Notes for real analysis

John Kerl

February 3, 2008

Abstract

The following are notes to help me prepare for the University of Arizona math department's Real Analysis qualifier in August 2006. Disclaimer: Nothing in this paper is claimed to be true. Rather, what is written here reflects my current understanding, however erroneous that may be. *This paper is under construction.*

Contents

1 Principles

Royden ([**Roy**], section 3.6) attributes to Littlewood the following three principles of analysis:

(i) Every measurable set is nearly a finite union of intervals;

(ii) every measurable function is nearly continuous;

(iii) every convergent sequence of measurable functions is nearly uniformly convergent.

The point here is that when we don't know what to do, we can *approximate*.

Here are some themes I see in the qual packets and qual review sessions:

- Know the main theorems with **precise statements** of their **hypotheses** and **conclusions**. On the qual we will not only need to state the name of the theorem we're using (e.g. when interchanging the order of iterated integrals) but also we will need to write little lemmas to *prove* that the theorem applies.
- **Name-dropping**: If I've written up a response which does not make use of a theorem named after someone, I've probably missed something: either I've done something incorrect, or I've done a step without sufficient justification.
- **Work avoidance**: Before performing a computation, see if it's necessary. For example (this is a geometry example), don't compute the projection of a vector onto a plane if a quick dot-product computation reveals that the vector is already parallel or perpendicular to the plane. Don't worry about computing finite change-of-variable scale factors in an integral if all you need to show is that the integral is infinite. This save valuable time, and will also demonstrate competence.
- **Memorization**: Know some standard integrals, power-series expansions, etc. so that they may be recognized on sight. [xxx insert xrefs].
- **Association**: A lower-level exam might tell you the name of a theorem, merely asking you to state it. A qual requires you to free-associate among many possibilities. Furthermore, some experimentation is often required in a qual response. Throughout my qual-prep notes I look for opportunites to populate my free-association list. For example, whenever I see a double integral, I should think Tonelli and Fubini (theorems 7.11 and 7.12); whenever I see an integral inequality I should think of Hölder (proposition 9.9); etc.

2 Topology

Definition 2.1. Let X be a set. Let $\mathcal{T}$ be a subset of $P(X)$. Then $\mathcal{T}$ is a **topology** if it is closed under:

- arbitrary (even uncountable) unions and
- finite intersections

and if

- $X \in \mathcal{T}$ and
- $\emptyset \in \mathcal{T}$.

If X has a topology $\mathcal{T}$, then $(X, \mathcal{T})$ is called a **topological space**.

Definition 2.2. Let Γ be a collection of closed subsets of a topological space X. Then Γ has the **finite intersection property** if every finite subcollection of Γ has non-empty intersection.

There are multiple characterization of **compactness**. A topological space X is compact if:

- Every open cover has a finite subcover.
- Every sequence has a convergent subsequence.
- Every collection of closed sets with the finite intersection property has a common point.
- For every collection of sets with the finite intersection property, there exists a point near (in the closure of) each set.
- X is the continuous image of another compact topological space.
- it is a subset of a metric space, and is **complete** and **totally bounded**. (The latter means is can be covered by finitely many sets of finite diameter.)

Theorem 2.3 (Tychonoff product theorem)**.** *An arbitrary product of compact topological spaces is compact.*

Theorem 2.4. *cts image of cpct is cpct*

Theorem 2.5. *cts image of connd is connd*

Definition 2.6. A topological space is **separable** if it contains a countable dense subset.

Example 2.7. The rationals, which are countable, are dense in the reals (with the usual topology). Likewise $\mathbb{R}^n$ (finite-dimensional Euclidean space). (Triv topo on any set since any singleton is dense.) (Cite Wikipedia here . . .)

xxx need a non-example as well. Use the reals with the discrete topology? Infinite-dimensional Hilbert space — separable iff it has a countable orthonormal basis?

3 Metric spaces

Definition 3.1. metric space . . . .

Definition 3.2. A **Borel function** is a pointwise limit of a sequence of continuous real-valued functions on a metrizable space.

4 Calculus

Theorem 4.1 (Fundamental theorem of calculus)**.** *If F is an antiderivative of f, then*

$$\int_a^b f(x)dx = F(b) - F(a).$$

Corollary 4.2. *Writing $F' = f$ and applying the FTC to f', we have*

$$\begin{aligned} \int_a^x f'(t)dt &= F'(x) - F'(a) \\ &= f(x) - f(a). \end{aligned}$$

In particular, if $f(0) = 0$, then

$$\int_0^x f'(t)dt = f(x).$$

Theorem 4.3 (Second fundamental theorem of calculus)**.** *If*

$$F(x) = \int_a^x f(t)dt,$$

then

$$F'(x) = f(x).$$

Corollary 4.4.

$$\frac{d}{dx}\int_a^x f(t)dt = f(x).$$

The following is called the Extreme Value Theorem in [**Ant**]; at the graduate level, it becomes topological.

Theorem 4.5 (Extreme Value Theorem)**.** *A continuous real-valued function from a closed interval in $\mathbb{R}$ attains its maximum and minimum values.*

Proof. This is an elementary statement of the fact (theorem 2.4) that the continuous image of a compact set is compact. □

Theorem 4.6 (Intermediate Value Theorem)**.** *Let f be a continuous function from the closed interval $[a, c]$ to $\mathbb{R}$. Let $x = f(a)$ and $z = f(c)$. If $x < y < z$ or $x > y > z$ then there is $a < b < c$ such that $y = f(b)$.*

Proof. This is an elementary statement of the fact (theorem 2.5) that the continuous image of a connected set is connected. □

Rolle

Cor: MVT

xxx 1st, series $\sum a_n$ and tests. Then, power series.

Definition 4.7. A **power series** is a formal expression of the form

$$\sum_{n=0}^{\infty} a_n x^n.$$

Coefficients are taken to be in $\mathbb{R}$ or $\mathbb{C}$. If we think of a power series as a *function* to $\mathbb{R}$ or $\mathbb{C}$, respectively, we immediately ask for what x's the series converges to a real number. It is shown in [**CB**], for the complexes, that there is a **radius of convergence** (perhaps 0, perhaps infinite) within which the series converges for all x and outside of which it converges for no x. Behavior on the border circle is less simple and is not addressed here.

Criteria [**Ros**] for finding the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$

- When the limit exists:
$$\lim \left| \frac{a_n}{a_{n+1}} \right|.$$
- With the conventions that $1/\infty = 0$ and $1/0 = \infty$:
$$\frac{1}{\limsup \sqrt[n]{|a_n|}}.$$

You can **integrate** and **differentiate term by term** anywhere inside the radius of convergence.

Three handy facts:

(1) The **harmonic series** diverges:
$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty$$

(2) The **alternating harmonic series** converges. Specifically:
$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = \ln(2).$$

(3) The **square harmonic series** (for lack of a better name) converges. Specifically:
$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Remark 4.8. Here are some power series (all Maclaurin series) which we should recognize on sight:

- Exponential function:
$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$
Its radius of convergence is infinite.
- Cosine:
$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos(x).$$
Its radius of convergence is infinite.
- Sine:
$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sin(x).$$

- The **geometric series**:
$$1 + x + x^2 + x^3 + \ldots = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$
Its radius of convergence is 1.
- Log:
$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = \ln(1+x).$$
Its radius of convergence is 1.
- Log:
$$x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots = \sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x).$$
Its radius of convergence is 1.

Theorem 4.9 (Implicit function theorem)**.** *Let E be an open subset of $\mathbb{R}^{m+n}$, and let $\mathbf{f} : E \to \mathbb{R}^n$ be a vector-valued function which is continuously differentiable, i.e. of type C^2. Let*
$$(\mathbf{a}, \mathbf{b}) = (a_1, \ldots, a_m, b_1, \ldots, b_n) \in E$$
be such that
$$\mathbf{f}(\mathbf{a}, \mathbf{b}) = 0.$$
Let D be the $n \times n$ submatrix of the Jacobian of f given by $\partial f_i / \partial x_j$, for $i = 1, \ldots, n$ and $j = m+1, \ldots, n$, evaluated at $\mathbf{b}$. If $\det(D) \neq 0$, then there exists a neighborhood U of $\mathbf{a}$ and a unique C^2 function $\mathbf{g} : U \to \mathbb{R}^n$ such that
$$\mathbf{g}(\mathbf{a}) = \mathbf{b}, \quad \textit{i.e.} \quad \mathbf{f}(\mathbf{a}, \mathbf{g}(\mathbf{a})) = 0,$$
and for all $\mathbf{a}' \in U$,
$$\mathbf{f}(\mathbf{a}', \mathbf{g}(\mathbf{a}')) = 0.$$
That is, we can **solve for** *the $\mathbf{b}$ variables at and near $\mathbf{a}$.*

Remark 4.10. The point is that, given a system of equations (for this course, usually a single equation), we have an easy criterion for when we can solve for some variables in terms of the others. Note however that the implicit function theorem ensures existence and uniqueness; actually *finding* the function $\mathbf{g}$ is another matter.

Remark 4.11. The theorem, as stated, has the to-be-solved-for variable(s) written last. In practice, this may not be the case. E.g. given a function $f(v, w, x, y, z) : \mathbb{R}^5 \to \mathbb{R}^2$, we might want to solve for, say, v and x. In that case, we would need to check the submatrix formed by the first and third columns of the Jacobian of f. Furthermore, we might not know ahead of time which variables to solve for, until we apply the implicit function theorem to various submatrices of the Jacobian.

Example 4.12. Let
$$f : \mathbb{R}^3 \to \mathbb{R} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x^2 + y^2 + z^2 - 1.$$
Here we have $m = 2$ and $n = 1$. Then the kernel of f is the sphere S^2. Consider the north pole, $(0, 0, 1)$, written as
$$(\mathbf{a}, \mathbf{b}) = (a_1, a_2, b_1) = (x, y, z) = (0, 0, 1).$$

The Jacobian of f is

$$\begin{pmatrix} 2x & 2y & 2z \end{pmatrix}$$

which evaluated at the north pole is

$$\begin{pmatrix} 0 & 0 & 1. \end{pmatrix}$$

Now, there is only one 1×1 submatrix of this which is non-zero, namely, the last. So, there is a neighborhood U of $(0, 0)$ and a unique function $g : U \to \mathbb{R}$ such that $z = g(x, y)$. Here, it's clear what this is: take

$$z = g(x, y) = \sqrt{1 - x^2 - y^2}.$$

Theorem 4.13 (Inverse function theorem)**.** *Let E be an open subset of $\mathbb{R}^m$, and let $\mathbf{f} : E \to \mathbb{R}^m$ be a vector-valued function which is continuously differentiable, i.e. of type C^2. If the Jacobian of $\mathbf{f}$ is nonsingular at a point $\mathbf{q}$ of E, then there exist open neighborhoods U of $\mathbf{q}$ and V of $\mathbf{f}(\mathbf{q})$ such that $\mathbf{f}$ is a diffeomorphsim from U into V.*

Remark 4.14. The point is that, even if such a function $\mathbf{f}$ is wildly non-linear, to check for invertibility at a point it suffices to make the much simpler check of the invertibility of the linearization of $\mathbf{f}$.

[xxx where to file] interchange limit and riemann integral

5 Measure spaces and measures of sets

Definition 5.1. Let X be a set. Let $\mathcal{A}$ be a subset of $P(X)$. Then $\mathcal{A}$ is a **σ-algebra** if it is closed under:

- countable unions,
- countable intersections, and
- complements,

and if

- $X \in \mathcal{M}$ and
- $\emptyset \in \mathcal{M}$.

Note that a σ-algebra needs to tolerate *more* intersections than a topology on the same space (countably infinitely many, not just finitely many), but needs to tolerate *fewer* unions (countably infinitely many, not uncountably many).

Definition 5.2. A **σ-ring** is similar to a σ-algebra but need only be closed under countable unions. Also it need not contain all of X.

Definition 5.3. Let $(X, \mathcal{T})$ be a topological space. Let $\mathcal{B}$ be the smallest σ-algebra containing $\mathcal{T}$. Then $\mathcal{B}$ is the **Borel σ-algebra** for $(X, \mathcal{T})$. (We say $\mathcal{B}$ is the **Borel σ-algebra** *generated by* the topology $\mathcal{T}$.)

Definition 5.4. A **Borel set** is an element of a Borel σ-algebra. That is, a Borel set is obtained from open sets by countable intersections, countable unions, and complements (in any order).

Definition 5.5. Let X be a set with a σ-algebra $\mathcal{M}$. A (set) **measure** on $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$, and
- If $\{E_j\}_{j=1}^{\infty}$ is a sequence of *disjoint* sets in $\mathcal{M}$ then

$$\mu\left(\cup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j).$$

  This is the **countable additivity** property.

Remark 5.6. If $\{E_j\}_{j=1}^{\infty}$ is a sequence of sets in $\mathcal{M}$ which is *not* necessarily disjoint, then

$$\mu\left(\cup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu(E_j).$$

This is the **countable subadditivity** property.

Mnemonic 5.7. Measure closed intervals $[a, b]$ in $\mathbb{R}$ by $\mu([a, b]) = b - a$. Let $E_1 = E_2 = [0, 1]$. Then

$$\mu(E_1 \cup E_2) = 1 \leq 2 = 1 + 1 = \mu(E_1) + \mu(E_2).$$

The subadditivity of measure of non-disjoint sets comes from the *overcounting* that happens when sets *overlap.*

Definition 5.8. If

- X is a set, and
- $\mathcal{M}$ is a σ-algebra on X, and
- μ is a non-negative countably additive set function,

then $(X, \mathcal{M}, \mu)$ is called a **measure space**.

Remark 5.9. Ideally we'd want to be able to "measure" the size of *any* subset of *any* set X. That is, instead of σ-algebras of X, we'd simply use all of $P(X)$. However, it can be shown ([**Fol**], chapter 1) that this is too much to hope for. Thus we restrict ourselves to measuring elements of something (much) smaller than all of $P(X)$.

Definition 5.10. Let $X = \mathbb{R}$ with the Borel σ-algebra. Then **Lebesgue measure** of a closed interval $[a, b]$ is $b - a$, written

$$\lambda([a, b]) = b - a.$$

Likewise, on $\mathbb{R}^n$, Lebesgue measure of a cube is given by its volume.

Remark 5.11. Obviously this only tells us how to measure intervals and cubes. Other sets are measured by approximating them from within and without. (Making this simple, intuitive statement precise occupies entire chapters of [**Jon**], and also takes up significant space in the other cited references. The most concise treatment of the four authors is Rudin's.) More precisely ([**Jon**], ch. 2), we measure open sets by approximating them from within by combinations of cubes. Then a set E of X is measurable iff for all $\varepsilon > 0$ there are closed and open sets F, G respectively such that

$$F \subseteq E \subseteq G$$

and such that $\lambda(G \backslash F) < \varepsilon$. Always remember that at its simplest — details and distractions notwithstanding — on Euclidean spaces, **Lebesgue measure is volume**.

Remark 5.12. For $\mathbb{R}^n$, the class of Borel sets of X is contained in the set of Lebesgue-measurable sets. See [**Jon**], ch. 5, for an example. That is, for Euclidean space, *all Borel sets are measurable. Not all measurable sets are Borel.*

Definition 5.13. Let $(X, \mathcal{M}, \mu)$ be a measure space. Then μ is said to be a **finite measure** if $\mu(X) < \infty$.

For example, take X to be the unit interval, with Lebesgue measure. Then $\mu(X) = 1$. For a non-example, take X to be $\mathbb{R}$, with Lebesgue measure.

Definition 5.14. Let $(X, \mathcal{M}, \mu)$ be a measure space. Then μ is said to be a **σ-finite measure** if there is an increasing sequence of sets $E_j \uparrow X$ (in particular I mean $\cup_j E_j = X$) such that $\mu(E_j) < \infty$.

Example 5.15. For an example, take X to be $\mathbb{R}$, with Lebesgue measure, and $E_j = [-j, j]$. For a non-example, take $X = \mathbb{R}$ with counting measure, i.e. $\mu(\{q\}) = 1$ for all singletons. It takes uncountably many points q to cover $\mathbb{R}$.

6 Integration (measures of functions)

Definition 6.1. The **extended real number system** is $\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Definition 6.2. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \tilde{\mathbb{R}}$. Note that while X has its σ-algebra $\mathcal{M}$, $\tilde{\mathbb{R}}$ has the standard Borel σ-algebra $\mathcal{B}$. The function f is said to be **Lebesgue measurable** if any of the following (which are equivalent) hold:

- For all B in $\mathcal{B}$, $f^{-1}(B)$ is a measurable set in X. (That is, preimages of measurable subsets are measurable subsets.)
- For all closed intervals $[a, b]$ of $\tilde{\mathbb{R}}$, $f^{-1}([a, b])$ is a measurable set in X.
- For all $a \in \mathbb{R}$, $\{x \in X : f(x) > a\}$ is a measurable set in X.
- For all $a \in \mathbb{R}$, $\{x \in X : f(x) \geq a\}$ is a measurable set in X.
- For all $a \in \mathbb{R}$, $\{x \in X : f(x) < a\}$ is a measurable set in X.
- For all $a \in \mathbb{R}$, $\{x \in X : f(x) \leq a\}$ is a measurable set in X.

Remark 6.3. Lebesgue measurability is distinct from Lebesgue integrability (below).

Remark 6.4. The set of measurable functions is closed under pointwise sums, differences, and products, pointwise absolute value, pointwise max and min, pointwise limits, but not composition.

Definition 6.5. Let $(X, \mathcal{M}, \lambda)$ be as above. A function $f : X \to \tilde{\mathbb{R}}$ is a **simple function** iff $f(X)$ is finite. (Think of step functions, and indicator functions of sets.)

Definition 6.6. Let $(X, \mathcal{M}, \lambda)$ be as above, and let $E : X \to \tilde{\mathbb{R}}$ be a simple function. Let $e_1, \ldots, e_n$ be the finite number of points in $E(X)$. The **Lebesgue integral** of E is defined to be

$$\lambda(E) = \int_X E d\lambda = \sum_{j=1}^{n} e_j \lambda(f^{-1}(e_j))$$

where the λ on the left-hand side is a function measure, and the λ on the right-hand side is a set measure. (Think of computing the areas of step functions. This seems not too different from the Riemann integral, until you consider, say, the simple functions which is the indicator of the rationals on the unit invterval.)

Definition 6.7. Let $(X, \mathcal{M}, \lambda)$ be as above and let $f : X \to \tilde{\mathbb{R}}$. The **Lebesgue integral** of f is obtained by approximating f by simple functions:

$$\sup_{L \leq f} \lambda(L) = \inf_{U \geq f} \lambda(U),$$

with the sup and inf taken over *all* simple functions L and U, then

$$\lambda(f) = \int_X f d\lambda$$

is equal to this common value. More practically, Namely, let $L_k \uparrow f$ and $U_k \downarrow f$. If

$$\lim_{k \to \infty} \lambda(L_k) = \lim_{k \to \infty} \lambda(U_k)$$

then we write

$$\lambda(f) = \lim_{k \to \infty} \lambda(L_k) = \lim_{k \to \infty} \lambda(U_k).$$

I believe that this is the essential picture, modulo some fuss about details and corner cases. Note in particular, though, that we only attempt any of this for functions which are Lebesgue measurable.

Construction of the Lebesgue integral (Assane Lo calls this the **standard machine**):

- Characteristic functions
- Simple functions
- Non-negative measurable functions
- Arbitrary measurable functions: $\lambda(f) = \lambda(f_+) - \lambda(f_-)$. This is defined if at most one of $\lambda(f_+)$ and $\lambda(f_-)$ are finite. Otherwise, we have an $\infty - \infty$ situation.

Remark 6.8. When the qual problems ask to show that something is *not* Lebesgue integrable, typically one of two techniques will suffice:

- Show that $\int f_+$ and $\int f_-$ are both infinite. (An example of this is below.)
- For two-variable integrals, if you can show that $\int\int f(x,y)dx\,dy$ and $\int\int f(x,y)dy\,dx$ are different, then f is not Lebesgue integrable.

Example 6.9. The function $f(x) = x$ is Lebesgue measurable but not Lebesgue integrable: $f^{-1}([a,b]) = [a,b]$ which is measurable in $\mathbb{R}$. However, $\int_{-\infty}^{\infty} f_+(x)dx$ and $\int_{-\infty}^{\infty} f_-(x)dx$ are both infinite.

Remark 6.10. Why isn't

$$\int_{-\infty}^{\infty} x\,dx = 0?$$

Isn't this just

$$\lim_{a\to\infty} \int_{-a}^{a} x\,dx,$$

which is clearly zero? The problem is in *how* we go to infinity; the integral balances on a knife edge. If we take

$$\lim_{a\to\infty} \int_{-a}^{2a} x\,dx,$$

then we get something different. Going to infinity symmetrically, via

$$\lim_{a\to\infty} \int_{-a}^{a} x\,dx,$$

is the **Cauchy principal value** of the integral. This becomes more clear when we consider

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0. \end{cases}$$

Clearly the principal value of the integral is zero, via

$$\lim_{a\to\infty} \int_{-a}^{a} \operatorname{sgn}(x)dx,$$

but it could be any t via

$$\lim_{a\to\infty} \int_{-a}^{a+t} \operatorname{sgn}(x)dx.$$

Specifically, this example shows that the Cauchy principal value is not **translation invariant**.

Example 6.11. The function $f(x) = \sin(x)/x$ is not Lebesgue integrable on $(0,\infty)$. We need f_+ and f_- to each be integrable. [xxx type up the details about bounding $\sin(x)/x$ below by $\sin(x)/2n\pi$ on $[2n\pi,(2n+1)\pi]$ and obtaining the (divergent) harmonic series.]

7 Theorems and inequalities for integration

Fact:

$$|\mu(f)| \leq \mu(|f|).$$

Mnemonic 7.1. How to remember which way the signs go? Use sine on $[0, 2\pi]$:

$$\left|\int_0^{2\pi} \sin(x)dx\right| = 0 \leq \int_0^{2\pi} |\sin(x)|dx = 2\int_0^{\pi} \sin(x)dx = 4.$$

On the left, the up-hump and the down-hump cancel to zero. On the right, the pair of up-humps gives a positive area.

Mnemonic 7.2. Another mnemonic: I call this the **toothpaste theorem**, since when you squeeze it, more comes out:

$$\left|\int f(x)dx\right| \leq \int |f(x)|\, dx.$$

xxx need **Chebyshev inequality**?

Fundamental norm identity:

$$\begin{aligned} \|u+v\|^2 &= \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2; \\ \|u-v\|^2 &= \|u\|^2 - \langle u, v\rangle - \langle v, u\rangle + \|v\|^2. \end{aligned}$$

Real case:

$$\begin{aligned} \|u+v\|^2 &= \|u\|^2 + 2\langle u, v\rangle\rangle + \|v\|^2; \\ \|u-v\|^2 &= \|u\|^2 - 2\langle u, v\rangle\rangle + \|v\|^2. \end{aligned}$$

Triangle inequality:

$$\|u \pm v\| \leq \|u\| + \|v\|.$$

Reverse triangle inequality:

$$\|u \pm v\| \geq |\ \|u\| - \|v\|\ |.$$

Together:

$$|\ \|u\| - \|v\|\ | \leq \|u \pm v\| \leq \|u\| + \|v\|.$$

Parallelogram law:

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

Cauchy-Schwartz inequality:

$$\begin{aligned} |\langle \mathbf{u}, \mathbf{v}\rangle|^2 &\leq \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle \\ |\langle \mathbf{u}, \mathbf{v}\rangle| &\leq \|\mathbf{u}\|_2\, \|\mathbf{v}\|_2 \end{aligned}$$

See also Cauchy-Schwarz for functions: proposition 9.10.

An elementary but handy inequality:

$$ab \leq a^2 + b^2.$$

Proof.

$$\begin{aligned} 0 &\leq (a-b)^2 \\ 0 &\leq a^2 - 2ab + b^2 \\ 2ab &\leq a^2 + b^2 \\ ab &\leq \frac{a^2+b^2}{2} \leq a^2 + b^2. \end{aligned}$$

□

Definition 7.3. A function is **absolutely continuous** if it is continuous and is an indefinite integral.

Theorem 7.4 (Monotone convergence theorem)**.** *If $f_n \geq 0$ and $f \geq 0$, and if $f_n \uparrow f$ pointwise, then $\mu(f_n) \uparrow f$.*

xxx note rel'nship with countable additivity (1-11 notes).

Theorem 7.5 (Improved monotone convergence theorem)**.** *If $-\infty < f_1$, and if $f_n \uparrow f$ pointwise, then $\mu(f_n) \uparrow f$.*

Theorem 7.6 (Fatou's lemma)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. [xxx do we assume μ is Lebesgue measure λ?] Let f_n be a sequence of non-negative measurable functions. If*

$$f(x) = \liminf_{n\to\infty} f_n(x)$$

then

$$\liminf_{n\to\infty} \int_X f_n d\mu \geq \int_X f d\mu$$

i.e.

$$\liminf_{n\to\infty} \int_X f_n d\mu \geq \int_X \liminf_{n\to\infty} f_n d\mu.$$

Mnemonic 7.7. One can only *lose mass* by moving the limit inside the integral. One may make this mnemonic embarrassingly more memorable by matching the *fat* in *Fatou* with *weight loss.*

Mnemonic 7.8. To remember which way the signs go, let

$$f_n = 1_{[n,n+1]}.$$

These are little squares with height one, starting at $x = n$. The integral of each is of course one, but the pointwise limit of the f_n's is $f = 0$. (The unit mass gets shifted off to infinity and so is lost in the limit.)

Theorem 7.9 (Dominated convergence theorem (DCT), or Lebesgue dominated convergence theorem (LDCT))**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Let f_n be a sequence of measurable functions on X. Let $g \geq 0$ be in $L^1(X)$. Suppose the pointwise limit of f_n exists for all $x \in X$, say to f. If*

$$|f_n(x)| \leq g(x)$$

for all $x \in X$, then $f \in L^1(X)$ and

$$\int f d\mu = \lim \int f_n d\mu.$$

Remark 7.10. That is, if the f_n's are dominated in absolute value by a function in L^1, then we can move the limit through the integral.

Theorem 7.11 (Tonelli's theorem). *Let (X, μ) and (Y, ν) be σ-finite measure spaces. If f is a non-negative real-valued function on $X \times Y$, then*

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

Theorem 7.12 (Fubini's theorem). *Let (X, μ) and (Y, ν) be complete measure spaces. If $f \in L^1(X \times Y)$, i.e. if*

$$\int_{X\times Y} |f d(\mu \times \nu)| < +\infty,$$

then

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

When can we write

$$\frac{d}{dt}\int_a^b \phi(x,t)\, dx = \int_a^b \frac{\partial \phi(x,t)}{\partial t}\, dx?$$

Theorem 7.13 (Differentiation under the integral sign: [**Rud**] p. 237). *Let $\phi(x,t) : [a,b] \times [c,d] \to \mathbb{R}$. If the following are true:*

- *for all t, ϕ is integrable along x, and*
- *there is an $s \in [c,d]$ such that $\partial\phi/\partial t$ is continuous for all $x \in [a,b]$ and for all t in some open interval around s,*

then

$$\left[\frac{d}{dt}\int_a^b \phi(x,t)\, dx\right]_{t=s} = \int_a^b \left[\frac{\partial \phi(x,t)}{\partial t}\right]_{t=s} dx.$$

Here is an alternate characterization which Assane has been using.

Theorem 7.14 (Differentiation under the integral sign). *Let $\phi(x,t) : [a,b] \times [c,d] \to \mathbb{R}$. If the following are true:*

- *$\partial\phi/\partial t$ exists almost everywhere,*
- *$|\partial\phi/\partial t| < F$ for some integrable F, and*
- *$\phi(x,t)$ is integrable,*

then

$$\left[\frac{d}{dt}\int_a^b \phi(x,t)\, dx\right]_{t=s} = \int_a^b \left[\frac{\partial \phi(x,t)}{\partial t}\right]_{t=s} dx.$$

[xxx elaborate on the following heuristic: when we have an integral problem with two parameters, e.g. e^{-st} somewhere in the integrand, go ahead and differentiate it, then move the derivative inside and hope the view improves.]

8 Weighted integrals

Definition 8.1. The **Riemann integral** is familiar from calculus.

The **Riemann-Stieltjes integral** is a Riemann integral with a weight function. Example: Let $f(x) = x$ and $w(x) = 1/(1 + x^2)$. Then integrate f with w as weight. This is just

$$\int_{-\infty}^{\infty} f(x)w(x)dx.$$

Notation: Let $W(x)$ be an antiderivative of $w(x)$, namely

$$W(x) = \begin{cases} \lambda((0, x]), & x \geq 0 \\ -\lambda((x, 0]), & x < 0. \end{cases}$$

Then we write

$$\int_{-\infty}^{\infty} f(x)dW(x) = \int_{-\infty}^{\infty} f(x)w(x)dx.$$

Mnemonic 8.2. How to remember this? The usual (uniformly weighted) Riemann integral uses $w(x) = 1$, with $W(x) = x$. Then we think of

$$dW(x) = \frac{dW}{dx}dx = w\,dx = dx.$$

This is just a mnemonic for the special case of continuous $w(x)$, though — the Riemann-Stieltjes construction works for $W(x)$ that aren't even continuous. [xxx something — use an example — of W with steps. Use these to obtain point masses.]

Definition 8.3. The **Lebesgue-Stieltjes integral**: direct analogy?

xxx say how to construct it for any non-decreasing (LSC?) W.

9 Functional analysis

Definition 9.1. Let $1 \leq p \leq \infty$. Let $(X, \mathcal{M}, \mu)$ be a measure space. For a real-valued function $f : X \to \mathbb{R}$, we write

$$\|f\|_p = \left[\int_X |f|^p \, d\mu\right]^{1/p}.$$

for the **norm** of f.

Definition 9.2. Let $1 \leq p \leq \infty$. Let Y denote the set of infinite complex-valued (or complex-valued) sequences. For a sequence $\mathbf{x} = (x_1, x_2, x_3, \ldots) \in Y$, we write

$$\|\mathbf{x}\|_p = \left[\sum_{k=1}^{\infty} |x_k|^p\right]^{1/p}.$$

for the **norm** of f.

Mnemonic 9.3. This generalizes the familiar square-root-of-sum-of-squares computation in the Euclidean distance formula.

Definition 9.4. Let $1 \leq p \leq \infty$. Let $(X, \mathcal{M}, \mu)$ be a measure space. We write $\mathcal{L}^p(X)$ for the class of real-valued functions on X such that

$$\|f\|_p < \infty.$$

Define an equivalence relation whereby $f \sim g$ iff $f = g$ almost everywhere. We write

$$L^p(X) = \mathcal{L}^p(X)/\sim.$$

Definition 9.5. We write $\ell^p(\mathbb{R})$ for the set of real-valued infinite sequences $\mathbf{x} = (x_1, x_2, x_3, \ldots)$ such that

$$\|\mathbf{x}\|_p < \infty.$$

Definition 9.6. A **Banach space** is a vector space which is normed and complete.

Mnemonic 9.7. Q: What's yellow, normed, and complete? A: A Bananach space.

Definition 9.8. A **Hilbert space** is a Banach space which is also an inner product space.

Proposition 9.9 (Hölder's inequality)**.**

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q$$

with $1/p + 1/q = 1$.

With $p = q = 2$ we have the following:

Corollary 9.10 (Cauchy-Schwarz inequality)**.**

$$\|fg\|_1 \leq \|f\|_2 \, \|g\|_2.$$

Remark 9.11. To use Hölder, one needs a product of two functions. Whenever we see an integral inequality using L^p norms, and if there is a single function f on the left-hand side, we can expect that we will need to **insert** $g = 1$, so that $\|f\|_1$ becomes $\|f \cdot 1\|_1$.

Proposition 9.12. $\|f\|_p^p = \|f^p\|_1$.

Proof.

$$\|f\|_p^p = \left[\left(\int |f|^p\right)^{1/p}\right]^p = \int |f|^p = \int |f^p| = \|f^p\|_1.$$

□

Proposition 9.13 (Minkowski inequality)**.**

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

Theorem 9.14 (Projection theorem)**.** *Let H be a Hilbert space and let M be a closed subspace of H. If $u \in H$ then there is a unique vector v in M closest to u. In particular, $u - v$ is perpendicular to M.*

Definition 9.15. A linear functional λ is said to be **bounded** if there is $M \in \mathbb{R}$ such that for all $x \in H$, $\lambda(x) \leq M\|x\|$.

Theorem 9.16 (Riesz-Fréchet representation theorem)**.** *In a Hilbert space, all bounded linear functionals L come from a u^*.*

Example 9.17. How could a linear functional failed to be bounded? How "linear" could such a functional be? In finite dimensions, this isn't possible. However, it is possible in infinite dimensions it is. Take a separable Hilbert space H — it then has a countable basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$. Let $\lambda(e_k) = k$, and extend linearly from the basis to all of H. This is clearly unbounded.

Remark 9.18. This is a familiar fact from finite-dimensional vector spaces: all linear functionals may be represented by row vectors.

Definition 9.19. Let H be a Hilbert space. Let u_n be a sequence in H, and let $u \in H$. Then u_n converges to u **weakly** iff for all $v \in H$,

$$\langle u_n, v\rangle \to \langle u, v\rangle.$$

We say that u_n converges to u **strongly** if

$$\|u_n - u\| \to 0.$$

Example 9.20. An infinite orthonormal sequence converges weakly to zero, but not strongly.

Definition 9.21. strong topology ....

Definition 9.22. Let E be a Banach space. The **weak topology** on E is the coarsest topology such that every element of E^* is continuous.

Definition 9.23. Let E be a Banach space. The **weak* topology** on $E*$ is the coarsest topology such that every element of E^{**} is continuous.

Definition 9.24. A sequence of functions f_n is said to **converge in measure** if for all $\varepsilon > 0$ there is N such that for all $n \geq N$,

$$\mu\{x : |f(x) - f_n(x)| > \varepsilon\} < \varepsilon.$$

10 Fourier analysis

xxx needs elaboration

Definition 10.1. Write

$$\langle f, g\rangle = \int \overline{f} g.$$

Remark 10.2. Note that

$$\|f\|_2^2 = \langle f, f\rangle.$$

Definition 10.3. For $f(x) \in L^2$ (usually on $[0,1]$ or $\mathbb{R}$), and for basis functions $b_k(x)$,

$$\hat{f}(k) = \langle f, b_k\rangle.$$

Remark 10.4. For $f(x) \in L^2([0,1])$, basis functions $b_k(x)$ may be taken to be $e^{i2\pi kx}$ for $k \in \mathbb{Z}$. For $f(x) \in L^2(\mathbb{R})$, the k's may take on any real value.

Definition 10.5. We write

$$f(x) * g(x) = \int f(x-y)g(y)dy.$$

This is the **convolution** of f and g.

Proposition 10.6 (Convolution for Fourier integrals)**.** *Let $\mathcal{F}$ denote the Fourier transform. Then*

$$\mathcal{F}(g * h) = \mathcal{F}(f) \cdot \mathcal{F}(g).$$

That is,

$$\widehat{g * h} = \hat{g}\hat{h}.$$

Remark 10.7. The saying is that *convolution in the time domain is multiplication in the frequency domain.*

Proposition 10.8 (Bessel's inequality, Parseval's identity, and Plancherel's theorem)**.** *Let H be a Hilbert space. Let $f \in H$ and let b_k, $k = 1, 2, 3, \ldots$, be a sequence of orthonormal vectors in H. Then*

$$\sum_{k=1}^{\infty} |\langle f, b_k\rangle|^2 \le \|f\|^2.$$

*This is **Bessel's inequality**. If the b_k's are furthermore a basis for H, then we have **Parseval's identity**:*

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

*This is a special case of **Plancherel's theorem**. (This statement of Plancherel is due to Assane Lo. See also remark 10.11 below.) Namely, for $f, g \in H$,*

$$\sum_{k=1}^{\infty} \overline{\hat{f}(k)}\hat{g}(k) = \langle f, g\rangle.$$

(Bessel and Parseval then follow for $f = g$, recalling that $\|f\|^2 = \langle f, f\rangle$ and $\hat{f}(k) = \langle f, b_k\rangle$.)

Remark 10.9. The context of this remark is a closed interval of the real line (nominally, the unit interval). Then $f(x) \in L^2([0,1])$ and $\hat{f}(k) \in \ell^2(2\pi\mathbb{Z})$. Using Fourier series we reconstruct f as a linear combination of the orthonormal basis, where the kth coefficient is the dot with the kth basis vector. That is,

$$f = \sum_{k=1}^{\infty} \langle f, b_k\rangle b_k. \qquad (*)$$

Parseval's identity guarantees that this series converges when the b_k's form a basis for H. Parseval's identity may be paraphrased by saying that when we insert norms in $(*)$, we still have an equality.

Mnemonic 10.10. Engineers routinely paraphrase Parseval's identity by saying that the power in the time-domain spectrum is the same as the power in the frequency-domain spectrum. Briefly: *power in equals power out.*

Remark 10.11. Faris and Wikipedia both state Plancherel's theorem by saying that if $f \in L^1(\mathbb{R}, dx) \cap L^2(\mathbb{R}, dx)$, then $\hat{f} \in L_2(\mathbb{R}, dk/2\pi)$. Here, we are taking the Fourier transform from $\mathbb{R}$ to $\mathbb{R}$, so both x and k are allowed to take all real values.

Proposition 10.12 (Riemann-Lebesgue lemma)**.** *If $f \in L^1([0,1])$, then $\hat{f}(k) \to 0$ as $k \to \infty$.*

11 Radon-Nikodym derivatives

Definition 11.1. Let $(X, \mathcal{A}, \mu)$ and $(X, \mathcal{A}, \nu)$ be measure spaces (on the same set — same σ-algebra too?). Assume that μ and ν are σ-finite. Also assume that ν is signed and μ is positive. The measure ν is said to be an **absolutely continuous measure** with respect to μ, written

$$\nu \ll \mu,$$

if

$$\mu(E) = 0 \implies \nu(E) = 0.$$

Example 11.2. Let ν be Lebesgue measure on $X = [-1, 1]$ (with the Borel σ-algebra) and let μ be $\mu + \delta_0$. Let $E \subseteq X$. If $0 \in E$, then $\mu(E) = 1$. If $0 \notin E$, then $\mu(E)\nu(E)$. Thus $\nu \ll \mu$.

Definition 11.3. We say h is the **Radon-Nikodym derivative** of ν with respect to μ, written

$$h = \frac{\nu}{\mu},$$

if for all measurable subsets E of X,

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

Theorem 11.4 (Radon-Nikodym derivatives exist)**.** *If $\nu \ll \mu$ as above, then a Radon-Nikodym derivative exists.*

Example 11.5. Let μ and ν be as in example 11.2. Then take $h(x) = d\nu/d\mu$ to be 1 if $x \neq 0$ and 0 if $x = 0$. That is, h is designed to pluck out 0, the one spot at which μ and ν differ.

12 Radon measures

Radon measures

Riesz representation for Radon measures.

Pair Radon measure μ with continuous function f: $\langle \mu, f \rangle$.

Why Radon measures?

13 Farisisms

Definition 13.1. A **lattice** is poset P such that for all $a, b \in P$, $a \wedge b$ and $a \vee b$ are in P.

Definition 13.2. A **Stone vector lattice** is:

- A vector space L of real-valued functions (i.e. for $f, g \in L$ and $a \in \mathbb{R}$, $f + g$ and af are in L),
- which is a lattice ($f \wedge g$ is pointwise sup and $f \vee g$ is pointwise),
- with the *Stone property*, namely, for all $f \in L$, $f \wedge 1 \in L$ ("locally constant").

Definition 13.3. A **σ-ring** of real-valued functions on a set is a Stone vector lattice which is closed under pointwise convergence.

Definition 13.4. A **σ-algebra** of real-valued functions on a set is a σ-ring which also contains constant functions. (By linearity, if $1 \in L$, then all constant functions are in L.)

Definition 13.5. An **elementary integral** is linear, order-preserving, and satisfies monotone convergence. [xxx state this precisely.]

Mnemonic 13.6. L-O-M — eLOMentary integral — linear, order-preserving, monotone convergence.

Theorem 13.7 (Dini's theorem)**.** *Let K be a compact topological space. Let f_n be a sequence of continuous real-valued functions on K such that $f_n \to 0$. Then $f_n \to 0$ uniformly.*

14 Problem solutions

14.1 HW 15.1

HW 15.1 — two steps.

14.2 A02.6A

A02.6A — a beautiful example of many things.

14.3 A02.4B

Let f be a continuously differentiable real-valued function on $[0,1]$ with $f(0)=0$. Prove the inequalities:

$$\begin{array}{lll} (i) & |f(x)| \le x^{1/q}\|f'\|_p, \quad x \in (0,1) & \text{and} \\ (ii) & \|f\|_p^p \le \dfrac{1}{p}\|f'\|_p^p & \end{array}$$

where $1 < p < \infty$, $\|\cdot\|_p$ is the $L^p(0,1)$-norm, and $1/p + 1/q = 1$.

Answer. With regard to the free-association principle mentioned in section 1, several things should immediately come to mind. First, when we see f and f' together, especially with the $f(0)=0$ item, we should think of the corollary to the FTC (corollary 4.2)

$$\int_0^x f'(t)dt = f(x).$$

Second, integral inequalities with L^p norms in them should lead us to Hölder's inequality (proposition 9.9). Since there is only a single function f involved, we may need to insert a 1 at some point, as mentioned in remark 9.11. Lastly, the toothpaste theorem (see mnemonic 7.2) is used in many integral problems.

So, let's start:

$$\begin{aligned} f(x) &= \int_0^x f'(t)dt \\ |f(x)| &= \left|\int_0^x f'(t)dt\right| \le \int_0^x |f'(t)|\,dt = \int_0^x |1\cdot f'(t)|\,dt \\ &\le \left[\int_0^x 1^q\,dt\right]^{1/q}\left[\int_0^x |f'(t)|^p\,dt\right]^{1/p} \qquad \text{(we knew we needed a } p\text{)} \\ &\le x^{1/q}\left[\int_0^1 |f'(t)|^p\,dt\right]^{1/p} \qquad \text{(approximate to get the } p\text{-norm we need)} \\ &= x^{1/q}\|f'\|_p. \end{aligned}$$

For part (ii), we could start at the same spot as we did in part (i) and follow our noses. Or, since we need to get the p-norm of f on the left, and we have $|f|$, we can raise the left-hand side of the part (a) result to the pth power and integrate:

$$
\begin{aligned}
|f(x)| &\leq x^{1/q}\|f'\|_p \\
\int_0^1 |f(x)|^p \, dx &\leq \int_0^1 x^{p/q}\|f'\|_p^p \, dx.
\end{aligned}
$$

What is p/q? Since $1/p + 1/q = 1$, we have

$$p/q = p - 1.$$

Also, $\|f'\|_p^p$ is a constant and so comes out of the right-hand integral. We're left with

$$\|f\|_p^p \leq \int_0^1 x^{p-1} dx \, \|f'\|_p^p = \left[\frac{x^p}{p}\right]_0^1 \|f'\|_p^p = \frac{1}{p}\|f'\|_p^p$$

as desired.

References

[**Ant**] Anton, H.A. *Calculus with Analytic Geometry* (3rd ed.). Wiley, 1988.

[**CB**] Churchill, R.V. and Brown, J.W. *Complex Variables and Applications* (5th ed.). McGraw-Hill, 1990.

[**Fol**] Folland, G.B. *Real Analysis: Modern Techniques and Their Applications* (2nd ed.). Wiley-Interscience, 1999.

[**Jon**] Jones, F. *Lebesgue Integration on Euclidean Space* (revised ed.). Jones and Bartlett, 2001.

[**Ros**] Rosenlicht, M. *Introduction to Analysis.* Dover, 1986.

[**Roy**] Royden, H.L. *Real Analysis* (2nd ed.). MacMillan, 1968.

[**Rud**] Rudin, W. *Principles of Mathematical Analysis* (3rd ed.). McGraw-Hill, 1976.

Index