REAL ANALYSIS QUALS - HELPFUL INFORMATION

EVAN WARNER

1. Set theory

This stuff is pretty much all easy. Just recount the following, which is equivalent to the axiom of choice:

Theorem 1.1 (Zorn's Lemma)**.** *Let X be a poset for which every totally ordered subset has an upper bound. Then X has a maximal member.*

Zorn's Lemma is often used when X is the collection of subsets of a given set satisfying a given property to help show that there is a maximal collection of subsets satisfying that property.

2. Topology

Not going to review all of topology here, just some important bits for analysis.

2.1. **Basic properties: separation, countability, separability, compactness, connectedness.** A space X is *Tychonoff* if for $u, v \in X$, there exists a neighborhood of u not containing v (and by symmetry a neighborhood of v not containing u). It is *Hausdorff* if every two points can be separated by disjoint neighborhoods. It is *regular* if it is Tychonoff and every closed set and point not in the set can be separated by disjoint neighborhoods. It is *normal* if it is Tychonoff and every two disjoint closed sets can be separated by disjoint neighborhoods. To be Tychonoff precisely means that every singleton set is closed.

Every metric space is normal (easy to see using distance functions). If a space is Tychonoff, it is normal if and only if for U a neighborhood of a closed set F there is a neighborhood O such that

$$F \subseteq O \subseteq \overline{O} \subseteq U.$$

A topology is *first countable* if there is a countable base at each point and *second countable* if there is a countable base for the whole topology. Every metric space is clearly first countable. For first countable spaces, being closed is equivalent to being sequentially closed. A topology is *separable* if it has a countable dense subset.

Given a set X and a collection of mappings $f_\lambda : X \to X_\lambda$, where the X_λ are topological spaces, we can form the *weak topology* induced by the collection, which is defined to be the smallest topology such that all of the f_λ are continuous. The collection of all inverse images of open sets under all the mappings is a subbase for this topology.

A topology is *compact* if every open cover has a finite subcover. If we say that a collection of sets has the *finite intersection property* provided every finite subcollection has nonempty intersection, compactness is equivalent to the property that every collection of closed subsets having the finite intersection property has nonempty intersection. A closed subset of a compact space is clearly compact; conversely, a

compact subspace of a Hausdorff space is closed (use Hausdorff property to find a cover of the space).
Facts about compactness: for a second countable space, compactness is equivalent to sequential compactness. A compact Hausdorff space is normal (bootstrapping up to regular first). The continuous image of a compact space is compact; therefore a continuous real-valued function from a compact space takes a maximum and a minimum.
For Hausdorff topological spaces X and Y, we can define the *compact-open* topology on the set of continuous from $X \to Y$ by taking as a subbase sets of the form

$$U_{K,O} = \{f : X \to Y | f(K) \subseteq O\}$$

where K is compact in X and O is open in Y. A sequence of maps converging in the compact-open topology converges pointwise. Furthermore if Y is a metric space a sequence converges if and only if it converges uniformly on compact subsets.

Theorem 2.1 (Dini's theorem)**.** *Let $\{f_n\}$ be a sequence of continuous real-valued functions on a countable compact space such that $f_n(x)$ increases monotonically, converging to $f(x)$. Then f is continuous.*

Proof sketch. Given $\epsilon > 0$, let E_n be the set of all points for which $f_n(x)$ is strictly ϵ-close to $f(x)$. The E_n are open, (by monotonicity) increasing, and cover the space, so by countable compactness there is an N such that E_N is the whole space. □

A space is *connected* if there are no clopen sets except for the empty set and the whole space. The continuous image of a connected space is connected. If we let the *intermediate value property* mean that the image of every continuous real-valued function is an interval, then a topological space has the intermediate value property if and only if it is connected.

2.2. **Big theorems.** For metric spaces, we know we have an abundance of continuous real-valued functions taking prescribed values by using distance functions. Normal spaces also have similarly abundant continuous maps:

Theorem 2.2 (Urysohn's lemma)**.** *Let A and B be disjoint closed subsets of a normal topological space. Then for any real numbers a, b, there is a continuous real-valued function defined on X that takes values in $[a, b]$, such that $f = a$ on A and $f = b$ on B.*

Proof sketch. Take $a = 0$, $b = 1$ for simplicity. Using normality, we can construct a lot of neighborhoods "interpolating" between A and B such that the closure of one is contained in any larger one. Specifically, we construct a neighborhood O_λ for every dyadic rational λ between 0 and 1. Then we define $f(x) = 1$ on the complement of the union of all the O_λ and

$$f(x) = \inf\{\lambda \text{ a dyadic rational} : x \in O_\lambda\}$$

otherwise, which is what we want. □

This is sort of an extension result, saying we can take the obviously continuous mapping from the unconnected space $A \cup B$ and extend it to a continuous mapping from the whole space. The following is a strengthening:

Theorem 2.3 (Tietze extension theorem)**.** *Let X be normal, F a closed subset, and f a continuous real-valued function on F that takes values in the bounded interval $[a, b]$. Then f has a continuous extension to all of X that also takes values in $[a, b]$.*

Proof sketch. Suffices to consider $[-1,1]$. We want to construct a sequence of functions g_n that are bounded $|g_n| \le (2/3)^n$ on X and whose sum converges uniformly to f on F. If we can do this, the infinite sum is well-defined, continuous, bounded by 1 in absolute value, and equals f on F, so it is what we want. We find the g_n inductively, using Urysohn at each step, letting A be the closed set where $h = f - [g_1 + \ldots + g_{n-1}]$ is less than, say, $-a_n = -(1/3)(2/3)^{n-1}$ and B be the closed set where h is larger than a_n. □

There is an easy extension to unbounded real-valued functions, by taking the function $f/(1+|f|)$, extending, and mapping back.

Theorem 2.4 (Urysohn metrization theorem). *Let X be second countable. Then X is metrizable if and only if it is normal.*

The proof uses Urysohn's lemma and is relatively unenlightening.
Recall that the product topology is defined to have a base consisting of products of open sets where cofinitely many terms are the whole space. It is the weak topology with respect to the collection of all projections. Convergence of a sequence with respect to the product topology is pointwise convergence (that is, convergence on each element of the indexing set). An arbitrary product of Hausdorff spaces is Hausdorff; a product of Tychonoff spaces is Tychonoff.

Theorem 2.5 (Tychonoff product theorem). *An arbitrary Cartesian product of compact spaces is compact.*

The proof is not at all important for our purposes.
For the next theorem, an *algebra* of continuous real-valued functions is a linear subspace of the continuous functions that is closed under multiplication. An algebra *separates points* if for any two points, there is an element of the algebra mapping them to different values (the algebra can "distinguish between the two points").

Theorem 2.6 (Stone-Weierstrass theorem). *Let X be compact Hausdorff. Suppose $\mathcal{A}$ is an algebra of continuous real-valued functions on X that contains the constant functions and separates points. Then $\mathcal{A}$ is dense in $C(X)$.*

The method of proof is not terribly relevant. The Urysohn lemma will have to be used at some stage (which is permissible because compact and Hausdorff implies normal). Noting that the polynomials on a closed bounded interval form an algebra that separates points and contains the constant functions, we have the following classical theorem:

Theorem 2.7 (Weierstrass approximation theorem). *For $f : [a,b] \to \mathbb{R}$ and any $\epsilon > 0$, there is a polynomial p such that*

$$|f(x) - p(x)| < \epsilon$$

for all $x \in [a,b]$.

Another consequence is the following:

Theorem 2.8 (Riesz's theorem). *Let X be compact Hausdorff. Then $C(X)$ is separable iff X is metrizable.*

Proof sketch. If X is metrizable, it is separable, so choose a countable dense subset $\{x_n\}$ and let $f_n(x) = \rho(x, x_n)$. The set of polynomials with real coefficients in

finitely many of the f_k satisfies the hypotheses of Stone-Weierstrass, and is countable.
Conversely, if $C(X)$ is separable, let $\{g_n\}$ be a countable dense subset, and let $O_n = \{x \in X | g_n(x) > 1/2\}$. Using normality, Urysohn's lemma, and the denseness of g_n, we can verify that this is a base of the topology on X. Therefore X is second countable and, by the Urysohn metrization theorem, metrizable. □

3. Metric spaces

3.1. **Basics.** Everyone knows what a metric space is. All normed linear spaces are metric spaces with the obvious metric. Two metrics on the same space are *equivalent* if they agree up to universal multiplicative constants; i.e. ρ and σ are equivalent if there exist positive c_1, c_2 such that

$$c_1 \cdot \sigma(x_1, x_2) \leq \rho(x_1, x_2) \leq c_2 \cdot \sigma(x_1, x_2)$$

for all points x_1, x_2 in space. If we have a *pseudometric* (a metric ρ where $\rho(x, y) = 0$ does not necessarily imply $x = y$) then the relation $\rho(x, y) = 0$ is an equivalence relation on the space, and the metric descends to the corresponding quotient space, forming a legitimate metric space. This is what happens, for example, in the construction of the L^p spaces as metric spaces.
A metric space determines a topology with basis

$$B(x, r) = \{x' \in X : \rho(x, x') < r\}.$$

Equivalent metrics give rise to equivalent topologies. For a metric space, $\lim x_n \to x$ in the topology if and only if $\lim \rho(x_n, x) \to 0$. For a metric space, a function is continuous if and only if it is sequentially continuous; i.e. if $x_n \to x$ then $f(x_n) \to f(x)$. A *uniformly continuous* map $f : (X, \rho) \to (Y, \sigma)$ is such that for all $\epsilon > 0$, there is a $\delta > 0$ such that for all $u, v \in X$ we have

$$\rho(u, v) < \delta \implies \sigma(f(u), f(v)) < \epsilon.$$

For a *Lipschitz* map, there exists a constant c such that

$$\sigma(f(u), f(v)) \leq c \cdot \rho(u, v).$$

We clearly have Lipschitz $\implies$ uniformly continuous $\implies$ continuous; the reverse implications are false. Consider $f : [0, 1] \to [0, 1]$ given by $f(x) = \sqrt{x}$, which is uniformly continuous but not Lipschitz, and $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^2$, which is continuous but not uniformly continuous.

3.2. **Complete metric spaces.** A *complete* metric space is one where every Cauchy sequence converges to a point. A subset of a complete metric space is complete if and only if it is closed.

Theorem 3.1 (Cantor intersection theorem)**.** *Let X be a metric space. Then X is complete if and only if whenever $\{F_n\}$ is a sequence of closed sets such that*

$$\lim_{n \to \infty} \operatorname{diam}(F_n) = 0,$$

there is a point $x \in X$ such that $\bigcap_{n=1}^{\infty} F_n = \{x\}$.

Proof sketch. Straightforward. The converse uses the fact that if a subsequence of a Cauchy sequence converges if and only if the whole sequence converges. □

This is not true if the diameters do not converge to zero. Counterexample: $F_n = [n, \infty)$ as subsets of $\mathcal{R}$ have empty intersection.
We can complete any metric space canonically in the usual fashion by taking equivalence classes of sequences of points and quotienting by the equivalence relation of sequences whose distance goes to zero.

3.3. **Compact metric spaces.** A *compact* metric space is one that is compact as a topological space; i.e., every open cover has a finite subcover. As before, this is the case, by de Morgan's identities, if and only if every collection $\mathcal{F}$ of closed subsets with the *finite intersection property* (any finite subcollection of $\mathcal{F}$ has nonempty intersection) has a nonempty intersection. Also recall that the image of a compact space under a continuous mapping is compact, and this does not in general hold for preimages (such mappings are called *proper*).

Theorem 3.2 (Characterization of compactness for metric spaces)**.** *For X a metric space, the following are equivalent:*

(i) X is complete and totally bounded,
(ii) X is compact,
(iii) X is sequentially compact.

Proof sketch. (i) $\implies$ (ii): An argument by contradiction using the Cantor intersection theorem. (ii) $\implies$ (iii): Let F_n be the closure of $\{x_k | k > n\}$; by Cantor intersection theorem there is a point belonging to the intersection of all the F_n. This sequence converges to this point. (iii) $\implies$ (i): Easy to construct a sequence with no convergent subsequence if a space is not totally bounded. □

The equivalence of "complete and totally bounded" and "compact" is called the Heine-Borel theorem. The equivalence of "complete and totally bounded" and "sequentially compact" is called the Bolzano-Weierstrass theorem. Remember that totally bounded and bounded are equivalent for subsets of $\mathbb{R}^n$.

Theorem 3.3 (Extreme value theorem)**.** *A metric space X is compact if and only if every continuous real-valued function on X takes a minimum and a maximum value.*

Proof sketch. The forward (and most useful) direction is clear, using that the image must be compact, hence closed and bounded. The converse proceeds by showing that X is totally bounded (by contradiction, constructing a function) and complete. □

Theorem 3.4 (Lebesgue covering lemma)**.** *For any open cover of a compact metric space X, there is an $\epsilon > 0$ such that for each $x \in X$, $B(x, \epsilon)$ is contained in some member of the cover.*

Proof sketch. Argue by contradiction: construct a sequence $\{x_n\}$ such that $B(x_n, 1/n)$ is not contained in a member of the cover. By sequential compactness we can derive a contradiction. □

The Lebesgue covering lemma can be used to quickly prove that a continuous mapping from a compact metric space into a metric space is uniformly continuous. In a compact metric space, the Cantor intersection theorem can be strengthened in that the diameters of the descending sequence of sets are no longer required to go to zero (although there then may be more than one point in the intersection).

If X is a compact metric space, then $C(X)$ is a complete metric space (with metric induced by the maximum norm, as usual).

3.4. **Separable metric spaces.** A metric space is *separable* if it contains a countable dense subset. A compact metric space is separable (using total boundedness). A metric space is separable if and only if it is second countable; that is, there is some countable collection of open subsets such that every open subset is a union of a subcollection (we can take this collection to be, for example, the balls $B(x_n, 1/m)$, where x_n ranges over a dense subset and m ranges over the positive integers). Using this characterization, every subspace of a separable metric space is separable (which is not true for general separable spaces).

3.5. **Big theorems about metric spaces.** The first, Arzelà-Ascoli, answers the question of when we can take a uniformly convergent subsequence of a uniformly bounded sequence (of real-valued continuous functions on a compact space). The key property that is needed is *equicontinuity*. A collection of functions $\mathcal{F}$ is equicontinuous at a point x provided that for each $\epsilon > 0$, there is a $\delta < 0$ such that for every x', if $\rho(x, x') < \delta$ then $|f(x) - f(x')| < \epsilon$ for each $f \in \mathcal{F}$. That is, the functions are all continuous at x and given ϵ we can take a single choice of δ for all of them. The collection is equicontinuous if it is equicontinuous at every point.
An equicontinuous collection on a compact space is uniformly equicontinuous, in that we can pick the same δ for all points.
Example of a collection of functions that is not equicontinuous: let $f_n = x^n$ on $[0, 1]$; the whole collection is not equicontinuous at 1. Examples of a collection of functions that is equicontinuous: all continuous functions on $[a, b]$ that are differentiable on (a, b) with uniformly bounded derivative (via mean value theorem), or all Lipschitz functions with bounded constant (directly).

Lemma 3.5 (Arzelà-Ascoli lemma)**.** *Let X be a separable metric space and $\{f_n\}$ an equicontinuous sequence in $C(X)$ that is pointwise bounded. There is a subsequence that converges pointwise everywhere to a real-valued function f.*

Proof sketch. Take a countable dense subset and diagonalize to find a sequence convergent at those points, which by continuity will be pointwise convergent everywhere. □

Theorem 3.6 (Arzelà-Ascoli theorem)**.** *Let X be a compact metric space and $\{f_n\}$ a uniformly bounded equicontinuous sequence of real-valued functions. Then $\{f_n\}$ has a subsequence that converges uniformly to a continuous function on X.*

Proof sketch. Compact metric implies separable, so use above lemma to find a pointwise-converging subsequence. Use total boundedness and (uniform) equicontinuity to prove that this subsequence is in fact Cauchy, so by completeness of $C(X)$ we're done. □

A consequence/reformulation:

Theorem 3.7. *Let X be a compact metric space. A subset of X is compact if and only if it is closed, uniformly bounded, and continuous.*

Proof sketch. Remember that compactness and sequential compactness are equivalent for metric spaces.

Assuming closed, uniformly bounded, and continuous, Arzelà-Ascoli shows that the subset is sequentially compact, therefore compact.
Assuming compact, closed and bounded is true for all metric spaces (proof easy enough). Assume it is not equicontinuous to derive a contradiction, using sequential compactness. □

Checking that a subset is equicontinuous if and only its closure is, we can also state that a subset of $C(X)$ has compact closure if and only if it is bounded and equicontinuous.
There is a similar criterion for compact subsets of l^p; a subset there is compact if and only if it is closed, bounded, and equisummable in that for each $\epsilon > 0$ there is an index N such that the tails of all of the sequences beyond N are bounded in l^p norm by ϵ.
The following theorem has surprisingly many consequences:

Theorem 3.8 (Baire category theorem). *Let X be a complete metric space and $\{O_n\}$ a countable collection of open dense subsets of X. Then the intersection $\bigcap O_n$ is still dense.*

Proof sketch. Relatively unenlightening and straightforward. We use the Cantor intersection theorem at one point, which is why completeness is necessary. □

Call a subset *hollow* if it has empty interior; equivalently, its complement has empty exterior (i.e. is dense). Using de Morgan's identities, the above theorem is equivalent to the statement that if X is a complete metric space and $\{F_n\}$ a countable collection of closed hollow subsets of X, then the union $\bigcap F_n$ is still hollow.
We have the following immediate corollary: if $\{F_n\}$ is a countable collection of closed sets such that $\bigcup F_n$ has nonempty interior, then one of the F_n must have nonempty interior.
Because boundaries of closed or open sets are always hollow, by Baire a countable union of boundaries of such sets are hollow.
Here are two consequences of the Baire category theorem:

Theorem 3.9. *Let $\mathcal{F}$ be a family of continuous real-valued functions on a complete metric space X that is pointwise bounded. Then there is a nonempty open subset of X on which the family is uniformly bounded.*

Proof sketch. Let $E_n = \{x \in X : f(x) \leq n \text{ for all } f \in \mathcal{F}\}$. The E_n are closed because the f are continuous, and their union is all of X. Therefore by Baire one of them, say E_m, has nonempty interior, so there is a nonempty open subset of X on which $\mathcal{F}$ is bounded by m. □

Theorem 3.10. *Let $\{f_n\}$ be a sequence of continuous functions on a complete metric space X that converge pointwise to a function f. Then there is a dense subset of X on which f is continuous.*

Proof sketch. Let

$$E(m,n) = \{x \in X : |f_j(x) - f_k(x)| \leq 1/m \text{ for all } j,k \geq n\}.$$

Each such set is closed, so the union of their boundaries is hollow. The sequence can be checked to be equicontinuous at each point not in this union, which is a dense subset. The equicontinuity implies that the limit function is continuous at those points. □

Another nice consequence: the collection of nowhere differentiable functions in $C[a,b]$ is dense, because if F_n is the collection of continuous functions whose Lipschitz constant at a some point is $\leq n$, then F_n is hollow, and every function not in $\bigcup F_n$ (which is hollow by Baire) is nowhere differentiable.
For terminological purposes, a set of the *first category* is the union of a countable collection of nowhere dense subsets.
The last "big theorem" about metric spaces is the following:

Theorem 3.11 (Banach contraction principle). *Let X be an complete metric space and $T : X \to X$ be a contraction mapping; i.e., a Lipschitz mapping with a constant strictly less than one. Then T has exactly one fixed point.*

The proof is not difficult. From this theorem we can prove the following fundamental result in differential equations:

Theorem 3.12 (Picard local existence theorem). *Let O be an open set of the plane containing (x_0, y_0) and let $g : O \to \mathbb{R}^2$ be continuous and Lipschitz in the second variable. Then there is an open interval I containing x_0 and a function $f : I \to \mathbb{R}$ such that $f'(x) = g(x, f(x))$ and $f(x_0) = y_0$.*

[proof omitted]

4. Normed linear spaces

4.1. **Basics.** A *normed linear space* is a vector space with a norm; we assume the basic terminology of vector spaces (subspaces, span, direct sum, linear operators, kernels, images). The closure of a subspace is also a subspace. Linear operators T are continuous if and only if they are bounded, which in this context means that there is a constant M such that $||Tu|| \leq M||u||$ for all u. The infimum of all such M is the operator norm $||T||$. The collection of bounded linear operators from X to Y (two normed linear spaces) is itself a normed linear space $\mathcal{L}(X, Y)$ under the operator norm and with the obvious definitions. Any two norms on a finite dimensional space are equivalent (essentially, using Cauchy-Schwarz and the compactness of the unit ball in $\mathbb{R}^n$).
A *Banach* space is a complete normed linear space. If Y is a Banach space, so is $\mathcal{L}(X, Y)$.

Theorem 4.1 (Riesz's theorem). *The closed unit ball of a normed linear space X is compact if and only if X is finite dimensional.*

Proof sketch. If X is finite dimensional, then the closed unit ball is compact because the space is isomorphic to $\mathbb{R}^n$.
If X is infinite dimensional, we need a lemma (Riesz's lemma) telling us that given any closed proper linear subspace Y, we can find a unit vector x_0 that is far away from everything in Y in the sense that given an $\epsilon > 0$, we can choose x_0 such that

$$||x_0 - y|| > 1 - \epsilon.$$

We can then use this (with, say, $\epsilon = 1/2$) to construct a sequence of unit vectors $\{x_n\}$ such that $||x_n - x_m|| > 1/2$ for $n \neq m$, which implies immediately that the sequence has no convergent subsequence and the unit ball of X is not (sequentially) compact. □

4.2. **Slightly less basic.** The following theorems are important enough to mention, but not central enough to provide proofs of. A mapping is *open* if it takes open sets to open sets in the image (with the subspace topology). In particular, an injective continuous mapping is open if and only if it is a homeomorphism.

Theorem 4.2 (Open mapping theorem)**.** *Let X and Y be Banach spaces and $T : X \to Y$ be continuous linear. The image $T(X)$ is closed if and only if T is an open mapping.*

As an immediate corollary, a bijective continuous linear map has a continuous inverse (and is therefore an isomorphism). Another corollary is that two norms on the same space, both of which give rise to Banach spaces, are equivalent if one dominates the other (up to a fixed constant multiple); simply apply the theorem to the identity mapping taking one Banach space to the other.
A linear operator $T : X \to Y$ is *closed* provided whenever $x_n \to x_0$ and $Tx_n \to y_0$, we have $T(x_0) = y_0$. This is true if and only if the graph $\{(x, Tx) \in X \times Y : x \in X\}$ is a closed set in the product space.

Theorem 4.3 (Closed graph theorem)**.** *Let X and Y be Banach spaces and $T : X \to Y$ be linear. Then T is continuous if and only if it is closed.*

As a consequence of these, it can be shown that a closed subspace of a Banach space has a closed linear complement if and only if there exists a continuous projection onto it. It is true that if $T(X)$ has a closed linear complement, then $T(X)$ is a closed set.

Theorem 4.4 (Uniform boundedness principle)**.** *For X a Banach space and Y a normed linear space, let $\mathcal{F} \subseteq \mathcal{L}(X, Y)$. Suppose the family is pointwise bounded in the sense that for each x there is an M_x such that*

$$||T(x)|| \le M_x \quad \text{for all } T \in \mathcal{F}.$$

Then there is an M such that $||T|| \le M$ for all $T \in \mathcal{F}$.

Proof sketch. Let $f_T : X \to \mathbb{R}$ be defined by $f_T(x) = ||Tx||$. This family of functions is pointwise bounded, so by the above consequence of the Baire category theorem there is an open set on which the family is uniformly bounded; i.e. $||Tx|| \le C$ on some open ball. We can then use linearity/triangle inequality to bound $||T||$. □

The point of the following is that we can take pointwise limits of continuous operators and get another continuous operator. A caution: in general, the limit will not hold in the operator norm!

Theorem 4.5 (Banach-Saks-Steinhaus theorem)**.** *Let X be a Banach space and Y a normed linear space. Let $T_n : X \to Y$ be a sequence of continuous operators such that $\lim_{n\to\infty} T_n(x)$ exists for each $x \in X$. Then the sequence is uniformly bounded, and the limit operator*

$$T(x) = \lim_{n\to\infty} T_n(x)$$

is linear, continuous, and $||T|| \le \liminf ||T_n||$.

Proof sketch. The pointwise limit of linear operators is linear, so T is linear. By uniform boundedness, $\{T_n\}$ is uniformly bounded. The inequality is a simple deduction, from which continuity of T follows. □

4.3. **Hahn-Banach.** The normed linear space of continuous real-valued linear functions on a normed linear space X is called the *dual space* and denoted by X^*. There is a natural pairing $X \times X^* \to \mathbb{R}$ given by $(x, \phi) \mapsto \phi(x)$.

An obvious but useful observation: given any nonzero real-valued linear function ϕ on X (not necessarily even continuous) and some x_0 such that $\phi(x_0) \neq 0$, we can write

$$X = [\ker \phi] \oplus \operatorname{span}[x_0].$$

In particular, the kernel is a codimension-one subspace; i.e. a *hyperplane.*

We have the following geometric characterization of linear dependence:

Theorem 4.6 (Linear combinations of functionals)**.** *Let ϕ and $\{\phi_i\}_i^n$ be linear functionals on X. Then ϕ is a linear combination of the ϕ_i if and only if*

$$\bigcap_{i=1}^n \ker \phi_i \subseteq \ker \phi.$$

Proof sketch. Algebra. □

Given any collection $\mathcal{F}$ of real-valued functions on a topological space, the *$\mathcal{F}$-weak topology* on that space is the weakest topology such that each function in $\mathcal{F}$ is continuous. A base at $x \in X$ for the $\mathcal{F}$-weak topology is given by the sets

$$\mathcal{N}_{\epsilon, f_1, \ldots, f_n}(x) = \{x' \in X : |f_i(x) - f_i(x')| < \epsilon, 1 \leq i \leq n\},$$

where $f_1, \ldots, f_n$ runs over all finite subcollections of $\mathcal{F}$ and $\epsilon > 0$. We have $x_n \to x$ with respect to the $\mathcal{F}$-weak topology if and only if

$$f(x_n) \to f(x)$$

for each $f \in \mathcal{F}$.

Given a subspace of linear functionals W, a functional is W-weakly continuous if and only if it lies in W (the nontrivial direction uses the explicit description of neighborhood bases and the above proposition).

The X^*-weak topology is just called the *weak topology.*

Define $J(x) : X^* \to \mathbb{R}$ by

$$J(x)[\phi] = \phi(x).$$

Then $J : X \to (X^*)^*$ is linear. The weak topology on X^* induced by $J(X) \subseteq (X^*)^*$ is called the *weak-$*$ topology.* It has a neighborhood base at $\phi \in X^*$ given by sets of the form

$$\mathcal{N}_{\epsilon, x_1, \ldots, x_n}(\phi) = \{\phi' \in X^* : |(\phi - \phi')(x_i)| < \epsilon, 1 \leq i \leq n\},$$

where $x_1, \ldots, x_n$ runs over all finite subsets of points of X and $\epsilon > 0$. We have $\phi_n \to \phi$ with respect to the weak-$*$ topology if and only if

$$\phi_n(x) \to \phi(x)$$

for every $x \in X$.

For the following theorem, positively homogeneous means commuting with positive scalar multiplication and subadditive means satisfying the triangle inequality.

Theorem 4.7 (Hahn-Banach theorem)**.** *Let p be a positively homogeneous, subadditive functional on a linear space X and Y a subspace of X for which there is defined a linear functional ψ such that $\psi \leq p$ on Y. Then there is an extension of ψ to a linear functional on all of X satisfying the same inequality.*

The proof uses only very simple tools, bootstrapping up from a one-dimensional version via Zorn's lemma. The Hahn-Banach theorem is often used in the following context:

Theorem 4.8 (Extension of bounded linear functionals)**.** *Let X_0 be a linear subspace of a normed linear space X. Then each bounded linear functional ψ on X_0 has an extension to a bounded linear functional on all of X that has the same norm as ψ.*

Proof sketch. Define $p : X \to \mathbb{R}$ by $p(x) = ||\psi||_{X_0} \cdot ||x||$. By the definition of the operator norm, $\psi \leq p$ on X_0, so there is an extension of ψ such that $\psi(x) \leq ||\psi|_{X_0} \cdot ||x||$ for all x. Replacing x by $-x$ we get that the absolute value is so bounded. Therefore the extension is bounded and has the same norm. □

As a corollary, for each $x \in X$ we can define $\eta(\lambda x) = \lambda \cdot ||x||$, which is a linear functional on the span of x having norm one. By the theorem, it has an extension to a bounded linear functional on X that has norm one. Therefore for any x in any normed linear space we can construct a bounded linear functional such that $\eta(x) = ||x||$ and η has norm one.
This theorem also yields the following trick to show that the dual of $L^\infty[a, b]$ is bigger than $L^1[a, b]$: Define $\psi : C[a, b] \to \mathbb{R}$ by

$$\psi(f) = f(x_0),$$

which is a linear functional. By Hahn-Banach, it has an extension to L^∞, even though we cannot explicitly describe that extension. It is then easy to show that no element of L^1 can represent it, so the dual of L^∞ is larger than L^1.

Theorem 4.9 (Finite dimensional subspaces have closed linear complements)**.** *Let X be a normed linear space and X_0 a finite dimensional subspace. Then there is a closed linear subspace X_1 of X such that $X = X_1 \oplus X_0$.*

Proof sketch. Defining a basis for X_0, we define $\psi_k : X_0 \to \mathbb{R}$ by projection to the kth basis element. The ψ_k are obviously bounded so we can extend them to all of X. Bounded linear functionals have closed kernels, so

$$X_1 = \bigcap_{k=1}^{n} \ker \psi_k$$

is also closed. It is the complement we want. □

Theorem 4.10 (Natural map to bidual is an isometry)**.** *Let X be a normed linear space. Then $J : X \to (X^*)^*$ is an isometry, and therefore an injection.*

Proof sketch. By the definition of J and the norm of ψ, we have

$$|J(x)(\psi)| = |\psi(x)| \leq ||\psi|| \cdot ||x||.$$

Therefore $J(x) : X^* \to \mathbb{R}$ is bounded with norm $\leq ||x||$. On the other hand, since there exists a $\psi : X \to \mathbb{R}$ such that $||\psi|| = 1$ and $J(x)(\psi) = \psi(x) = ||x||$, we must have $||J(x)|| = ||x||$. □

Theorem 4.11 (Closure of a subspace)**.** *Let X_0 be a subspace of a normed linear space X. Then x belongs to the closure of X_0 if and only if whenever a bounded linear functional vanishes on X_0, it also vanishes at x.*

Proof sketch. One direction clear by continuity. In the other direction, use Hahn-Banach to construct an element of X^* that vanishes on X_0 but not at x. □

Theorem 4.12 (Weakly convergent sequences are bounded). *Let X be a normed linear space. Then if $\{x_n\} \rightharpoonup x$, $||x|| \leq \liminf ||x_n||$.*

Proof. By the definition of weak convergence, $J(x_n)$ is a sequence of functionals on X^* that converges pointwise to $J(x)$. By Banach-Steinhaus, they are uniformly bounded. But J is an isometry, so x_n is uniformly bounded as well.
For the inequality, we know there is a $\psi \in X^*$ such that $||\psi|| = 1$ and $\psi(x) = ||x||$, so we have

$$|\psi(x_n)| \leq ||\psi|| \cdot ||x_n|| = ||x_n||.$$

But the left hand side converges to $\psi(x) = ||x||$, and the inequality follows. □

We now know that J is always injective. If it is also surjective, we call X a *reflexive* space. The weak and weak-$*$ topologies on X^* are the same if and only if X is reflexive. It is not difficult to show that if X^* is separable, so is X, which implies immediately that reflexive Banach space is separable if and only if its dual is separable. It is also true that a closed subspace of a reflexive Banach space is reflexive.

Theorem 4.13 (Helley's selection theorem). *Let X be a separable normed linear space. Then every bounded sequence $\{\phi_n\}$ in X^* has a subsequence that converges in the weak-$*$ topology to some $\phi \in X^*$.*

Proof sketch. Choose a countable dense subset and use Bolzano-Weierstrass on each point to select subsequences that converge on that point. Then diagonalize. □

Theorem 4.14 (Helley's theorem in reflexive spaces). *Let X be a reflexive Banach space. Then every bounded sequence in X has a weakly convergent subsequence.*

Proof sketch. We can consider only the closure of the span of the sequence, which is separable and by the above facts still reflexive. Then just apply Helley's selection theorem and note that weak convergence is equivalent to weak-$*$ convergence for reflexive spaces. □

As a corollary, every continuous linear functional on a reflexive Banach space takes a maximum value on the closed unit ball of X, using the fact that weakly convergent sequences converge to a point bounded by the limit infimum.
One can often prove that spaces are not reflexive by exhibiting a bounded sequence with no weakly convergent subsequence. For example, on $C[0,1]$ the example of the evaluation operator shows that a weakly convergent sequence must converge pointwise. But $f_n(x) = x^n$ is a bounded sequence for which no subsequence converges pointwise; hence, $C[0,1]$ is not reflexive. Hence L^∞ spaces are not reflexive, because the continuous functions are a closed subspace. Likewise, L^1 spaces are not reflexive because they are separable but have non-separable duals. On the other hand, using the Riesz representation theorem, L^p spaces when $1 < p < \infty$ are reflexive.

4.4. **Fun with convexity.** Taking a step back from norms, we want to consider a large class of topologies on linear spaces that includes both the strong and weak topologies. A *locally convex topological vector space* is a linear space with a Hausdorff topology such that scalar multiplication and vector addition are continuous (from the relevant product topologies) and there is a base at the origin for the topology consisting of convex sets. A *convex* set K in a linear space is a set such that if $u, v \in K$, then $\lambda u + (1-\lambda)v \in K$ for all $0 \leq \lambda \leq 1$.
A subspace W of X^* is said to *separate points* if for every two points in X there is an element of W evaluating differently on those two points. This is clearly necessary if the W-weak topology on X is to separate points. In fact, X is a locally convex topological vector space with respect to any W-weak topology, where W separates points, as well as with respect to the norm (strong) topology.
An *internal point* of a subset E is a point x_0 such that for each direction $x \in X$, there is some $\lambda_0 > 0$ such that $x_0 + \lambda \cdot x \in E$ for all $|\lambda \leq \lambda_0$. Note that this does not necessarily mean that we can place a ball around x_0 in E, as λ_0 might change depending on x.
In a locally convex topological vector space, the closure of a convex subset is convex, and every point in an open subset is an internal point of that subset. A linear functional ϕ is continuous with respect to a locally convex topological vector space if and only if there is a neighborhood of the origin on which $|\phi|$ is bounded. Keep in mind that, as these spaces may not be metrizable, sequential convergence/compactness is not equivalent to convergence/compactness.

Theorem 4.15 (Hyperplane separation theorem)**.** *Let X be a locally convex topological vector space, K a nonempty closed convex subset of X, and x_0 a point in X that lies outside of K. Then there is a continuous linear functional ψ for which*

$$\psi(x_0) < \inf_{x \in K} \psi(x).$$

Proof omitted. In particular, if we choose the strong topology, we can choose a functional in X^*, or if we choose an appropriate weak topology we can actually make sure the functional comes from a subspace of X^* that separates points.

Theorem 4.16 (Mazur's theorem)**.** *Let K be a convex subset of a normed linear space X. Then K is strongly closed if and only if it is weakly closed.*

Proof sketch. Weakly closed implies strongly closed always. In the other direction, given K strongly closed, any point x_0 not in K can be separated by a hyperplane, therefore is contained in a weak neighborhood disjoint from K. Therefore K is weakly closed. □

As a further corollary, if a sequence in a convex, strongly closed K weakly converges to another point, then that point must be in K as well. Using Helley's selection theorem, we can conclude that if X is a reflexive Banach space, then each bounded convex subset of X is weakly sequentially compact: every bounded sequence in K has a weakly convergent subsequence, and its limit must be in K.
Let K be a nonempty convex subset of a locally convex topological vector space X. An *extreme point* of K is a point that cannot be expressed as a nontrivial convex combination of points in K. The *closed convex hull* of a subset K is defined to be the intersection of all closed convex subsets of X containing K.

Theorem 4.17 (Krein-Milman theorem). *Let K be a nonempty compact convex subset of a locally convex topological vector space. Then it is the closed convex hull of its extreme points.*

Proof omitted. Note the very important word "compact."

4.5. **This seems important enough to include.**

Theorem 4.18 (Alaoglu's theorem). *Let X be a normed linear space. The closed unit ball of the dual space X^* is compact in the weak-$*$ topology.*

This does not imply weak-$*$ sequential compactness!

Theorem 4.19 (Kakutani's theorem). *A Banach space is reflexive if and only if its closed unit ball is weakly compact.*

This follows from the previous because the embedding J is a homeomorphism between the weak topology on X and the weak-$*$ topology on $J(X)$.

Theorem 4.20 (Eberlein-Smulian theorem). *The closed unit ball of a Banach space is weakly compact if and only if it is weakly sequentially compact.*

This is notable because in general the weak and weak-$*$ topologies are not metrizable. [However, if X is reflexive, then the weak topology is metrizable if and only if X is separable.]

4.6. **Hilbert spaces.** Knowledge of definition assumed: a Hilbert space is an inner product space that is also a Banach space with respect to the norm topology. The most important basic fact:

Theorem 4.21 (Cauchy-Schwarz inequality). *In an inner product space H, $|\langle u, v\rangle| \leq ||u|| \cdot ||v||$ for any $u, v \in H$.*

Another important fact:

Theorem 4.22 (Parallelogram identity). *For any u, v in an inner product space H,*

$$||u - v||^2 + ||u + v||^2 = 2||u||^2 + 2||v||^2.$$

Finally, a fact that is quite useful for transferring results from norm to inner product. It is slightly more complicated if the underlying field is $\mathbb{C}$ rather than $\mathbb{R}$, because this only gives the real part:

Theorem 4.23 (Polarization identity). *For any u, v in an inner product space H,*

$$\langle u, v\rangle = \frac{1}{4}[||u + v||^2 - ||u - v||^2].$$

For a Hilbert space, all closed subspaces have closed linear complements. Orthogonal projections have norm one and obey the relations $\langle Pu, v\rangle = \langle Pu, Pv\rangle = \langle u, Pv\rangle$ for any vectors.

Theorem 4.24 (Riesz-Fréchet representation theorem). *Let H be a Hilbert space. Define $T : H \to H^*$ by defining*

$$T(h)[u] = \langle h, u\rangle \quad \text{for all } u \in H.$$

Then T is a linear isometry of H onto H^*.

This describes fully the continuous linear functionals on a Hilbert space. A consequence:

Theorem 4.25. *Every bounded sequence in a Hilbert space has a weakly convergent subsequence.*

Another consequence is that a linear operator between Hilbert spaces is bounded if and only if it maps weakly convergent sequences to weakly convergent sequences. The proof of the following is just an application of the parallelogram identity:

Theorem 4.26 (Radon-Riesz theorem)**.** *Let $u_n \rightharpoonup u$. Then $u_n \to u$ if and only if $||u_n|| \to ||u||$.*

Using Riesz-Fréchet, it is easy to see that Hilbert spaces are reflexive. Therefore by Kakutani its closed unit ball is weakly compact, in addition to being weakly sequentially compact.

Theorem 4.27 (Bessel's inequality)**.** *For any orthonormal sequence ϕ_k in an inner product space H, we have*

$$\sum_{k=1}^{\infty} \langle \phi_k, h \rangle^2 \leq ||h||^2.$$

The proof is the Pythagorean identity followed by taking a limit. An orthonormal sequence is *complete*, or a *basis*, if the only vector orthogonal to every element is the zero vector. For a complete orthonormal sequence, we have the Parseval's identity

$$\sum_{k=1}^{\infty} \langle \phi_k, h \rangle^2 = ||h||^2.$$

Of course, a Hilbert space with an orthonormal basis must be separable. The converse is also true: a separable Hilbert space has a complete orthonormal basis.
Consider a continuous linear mapping $T : H \to H$. For a fixed v, the mapping $u \mapsto \langle Tu, v \rangle$ is a bounded linear functional, so it must be represented by a vector h such that $\langle Tu, v \rangle = \langle u, hv \rangle$. Denoting this vector by T^*v, it is clear that $T^* : H \to H$ is a linear map. By Cauchy-Schwarz, we have $||T|| = ||T^*||$, so it is bounded linear. This map is called the *adjoint* of T. Untangling the definitions, we see that $\ker T^*$ is the perpendicular complement of $\operatorname{Im} T$. Therefore if the image of T is closed, we have a decomposition

$$H = \operatorname{Im} T \oplus \ker T^*.$$

An operator is *self-adjoint* if it equals its own adjoint. It is *positive semidefinite* if $\langle Tu, u \rangle \geq 0$ for all $u \in H$. For a self-adjoint operator,

$$||T|| = \sup_{||u||=1} |\langle Tu, u \rangle|.$$

If a symmetric operator maps V into V, where V is a subspace of H, then it also maps the orthogonal complement of V into the orthogonal complement of V.
A bounded operator is said to be *finite rank* if its images is finite dimensional. A bounded operator is said to be *compact* if the image of the unit ball has compact closure. Finite rank operators are compact. Invertible operators are not compact. Because Hilbert spaces are metric spaces, compactness is equivalent to sequential compactness is equivalent to being closed and totally bounded. Therefore an operator K is compact if and only if its image is totally bounded, if and only if for every bounded sequence h_n in H then $\{Kh_n\}$ has a (strongly) convergent subsequence.

A good source of examples: consider a separable Hilbert space with an orthonormal basis $\{e_n\}$, and let T be the operator mapping e_n to $\lambda_n e_n$ and extending by linearity. Then T is compact if and only if $\lambda_n \to 0$, which can be shown using total boundedness in one direction and a simple construction in the other.

Theorem 4.28 (Characterization of compact operators)**.** *A bounded linear operator T on a Hilbert space H is compact if and only if it maps weakly convergent sequences to strongly convergent sequences.*

Proof sketch. Using adjoints, it is easy to show that any bounded linear operator maps weakly convergent sequences to weakly convergent sequences. By compactness, every subsequence of the image has a further subsequence converging strongly, and because it is already weakly convergent to the "right thing" the whole sequence must be in fact strongly convergent.
Conversely, if T maps weakly convergent sequences to strongly convergent ones, pick any bounded sequence, which must have a weakly convergent subsequence. The image of this subsequence converges strongly. Therefore the image of the unit ball is sequentially precompact. □

Theorem 4.29 (Schauder's theorem)**.** *A compact linear operator has a compact adjoint.*

Proof sketch. Use the above characterization. □

Even if a bounded linear operator on a Hilbert space is symmetric, it may not have eigenvectors (let alone an orthonormal basis of them). This is unlike what occurs in the finite dimensional case. For example, let $T : L^2[0,1] \to L^2[0,1]$ be defined by $[Tf](x) = x \cdot f(x)$. Then T is bounded and symmetric but there are no eigenvectors. The criterion we add for operators on a Hilbert space is compactness. As a relatively ugly lemma (using the maximizing property of the Raleigh quotient $\langle Th, h\rangle/\langle h, h\rangle$, for instance), we find that a compact symmetric operator on a Hilbert space has an eigenvalue of absolute value equal to $||T||$. Then we proceed iteratively, pulling off one-dimensional subspaces, to prove the following theorem:

Theorem 4.30 (Hilbert-Schmidt theorem)**.** *Let H be a Hilbert space and K a compact symmetric bounded linear operator that is not of finite rank. Then there is an orthonormal basis ψ_k to $[\ker K]^\perp$ together with a sequence of real nonzero numbers λ_k that tend to zero, such that $K(\psi_k) = \lambda_k \psi_k$ for each k. Thus by linearity*

$$K(h) = \sum_{k=1}^{\infty} \lambda_k \langle \psi_k, h\rangle \psi_k$$

for all $h \in H$.

An operator on a separable Hilbert space is compact if and only if it is the strong limit of finite rank operators (we can take the finite rank operators in question to be the projections onto larger and larger finite subspaces). This is proved in one direction via total boundedness, and in the other direction by Arzelá-Ascoli.
If K is compact, $\mathrm{Id} + K$ has a finite dimensional kernel and closed image. If $\mathrm{Id} + K$ is injective, it is surjective. These two facts can be used to prove the following:

Theorem 4.31 (Riesz-Schauder theorem)**.** *Let H be a Hilbert space and K a compact bounded linear operator. Then $Im(Id + K)$ is closed and*

$$\dim\ker(\mathrm{Id} + K) = \dim\ker(\mathrm{Id} + K^*) < \infty,$$

Using the aforementioned fact that if an operator T has closed image then we can write $H = \operatorname{Im} T \oplus \ker T^*$, this means that

$$\dim \ker(\operatorname{Id} + K) = \operatorname{codim} \operatorname{Im}(\operatorname{Id} + K).$$

Therefore $\operatorname{Id} + K$ is injective if and only if it is surjective. This gives rise to the more famous version of this result:

Theorem 4.32 (Fredholm alternative)**.** *Let H be a Hilbert space, K a compact bounded linear operator, and $\mu \neq 0$ a real number. Then exactly one of the following holds:*

(i) There is a nonzero solution to the equation $\mu h - Kh = 0$.
(ii) For every $h_0 \in H$, there is a unique solution of the equation $\mu h - Kh = h_0$.

The first situation corresponds to h lying in the kernel of $\mu \operatorname{Id} + K$, while the second corresponds to when it does not.
In general if $\dim \ker T$ and $\operatorname{codim} \operatorname{Im} T$ are finite we say that T is *Fredholm* and we let the Fredholm index be the difference $\dim \ker T - \operatorname{codim} \operatorname{Im} T$. The Riesz-Schauder theorem shows that when K is compact, $\operatorname{Id} + K$ is Fredholm of index zero. It is not hard to show that T is Fredholm of index zero if and only if it can be written as $S + K$, where S is invertible and K is compact (in fact, we may take K to be of finite rank).

5. Measure and integration

5.1. **Basics and decomposition of signed measures.** A *measurable space* consists of a space X and a σ-algebra of subsets of X. A *measure* on that space is a nonnegative set function taking the σ-algebra to the extended reals which is countably additive and such that the measure of the empty set is zero. From this we have the usual properties, including the very useful *countable monotonicity*; i.e. if E_k is a countable collection of measurable sets that covers E, we have

$$\mu(E) \leq \sum_{k=1}^{\infty} \mu(E_k).$$

We also have continuity of measure: if A_k is an ascending sequence of measurable sets, then

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

If B_k is a descending sequence of measurable sets all of finite measure, then

$$\mu\left(\bigcap_{k=1}^{\infty} B_k\right) = \lim_{k\to\infty} \mu(B_k).$$

This does not hold if the B_k all have infinite measure; consider $B_k = [k, \infty) \subset \mathbb{R}$, for instance.

Theorem 5.1 (Borel-Cantelli lemma)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{E_k\}$ a countable collection of measurable sets for which $\sum_{k=1}^{\infty} \mu(E_k) < \infty$. Then all most all x in X belong to at most a finite number of the E_k.*

Proof sketch. We calculate by continuity and countable monotonicity that

$$\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_k\right) = \lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty} E_k\right) \le \lim_{n\to\infty}\sum_{k=n}^{\infty}\mu(E_k) = 0.$$

The set on the left hand side is precisely the set of all x in X that belong to an infinite number of the E_k. □

A measure space is *finite* if $\mu(X) < \infty$ and *σ-finite* if it is the countable union of finite sets, which may be taken to be disjoint. A measure space is *complete* if every subset of a set of measure zero is measurable. Every measure space can be completed in an obvious canonical way.
A *signed measure* ν is not required to be nonnegative but is required to take on at most one of the values $-\infty, +\infty$, and series are required to converge absolutely for countable additivity. Call a set *positive* if it is measurable and for every measurable subset E we have $\nu(E) \ge 0$; likewise *negative* and *null*. It can be shown (Hahn's lemma) that if E is a subset such that $0 < \nu(E) < \infty$, there is a positive measurable subset of E. This can be sued to prove the following:

Theorem 5.2 (Hahn decomposition theorem)**.** *Let $(X, \mathcal{M})$ be a measurable space and ν a signed measure on it. Then there is a positive set A and a negative set B for which $X = A \cup B$ and $A \cap B = \emptyset$.*

Note that this decomposition is not quite unique, because we could excise a null set from A and put it on B or vice versa. If $\{A, B\}$ is a Hahn decomposition for ν, define

$$\nu^+(E) = \nu(E \cap A), \quad \nu^-(E) = -\nu(E \cap B),$$

each of which will be be *mutually singular* measures such that $\nu = \nu^+ - \nu^-$, which means there are two sets A, B with $A \cup B = X$ and $\nu^+(A) = \nu^-(B) = 0$. It is then easy to prove the following:

Theorem 5.3 (Jordan decomposition theorem)**.** *Let $(X, \mathcal{M})$ be a measurable space and ν a signed measure on it. Then there are two mutually singular measures ν^+ and ν^- on $(X, \mathcal{M})$ for which $\nu = \nu^+ - \nu^-$, and these two measures are unique.*

If we let $|\nu|(E) = \nu^+(E) + \nu^-(E)$ and call $|\nu|$ the *total variation* of ν, then we have

$$|\nu|(X) = \sup \sum_{k=1}^{n} |\nu(E_k)|,$$

where $\{E_k\}$ runs over all finite disjoint collections of measurable subsets of X.

5.2. **The Carathéodory constructions.** A set function $\mu^* : 2^X \to [0, \infty]$ is called an *outer measure* provided $\mu^*(\emptyset) = 0$ and μ^* is countably monotone. Given an outer measure, call a set E *measurable* if for every subset A of X, we have

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

By monotonicity we know that we always have $\le$ in the above expression, so it suffices to check $\ge$ for all sets of finite outer measure. After some annoying set chasing, it turns out that the collection of sets that are measurable with respect to μ^* is a σ-algebra, and the restriction of μ^* to this σ-algebra yields a complete measure space.

How do we get outer measures, then? If we start with any set function $\mu' : \mathcal{S} \to [0, \infty)$, where $\mathcal{S}$ is any collection of subsets of X, we can define the *outer measure induced by* μ' to be defined by $\mu^*(\emptyset) = 0$ and

$$\mu^*(E) = \inf \sum_{k=1}^{\infty} \mu'(E_k),$$

where $\{E_k\}$ runs over all countable subfamilies of $\mathcal{S}$ that cover E (if no elements of $\mathcal{S}$ cover E, set $\mu^*(E) = \infty$). Then this is an outer measure, as can be verified by checking countable monotonicity. Its induced measure μ, as described above, is called the Carathéodory measure induced by the set function μ'. As constructed now, there is no guarantee that the sets in $\mathcal{S}$ will be measurable!
The following theorem is very useful. Let $\mathcal{T}$ be the collection of all countable intersections of countable unions of elements of $\mathcal{S}$ (so, for example, if $\mathcal{S}$ consists of open intervals on the line, then $\mathcal{T}$ consists of the G_δ sets, countable intersections of open sets).

Theorem 5.4 (Approximation by simple sets)**.** *Let $\mu' : \mathcal{S} \to [0, \infty)$ be a set function and μ the induced Carathéodory measure, and let E be such that $\mu^*(E) < \infty$. Then there is a set $A \in \mathcal{T}$, with $\mathcal{T}$ defined as above, such that $E \subseteq A$ and $\mu^*(E) = \mu^*(A)$. If in addition E and each set in $\mathcal{S}$ is measurable, then so is A and we have $\mu(A \setminus E) = 0$.*

Proof sketch. We first find a set A_ϵ that is a countable union of elements of $\mathcal{S}$ that is ϵ-close to satisfying the proposition, and take the intersection of all the $A_{1/k}$. □

We now ask what restrictions we need to ensure that the elements of $\mathcal{S}$ are measurable under the induced measure. It is easy enough to show that a necessary condition is that μ' is a *premeasure* - finitely additive, countably monotone, and if $\emptyset \in \mathcal{S}$, $\mu'(\emptyset) = 0$. This is also a sufficient condition provided we make some restrictions on $\mathcal{S}$ itself. We say that $\mathcal{S}$ is closed with respect to the formation of relative complements provided whenever $A, B \in \mathcal{S}$, we have $A \setminus B \in \mathcal{S}$.

Theorem 5.5. *If $\mathcal{S}$ is closed with respect to the formation of relative complements and $\mu' : \mathcal{S} \to [0, \infty)$ is a premeasure, then the induced Carathéodory measure is an extension of μ'.*

Proof sketch. Use finite additivity of μ' to show that every set in $\mathcal{S}$ is measurable and countable monotonicity to ensure that μ extends μ'. □

We can take one further step back. Call a collection $\mathcal{S}$ a *semiring* if it is closed under intersections and for every $A, B \in \mathcal{S}$, there is a finite disjoint collection C_k of sets in $\mathcal{S}$ such that

$$A \setminus B = \bigcup_{k=1}^{n} C_k.$$

Given any semiring $\mathcal{S}$, if we define $\mathcal{S}'$ to be the collection of unions of finite disjoint collections of sets in $\mathcal{S}$ then $\mathcal{S}$ is closed with respect to relative complements. Furthermore, any premeasure on $\mathcal{S}$ has a unique extension to a premeasure on $\mathcal{S}'$. Putting all this together with one last uniqueness statement, proven using the approximation by simple sets theorem above, we have the following:

Theorem 5.6 (Carathéodory-Hahn theorem)**.** *Let $\mu' : \mathcal{S} \to [0, \infty)$ be a premeasure on a semiring. Then the Carathéodory measure μ induced by μ' is an extension of μ'. If μ' is σ-finite, then so is μ and the extension to the σ-algebra of μ^*-measurable sets is unique.*

This theorem implies that we can check whether or not two σ-finite measures are equal on the smallest σ-algebra containing a semiring by just checking on the semiring itself.

5.3. **Measurable functions and integration.** We call an extended real-valued function on a measure space X *measurable* provided that for all $c \in \mathbb{R}$, the set $\{x \in X : f(x) < c\}$ is measurable. This is equivalent to replacing $<$ with $\leq$, $>$, or $\geq$. If the function is real-valued (nowhere infinite), it is equivalent to asserting that for every open set O of real numbers, $f^{-1}(O)$ is measurable. Using this criterion, it is clear that the composition $\phi \circ f$, where ϕ is continuous and f is measurable, is measurable. Restrictions of functions to measurable sets are measurable, and if a function is measurable on some measurable set E and $X \setminus E$ it is measurable on X.

If we are dealing with a complete measure space, a function is measurable if and only if it is measurable almost everywhere (that is, when restricted to a set X_0 such that $\mu(X \setminus X_0) = 0$). In this case, if $g = h$ a.e. and g is measurable, so is h. The sum and product of measurable functions are measurable. If we are adding two functions that are finite a.e. and take on opposite infinite values at the same point, we regard their sum as being defined on a slightly smaller set X_0 such that $\mu(X \setminus X_0) = 0$. If X is a complete measure space, this is equivalent to the assertion that any extension of the sum to all of X is a measurable function.

Theorem 5.7 (Pointwise limits of measurable functions are measurable)**.** *Let $\{f_n\}$ be a sequence of measurable functions on $(X, \mathcal{M}, \mu)$ such that either X is complete and $\{f_n\} \to f$ pointwise almost everywhere, or $\{f_n\} \to f$ pointwise everywhere. Then f is measurable.*

Proof sketch. Can assume pointwise everywhere convergence. By set manipulation,

$$\{x \in X : f(x) < c\} = \bigcup_{k,n=1}^{\infty} \bigcap_{j=k}^{\infty} \{x \in X : f_j(x) < c - 1/n\},$$

which is therefore measurable. □

This immediately implies that things like $\sup(f_n)$, $\text{limsup}(f_n)$, etc., are measurable as well.

A *simple function* is a measurable real-valued function that takes on finitely many values. We can write every simple function ψ as

$$\psi = \sum_{k=1}^{n} c_k \cdot \chi_{E_k}$$

for some real numbers c_k and some measurable sets E_k. By the definition of measurability and some legwork, it is easy to see that any bounded measurable function can be uniformly approximated in both directions by simple functions. That is, for every $\epsilon > 0$ there are simple functions ϕ_ϵ and ψ_ϵ such that $\phi_\epsilon \leq f \leq \psi_\epsilon$ and $0 \leq \psi_\epsilon - \phi_\epsilon < \epsilon$ on X. It is then simple to prove the following:

Theorem 5.8 (Simple approximation theorem)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space and f a measurable function on X. Then there is a sequence of simple functions $\{\phi_n\}$ that converge pointwise to f on X and such that $|\phi_n| \leq |f|$ on X for all n. If X is σ-finite, we can make sure that the ϕ_n vanish outside a set of finite measure. If f is nonnegative, we can choose ϕ_n to be nonnegative and increasing.*

The next theorem is a statement that pointwise convergence on finite measure spaces to finite a.e. functions is uniform convergence except on a set of arbitrarily small positive measure. Its proof uses the continuity of measure.

Theorem 5.9 (Egoroff's theorem)**.** *Let $(X, \mathcal{M}, \mu)$ be a finite measure space and $\{f_n\}$ a sequence of measurable functions converging pointwise a.e. on X to a function f that is finite a.e. on X. Then for each $\epsilon > 0$, there is a measurable set X_ϵ such that $\{f_n\} \to f$ uniformly on X_ϵ and $\mu(X \setminus X_\epsilon) < \epsilon$.*

We define the integral of a simple function in the obvious way, and define the integral of a nonnegative measurable function f to be the supremum of the integrals of the simple functions ϕ for which $0 \leq \phi \leq f$.

Theorem 5.10 (Chebychev's inequality)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space, f a nonnegative measurable function on X, and λ a positive real number. Then*

$$\mu\{x \in X : f(x) \geq \lambda\} \leq \frac{1}{\lambda} \int_X f \, d\mu.$$

Proof sketch. Let ϕ be λ multiplied by the characteristic function of the set on the left hand side. Then since $\phi \leq f$, we can integrate to get the desired inequality. □

Using Chebychev's inequality, it is easy to see that if $\int_X f \, d\mu < \infty$, then f is finite a.e. and the set on which it is nonzero is σ-finite.
If $|f|$ has a finite integral, we say f is integrable. Its integral is defined to be the difference of the integrals of its positive and negative parts. It is immediate that $\left|\int_X f \, d\mu\right| \leq \int_X |f| \, d\mu$. The integral is linear, monotonic (if $f \leq g$ a.e. then $\int_X f \, d\mu \leq \int_X g \, d\mu$) and countably additive over domains: if $\{X_n\}$ are countably many measurable sets whose disjoint union is X, then

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_{X_n} f \, d\mu.$$

(To prove this, use monotone convergence, below.)
It is then easy to show that any bounded measurable function zero outside a set of finite measure is integrable, as we would expect, and that if X is a compact topological space and a finite measure space that contains the topology on X, all continuous functions are integrable.
These properties imply that, for any nonnegative integrable function g on X, the set function $E \mapsto \int_E g \, d\mu$ is a finite measure on the same measurable space $(X, \mathcal{M})$. Applied separately to the positive and negative parts of a function, we get the following (noting that, since we are dealing with a finite measure, no restrictions on the descending collection are necessary):

Theorem 5.11 (Continuity of integration)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space and f an integrable function on X. If $\{X_n\}$ is an ascending countable collection of* *measurable subsets whose union is X, then*

$$\lim_{n\to\infty} \int_{X_n} f \, d\mu = \int_X f \, d\mu.$$

If $\{X_n\}$ is a descending countable collection of measurable subsets, then

$$\lim_{n\to\infty} \int_{X_n} f \, d\mu = \int_{\cap X_n} f \, d\mu.$$

5.4. **Integral convergence theorems.**

Theorem 5.12 (Fatou's lemma). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of nonnegative measurable functions that converge to f pointwise a.e., and assume f is measurable (which is unnecessary if X is complete). Then*

$$\int_X f \, d\mu \leq \liminf \int_X f_n \, d\mu.$$

The proof is sort of annoying, but proceeds by showing that any simple function dominated by f obeys the inequality $\int_X \phi \, d\mu \leq \liminf \int_X f_n \, d\mu$. From Fatou's lemma, the following two theorems are easy:

Theorem 5.13 (Monotone convergence theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ an increasing sequence of nonnegative measurable functions on X. Define $f(x) = \lim_{n\to\infty} f_n(x)$. Then*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Theorem 5.14 (B. Levi's lemma). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ an increasing sequence of nonnegative measurable functions on X. If the sequence $\{\int_X f_n \, d\mu\}$ is bounded, then $\{f_n\}$ converges pointwise on X to a measurable function f that is finite a.e. on X and*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu < \infty.$$

The monotone convergence theorem can also be applied quickly to show that for every nonnegative function there exist an increasing sequence of simple functions that converge both pointwise and in the integral to that function. This is then used to show that the integral is linear for nonnegative functions.

Theorem 5.15 (Lebesgue dominated convergence theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of measurable functions on X for which $\{f_n\} \to f$ pointwise a.e., and f is measurable (again, this last assumption is unnecessary if X is complete). Assume there is a function g that is integrable over X such that $|f_n| \leq g$ a.e. on X. Then f is also integrable and*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proof sketch. Apply Fatou's lemma to $\{g - f_n\}$ and $\{g + f_n\}$. □

Let $\{f_n\}$ be a sequence of functions on X, each of which is integrable. The sequence is *uniformly integrable* if for each $\epsilon > 0$ there is a $\delta > 0$ such that for any n and measurable E we have that

$$\text{if } \mu(E) < \delta, \text{ then } \int_E |f_n| \, d\mu < \epsilon$$

The sequence is *tight* provided that for each $\epsilon > 0$ there is a subset X_ϵ that has finite measure and such that for any n,

$$\int_{X \setminus X_\epsilon} |f_n| \, d\mu < \epsilon.$$

Approximation by simple functions shows quickly that a single integrable function by itself is uniformly integrable and tight. The proof of the following uses Egoroff's theorem but is otherwise a simple application of the definitions:

Theorem 5.16 (Vitali convergence theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of functions that is both uniformly integrable and tight. Assume $\{f_n\} \to f$ pointwise a.e. and f is integrable over X. Then*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Unlike in the case of the real line, we must *a priori* assume that f is integrable. Consider the σ algebra $\mathcal{M} = \{\emptyset, E, X \setminus E, X\}$ on some space and define $\mu(E) = \mu(X \setminus E) = 1/2$. Then $f_n = n \cdot \chi_E - n \cdot \chi_{X \setminus E}$ is obviously tight, and is uniformly integrable because there are no small measurable sets. It converges pointwise to a function that is clearly not integrable.
As a corollary, if the f_n are nonnegative and $\{f_n\} \to 0$ pointwise a.e., then the limit of the integrals is zero if and only if $\{f_n\}$ is uniformly integrable and tight.
A sequence of measurable functions $\{f_n\}$ is said to *converge in measure* to f provided that for each $\eta > 0$,

$$\lim_{n\to\infty} \mu\{x \in X : |f_n(x) - f(x)| > \eta\} = 0.$$

Using Egoroff's theorem, if X is a finite measure space then pointwise a.e. convergence implies convergence in measure. In general, using the Borel-Cantelli lemma, convergence in measure implies that there is a subsequence that converges a.e. pointwise. Together with the Vitali convergence theorem, it can then be proven that if $\{f_n\}$ is a sequence of measurable nonnegative functions, the integrals tend to zero if and only if the sequence is uniformly integrable, tight, and convergent in measure to the zero function. If X is a finite measure space, in fact, $\{f_n\} \to f$ in measure if and only if every subsequence of $\{f_n\}$ has a further subsequence that converges a.e. pointwise. Fatou's lemma, the monotone convergence theorem, the dominated convergence theorem, and the Vitali convergence theorem all hold if pointwise a.e. convergence is replaced with convergence in measure.

5.5. **Radon-Nikodym and friends.** Given two measures ν and μ on a single measurable space, say that ν is *absolutely continuous* with respect to μ if whenever $\mu(E) = 0$, $\nu(E) = 0$. Example: if f is a nonnegative integrable function, and $\nu(E) = \int_X f \, d\mu$, then ν is absolutely continuous with respect to μ. A fairly simple argument can be used to provide an alternative characterization: ν is absolutely continuous with respect to μ if and only if for every $\epsilon > 0$, there is a $\delta > 0$ such that whenever $\mu(E) < \delta$, we have $\nu(E) < \epsilon$.

Theorem 5.17 (Radon-Nikodym theorem). *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and ν a σ-finite measure defined on $(X, \mathcal{M})$ and absolutely continuous with respect to μ. Then there is a nonnegative measurable function f such that for all*

$E \in \mathcal{M}$,

$$\nu(E) = \int_E f \, d\mu.$$

If g is another such function, then $f = g$ a.e. on with respect to μ.

The proof in the finite case is in two parts; first it is shown that there exists a nonnegative measurable function with a positive integral such that $\int_E f \, d\mu \leq \nu(E)$ for all measurable E by using an argument by contradiction and the Hahn decomposition of the signed measure $\nu - \lambda\mu$, where $\lambda > 0$. Then one takes the supremum of all such functions and proves that this supremum is achieved and the desired equality holds. For the σ-finite case a simple union of sets is taken.
The function found by the Radon-Nikodym theorem is called the Radon-Nikodym derivative of ν with respect to μ, and denoted $\frac{d\nu}{d\mu}$.
There is an immediate extension for signed measures:

Theorem 5.18. *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and ν a finite signed measure defined on $(X, \mathcal{M}, \mu)$ that is absolutely continuous with respect to μ. Then there is a measurable function f such that for all $E \in \mathcal{M}$,*

$$\nu(E) = \int_E f \, d\mu,$$

and if g is another such function then $f = g$ a.e. with respect to μ.

Theorem 5.19 (Lebesgue decomposition theorem)**.** *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and ν a σ-finite measure on $(X, \mathcal{M})$. There is a measure ν_0 mutually singular with respect to to μ and a measure ν_1 absolutely continuous with respect to μ such that $\nu = \nu_0 + \nu_1$. The given measures are unique.*

Proof sketch. Consider $\lambda = \mu + \nu$, use Radon Nikodym on μ which is absolutely continuous with respect to λ to find an f such that

$$\mu(E) = \int_E f \, d\lambda = \int_E f \, d\mu + \int_E f \, d\nu.$$

Let X_+ be the set on which f is positive and X_0 the set on which f is zero. Then setting

$$\nu_0 = \nu(E \cap X_0), \quad \nu_1 = \nu(E \cap X_+)$$

is the desired decomposition. □

Given a finite measure space $(X, \mathcal{M}, \mu)$, there is a natural equivalence relation on the set of measurable sets given by $A \simeq B$ if $\mu(A \triangle B) = 0$. Taking the quotient with respect to this equivalence relation and setting $\rho([A], [B]) = \mu([A] \triangle [B])$, which is well-defined, we get the *Nikodym metric space* associated to the finite measure space. This turns out to be a complete measure space (using the Riesz-Fischer theorem). If ν is a finite measure on the same measurable space that is absolutely continuous with μ, then it is well-defined to set $\nu([A]) = \nu(A)$, and ν is a uniformly continuous function on the Nikodym metric space.
Now let $\{\nu_n\}$ be a sequence of finite measures on the same space, each of which are absolutely continuous with respect to μ. They are said to be *uniformly absolutely continuous* if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every n and measurable E, if $\mu(E) < \delta$, then $\nu_n(E) < \epsilon$. With this setup following three properties are equivalent: the $\{\nu_n\}$ are uniformly absolutely continuous if and only if the sequence of functions $\{\nu_n : \mathcal{M} \to \mathbb{R}\}$ is equicontinuous with respect to the Nikodym metric,

if and only if the sequence of Radon-Nikodym derivatives $\{d\nu_n/d\mu\}$ is uniformly integrable over X with respect to μ.
Setwise convergence of measures to a set function has the obvious meaning: for every measurable set, the measure approaches the desired value in the limit. The next theorem states that setwise convergence of absolutely continuous measures gives an absolutely continuous measure in the limit.

Theorem 5.20. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space and $\{\nu_n\}$ a sequence of uniformly absolutely continuous finite measures on $\mathcal{M}$. If $\{\nu_n\}$ converges setwise to ν, then ν is a measure, absolutely continuous with respect to μ.*

If $\{\nu_n(X)\}$ is a bounded sequence, we do not need to assume the sequence is uniformly absolutely continuous, which is a consequence of the Baire category theorem:

Theorem 5.21 (Vitali-Hahn-Saks)**.** *Let $(X, \mathcal{M}, \mu)$ be a finite measure space and $\{\nu_n\}$ a sequence of absolutely continuous finite measures on $\mathcal{M}$. if $\{\nu_n\}$ converges setwise to ν and $\{\nu_n(X)\}$ is a bounded sequence, then the sequence is uniformly absolutely continuous, so ν is therefore an absolutely continuous measure.*

This is a surprising fact on par with the fact that the pointwise limit of a sequence of continuous linear operators is continuous (Banach-Steinhaus).
If we are just given a sequence $\{\nu_n\}$ that converges setwise to a set function ν and such that $\{\nu_n(X)\}$ is a bounded sequence, we can define

$$\mu(E) = \sum_{n=1}^{\infty} \frac{1}{2^n} \nu_n(E)$$

and conclude from the above that ν is a measure.

5.6. **Product measures and Fubini's theorem.** Given two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{A}, \nu)$, we can form a product measure $\mu \times \nu$ by first defining the product on the semiring of measurable rectangles in the obvious way, namely

$$(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B),$$

and then extending with the Carathéodory construction.
For the purposes of repeated integration, a function that is defined almost everywhere will be integrated by restricting to where it is defined.

Theorem 5.22 (Fubini's theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two measure spaces and let ν be complete. Let f be integrable with respect to the product measure $\mu \times \nu$. Then for all most all $x \in X$, $f(x, \cdot)$ is integrable over Y with respect to ν and*

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X \left[\int_Y f(x, y) \, d\nu(y) \right] \, d\mu(x).$$

The proof is long and tedious, building up from characteristic functions of sets of measure zero and sets that are countable intersections of countable unions of rectangles (which together suffice to establish the result for general characteristic functions, due to properties of the Carathéodory extension). The monotone convergence theorem is used many times.
Often it is difficult to establish *a priori* that a function is integrable with respect to the product measure. In the case that both μ and ν are σ-finite, ν is complete, and f is nonnegative and measurable with respect to the product measure, then we do have equality:

Theorem 5.23 (Tonelli's theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ-finite measure spaces and let ν be complete. Assume f is a nonnegative $(\mu \times \nu)$-measurable function. Then for almost all $x \in X$, $f(x, \cdot)$ is measurable in ν, the resulting function on Y defined almost everywhere by integration with respect to $d\nu$ is measurable in μ, and*

$$\int_{X\times Y} f \, d(\mu \times \nu) = \int_X \left[\int_Y f(x,y) \, d\nu(y) \right] \, d\mu(x).$$

Both sides may, of course, be infinite.

As a corollary, if both μ and ν are complete and σ-finite and f is nonnegative, if one of the repeated integrals is finite then all three integrals are finite and equal. If a function on the product space is measurable with respect to the smaller σ-algebra defined as the smallest σ-algebra containing all measurable rectangles, then completeness is not necessary in the hypotheses of the above theorem.

6. Functions on the real line

6.1. **Real numbers and Lebesgue measure.** A closed and bounded set of real numbers is compact, a result which is often called the Heine-Borel theorem. The result that a closed and bounded set of real numbers is sequentially compact is often called the Bolzano-Weierstrass theorem. A monotone sequence of real numbers converges if and only if it is bounded. A monotone real-valued function always possesses right- and left-sided limits, so the only possible discontinuities it may have are jump discontinuities.

Theorem 6.1 (Intermediate value theorem)**.** *Let f be a continuous real-valued function on $[a, b]$ for which $f(a) < c < f(b)$. Then there is a point x_0 in (a, b) at which $f(x_0) = c$.*

Proof sketch. Inductively bisect the interval and use the Cantor intersection theorem. □

The smallest σ-algebra of sets on the real line containing the open sets is called the Borel σ-algebra. It is strictly smaller than the Lebesgue σ-algebra, which is the completion of the Borel σ-algebra. To define Lebesgue measure m, take the Carathéodory extension of the set function that assigns the value $b - a$ to every interval $[a, b)$, the collection of all such intervals forming a semiring (or alternatively to every interval, closed, open, or half closed half open). Lebesgue measure is translation invariant.

Lebesgue measure has strong inner and outer approximation properties. Recall that a G_δ set is a countable intersection of open sets (G for Gebiet, German for area or neighborhood, δ for Durchschnitt, German for intersection) and an F_σ set is a countable union of closed sets (F for fermé, French for closed, σ for somme, French for union). The following theorems are useful and have straightforward proofs.

Theorem 6.2 (Approximation of Lebesgue measurable sets)**.** *Let E be any set of real numbers. Any one of the following four properties is equivalent to the measurability of E:*

(i) For any $\epsilon > 0$, there is an open set O containing E for which $m^*(O \setminus E) < \epsilon$.
(ii) There is a G_δ set G containing E for which $m^*(G \setminus E) = 0$.
(iii) For any $\epsilon > 0$, there is a closed set F contained in E for which $m^*(E \setminus F) < \epsilon$.
(iv) There is an F_σ set F contained in E for which $m^*(E \setminus F) = 0$.

Theorem 6.3 (Approximation of Lebesgue measurable sets, II)**.** *Let E be a measurable set of finite measure. Then for each $\epsilon < 0$, there is a finite disjoint collection of intervals whose union O satisfies*

$$m^*(E \triangle O) < \epsilon.$$

There exist sets of real numbers that are not Lebesgue measurable. The standard construction takes a single point from every equivalence class of reals (in a certain set, say) stemming from the equivalence relation whereby two reals are equivalent if their difference is rational.

Theorem 6.4 (Vitali)**.** *Every set of real numbers of positive outer measure contains a nonmeasurable subset.*

Define $C_1 = [0, 1/3] \cup [2/3, 1]$, $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$, and so on, at each stage removing the middle third from each interval. Define the Cantor set to be the intersection of all such sets. The Cantor set is closed and has measure zero. It is also uncountable, which can be proven via a simple argument by contradiction.
Define the Cantor-Lebesgue function ϕ as follows. Let O_k be the union of intervals removed at up to the kth step, and let ϕ be constant on each of the $2^k - 1$ intervals of O_k, taking on the $2^k - 1$ values $\{1/2^k, 2/2^k, \ldots, [2^k - 1]/2^k\}$. This suffices to define ϕ on O, the complement of the Cantor set. Then define $\phi(0) = 0$ and

$$\phi(x) = \sup\{\phi(t) : t \in O \cap [0, x)\}$$

if x is in the Cantor set and $x \neq 0$. The Cantor-Lebesgue function is increasing, continuous, and is differentiable on O because it is constant there. By the intermediate value theorem, it maps $[0, 1]$ surjectively onto $[0, 1]$. Because $m(O) = 1$, one can conclude that any reasonable version of the fundamental theorem of calculus must fail for this function.
Now consider the function $\psi(x) = \phi(x) + x$. It is strictly increasing, maps the Cantor set onto a measurable set of positive measure 1, and maps a measurable set onto a nonmeasurable set. Strictly increasing continuous functions map Borel sets to Borel sets, so we can conclude that there exists a Lebesgue measurable set which is not Borel.

6.2. **Functions on the real line.** As noted in greater generality previously, continuous functions are Lebesgue measurable. Monotone functions defined on intervals are also measurable (one proves this by first proving it for strictly increasing functions, and then by approximating a monotone function pointwise by such functions).
The composition of two measurable functions need not be measurable, an example of which can be constructed easily using the function ψ above. If, however, f is continuous and g is measurable, then $f \circ g$ is measurable.
As a particular case of a general theorem about Radon measures to be proven, we have the following:

Theorem 6.5 (Lusin's theorem)**.** *Let f be a real-valued measurable function on $E \subseteq \mathbb{R}$. For each $\epsilon > 0$, there is a continuous function g and a closed set $F \subseteq E$ on which $f = g$ and $m(E \setminus F) < \epsilon$.*

Proof sketch. Finite case first; approximate simple functions appropriately. Then use Egoroff's theorem (restricting a bit further) to achieve uniform convergence of the approximations to simple functions. □

Because of the structure of the measure space on the real line, for the Vitali convergence theorem in this setting we do not have to assume that f is integrable over X; it is instead a conclusion of the theorem.
Finally, we might as well mention one thing about Riemann integration.

Theorem 6.6 (Lebesgue)**.** *Let f be a bounded function on a closed, bounded interval. Then f is Riemann integrable if and only if the set of points at which f fails to be continuous has measure zero.*

6.3. **Differentiation and integration.** A monotone function is continuous at all points except possibly countably many, which is a corollary to the aforementioned fact that the only discontinuities that can occur for monotone functions are jump discontinuities. A reasonably easy construction with a convergent series shows that for any countable subset of an interval, there exists an increasing function on that interval discontinuous on precisely that subset.
A collection $\mathcal{F}$ of closed, bounded, nondegenerate intervals is said to be a *Vitali covering* of a set E if for each $x \in E$ and $\epsilon > 0$ there is an interval $I \in \mathcal{F}$ such that $x \in I$ and $m(I) < \epsilon$.

Theorem 6.7 (Vitali covering theorem)**.** *Let E be a set of finite outer measure and $\mathcal{F}$ a Vitali covering of E. Then for each $\epsilon > 0$, there is a finite disjoint subcollection $\{I_k\}$ of $\mathcal{F}$ such that*

$$m^* \left(E \setminus \bigcup_{k=1}^{n} I_k \right) < \epsilon.$$

Proof sketch. The proof involves inductively selecting a disjoint sequence of intervals I_k such that

$$E \setminus \bigcup_{k=1}^{n} I_k \subseteq \bigcup_{k=n+1}^{\infty} 5 * I_k,$$

where $5 * I_k$ is the closed interval that has the same midpoint as I_k but five times the length. At each stage, pick an interval disjoint from those already picked and with length larger than half the supremum of the lengths of the other intervals disjoint from those already picked. After some thought, this will suffice for the set inclusion given above, which is enough to quickly prove the theorem. □

Given a real-valued function f, define the *lower and upper derivatives* at x to be

$$\overline{D}f(x) = \lim_{h \to 0} \left[\sup_{0<|t|\leq h} \frac{f(x+t) - f(x)}{t} \right],$$
$$\underline{D}f(x) = \lim_{h \to 0}] \left[\inf_{0<|t|\leq h} \frac{f(x+t) - f(x)}{t} \right],$$

respectively. If the two values are equal then we say f is differentiable at x and notate the value as $f'(x)$.

Theorem 6.8 (Mean value theorem for upper derivatives)**.** *Let f be an increasing function on $[a, b]$. For each $\alpha > 0$,*

$$m^*\{x \in (a, b) : \overline{D}f(x) \geq \alpha\} \leq \frac{1}{\alpha}[f(b) - f(a)]$$

and $m^*\{x \in (a, b) : \overline{D}f(x) = \infty\} = 0$.

Proof sketch. Let α' be some number such that $0 < \alpha' < \alpha$. Let $\mathcal{F}$ be the collection of closed, bounded intervals $[c, d]$ contained in (a, b) such that $f(d)-f(c) \geq \alpha'(d-c)$. Then $\mathcal{F}$ is a Vitali covering of the set

$$E_\alpha = \{x \in (a, b) : \overline{D}f(x) \geq \alpha\}.$$

By the Vitali covering theorem, we can extract a finite disjoint collection of intervals from $\mathcal{F}$ that come ϵ-close to covering E_α. Working through the details, we conclude that

$$m^*(E_\alpha) \leq \frac{1}{\alpha'}[f(b) - f(a)] + \epsilon,$$

which suffices. □

Theorem 6.9 (Lebesgue's theorem)**.** *If f is monotone on an open interval (a, b) then it is differentiable almost everywhere on (a, b).*

Proof sketch. It suffices to prove that

$$E_{\alpha,\beta} = \{x \in (a, b) : \overline{D}f(x) > \alpha > \beta > \underline{D}f(x)\}$$

has measure zero, which can be done with the Vitali covering lemma and an application of the above mean value theorem. □

Let f be integrable over $[a, b]$, and extend f to take the value $f(b)$ on $(b, b+1]$. For $0 < h \leq 1$, define the *divided difference function*

$$\mathrm{Diff}_h\, f(x) = \frac{f(x + h) - f(x)}{h}$$

and the *average value function*

$$\mathrm{Av}_h\, f(x) = \frac{1}{h}\int_x^{x+h} f$$

for all $x \in [a, b]$. A change of variables yields the relation

$$\int_u^v \mathrm{Diff}_h\, f = \mathrm{Av}_h\, f(v) - \mathrm{Av}_h\, f(u),$$

which is a sort of discrete version of the fundamental theorem of calculus.

Theorem 6.10. *Let f be increasing on $[a, b]$. Then f' is integrable over $[a, b]$ and*

$$\int_a^b f' \leq f(b) - f(a).$$

Proof sketch. By Lebesgue's theorem, $\{\mathrm{Diff}_{1/n}\, f\}$ converges to f' pointwise a.e., so by Fatou

$$\int_a^b f' \leq \liminf_{n\to\infty} \left[\int_a^b \mathrm{Diff}_{1/n}\, f\right].$$

By a change of variables and monotonicity, on the other hand,

$$\int_a^b \mathrm{Diff}_{1/n} f \leq f(b) - f(a),$$

and these estimates suffice. □

Recall that this inequality may in fact be strict, as the example of the Cantor-Lebesgue function demonstrates. In the absence of monotonicity, even if f' exists it may not be integrable. The standard example is $f(x) = x^2 \sin(1/x^2)$ on $(0, 1]$, $f(0) = 0$; its derivative exists on $(0, 1]$ but is not integrable there.
Let f be a real-valued function on the bounded interval $[a, b]$. Define the *variation* of f with respect to a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ to be

$$V(f, P) = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|.$$

Define the *total variation* $TV(f)$ of f to be the supremum of the above over all partitions P. If the total variation is finite, we say f is of *bounded variation*. It is easy to see that all Lipschitz functions are of bounded variation.

Theorem 6.11 (Jordan's theorem)**.** *A function f on $[a, b]$ is of bounded variation if and only if it is the difference of two increasing functions on $[a, b]$.*

Proof sketch. If f is of bounded variation, one can check that

$$f(x) = [f(x) + TV(f|_{[a,x]})] - TV(f|_{[a,x]})$$

is a desired decomposition. The other direction is a simple estimate. □

This decomposition is called a *Jordan decomposition*. As a corollary, by the work after Lebesgue's theorem above, if f is of bounded variation on $[a, b]$ then it is differentiable almost everywhere on (a, b) and f' is integrable on $[a, b]$.
A real-valued function is *absolutely continuous* on $[a, b]$ provided for each $\epsilon > 0$, there is a $\delta > 0$ such that for every finite disjoint collection $\{(a_k, b_k)\}_{k=1}^n$ of open intervals in (a, b),

$$\text{if } \sum_{k=1}^{n} [b_k - a_k] < \delta, \text{ then } \sum_{k=1}^{n} |f(b_k) - f(a_k)| < \epsilon.$$

Immediately from the definition, absolutely continuous functions are uniformly continuous. The converse is not true; the Cantor-Lebesgue function is continuous on a compact set, hence uniformly continuous, but easily seen to be not absolutely continuous. All Lipschitz functions are easily seen to be absolutely continuous; again, the simple example $f(x) = \sqrt{x}$ on $[0, 1]$ shows that the converse is false. Some standard estimates show that the definition of absolutely continuous can be phrased equivalently in terms of all measurable sets, not just finite disjoint unions of intervals. This can be used to show that an increasing absolutely continuous function maps sets of measure zero to sets of measure zero, and therefore (after some more work) that it maps measurable sets to measurable sets.

Theorem 6.12. *Let f be absolutely continuous on $[a, b]$. Then it is the difference of increasing absolutely continuous functions, and in particular f is of bounded variation.*

Proof sketch. First prove that f is of bounded variation, then prove that its total variation function is absolutely continuous. □

Theorem 6.13. *Let f be continuous on $[a, b]$. Then f is absolutely continuous if and only if the family of divided difference functions $\{\mathrm{Diff}_h f\}_{[0<h\leq 1]}$ is uniformly integrable over $[a, b]$.*

Proof sketch. A bit tricky but not all that enlightening. If the divided differences are uniformly integrable, use the discrete version of the fundamental theorem of calculus to bound the total variation of $\mathrm{Av}_h f(x)$, then take a limit as $h \to 0$ to bound the total variation of f. In the other direction, to show that a function is uniformly integrable one has to control its integral over all small measurable subsets, so approximate from the outside with a G_δ and use the continuity of integration to reduce the problem to sets that are finite disjoint unions of open intervals. □

By Lebesgue's theorem, the divided difference functions converge pointwise a.e., so by Vitali's theorem we can pass the limit under the integral sign in the "discrete fundamental theorem of calculus" and conclude that

$$\int_{[a,b]} f' = f(b) - f(a)$$

for absolutely continuous functions.

Theorem 6.14. *A function f on a closed bounded interval $[a, b]$ is absolutely continuous if and only if it is expressible as an indefinite integral over $[a, b]$.*

Proof sketch. If f is absolutely continuous, the fundamental theorem of calculus gives the result. If f is an indefinite integral, we can verify absolute continuity directly. □

If f is monotone, it is easy to check that f is absolutely continuous if and only if the fundamental theorem of calculus holds.
Using this form, we can prove the other as well, although the proof is relatively uninteresting:

Theorem 6.15. *Let f be integrable over $[a, b]$. Then*

$$\frac{d}{dx}\left[\int_a^x f\right] = f(x)$$

for almost all $x \in (a, b)$.

A function is *singular* if its derivative vanishes almost everywhere. An absolutely continuous function is singular if and only if it is constant. Given a function f of bounded variation, define

$$g(x) = \int_a^x f', \qquad h(x) = f(x) - \int_a^x f'.$$

The function g is an indefinite integral, hence absolutely continuous. The function h is singular by the above theorem applied to f'. This decomposition of a function of bounded variation into an absolutely continuous part and a singular part is called the *Lebesgue decomposition.*

6.4. **Convexity.** A real-valued function ϕ on (a, b) is *convex* provided for $x_1, x_2 \in (a, b)$ and $0 \leq \lambda \leq 1$,

$$\phi(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\phi(x_1) + (1-\lambda)\phi(x_2).$$

That is, each point on the chord connecting two points on the graph of ϕ lies above the graph of ϕ. We can also rewrite the condition as

$$\frac{\phi(x) - \phi(x_1)}{x - x_1} \leq \frac{\phi(x_2) - \phi(x)}{x_2 - x}$$

for all $x_1 < x < x_2$ in (a, b). That is, the slope of the chord from one point on the graph to a second is no larger than the slope from the second point to a third, if the points are ordered in increasing x-value. Using the mean value theorem, if ϕ is differentiable and its derivative ϕ' is increasing, then ϕ is convex; in particular, if ϕ is twice differentiable then it is convex if and only if ϕ'' is nonnegative.
Finally, if ϕ is convex and p_1, p_2, p_3 are three points on its graph in increasing x-value, then the slope of the chord connecting p_1 and p_2 is less than or equal to the slope of the chord connecting p_1 and p_3, which is less than or equal to the slope of the chord connecting p_2 and p_3. The proof of this last fact relies on the fact that a convex function has one-sided derivatives at every point. In particular, a convex function is Lipschitz (and therefore absolutely continuous) on each closed, bounded subinterval of its domain. It is differentiable except at countably many points (because the right and left hand derivatives exist, it can only fail to be defined at a jump discontinuity) and its derivative is an increasing function.

Theorem 6.16 (Jensen's inequality)**.** *Let ϕ be a convex function on $(-\infty, \infty)$, f integrable on $[0, 1]$, and $\phi \circ f$ integrable on $[0, 1]$. Then*

$$\phi\left(\int_0^1 f\right) \leq \int_0^1 \phi \circ f.$$

Proof sketch. We use the fact that there is at least one *supporting line* (line always lying beneath the graph) of the graph of ϕ at each point, and considering that supporting line at the point $(\alpha, \phi(\alpha))$, where $\alpha = \int_0^1 f$. □

If not applied to an interval of measure 1, some rescaling must be done.

6.5. **Lebesgue measure in $\mathbb{R}^n$.** The Lebesgue measure μ_n on $\mathcal{R}^n$ defined as the product measure of the Lebesgue measure on $\mathbb{R}$. It is therefore complete, as a Carathéodory extension. It is both outer regular:

$$\mu_n(E) = \inf\{\mu_n(O) : O \text{ open}, E \subseteq O\}$$

and inner regular:

$$\mu_n(E) = \sup\{\mu_n(K) : K \text{ compact}, K \subseteq E\}$$

if E is measurable. Therefore there is the same equivalence in terms of outer approximation by G_δ sets and inner approximation by F_σ sets as in the one-dimensional case.
The measure μ_n is translation invariant and all invertible linear maps send Lebesgue measurable sets to Lebesgue measurable sets, and therefore Lebesgue measurable functions to Lebesgue measurable functions. By building up from simple linear transformations, we can prove the following:

Theorem 6.17. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be invertible and $f : \mathbb{R}^n \to \mathbb{R}$ be integrable over $\mathbb{R}^n$. Then $f \circ T$ is integrable and we have*

$$\int_{\mathbb{R}^n} f \circ T \, d\mu_n = \frac{1}{|\det T|} \int_{\mathbb{R}^n} f \, d\mu_n.$$

Any mapping $\mathbb{R}^n$ to $\mathbb{R}^n$ that preserves Euclidean distance (a *rigid mapping*) must preserve Lebesgue measure.

6.6. **Lebesgue-Stieltjes integration.** Let μ be a *Borel measure* on $[a, b]$; that is, a finite measure on the Borel subsets of $[a, b]$. Its *cumulative distribution function* is given by

$$g_\mu(x) = \mu([a, x]).$$

A cumulative distribution is increasing and continuous on the right; conversely, every function on $[a, b]$ that is increasing and continuous on the right is the cumulative distribution function of a unique Borel measure. The measure μ is absolutely continuous with respect to Lebesgue measure if and only if g_μ is absolutely continuous. Given an increasing function g on $[a, b]$ that is continuous on the right and a Borel measurable function f, define the *Lebesgue-Stieltjes* integral of f with respect to g to be

$$\int_{[a,b]} f \, dg = \int_{[a,b]} f \, d\mu_g.$$

If g is absolutely continuous, then this is equal to the Lebesgue integral of $f \cdot g'$. To put this another way, in terms of the Radon-Nikodym derivative, we have $d\mu_g/dm = g'$.

7. L^p SPACES

7.1. **Basics and completeness.** The L^p spaces on a measure space X are defined in the usual way, as collections of equivalence classes of a.e.-equal functions whose pth powers are integrable:

$$\int_X |f|^p < \infty.$$

The norm of a function in L^p is this value. The L^∞ space is likewise a collection of equivalence classes of essentially bounded functions; i.e. those for which there is a constant M such that $|f| \leq M$ a.e. on X. The norm of an L^∞ function is the infimum of its essential upper bounds. All of these are linear spaces.

Theorem 7.1 (Basic inequalities)**.** *If $f \in L^p$ and $g \in L^q$, where p and q obey $1/p + 1/q = 1$ (and $1/\infty = 0$), then we have Hölder's inequality*

$$||f \cdot g||_1 \leq ||f||_p ||g||_q.$$

If $f \neq 0$, there is a function $f^* \in L^q$ such that $||f^*||_q = 1$ and

$$\int_X f f^* \, d\mu = ||f||_p;$$

this function is given explicitly by

$$f^* = \text{sgn}(f) \frac{|f|^{p-1}}{||f||_p^{p-1}}.$$

Finally, if $f, g \in L^p$, then we have Minkowski's inequality

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

so the L^p spaces are normed linear spaces.

Proof sketch. First establish Hölder's inequality using Young's inequality: for any nonnegative a, b and $p > 1$,

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q.$$

Use the existence of f^* to prove Minkowski's inequality. □

If X is a finite measure space, then Hölder's inequality proves that $L^{p_2}(X, \mu) \subseteq L^{p_1}(X, \mu)$ for $1 \leq p_1 < p_2 \leq \infty$. For a function f in L^{p_2}, it gives a specific bound for $||f||_{p_1}$ in terms of $||f||_{p_2}$ and $\mu(X)$. This allows us to prove that for $p > 1$, a bounded sequence of functions in L^p is uniformly integrable.

Theorem 7.2 (Riesz-Fischer)**.** *The L^p spaces are complete, hence Banach spaces. If we have a convergent sequence in L^p, there is a subsequence converging pointwise a.e. on X.*

Proof sketch. Use rapidly Cauchy sequences in L^p, which not only converge in L^p but also converge pointwise a.e. on X. □

As a corollary, simple functions vanishing outside a set of finite measure are dense in L^p. The following is a simple application of the usual Vitali convergence theorem:

Theorem 7.3 (Vitali convergence criterion in L^p)**.** *Let $1 \leq p < \infty$ and $\{f_n\}$ converge pointwise a.e. to a function f, all functions being in L^p. Then $\{f_n\} \to f$ in L^p if and only if $\{|f_n|^p\}$ is uniformly integrable and tight.*

7.2. **Riesz and Kantorovitch representation theorems.** Let $1 \leq p < \infty$. Given $f \in L^q(X, \mu)$, let T_f denote the functional $L^p(X, \mu) \to \mathbb{R}$ given by

$$T_f(g) = \int_X fg \, d\mu.$$

By Hölder, this is a bounded linear functional of norm $||f||_q$, so $T : L^q(X, \mu) \to (L^p(X, \mu))^*$ is an isometry. If X is σ-finite, it is also onto:

Theorem 7.4 (Riesz representation theorem)**.** *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space, $1 \leq p < \infty$, and q the conjugate of p. Then T, as defined above, is an isometric isomorphism of $L^q(X\mu)$ onto $(L^p(X, \mu))^*$.*

Proof sketch. Given a bounded linear function S, the proof proceeds by proving that $E \mapsto S(\chi_E)$ defines a measure absolutely continuous with respect to μ. Then by the Radon-Nikodym theorem there is a function f such that

$$S(\chi_E) = \int_E f \, d\mu.$$

Then it can be shown that f is in L^q and

$$S(g) = \int_X fg \, d\mu$$

generally for all $g \in L^p$. □

The dual of L^∞ is more complicated. If $(X, \mathcal{M})$ is a measurable space and $\nu : \mathcal{M} \to \mathbb{R}$ a finitely additive set function, define the total variation in the usual way:

$$|\nu|(E) = \sup \sum_{k=1}^{n} |\nu(E_k)|$$

where the supremum is taken over finite disjoint collections of sets in $\mathcal{M}$ that are contained in E. If $||\nu||_{var} = |\nu|(X) < \infty$, we say that ν is a *bounded finitely additive signed measure*. We can define integrals with respect to such a ν; they are well-defined although they lack some of the nice properties we are used to. Denote by $BFA(X, \mathcal{M}, \mu)$ the (normed linear) space of bounded finitely additive signed measures ν on $\mathcal{M}$ that are absolutely continuous with respect to ν.

Theorem 7.5 (Kantorovitch representation theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space. For $\nu \in BFA(X, \mathcal{M}, \mu)$, define $T_\nu : L^\infty(X, \mu) \to \mathbb{R}$ by*

$$T_\nu(f) = \int_X f \, d\nu.$$

Then T is an isometric isomorphism of $BFA(X, \mathcal{M}, \mu)$ onto $(L^\infty(X, \mu))^*$.

By the Riesz representation theorem, for σ-finite measure spaces the Banach space L^p, $1 < p < \infty$, must be reflexive. Every bounded sequence in a reflexive Banach space has a weakly convergent subsequence, so this property holds for L^p, $1 < p < \infty$.

Two quick results about weak compactness:

Theorem 7.6 (Radon-Riesz theorem). *Let $(X, \mathcal{M}, \mu)$ be σ-finite, $1 < p < \infty$, and $\{f_n\}$ a sequence in L^p converging weakly to f. Then $\{f_n\}$ converges strongly if and only if $\lim_{n\to\infty} ||f_n||_p = ||f||_p$.*

Because L^1 is not in general reflexive, there are bounded sequences without weakly convergent subsequences. However:

Theorem 7.7 (Dunford-Pettis theorem). *For a finite measure space $(X, \mathcal{M}, \mu)$ and a bounded sequence $\{f_n\}$ in $L^1(X, \mu)$, $\{f_n\}$ is uniformly integrable if and only if every subsequence of $\{f_n\}$ has a further subsequence that is weakly convergent in L^1.*

In particular, this holds if the f_n are all dominated by some $g \in L^1$.

8. Basic harmonic analysis

See Katznelson.

9. Miscellaneous

In blurb form.

A Borel measure on a topological space is a measure on the Borel σ-algebra such that every compact subset has finite measure. A Radon measure is a Borel measure that also possesses outer regularity (by open sets) and inner regularity (by compact sets).

The Riesz-Markov theorem says that every positive linear functional on $C_c(X)$, X locally compact, is given by integration against a unique Radon measure.

The Riesz representation theorem for $C(X)$ says that if X is compact Hausdorff, then the dual space of $C(X)$ is given by the space of signed Radon measures. If $X = [a, b]$, we can rephrase this: let $\mathcal{F}$ be the family of functions on $[a, b]$ that are of bounded variation, continuous on the right, and vanish at a. Then for each element ϕ of $(C([a, b]))^*$, there is a unique $g \in \mathcal{F}$ such that

$$\phi(f) = \int_{[a,b]} f \, dg.$$