \int_X fgd\mu = \lim_{n \to \infty} \int f_n gd\mu = \lim_{n \to \infty} \int f g \chi_{X_n} d\mu = \lim_{n \to \infty} \int f g_n d\mu \]

\[ = \lim_{n \to \infty} \phi_n g f) = \lim_{n \to \infty} \phi(f \chi_{X_n}) = \phi(f), \]

since \( \|f \chi_{X_n} - f\|_p \to 0 \) and by the continuity of \( \phi \). Uniqueness of \( g \) follows from the uniqueness of \( g_n \).

This shows the theorem holds for \( 1 < p < \infty \) and \( \mu \) \( \sigma \)-finite.

Now if \( p = 1 \) and \( \mu \) is \( \sigma \)-finite, argue as above to obtain unique \( g_n \in L^\infty \) and construct \( g \) as well. Then, \( \|g_n\|_\infty \leq \|\phi\| \) for all \( n \), and hence \( \|g\|_\infty = \sup_n \|g_n\|_\infty \leq \|\phi\| \).
Now, let $\mu$ be an arbitrary measure. For each set $E \subset X$ with $E$ of $\sigma$-finite measure, let $g_E \in L^q$ be such that

$$\phi(f \chi_E) = \int_X f g_E \chi_E d\mu.$$  

Notice that if $E$ and $F$ are both of $\sigma$-finite measure and $E \subset F$, then $g_F|_E = g_E$ and

$$\|g_E\|_q \leq \|\phi\|.$$  

Let

$$M := \sup\{\|g_E\|_q : E \text{ has } \sigma\text{-finite measure}\}.$$  

So $M \leq \|\phi\|$. Let $\{E_N\}$ be a sequence of $\sigma$-finite measure such that $\|g_{E_n}\|_q \to M$. Set

$$F = \bigcup_{n=1}^{\infty} E_n.$$  

Then $F$ is of $\sigma$-finite measure. Since $F \supset E_n$ for all $n$, and $\|g_F\|_q \geq \|g_{E_n}\|_q$ for all $n$, we have $\|g_F\|_q = M$.

If $A \supset F$ and $A$ has $\sigma$-finite measure, then $M \geq \|g_A\|_q \geq \|g_F\|_q = M$, so $\|g_A\|_q = M$ too.

Also, $g_A$ extends $g_F$ and so

$$M^q = \int_A |g_A|^q d\mu = \int_{A \setminus F} |g_A|^q + \int_F |g_A|^q$$

Therefore, $g_A = 0$ almost everywhere on $A \setminus F$ and $g_A = g_F$ almost everywhere.

Finally, for $f \in L^p$,

$$N := \{x \in X : |f(x)| > 0\}$$

is of $\sigma$-finite measure. So, take $A = N \cup F$. Thus,

$$\phi(f) = \int_X f g_A d\mu = \int_X f g_F d\mu$$

Thus, if $g = g_F$, we get

$$\phi(f) = \int_X f g d\mu$$
(3) If \( \tau = \mathcal{P}(X) \), then every set is open and \( \tau \) is the **discrete topology** on \( X \).

**Definition 6.2.** A set is **closed** if \( F^C \in \tau \).

**Definition 6.3.** If \( x \in X \), a **neighborhood of** \( x \) is any open set \( G \in \tau \) with \( x \in G \).

**Definition 6.4.** If \((X, \tau)\) and \((Y, \sigma)\) are topological spaces and \( f : X \to Y \) is a function, then \( f \) is **continuous** if whenever \( H \in \sigma \), \( f^{-1}(H) \in \tau \).

**Remark:** \( f \) is continuous at \( x \in X \) if and only if for all open subsets \( H \in \sigma \) with \( f(x) \in H \), there exists a \( G \in \tau \) such that \( x \in G \) and \( f(G) \subset H \).

**Example:** Let \( \text{id} \) be the identity map. Then, 
\[ \text{id} : (\mathbb{R}, \text{metric topology}) \to (\mathbb{R}, \text{trivial}) \]
is continuous, and 
\[ \text{id} : (\mathbb{R}, \text{metric top.}) \to (\mathbb{R}, \text{discrete}) \]
is not.

**Definition 6.5.** Let \( \sigma \) and \( \tau \) be topologies on \( X \). We say that \( \sigma \) is **weaker** (or **coarser**) than \( \tau \) if \( \sigma \subset \tau \). We also say that \( \tau \) is **stronger** (or **finer**) than \( \sigma \) if \( \sigma \subset \tau \).

**Definition 6.6.** Let \( (X, \tau) \) be a topological space and \( E \subset X \). The **closure** of \( E \) is 
\[ \bar{E} := \bigcap_{F \text{ closed }, F \supset E} F \]
The interior \( E^0 \) of \( E \) is 
\[ E^0 := \bigcap_{G \in \tau, G \subset E} G \]

**Proposition 6.1.** Let \( X \neq \emptyset \) be any nonempty set, and let \( C \subset \mathcal{P}(X) \). Then there is a weakest topology \( \tau \) such that \( C \subset \tau \).

**Definition 6.7.** Let \( (X, \tau) \) be a topological space. A collection \( B \subset \tau \) is a base for \( \tau \) if whenever \( G \in \tau \) and \( x \in G \), then there exists \( H \in B \) such that 
\[ x \in H \subset G. \]

**Monday, 3-28-2005:**

**Definition 6.8.** A **local base** (or a **neighborhood base**) at \( x \in X \) is a family \( B_x \subset \tau \) such that

(a) \( \forall G \in B_x, x \in G \).

(b) If \( H \subset X \) is open and \( x \in H \), then there exists a \( G \in B_x \) such that \( G \subset H \).

**Remark:** Suppose \( B \) is a base for \( \tau \). Then, \( G \in \tau \iff \forall x \in G, \exists H \in B \) such that \( x \in H \) and \( H \subset G \).

**Definition 6.9.** A topological space is **1st countable** if every point \( x \in X \) has a countable local base.
Definition 6.10. A topological space is 2nd countable if \( X \) has a countable base.

Examples:
Suppose \( X \) is a metric space. Using balls of rational radius, we get that \( X \) is 1st countable. If, in addition, \( X \) is separable, then \( X \) is second countable.

Proposition 6.2. A family \( B \) of subsets of a set \( X \) is a base for a topology \( \tau \) on \( X \) iff

(a) Each \( x \in X \) belongs to some element of \( B \) (i.e., \( \bigcup_{H \in B} H = X \)).

(b) If \( H_1, H_2 \in B \) and \( x \in H_1 \cap H_2 \), then \( \exists H_3 \in B \) such that \( x \in H_3 \subset H_1 \cap H_2 \).

Definition 6.11. A collection \( C \subset P(X) \) is a subbase for the topology \( \tau \) if

\[
B := \{ G_1 \cap \cdots \cap G_n : n \in \mathbb{N}, G_i \in C, 1 \leq i \leq n \}
\]

is a base for \( \tau \).

Example: Let \( X = \mathbb{R} \) and \( B = \{ [a, b) : a < b \} \).

Claim: \( B \) is a base for a topology \( \tau \) -

(a) holds as \( \bigcup_{H \in B} H = \mathbb{R} \).

(b) If \( x \in H_1 \cap H_2, H_i \in B \), then

\[
x \in [a_1, b_1) \cap [a_2, b_2)
\]

and

\[
x \in [\max\{a_1, a_2\}, \min\{b_1, b_2\}) \in B.
\]

Hence, \( B \) is the base for a topology by our proposition.

This is called the “half-open interval” topology. This topology is stronger than the usual Euclidean topology on \( \mathbb{R} \) (it contains the Euclidean topology).

Notice

- The \( \frac{1}{2} \)-open interval topology is separable - the rational numbers are dense in this topology.
- This topology is first countable.
- This topology is not second countable. Suppose \( \mathcal{A} \) is a base. Given \( t \in \mathbb{R} \), there exists a \( G_t \in \mathcal{A} \) such that \( G_t \subset [t, t+1) \). Suppose \( t < s \). Then, \( t \notin [s, s+1) \), so \( G_t \notin G_s \). Thus, the map \( t \rightarrow G_t \) is injective. Hence, \( B \) is uncountable, and this topology is not second countable. Thus, \( \tau \) is not metrizable.

Friday, 4-1-2005

Example: On the Banach space \( C[0,1] \), for \( x \in [0,1] \), let

\[
\rho_x(f, g) = |f(x) - g(x)|.
\]

Note that \( \rho_x \) is a pseudometric (satisfies all the hypothesis of a metric except that \( \rho_x(f, f) \) may be zero even if \( f \neq 0 \).

Let \( B \) consist of all sets of the form

\[
G_{x_1, \ldots, x_n, \epsilon}(f) := \{ g \in C[0,1] : \rho_{x_i}(f, g) < \epsilon, i = 1, \ldots, n \}
\]


Claim: $B$ is the base for a topology. Let
\[ B_f = G_{x_1, \ldots, x_n, \epsilon_1}(f) \]
and
\[ B_g = G_{y_1, \ldots, y_m, \epsilon_2}(g) \]
and suppose that $h \in B_f \cap B_g$. Let
\[ \{z_1, \ldots, z_k\} = \{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_m\} \]
and
\[ \epsilon = \min\{\epsilon_1, \epsilon_2\} \]
Define
\[ \delta = \max_{i, j} \{\epsilon - |f(x_i) - h(x_i)|, \epsilon - |g(y_j) - h(y_j)|\} \]
Let
\[ B_h = G_{z_1, \ldots, z_n, \delta}(h) \]
If $p \in B_h$,
\[ |p(x_i) - f(x_i)| \leq |p(x_i) - h(x_i)| + |h(x_i) - f(x_i)| \leq \epsilon_1 - |f(x) - h(x)| + |h(x) - f(x)| < \epsilon_1 \]
So, $p \in B_f$. A parallel argument will show that $p \in B_g$. Hence, $B_h \subset B_f \cup B_g$. Now, if $f \in C[0, 1]$, then certainly $\in B_{0,1}(f)$. Hence,
\[ \bigcup_{B \in \mathcal{B}} B = C[0, 1] \]

**Definition 6.12.** Let $(X, \tau)$ be a topological space. Suppose $\{x_n\}$ is a sequence in $X$ and $x \in X$. We say that $x_n$ converges to $x$ if given any open set $G$ with $x \in G$, there exists $N \in \mathbb{N}$ such that $x_n \in G$ for all $n \geq N$.

**Remark:** A subbase for the topology on $C[0, 1]$ described above is the collection
\[ \{B_{x, \epsilon} : f \in C[0, 1], \epsilon > 0, x \in [0, 1]\} \]
The topology on $C[0, 1]$ that we’ve just described is also called the topology of pointwise convergence. The reason for this name is that a sequence $f_n$ converges in $C[0, 1]$ in this topology if and only if $f_n$ converges pointwise.

**Claim:**
This topology $(C[0, 1], \mathcal{B})$ is not first countable. In particular, this means that $(C[0, 1], \mathcal{B})$ is not metrizable.

**Proof.** Let $f \in C[0, 1]$ and suppose $\{B_i\}_{i=1}^\infty$ is a countable local base at $f$. For each $i$, we may find $\epsilon_i > 0$ and $x_{1,i}, x_{2,i}, \ldots, x_{n_i,i} \in [0, 1]$ such that
\[ G_{x_{1,i}, \ldots, x_{n_i,i}, \epsilon_i}(f) \subset B_i \]
Let $t \in [0, 1]$ be such that
\[ t \notin \{x_{1,1}, \ldots, x_{n_1,1}, x_{1,2}, \ldots, x_{n_2,2}, \ldots\} \]
Find a function $g_i \in C[0, 1]$ such that $|g_i(t) - f(t)| = 1$ and
\[ g_i(x_{1,i}) = f(x_{1,i}) \]
\[ \vdots \]
\[ g_i(x_{n_i,i}) = f(x_{n_i,i}) \]
Then, \( G_i \notin G_{t, \frac{1}{2}}(f) \), but
\[
g_i \in G_{x_1, \ldots, x_n, i, \epsilon_i}(f) \subset B_i
\]
Therefore, \( g_i \in B_i \setminus G_{t, \frac{1}{2}}(f) \) for all \( i \). Hence, there is no \( i \) such that \( f \subset B_i \subset G_{t, \frac{1}{2}}(f) \). Hence, \( \{B_i\} \) is not a local base at \( f \). \( \square \)

**Monday, 4-4-2005:**

**Motivation:** Recall that a set \( E \) in a metric space is closed if and only if whenever \((x_n) \subset E\) is a sequence in \( E \) and \( x_n \rightarrow x \in X \), then \( x \in E \).

**Example:** Let \( \tau \) be the topology of pointwise convergence in \( C[0,1] \).

Let 
\[
E := \{g \in [0,1] : 0 \leq g(x) \leq \int_0^1 g(x) dx > 1\}
\]
Notice that if \( g_n \) is a sequence in \( E \) and \( g_n \rightarrow g \). Then, by the Dominated Convergence Theorem, \( g \in E \). Hence, from the metric space point of view, \( E \) is closed.

However, this is not true in \( \tau \); we claim that \( 0 \in \overline{E} \). Let \( G \) be a \( \tau \)-open set such that \( 0 \in G \). Then, there exists \( \epsilon > 0 \) and \( x_1, \ldots, x_n \in [0,1] \) such that
\[
0 \in G_{x_1, \ldots, x_n, \epsilon}(0) \subset G.
\]
It is easy to see that we can find a function \( f \in E \) such that \( f \in G_{x_1, \ldots, x_n, \epsilon}(0) \) (take \( f \equiv 2 \), then modify it so that it remains continuous and \( f(x_i) = 0 \), making sure to keep the integral above 1).

Hence, \( f \in G \cap E \). Thus, \( 0 \in \overline{E} \), and \( E \) is not closed in this metric.

Hence, the set \( E \) is sequentially closed, but not actually closed!

We would like to have an analogue of the above motivation, so something like it will hold true for all topological spaces.

**Definition 6.13.** Let \( \Lambda \) be a nonempty set. A relation \( \leq \) on \( \Lambda \) is a direction if
(a) \( \forall \lambda \in \Lambda, \lambda \leq \lambda \) (reflexive).
(b) If \( \lambda_1, \lambda_2, \lambda_3 \in \Lambda \) and \( \lambda_1 \leq \lambda_2, \lambda_2 \leq \lambda_3 \), then \( \lambda_1 \leq \lambda_3 \) (transitive).
(c) If \( \lambda, \mu \in \Lambda \), then \( \exists \gamma \in \Lambda \) such that \( \lambda \leq \gamma \) and \( \mu \leq \gamma \).

The pair \((\Lambda, \leq)\) is called a **directed set**

**Examples:**
(a) \( \mathbb{Z}, \mathbb{N}, \mathbb{R}, [0, \infty) \) are all directed sets with the usual \( \leq \).
(b) Let \((X, \tau)\) be a topological space. Then, for \( x \in X \), let \( \Lambda = \{G \in \tau : x \in G\} \). We define \( G_1 \leq G_2 \) if \( G_2 \subset G_1 \). It is easy to see that the axioms for a directed set hold.

**Definition 6.14.** A **net** is a function \( f : \Lambda \rightarrow X \), where \( \Lambda \) is a directed set.

For notation, we usually write \((x_\lambda)_{\lambda \in \Lambda}\) instead of \( \{f(\lambda) : \lambda \in \Lambda\} \).

**Examples:** Any sequence is a net, where the domain is the natural numbers.

**Example:** Let \((X, \tau)\) be a topological space and let \( \Lambda \) be the family of all neighborhoods of the given point \( x \in X \) directed by reverse inclusion. For every
\( \lambda \in \Lambda \), let \( x_\lambda \in \lambda \). Then, \( (x_\lambda)_{\lambda \in \Lambda} \) is a net.

**Example:** Let \( f : [0, 1] \to \mathbb{R} \) be a function. Let \( \mathcal{P} \) be the set of all partitions of the unit interval. Given a partition \( P = \{0 = x_0 < \cdots < x_n = 1\} \), an \( n \)-tuple \( \tau = \{c_1, \ldots, c_n\} \) is compatible with \( p \) if \( c_j \in [x_{j-1}, x_j] \) for all \( j \). Define

\[
\Lambda = \{(P, \tau) : P \in \mathcal{P} \text{ and } \tau \text{ is compatible with } P \}
\]

Define \( (P_1, \tau_1) \leq (P_2, \tau_2) \) if \( P_2 \) refines \( P \). Then, \( (\Lambda, \leq) \) is a directed set. Define a net in \( \mathbb{R} \) by

\[
x_{(P, \tau)} = \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1})
\]

We would like to define convergence on nets such that the net converges if and only if the function \( f \) is Riemann Integrable.

**Definition 6.15.** Let \( (X, \tau) \) be a topological space. Let \( (x_\lambda)_{\lambda \in \Lambda} \) be a net. We say that \( (x_\lambda) \) converges to \( x \in X \) if, for all neighborhoods \( G \) of \( x \), there is a \( \lambda_0 \in \Lambda \) such that \( (x_\lambda)_{\lambda \geq \lambda_0} \subset G \). Therefore, \( x \in \bar{G} \).

**Wednesday, 4-6-2005:**

**Fact:** Let \( (X, \tau) \) be a topological space such that \( E \subset X \). Then,

\[
\bar{E} = \{x \in X : \text{ whenever } G \text{ is a neighborhood of } x, G \cap E \neq \emptyset\}
\]

**Proof.** Let \( x \in \bar{E} \). Let \( G \) be a neighborhood of \( x \). Then, \( x \notin G^c \), which is closed. Therefore, \( G^c \not\supset E \), and \( G \cap E \neq \emptyset \). \( \square \)

**Definition 6.16.** Suppose \( (X, \tau) \) is a topological space and \( (x_\lambda)_{\lambda \in \Lambda} \) is a net in \( X \). We say that \( (x_\lambda) \) is eventually in the set \( E \subset X \) if there exists \( \lambda_0 \in \Lambda \) such that \( \lambda \geq \lambda_0 \) implies \( x_\lambda \in E \).

We say that \( (x_\lambda) \) is frequently in \( E \) if for every \( \lambda \in \Lambda \), there exists \( \mu \in \Lambda, \mu \geq \lambda \) and \( x_\mu \in E \).

**Proposition 6.3.** Let \( (X, \tau) \) be a topological space. Let \( E \subset X \). Then,

\[
\bar{E} = \{x \in X : \exists \text{ a net } (x_\lambda)_{\lambda \in \Lambda} \text{ s. t. } x_\lambda \in E \forall \lambda \text{ and } x_\lambda \to x\}
\]

**Proof.** Suppose \( (x_\lambda) \) is a net in \( E \) and \( x_\lambda \to x \).

Let \( H \) be a neighborhood of \( x \). Then there exists some \( \lambda_0 \in \Lambda \) such that \( \lambda \geq \lambda_0 \) implies \( x_\lambda \in H \). Then,

\[
H \cap E \neq \emptyset
\]

Hence, by the fact from the beginning of the class, we get that \( x \in \bar{E} \).

Conversely, suppose \( x \in \bar{E} \). Let \( \Lambda \) be the set of all neighborhoods of \( x \) directed by reverse inclusion. Then, for \( \lambda \in \Lambda \), we have \( \lambda \cap E \neq \emptyset \). So, pick \( x_\lambda \in \lambda \cap E \). Then, \( (x_\lambda)_{\lambda \in \Lambda} \) is a net in \( E \).

If \( H \) is a neighborhood of \( x \), then \( H \in \bar{E} \). If \( \lambda \geq H \), then \( \lambda \subset H \). Hence, \( (x_\lambda) \) is eventually in \( H \). Therefore, \( x_\lambda \in x \).

**Proposition 6.4.** Let \( (X, \tau) \) and \( (Y, \sigma) \) be topological spaces. Then a function \( f : X \to Y \) is continuous if and only if whenever \( (x_\lambda)_{\lambda \in \Lambda} \) is a convergent net \( x_\lambda \to x \) we have \( f(x_\lambda) \to f(x) \).
Proof: Suppose $f$ is continuous and $x_\lambda \to x$ be a convergent net in $X$. Let $H \subset Y$ be a neighborhood of $f(x) \in Y$. Then, $f^{-1}(H)$ is open in $X$ and $x \in f^{-1}(H)$. Because the net $(x_\lambda)$ is eventually in $f^{-1}(H)$, we get that $f(x_\lambda)$ is eventually in $H$. Therefore, $f(x_\lambda) \to f(x)$.

Conversely, suppose $f$ is not continuous. Then, there exists an open set $H \subset Y$ such that $f^{-1}(H)$ is not open. Hence, we can find $x_\lambda \notin f^{-1}(H)$ and $x_\lambda \in \Lambda$. Then, $(x_\lambda)$ is a net in $X$ and as before $x_\lambda \to x$.

But $x_\lambda \notin f^{-1}(H)$, so that $f(x_\lambda) \notin H$ for all $\lambda$. This means that $f(x_\lambda) \not\to f(x)$.

\[\square\]

**Example:** Let $C[0,1]$ be equipped with the topology of pointwise convergence. For $f \in C[0,1]$, let

$$I(f) = \int_0^1 f(x)dx$$

Is $I$ continuous? No. Let

$$E = \{ f \in C[0,1] : 0 \leq f \leq 2 \text{ and } \int_0^1 f(x)dx \geq 1 \}.$$ 

Since $0 \in \bar{E}$, there is a net $(f_\lambda) \subset E$ that converge to $E$. However, $I(f_\lambda) \geq 1$, so $I(f_\lambda) \not\to 0 = I(0)$. Hence, $I$ is not continuous.

**Example:** Let

$$B := \{ f \in C[0,1] : \|f\|_\infty \leq 1 \}$$

It is easy to see that we can pick out an infinite sequence of functions in $B$ such that $\|f_i - f_j\| = 1$ for all $i \neq j$. Hence, $B$ is not compact under the metric induced by $\|\cdot\|_\infty$. However, we will later show that $B$ is $\tau$ compact!

**Remark:** Convergence can be used to define topologies. For example, let $X$ be a set and consider $\mathcal{P}(X)$ as a net in $\mathcal{P}(X)$, we can define the

$$\limsup E_\lambda = \bigcap_\lambda \bigcup_{\lambda \leq \mu} E_\mu$$

This is what we think it “ought” to be. Then,

$$\liminf E_\lambda = \bigcup_\lambda \bigcap_{\lambda \leq \mu} E_\mu$$

We want to say that a sequence converges if and only if $\limsup E_\lambda = \liminf E_\lambda$. This is the star topology.

**Friday, 4-8-2005:**

**Examples** of directed sets and nets:

(a) $\Lambda = \mathbb{R} \times \mathbb{R}$

$$(x_1, y_1) \leq (x_2, y_2) \iff x_1 \leq x_2 \text{ and } y_1 \leq y_2$$
(b) $\Lambda = \mathbb{R}$, $\leq$ is usual order. Let $x_\lambda = e^{-\lambda}$, $\lambda \in \mathbb{R}$. Then,
\[
\lim_{\lambda} x_\lambda = 0
\]
Notice that this is a convergent net in $\mathbb{R}$ which is not bounded. On the other hand, if a sequence in $\mathbb{R}$ is convergent, then it is always bounded.

(c) Let $\Lambda = \mathbb{R}$; define $x \leq A y \iff y \leq x$.
Define $x_\lambda = e^{-\lambda}$ as before. Then, this net is not convergent. Hence, even though the set is the same, the convergence is different depending on $\Lambda$.

**Definition 6.17.** Let $(x_\lambda)_{\lambda \in \Lambda}$ be a net in the topological space $(X, \tau)$. A point $x \in X$ is a **cluster point** for $(x_\lambda)_{\lambda \in \Lambda}$ if given any neighborhood $G$ of $x$, $(x_\lambda)_{\lambda \in \Lambda}$ is frequently in $G$.

**Definition 6.18.** Let $A$ be a directed set and let $\Lambda$ be a directed set. A function $\theta : A \to \Lambda$ is **cofinal** for $\Lambda$ if for all $\lambda_0 \in \Lambda$, there exists an $a_0 \in A$ such that whenever $a \geq a_0$, we have $\theta(a) \geq \lambda_0$.

**Definition 6.19.** Let $(x_\lambda)_{\lambda \in \Lambda}$ be a net in $X$. A subnet of $(x_\lambda)_{\lambda \in \Lambda}$ is the composition $A \to \Lambda \to X$ of the net with a cofinal function $\theta : A \to \Lambda$. We sometimes write $(x_{\lambda a})_{a \in A}$.

**Remark:** Suppose $(X, \tau)$ is a topological space and $(x_\lambda)$ is a net which converges, $x_\lambda \to x$. Then every subnet of $(x_\lambda)$ converges to $x$.

**Example:** Let $(e^{-t})_{t \in \mathbb{R}}$ (usual ordering) be the net described earlier. Then, $(e^{-n})_{n \in \mathbb{N}}$ is a subnet. Here, the map $\mathbb{N} \to \mathbb{R}$ is the inclusion map. However, $(e^{1-\frac{1}{n}})_{n \in \mathbb{N}}$ is not a subnet, as $n \mapsto 1 - \frac{1}{n}$ is not cofinal.

**Example:** Let $(e^{-n})_{n \in \mathbb{N}}$ be viewed as a net. Define $\theta : \mathbb{R} \to \mathbb{N}$ by $\theta(t) = \max\{0, [t]\}$

Then, $(e^{-\theta(t)})_{t \in \mathbb{R}}$ is a subnet.

Moral: A subnet of a sequence need not be a sequence.

**Proposition 6.5.** Let $(x_\lambda)_{\lambda \in \Lambda}$ be a net in a topological space $X$. Then $x \in X$ is a cluster point of $(x_\lambda)_{\lambda \in \Lambda}$ if and only if there exists a subnet of $(x_{\lambda_a})_{a \in A}$ which converges to $x$.

**Proof.** Suppose $(x_{\lambda_a})_{a \in A}$ is a subnet such that $x_{\lambda_a} \to x$. Then let $G$ be a neighborhood of $x$. Let $\lambda_0 \in \Lambda$. Since $a \mapsto \lambda_a$ is cofinal, there exists a $a_0 \in A$ such that whenever $a \geq a_0$, $\lambda_a \geq \lambda_0$.

Since $x_{\lambda_a} \to x$, there exists an $a_1 \in A$ such that $x_{\lambda_a} \in G$. Pick $a_2 \geq a_1$ and $a_2 \geq a_0$. Then if $a \geq a_2$, then $x_{\lambda_a} \in G$ and $\lambda_a \geq \lambda_0$.

Hence, $(x_\lambda)$ is frequently in $G$, so $x$ is a cluster point.

**Monday, 4-11-2005:**

We would now like to prove the converse. Let $\mathcal{N}$ be the collection of all neighborhoods of $x$. Let $A = N \times \Lambda$ and define $(G, \lambda) \leq_A (H, \mu)$ if and only if $\lambda \leq \Lambda \mu$ and $H \subset G$. Then $A$ becomes a directed set. Define $\theta : A \to \Lambda$ as follows:
Given \((N, \lambda) \in A\), we know that since \(x\) is a cluster point of \((x_\lambda)\), \((x_\lambda)\) is frequently in \(N\). Hence, there exists \(\mu \in \Lambda\) such that \(x_\mu \in N\) and \(\mu \geq \lambda\).

Define \(\theta(N, \lambda) = \mu\) where \(\mu\) is chosen such that \(\mu \geq \lambda\) and \(x_\mu \in N\). To see that
\[\theta : A \to \Lambda\]
is a cofinal map, let \(\lambda_0 \in \Lambda\). Then if \((N, \lambda) \geq_A (X, \lambda_0)\), we have
\[\theta(N, \lambda) \geq \lambda \geq \lambda_0.\]
Hence, \(\theta\) is a cofinal function. Hence,
\[\left(x_{\theta(N, \lambda)}\right)_{(N, \lambda) \in A}\]
is a subnet of \((x_\lambda)_{\lambda \in \Lambda}\).

Let \(U\) be any neighborhood fo \(x\), and let \(\lambda_0 \in \Lambda\) be fixed. If \((N, \lambda) \geq (U, \lambda_0)\), then \(x_{\theta(N, \lambda)} = x_\mu\), where \(\mu \in \Lambda\) satisfies \(x_\mu \in N\) and \(\mu \geq \lambda \geq \lambda_0\).

So \(x_\mu \in N \subseteq U\) and hence
\[x_{\theta(N, \lambda)} \in U, \forall (N, \lambda) \geq (U, \lambda_0)\]
Therefore,
\[\left(x_{\theta(N, \lambda)}\right)_{(N, \lambda) \in A}\]
converges to \(x\).

\[\square\]

6.1. **Separation Axioms.** (See Section 8.3 of Royden)

**Motivation:** Let \((X, \tau) = (\mathbb{R}, \text{trivial topology})\)

**Question:** Does the sequence \((\frac{1}{n})_{n \in \mathbb{N}}\) converge? If so, to what?

Yes, and \(\frac{1}{n} \to t\) for all \(t \in \mathbb{R}\).

We would like the limit to be unique!

Suppose \((X, \tau)\) is a general topological space. Let \((z_\lambda)_{\lambda \in \Lambda}\) be a net such that \(z_\lambda \to x\) and \(z_\lambda \to y\). If \(x \neq y\), then given neighborhoods \(H\) of \(x\) and \(G\) of \(2y\), we have \(G \cap H \neq \emptyset\).

Conversely, suppose that \(x \neq y\) and that whenever \(G\) is a neighborhood of \(x\) and \(H\) is a neighborhood of \(y\), then \(G \cap H \neq \emptyset\). Then, there exists a net \((x_\lambda)\) such that \(x_\lambda \to x\) and \(x_\lambda \to y\).

**Proof.** Let
\[
\Lambda = \{(G, H) : G, H \in \tau, x \in G, y \in H\}
\]
Direct \(\Lambda\) by \((G_1, H_1) \leq (G_2, H_2)\) if and only if \(G_2 \subset G_1\) and \(H_2 \subset H_1\).

Then, \(\Lambda\) is a directed set. For \((G, H) = \lambda \in \Lambda\), choose \(x_\lambda \in G \cap H\).

**Claim:** \((x_\lambda)\) converges to \(x\) and to \(y\). Let \(U\) be a neighborhood of \(x\). Then, if \((G, H) \geq (U, X)\), we have \(x_{(G,H)} \in G \cap H \subset G \subset U\). Therefore, \(x_{(G,H)} \to x\).

Similarly, we get that \(y_{(G,H)} \to y\).

\[\square\]

**Definition 6.20.** Let \((X, \tau)\) be a topological space. We say that \(X\) is **Hausdorff** or \(T_2\) if given distinct \(x, y \in X\), there are neighborhoods \(G\) of \(x\) and \(H\) of \(y\) such that \(G \cap H = \emptyset\).

**Fact:** If \((X, \tau)\) is Hausdorff, then limits are unique.

**Examples:**
1. Any metric space is Hausdorff.
2. Let \(\tau\) be the topology of pointwise convergence on \(C[0,1]\). \(\tau\) is Hausdorff.
(3) Let $\mathcal{P}$ be the family of all polynomials with real coefficients in two variables. Given a polynomial $p \in \mathcal{P}$, let

$$S_p = \{(x, y) \in \mathbb{R}^2 | p(x, y) \neq 0\}$$

Then,

$$\{S_p : p \in \mathcal{P}\}$$

forms a subbase for a topology on $\mathbb{R}^2$, called the Zariski topology.

Wednesday, 4-13-2005:

This topology is not Hausdorff. Let $G$ be a basic neighborhood of $(1, 0)$ and $H$ be a basic neighborhood of $(0, 1)$. Write

$$G \supset S_{g_1} \cap \cdots \cap S_{g_n} = S_{g_1 \cdots g_n}$$

$$H \supset S_{h_1} \cap \cdots \cap S_{h_n} = S_{h_1 \cdots h_n}$$

So,

$$G \cap H \supset S_{g_1 \cdots g_n \cdot h_1 \cdots h_n} \neq \emptyset$$

So the Zariski topology isn’t Hausdorff

**Definition 6.21.** A topological space $(X, \tau)$ is $T_1$ if given any $x \in X$ and $y \in X$, there exists an open set $U$ containing $x$ but not $y$.

**Fact:** $(X, \tau)$ is $T_1$ if and only if every point satisfies $\{x\}$ is closed.

**Example:** The Zariski topology is $T_1$. If $(x_1, y_1) \neq (x_2, y_2)$, choose $A,B$, and $C$ in $\mathbb{R}$ such that $Ax_1 + By_1 + C = 0$, yet $Ax_2 + By_2 + C \neq 0$. Then, if $p(x, y) = Ax + Bx + C$, we get $(x_2, y_2) \in S_p$ yet $(x_1, y_1) \notin S_p$.

**Definition 6.22.** If $(X, \tau)$ is $T_1$ and given any $x \in X$, $F \subset X$ closed, there exists an open set $G$ with $G \supset F$, yet $x \notin G$, we say that $(X, \tau)$ is $T_3$ (or regular).

**Definition 6.23.** If $(X, \tau)$ is $T_1$ and given any closed sets $F_1, F_2 \subset X$ with $F_1 \cap F_2 = \emptyset$, there exists open sets $G_1$ and $G_2$ with $G_i \supset F_i$ and $G_1 \cap G_2 = \emptyset$, we say that $(X, \tau)$ is $T_4$ (or normal).

**Example:** Any metric space is normal.

**Proof.** Clearly, metric spaces are $T_1$. Let $F \subset X$ be closed. Define, for $x \in X$,

$$d(x, F) = \inf\{d(x, f) : f \in F\}$$

Check that $x \mapsto d(x, F)$ is continuous.

Given $F_1, F_2$ closed, $F_1 \cap F_2 = \emptyset$, define

$$G_1 = \{x \in X : d(x, F_2) > d(x, F_1)\}$$

$$G_2 = \{x \in X : d(x, F_2) < d(x, F_1)\}$$

Then, $G_1 \supset F_1$, $G_2 \supset F_2$, and $G_1 \cap G_2 = \emptyset$.

**Example:** This is left as an exercise. Any compact $T_2$ space is normal.
Theorem 28. **Urysohn’s Lemma**

If \((X, \tau)\) is normal, then given \(F_1, F_2\) disjoint and closed, there exists a continuous function \(f : X \to [0, 1]\) such that \(f(x) = 0\) for all \(x \in F_1\) and \(f(x) = 1\) for all \(x \in F_2\).

Theorem 29. **Tietze Extension**

Let \((X, \tau)\) be normal and \(F \subseteq X\) closed. If \(f : F \to \mathbb{R}\) is continuous, then there exists a continuous function \(\tilde{f} : X \to \mathbb{R}\) such that \(\tilde{f}|_F = f\).

Here’s an application of Tietze’s Extension Theorem. Let \(Y\) be a normal space and \(X \subseteq Y\), \(X\) closed. The map \(\phi : C(Y) \to C(X)\) given by \(\phi(f) = f|_X\) is onto (by Tietze)

See problems 22 and 23 of Royden chapter 8 for outlines of the proofs of these theorems.

6.2. **Product Spaces.** We want to take Cartesian product of a bunch of topological spaces, and give them a topology in some standard way.

Let \(A\) be a non-empty set and for each \(\alpha \in A\), let \((X_\alpha, \tau_\alpha)\) be a topological space. Recall that the product space

\[
\prod_{\alpha \in A} X_\alpha = \{ f : A \to \bigcup_{\alpha \in A} X_\alpha : f(\alpha) \in X_\alpha \forall \alpha \in A \}
\]

**Example:** Suppose first that \(A = \{1, 2, \ldots, n\}\). Then,

\[
\prod_{j=1}^{n} X_j = \{ f : \{1, \ldots, n\} \to \bigcup_{j=1}^{n} X_j : f(j) \in X_j \forall j \}
\]

Write \(x_j := f(j)\). Then, we may view \(f\) as the ordered \(n\)-tuple \((x_1, \ldots, x_n)\), where \(x_j \in X_j\). Hence,

When \((X_\alpha, \tau_\alpha)\) is a topological space, we give \(\prod_{\alpha \in A} X_\alpha\) a topology from the following basis:

\[
\mathcal{B} = \{ \prod_{\alpha \in A} G_\alpha : G_\alpha = X_\alpha \text{ except for at most finitely many } \alpha, \text{ and } G_\alpha \in \tau_\alpha \forall \alpha \in A \}
\]

The proof that this is a base is very similar to the proof of that for the topology of pointwise convergence.

**Definition 6.24.** Given the product \(\prod_{\alpha \in A} X_\alpha\) and \(\beta \in A\), define the \(\beta\)th projection map

\[
\pi_{\beta} : \prod_{\alpha \in A} X_\alpha \to X_\beta
\]

by

\[
\pi_{\beta}(f) = f(\beta)
\]
Proposition 6.6. Each $\pi_\beta$ is continuous and moreover, the product topology is the smallest topology making each $\pi_\beta$ continuous.

Friday, 4-15-2005:

Example: Let $\prod_{\alpha \in A} X_\alpha = X$. Explain why a net $(x_\lambda)_{\lambda \in \Lambda}$ in $X$ converges to $x \in X$ if and only if $\forall \beta \in A$, $\pi_\beta(x_\lambda) \to \pi_\beta(x)$.

Forward direction: follows from the continuity of $\pi_\beta$.

Converse direction. Let $G$ be an open subset of $X$ such that $x \in G$. Then, there exists $G_\alpha \in X_\alpha, G_\alpha \in \tau_\alpha$ with $x \in \prod_{\alpha \in A} G_\alpha \subset G$ and, for all but finitely many $\alpha$, $G_\alpha = X_\alpha$. Let $\{\alpha_1, \ldots, \alpha_n\} = \{\alpha \in A : G_\alpha \notin X_\alpha\}$. There exists a $\lambda_0 \in \Lambda$ such that for all $i = 1, \ldots, n$ and $\lambda \geq \lambda_0$, $\pi_{\alpha_i}(x_\lambda) \in G_{\alpha_i}$.

Thus, $x_\lambda(\alpha) \in G_\alpha$ for all $\alpha \in A$, provided that $\lambda \geq \lambda_0$. Hence, $x_\lambda \in \prod G_\alpha \subset G$ for all $\lambda \geq \lambda_0$.

6.3. Compactness.

Theorem: (Riesz Representation Theorem) Motivation.

Let $X$ be a compact Hausdorff space. Let $C(X)$ be the continuous complex valued functions on $X$. If $\phi : C(X) \to \mathbb{C}$ is bounded and linear, then there exists a unique complex Borel measure $\mu$ on $X$ such that for all $f \in C(X)$,

$$\phi(f) = \int_X f d\mu.$$

Definition 6.25. Let $(X, \tau)$ be a topological space. We say that $X$ is **compact** if every open cover has a finite subcover.

Equivalently, $X$ is **compact** if and only if every collection of closed sets that have the finite intersection property has nonempty intersection.

A collection is said to have the finite intersection property if every finite subcollection has nonempty intersection.

Proof. First observe that if $\{H_\alpha\}_{\alpha \in A}$ is any family of closed sets, then

$$\bigcap_{\alpha \in A} H_\alpha = \emptyset \iff \bigcup_{\alpha \in A} H_\alpha^c = X$$

Suppose that $X$ is not compact. Then, there is an open cover $\{G_\alpha\}_{\alpha \in A}$ without a finite subcover. Consider $F_\alpha := G_\alpha^c$. Then for any finite subset, $\{\alpha_1, \ldots, \alpha_n\}$, we have

$$\bigcap_{i=1}^n F_{\alpha_i} \neq \emptyset$$

Hence, $\{F_\alpha\}_{\alpha \in A}$ has the finite intersection property. However, as $\{G_\alpha\}_{\alpha \in A}$ is an open cover, $\{F_\alpha\}_{\alpha \in A}$ has empty intersection.
For the other direction, suppose that there exists a $F := \{F_\alpha\}_{\alpha \in A}$ with the finite intersection property and $\bigcap_{F \in F} F = \emptyset$. Then, $G_\alpha := F_\alpha^c$ is an open cover which satisfies

$$\bigcup_{\alpha \in A} G_\alpha = X.$$  

However, if $\alpha_1, \ldots, \alpha_n \in A$, then

$$\bigcup_{j=1}^n G_{\alpha_j} \neq X$$

since

$$\bigcup_{j=1}^n F_{\alpha_j} \neq \emptyset$$

$\{G_\alpha\}$ is an open cover with no finite subcover, and $X$ is not compact. 

\[ \square \]

**Proposition 6.7.**

(a) Closed subsets of compact spaces are compact.

(b) Compact subsets of Hausdorff spaces are closed.

**Proof.**

(a) Let $X$ be compact. Let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of the closed set $F \subset X$. Then, $\{G_\alpha\}_{\alpha \in A} \cup \{F^c\}$ is an open cover of $X$. Now, pass to finite subcovers.

Monday, 4-18-2005:

(b) Let $x_0 \in F^c$. For each $f \in F$, there exists $G_f, H_f \in \tau$ such that $G_f \cap H_f = \emptyset$, $x_0 \in G_f$, and $f \in H_f$. Then, $\{H_f\}_{f \in F}$ is an open cover for $F$, so there exists $f_1, \ldots, f_n \in F$ such that

$$F \subset \bigcup_{j=1}^n H_{f_j}$$

Let $G = \bigcap_{j=1}^n G_{f_j}$. Then, $x_0 \in G$ and

$$G \cap \left( \bigcup_{j=1}^n H_{f_j} \right) = \emptyset,$$

so $G \subset F^c$. Then, every point of $F^c$ is interior and $F^c$ is open, i.e., $F$ is closed.

\[ \square \]

**Example:** $(\mathbb{R}, \{\emptyset, \mathbb{R}\})$. Then, $\{x\}$ is compact, but not closed.

**Proposition 6.8.** Let $X,Y$ be topological spaces and suppose $f : X \to Y$ is continuous. If $F \subset X$ is a compact set, then $f(F)$ is also compact.

**Definition 6.26.** Let $X,Y$ be topological spaces. A map $f : X \to Y$ is a **homeomorphism** if $f$ is bijective and both $f$ and $f^{-1}$ are continuous.
Proposition 6.9. Suppose that \((X, \tau)\) is a compact space and \((Y, \sigma)\) is Hausdorff. If \(f : X \to Y\) is continuous and bijective, then \(f\) is a homeomorphism.

Proof. Let \(G \subseteq X\) be open. \(G^c\) is closed, so it is compact. Then, \(f(G^c)\) is compact in \(Y\). Since \(Y\) is Hausdorff, \(f(G^c)\) is closed. Since \(f\) is bijective, \(f(G) = f(G^c)^c\), so \(f(G)\) is open. Hence, the inverse will be continuous, and \(f\) is a homeomorphism.

Proposition 6.10. Suppose \((X, \tau)\) is compact and Hausdorff.

- Suppose \(\tau_1 \subset \tau\) is a topology. Then, \((X, \tau_1)\) is compact but not Hausdorff.
- Suppose \(\tau \subset \tau_2\) is a topology. Then, \((X, \tau_2)\) is Hausdorff but not compact.

Proof. Consider \(Id : (X, \tau) \to (X, \tau_1)\), the identity map. By hypotheses, \(Id\) is continuous, so \((X, \tau_1)\) is compact. If \((X, \tau_1)\) was Hausdorff, then \(Id\) is a homeomorphism between \((X, \tau)\) and \((X, \tau_1)\). However, because of the proper containment, \(Id^{-1}\) is not continuous.

For 2, note that \(Id : (X, \tau_2) \to (X, \tau)\) is continuous (as \(\tau \subset \tau_2\)). Also, \((X, \tau_2)\) is Hausdorff since \((X, \tau)\) is. If \((X, \tau_2)\) is compact, then again \(Id : (X, \tau_2) \to (X, \tau)\) is a homeomorphism which is impossible as \(\tau_2 \supset \tau\).

Definition 6.27. A point \(x \in X\) is a limit point of a set \(E \subset X\) if for all \(\epsilon > 0\), \(B_\epsilon(x) \cap (E \setminus \{x\}) \neq \emptyset\).

Theorem 30. A metric space is compact if and only if every infinite subset has a limit point.

Proof. (Sketch of proof of backwards direction):
Assume every infinite subset has a limit point.
Show \(X\) is separable (use hypotheses show that we can cover with a finite number of \(\delta\) balls for each \(\delta\), and then union over all the balls of \(\delta = \frac{1}{k}\) with \(k \in \mathbb{N}\).
Now, show that any separable metric space is second countable (done already).
Notice that step 2 implies that every open cover of \(X\) has a countable subcover. If not compact, then there exists a countable cover with no finite subcover. Suppose \(\mathcal{F}\) is a family of closed subsets of \(X\) with the finite intersection property. Suppose that \(\bigcap_{F \in \mathcal{F}} F = \emptyset\). Then,

\[\{F^c : F \in \mathcal{F}\}\]

is an open cover. Pass to a countable subcover \(F_n \in \mathcal{F}\) and

\[\bigcup_{i=1}^{\infty} F_n^c = X.\]

For \(k \in \mathbb{N}\), let \(x_k \in F_1 \cap F_2 \cap \cdots \cap F_k\). Then \(\{x_k\}_{k=1}^{\infty}\) is an infinite set (otherwise \(\exists x \in \bigcap_{k=1}^{\infty} F_k\)). Hence, \(\{x_k\}\) is a convergent subsequence; say \(x_{k_j} \to x \in X\).
Fix \(\epsilon > 0\). Then there exists an \(N \in \mathbb{N}\) such that \(d(x_{k_j}, x) < \epsilon\) if \(k_j \geq N\). So, if \(k_j > N_1 > N\), we have

\[x_{k_j} \in \bigcap_{i=1}^{k_j} F_i\]

and \(x_{k_j} \in B_\epsilon(x)\). Therefore, \(x \in F_{N_1} \cap B_\epsilon(x)\) and \(x \in \bigcap_{n=1}^{\infty} F_n\), a contradiction. \(\square\)
Wednesday, 4-20-2005:

Definition 6.28. Let \((X, \tau)\) be a topological space, and pick \(Y \subset X\). Then,

\[
\tau_Y := \{G \cap Y : G \in \tau\}
\]

Then \((Y, \tau_Y)\) is a topological space; \(\tau_Y\) is called the relative or subspace topology.

Definition 6.29.

- A topological space is countably compact if every countable open cover has a finite subcover.
- \(X\) has the Bolzano-Weierstrass property if every sequence \((x_n)_{n=1}^{\infty}\) in \(X\) has a cluster point.
- \(X\) is sequentially compact if every sequence has a convergent subsequence.

Remarks:

- Compact and sequentially compact are equivalent for metric spaces.
- Compactness implies countably compact.
- Sequentially compact implies the Bolzano-Weierstrass property.
- Countably compact and Bolzano-Weierstrass property are equivalent by Proposition 7 of Royden.

Theorem 31. For a topological space, \((X, \tau)\), the following are equivalent:

(a) \(X\) is compact.
(b) Every net in \(X\) has a cluster point.
(c) Every net in \(X\) has a convergent subsequence.

Proof. Note that (b) \(\iff\) (c) was done during our study of nets. Suppose \((X, \tau)\) is compact. Let \((x_\lambda)_{\lambda \in \Lambda}\) be a net in \(X\). Let

\[
T_\lambda := \{x_\mu : \mu \geq \lambda\}.
\]

Notice that the collection \(\{T_\lambda : \lambda \in \Lambda\}\) has the finite intersection property; indeed if \(\lambda_1, \ldots, \lambda_j \in \Lambda\), we can find \(\lambda_0 \geq \lambda_i, 1 \leq i \leq j\) so

\[
T_{\lambda_0} \subset \bigcap_{i=1}^{j} T_{\lambda_i}.
\]

Thus,

\[
\bigcap_{\lambda \in \Lambda} \bar{T}_\lambda \neq \emptyset.
\]

Let \(x \in \bigcap_{\lambda \in \Lambda} \bar{T}_\lambda\) and let \(U\) be any neighborhood of \(x\). Given \(\lambda_0 \in \Lambda, x \in \bar{T}_{\lambda_0}\), so in particular, \(U \cap T_{\lambda_0} \neq \emptyset\). Therefore, there exists a \(\lambda \geq \lambda_0\) such that \(x_\lambda \in U, i.e. (x_\lambda)\) is frequently in \(U\) and \(x\) is a cluster point of \((x_\lambda)\).

Now, suppose that \(X\) is not compact. Let \(\{U_\alpha\}_{\alpha \in A}\) be an open cover with no finite subcover. Let \(A\) be the set of all finite subsets of \(A\). Make \(A\) into a directed set using inclusion; i.e., if \(F, G \in A\) then \(F \leq G \iff F \subset G\). For each \(F \in \Lambda, X \neq \bigcup_{\alpha \in F} U_\alpha\), so we may choose \(x_F \in X \setminus \bigcup_{\alpha \in F} U_\alpha\).
Then \( \{x_F\}_{F \in \Lambda} \) is a net in \( X \). Given \( y \in X \), we’ll show \( y \) is not a cluster point of \( \{x_F\}_{F \in \Lambda} \). Find \( \alpha_0 \in A \) with \( y \in U_{\alpha_0} \). If \( F \geq \{\alpha_0\} \), i.e., \( \alpha_0 \in F \), then \( x_F \notin U_{\alpha_0} \). Thus, \( \{x_F\}_{F \in \Lambda} \) is not frequently in \( U_{\alpha_0} \). Therefore, \( y \) is not a cluster point for \( \{x_F\}_{F \in \Lambda} \).

\[ \square \]

Friday, 4-22-2005:

**Definition 6.30.** A partially ordered set is a set with a partial order, i.e., a relation \( \leq \) on \( P \) such that \( x \leq x \) and \( x \leq y \), \( y \leq z \) implies \( x \leq z \) (perhaps also need that \( x \leq y \), \( y \leq x \) implies \( x = y \)).

**Definition 6.31.** A chain is a totally ordered subset of \( P \), i.e., for any \( c_1, c_2 \in C \), we have either \( c_1 \leq c_2 \) or \( c_2 \leq c_1 \).

**Definition 6.32.** An element \( z \in P \) is a maximal element if \( y \geq z \) implies \( y = z \).

**Lemma 6.1.** Zorn’s Lemma:

Let \( P \) be a non-empty partially ordered set. Suppose that every chain \( C \in P \) has an upper bound in \( P \). Then, \( P \) has a maximal element.

**Theorem 32.** Tychanoff:

The product of compact spaces is compact.

**Proof.** (Paul Chernoff, 1992)

Let \( (x_\lambda)_{\lambda \in A} \) be a net in \( X := \prod_{\alpha \in A} X_\alpha \). We will prove that \( (x_\lambda) \) has a cluster point.

For each subset \( D \subset A \), let

\[ \pi_D : \prod_{\alpha \in A} X_\alpha \to \prod_{\alpha \in D} X_\alpha \]

be the restriction map; \( \pi_D(f) = f|_D \). For each \( D \subset A \), let

\[ C_D := \{ p \in \prod_{\alpha \in D} X_\alpha | p \text{ is a cluster point for } (\pi_D(x_\lambda))_{\lambda \in A} \} \]

and

\[ P = \bigcup_{D \subset A} C_D \]

Note that \( P \neq \emptyset \), for if \( \alpha_0 \in A \) and \( D = \{\alpha_0\} \) then the compactness of \( X_{\alpha_0} \) yields \( C_{\{\alpha_0\}} \neq \emptyset \). For \( p, q \in P \), define \( p \leq q \) if and only if the domain of \( p \) is a subset of the domain of \( q \) and \( q \) restricted to the domain of \( p \) is equal to \( p \).

Then \( P \) is a partially ordered set. Let \( C \subset P \) be a chain in \( P \). For \( p \in C \), let \( D_p \) be the domain of \( p \) and let \( D^* = \bigcup_{p \in C} D_p \). Let

\[ p^* \in \prod_{\alpha \in D^*} X_\alpha \]

be defined by

\[ p^*(\alpha) = p(\alpha) \]
if \( p \in C \) and \( \alpha \in D_p \). Then \( p^* \) extends every element of \( C \) to \( D^* \). We'll show that \( p^* \in C_{D^*} \), i.e., \( p^* \) is a cluster point for \( \pi_{D^*}(x_\lambda) \). Let \( U \in \prod_{\alpha \in D^*} X_\alpha \) be a neighborhood of \( p^* \). There exist \( \alpha_1, \ldots, \alpha_n \in D^* \) and open sets \( G_\alpha \subset X_\alpha \) such that \( p^* \in \prod_{\alpha \in D^*} G_\alpha \subset U \) and \( G_\alpha = X_\alpha \) if \( \alpha \notin \{\alpha_1, \ldots, \alpha_n\} \).

We may find \( p_1, \ldots, p_n \in C \) such that \( \alpha_j \in D_{p_j} \) for \( 1 \leq j \leq n \). Hence, there exists \( p \in C \) with \( \alpha_1, \ldots, \alpha_n \in D_p \). But \( p \) is a cluster point of \( \pi_{D_p}(x_\lambda) \), so \( \pi_{D_p}(x_\lambda) \) is frequently in \( \prod_{\alpha \in D^*} G_\alpha \). Hence, \( \pi_{D^*}(x_\lambda) \) is frequently in \( \prod_{\alpha \in D^*} G_\alpha \), and \( p^* \) is a cluster point for \( \pi_{D^*}(x_\lambda) \). We can then conclude that \( p^* \in C_{D^*} \), so \( p^* \in \mathcal{P} \). As \( p^* \) is an upper bound for the chain, we get that every chain in \( \mathcal{P} \) has an upper bound in \( \mathcal{P} \).

By Zorn’s Lemma, \( \mathcal{P} \) contains a maximal element, \( q \). We claim that the domain of \( q, D_q \), is all of \( A \).

If not, let \( \gamma \in A \setminus D_q \). We know that \( q \) is a cluster point for \( \pi_{D_q}(x_\lambda) \), and there is a subnet \( \pi_{D_q}(x_{\lambda_\gamma}) \to q \). Since \( X_\gamma \) is compact, there exists a subnet of \( x_{\lambda_\gamma}(\gamma) \) that converges. Let \( y \) be the limit of this subnet of \( x_{\lambda_\gamma}(\gamma) \). Define \( \bar{q} \) on \( D_q \cup \{\gamma\} \) by \( \bar{q}|_{D_q} = q \) and \( \bar{q}(\gamma) = y \in X_\gamma \). Then \( \bar{q} \neq q \), \( \bar{q} \in C_{D_q \cup \{\gamma\}} \), and \( \bar{q} \geq q \), contradicting the fact that \( q \) is maximal in \( \mathcal{P} \). Therefore, \( A = D_q \), and \( q \) is a cluster point of \( (x_\lambda) \). As \( (x_\lambda) \) is an arbitrary net, \( X \) is compact.

\[
\square
\]

\[\text{Monday, 4-25-205:} \]

Let \( I \) be a positive linear functional on \( C(X) \) (\( X \) is locally compact and \( T_2 \)), i.e., \( I : C(X) \to \mathbb{F} \) such that \( I(f) \geq 0 \) whenever \( f \geq 0 \).

For an open set \( G \subset X \), consider

\[
\sup\{I(f)|\text{supp}(f) \subset G \text{ and } 0 \leq f(x) \leq 1\}
\]

**Fact:** This supremum is bounded if \( G \) is compact. Show that

\[
\mu^*(E) := \inf \{\mu_0(G) : G \supset E, G \text{ open} \}
\]

is an outer measure. Then use Caratheodory to obtain a measure \( \mu \), and finally show that

\[
I(f) = \int f \, d\mu
\]

**Definition 6.33.** Baire measurable sets are the smallest \( \sigma \)-algebra such that each function \( f \in C_0(X) \) is measurable.

**Theorem 33.** If \( X \) is locally compact and \( C_0(X) \) is the space of all continuous functions vanishing at \( \infty \). Then \( C_0(X) \) may be identified with the Baire measures on \( X \); i.e., given a linear functional \( f \in C_0(X) \)

\[\text{7. Stone-Weierstrass Theorem}\]

**Theorem 34.** **Stone-Weierstrass** Let \( X \) be a compact Hausdorff space. Suppose \( \mathfrak{A} \subset C_0(X) \) is an algebra s.t.

\[\text{(i) } \mathfrak{A} \text{ separates points of } X \]
(ii) $\mathfrak{A}$ is norm closed; $\|f\| = \sup_{x \in X} |f(x)|$

Then, either $\mathfrak{A} = \mathcal{C}_R(X)$ or there exists $x_0 \in X$ such that

$$\mathfrak{A} = \{ f \in \mathcal{C}_R(X) : f(x_0) = 0 \}$$

Moreover, if $1 \in \mathfrak{A}$, then $\mathfrak{A} = \mathcal{C}_R(X)$.

**Definition 7.1.** A subset $\mathfrak{A} \subset \mathcal{C}_R(X)$ is an algebra if

(i) $\mathfrak{A}$ is a vector space (over $\mathbb{R}$).

(ii) If $f, g \in \mathfrak{A}$, then $f \cdot g \in \mathfrak{A}$.

**Definition 7.2.** A subset $\mathfrak{A} \subset \mathcal{C}_R(X)$ separates points of $X$ if whenever $x, y \in X$, then there exists a $f \in \mathfrak{A}$ such that $f(x) \neq f(y)$.

**Definition 7.3.** A subset $\mathcal{L} \subset \mathcal{C}_R(X)$ is a lattice if whenever $f, g \in \mathcal{L}$, then both $f \lor g := \max\{f, g\}$ and $f \land g := \min\{f, g\}$ belong to $\mathcal{L}$.

Two convenient formulas for the max / min are:

$$\max\{f, g\} = \frac{f + g}{2} + \frac{|f - g|}{2}$$

$$\min\{f, g\} = \frac{f + g}{2} - \frac{|f - g|}{2}$$

Notice that if $\mathcal{L}$ is a lattice, then $\mathcal{L}$ is a lattice, then $f_n \to f$ and $g_n \to g$ implies that $f_n \lor g_n \to f \lor g$ and $f_n \land g_n \to f \land g$. Similarly, if $\mathfrak{A} \subset \mathcal{C}_R(X)$ is an algebra, then $\mathfrak{A}$ is also an algebra.

**Example:** Let $X = \{1, 2\}$ be a 2-point space with the discrete topology. Then $\mathcal{C}_R(X) = \mathbb{R}^2$ and the only subalgebras of $\mathcal{C}_R(X)$ are $\mathbb{R}^2$, $\{(0, 0)\}$, $\{(t, 0) : t \in \mathbb{R}\}$, $\{(0, t) : t \in \mathbb{R}\}$, $\{(t, t) : t \in \mathbb{R}\}$.

Each of these are subalgebras. Suppose $\mathfrak{A} \subset \mathbb{R}^2$ is a subalgebra. Then, $\mathfrak{A} \neq \{(0, 0)\}$. Let $(a, b) \in \mathfrak{A}$, $(a, b) \neq (0, 0)$. Then if $a = 0$ or if $b = 0$, then $\mathfrak{A}$ is in the list. So suppose $a, b$ are both non-zero. If $a = b$, then $\mathfrak{A}$ is in the list. If $a \neq 0$, $b \neq 0$, and $a \neq b$, then $(a, b)$ and $(a^2, b^2)$ are linearly independent, so $\mathfrak{A} = \mathbb{R}^2$.

*Wednesday, 4-27-2005:*

**Proposition 7.1.** Let $g(x) = |x|$ for $x \in [-1, 1]$. Then given $\epsilon > 0$, there exists a polynomial $p(t)$ such that $p(0) = 0$ and $|p(x) - g(x)| < \epsilon$ for all $x \in [-1, 1]$.

**Proof.** See your Math 825 Notes. □

**Lemma 7.1.** Suppose $\mathfrak{A}$ is a closed subalgebra of $\mathcal{C}_R(X)$. If $f \in \mathfrak{A}$, so does $|f|$. Moreover, $\mathfrak{A}$ is also a lattice.

**Proof.** If $f \in \mathfrak{A}$ and $f \neq 0$, then let

$$h = \frac{f}{\|f\|}$$

Then $h(X) \subset [-1, 1]$. Given $\epsilon \geq 0$, we may find a polynomial $p$ with $p(0) = 0$ such that $||t| - p(t)| < \epsilon$ for all $t \in [-1, 1]$. Hence, $\forall x \in X$, $|h(x)| - p(h(x))| < \epsilon$. 


However, \( p(h(x)) \) belongs to \( A \) because \( A \) is an algebra. Therefore, \(|h(x)|\) may be uniformly approximated by elements of \( A \), and hence, \( A \) is closed so \(|h(x)| \in A \). Then, \( f(x) = \|f\| h(x) \in A \).

Finally, if \( f, g \in A \), then
\[
\begin{align*}
f \lor g &= \frac{f + g + |f - g|}{2} \\
f \land g &= \frac{f + g - |f - g|}{2}
\end{align*}
\]

**Lemma 7.2.** Suppose \( A \) is a closed lattice in \( C_\mathbb{R}(X) \) and suppose \( f \in C_\mathbb{R}(X) \). If for every \( x, y \in X \), there exists a \( g_{x,y} \in A \) such that \( g_{x,y}(x) = f(x) \) and \( g_{x,y}(y) = f(y) \), then \( f \in A \).

**Proof.** Let \( \epsilon > 0 \) and let
\[
U_{x,y} = \{ z \in X : f(z) < g_{x,y}(z) + \epsilon \} \\
V_{x,y} = \{ z \in X : g_{x,y}(z) - \epsilon < f(z) \}
\]
Then, \( U_{x,y} \) and \( V_{x,y} \) are both open in \( X \) and
\[\{x,y\} \subset U_{x,y} \cap V_{x,y}\]
Fix \( y \). Then, \( \{U_{x,y}\}_{x \in X} \) is an open cover for \( X \).

We may find \( x_1, \ldots, x_n \in X \) such that \( X = \bigcup_{j=1}^{n} U_{x_j,y} \). Let
\[
g_y = \max\{g_{x_1,y}, \ldots, g_{x_n,y}\}
\]
As \( A \) is a lattice, \( g_y \in A \). Given \( x \in X \), \( x \in U_{x_j} \), for some \( j \), so
\[f(x) < g_{x_j,y}(x) + \epsilon \leq g_y(x) + \epsilon\]
Put
\[
V_y = \bigcap_{j=1}^{n} V_{x_j,y}
\]
Then, for \( z \in V_y \),
\[g_{x_j,y}(z) - \epsilon < f(z)\]
Also, \( V_y \) is open in \( X \), \( y \in V_y \), and
\[g_y(z) - \epsilon < f(z), \forall z \in V_y\]
Since \( \{V_y\}_{y \in X} \) is an open cover for \( X \), find \( y_1, \ldots, y_m \) such that \( X = \bigcup_{j=1}^{m} V_{y_j} \). Let
\[
g = \min\{g_{y_1}, \ldots, g_{y_m}\} \in A
\]
Then \( g \in A \) and \( x \in X \), so \( f(x) < g(x) + \epsilon \).
As before, if \( x \in X \) and \( x \in V_{y_j} \) for some \( j \), so \( g_{y_j}(x) - \epsilon < f(x) \), so \( g(x) - \epsilon < f(x) < g(x) - \epsilon \). Hence, we have found a \( g \in A \) with
\[|f(x) - g(x)| < \epsilon\]
for all \( x \in X \). As \( \epsilon > 0 \) is arbitrary, we find that \( f \in A \). \(\square\)
Theorem 35. **Stone-Weierstrass:**

Let \( X \) be compact and Hausdorff, \( \mathcal{A} \subset C_\mathbb{R}(X) \) a closed subalgebra of \( C_\mathbb{R}(X) \) such that \( \mathcal{A} \) separates points. Then either \( \mathcal{A} = C_\mathbb{R}(X) \), or there exists a unique \( x_0 \in X \) such that

\[
\mathcal{A} = \{ f \in C_\mathbb{R}(X) : f(x_0) = 0 \}
\]

Also, \( \mathcal{A} = C_\mathbb{R}(X) \) if and only if \( 1 \in \mathcal{A} \).

**Proof.** Suppose that \( x \neq y \) and \( x, y \in X \). Put

\[
\mathcal{A}_{x,y} = \{ (f(x), f(y)) : f \in \mathcal{A} \}
\]

Then \( \mathcal{A}_{x,y} \) is a subalgebra of \( \mathbb{R}^2 \) (because \( \mathcal{A} \) is an algebra). If \( \mathcal{A}_{x,y} = \mathbb{R}^2 \) for all \( x, y \in X \) with \( x \neq y \), then given \( f \in C_\mathbb{R}(X) \), \( (f(x), f(y)) \in \mathbb{R}^2 \) so there exists \( g_{x,y} \in \mathcal{A} \) such that \( (g_{x,y}(x), g_{x,y}(y)) = (f(x), f(y)) \). Since \( \mathcal{A} \) is a closed algebra, it is also a lattice, we then have \( f \in \mathcal{A} \). Hence, \( \mathcal{A} = C_\mathbb{R}(X) \). If \( \mathcal{A}_{x,y} \neq \mathbb{R}^2 \) for some \( x, y \in X \) and \( x \neq y \). Then, \( \mathcal{A}_{x,y} = \{(0, t), t \in \mathbb{R} \} \) or \( \mathcal{A}_{x,y} = \{(t, 0) | t \in \mathbb{R} \} \). Note that \( \mathcal{A}_{x,y} \neq \emptyset \) because \( \mathcal{A} \) separates points. Hence, \( \exists x_0 \in X \) such that \( f(x_0) = 0 \) for all \( f \in \mathcal{A} \). Hence, \( x_0 \) is unique by the separation of points.

Let \( f \in C_\mathbb{R}(X) \) with \( f(x_0) = 0 \), then we want \( f \in \mathcal{A} \). If \( x, y \in X \) and \( x \neq y \). If \( x_0 \notin \{x, y\} \) then \( \mathcal{A}_{x,y} = \mathbb{R}^2 \) so \( (f(x), f(y)) \in \mathcal{A}_{x,y} \). If \( x_0 \in \{x, y\} \) then \( \mathcal{A}_{x,y} = \{(t, 0) \} \) or \( \{(0, t) \} \). In either case, there exists \( g_{x,y} \in \mathcal{A} \) such that \( (g_{x,y}(x), g_{x,y}(y)) = (f(x), f(y)) \). The lemma shows that \( f \in \mathcal{A} \). \( \square \)