Reed and Simon: Some typographic errors

P. 75, line -9: The extension of λ to $\tilde{Y}$, …

P. 82, line -1: $y - \sum_{j=1}^{n} Tx_j \in B_{\epsilon 2^{-n}}$

Reed and Simon, Chapter 1

Abstract Measure Theory

Section 1.2: Metric and normed linear spaces

Page 10, top: In the displayed set of equations, the fact that $\lim_{n\to\infty} \| x_n \| = \| x \|$ if $x_n \underset{n\to\infty}{\to} x$ follows from problem 8 as indicated, since

$$\lim_{n\to\infty} \|x_n\| = \lim_{n\to\infty} \|x_n - 0\| = \lim_{n\to\infty} d(x_n, 0) = d(x, 0) = \|x\|,$$

where we are using the fact that the norm can be viewed as a distance from 0.

Page 11, bottom: The Bolzano-Weierstrass theorem states, in one form, that any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Section I.3: The Lebesgue integral

Page 13, top: Since the space C[a,b] is incomplete, we can complete it, according to theorem I.3 above. However, this completion is an abstract normed linear space, and not necessarily a space of functions. The question here is whether this abstract space which is the completion can be identified in a one to one way with some collection of functions on [a,b].

Once we identify this completed space as a collection of functions, then we will be able to define the integral on this larger space (whatever collection of functions this new larger space turns out to be). The way to do this is to notice that the integral on the original space C[a,b] defined a norm, $\| f \|_1 = \int_a^b |f(x)|\, dx$. Thus if this norm is simply extended to the larger space of functions (all norms can be extended to the completion of the space they are defined on), it will define an integral on the larger space. In the end, this "larger space" turns out to be the space of measurable functions, and the norm or integral is the Lebesgue integral - a very economical and quick way to define the Lebesgue integral.

The question left then is whether it is indeed possible to identify the completion space as a space of functions. At this point the space L^1[a,b] (the space of integrable measurable functions) is defined in an independent way, and then by general arguments it is shown that L^1[a,b] is indeed the completion we discussed above.

Page 14, top: The function μ turns out to be nothing more than Lebesgue measure. If a set is an interval [a,b], its Lebesgue measure is just its size (or length), namely b - a. However, if a set is more complicated, its size is more difficult to define. This is where the definition of Lebesgue measure comes in.

Page 15, middle: Proving Theorem I.8 takes some technicalities, and at this point, we can take the theorem as given. This theorem establishes the basic properties of Lebesgue measure.

Page 16, middle: The proposition is a standard one in measure theory, but its proof has not been included in the text. This is true of most of the of the theorems stated in this section, as well as a number of assertions regarding measure theory.

Page 18, Theorem I.12: "so pass to a subsequence with

(1) $$\| f_n - f_{n+1} \|_1 \le 2^{-n}."$$

The reason there is a subsequence which satisfies this condition is as follows. Since $\{f_n\}$ is Cauchy, we know that for any k, there is an N such that for n,m $\geq$ N, $\| f_n - f_m \| \le 2^{-k}$. Thus choose the subsequence (i.e., a subset of the original sequence) to be $f_{n_1}, f_{n_2}, f_{n_3}, \ldots$, where n_1 represents the original index of the first function in our subsequence, n_2 represents the index of the second function in our subsequence, and so on. Thus we see that we can label the subsequence by $\{f_{n_i}\}_{i=1}^{\infty}$. How should we choose this subsequence? We can do it as follows. Choose n_1 first such that for n,m $\ge n_1$, we have $\| f_n - f_m \| \le 2^{-1}$. Then choose n_2 such that for n,m $\geq n_2$, we have $\| f_n - f_m \| \le 2^{-2}$. In general then, choose the sequence of numbers $n_1, n_2, n_3, \ldots$ such that for n, m $\geq n_i$, we have $\| f_n - f_m \| \le 2^{-i}$. Now pick a subsequence of the original one, consisting only of f's with the indices $n_1, n_2, n_3, \ldots$. The new subsequence has the form $f_{n_1}, f_{n_2}, f_{n_3}, \ldots$. Notice now that by our assumption, $\| f_{n_1} - f_{n_2} \| \le 2^{-1}$ by the above, since $n_1, n_2 \ge n_1$. For the same reason, $\| f_{n_2} - f_{n_3} \| \le 2^{-2}$, and in general $\| f_{n_i} - f_{n_{i+1}} \| \le 2^{-i}$. The book uses a somewhat different notation for this subsequence, namely, the subsequence we are calling $\{f_{n_i}\}$, the book is calling $\{f_n\}$ (there is not much possibility for confusion using the book's notation, even though this is also what the original sequence was called; essentially the book is replacing the original sequence by the subsequence). Thus the above statement is equivalent to (1), which is used in the book.

Page 18, middle: "Thus $|g_\infty(x)| < \infty$ a.e. . . . "

This follows from the standard fact in measure theory that any function with a finite integral must be finite almost everywhere. The fact that the functions $f_m(x)$ converge pointwise a.e. to a function f(x) follows from the fact that the sequence defining $f_m(x)$ is absolutely convergent to a function (as established just above), so that it must be convergent to some function.

The fact that $f_n \to f$ in L^1 is just the statement that

$$\| f_n - f \| = \int |f_n - f| dx \underset{n\to\infty}{\to} 0,$$

which follows from the dominated convergence theorem.

The corollary follows from the proof of the theorem.

Section I.4: Abstract measure theory

Page 19, bottom: The fact that the limits from the right and left of the function α exist follows from the fact that any bounded increasing (or decreasing) sequence of numbers a_n has a limit. The same is true for a function $\alpha(x)$ which is increasing (or decreasing) but bounded. Namely, under this assumption, $\lim_{x \uparrow y} \alpha(x)$ and $\lim_{x \downarrow y} \alpha(x)$ exist.

The fact that one can always extend the given measure on open intervals to all Borel sets is a standard fact of measure theory.

Page 20, top: "In particular, $\mu_\alpha([a,b]) = \alpha(b + 0) - \alpha(a - 0)$."

This follows from the fact that $[a,b] = (a,b) \cup \{a\} \cup \{b\}$. Thus

$$\mu_\alpha([a,b]) = \mu_\alpha((a,b)) + \mu_\alpha(\{a\}) + \mu_\alpha(\{b\}). \tag{2}$$

However, note that

$$\mu_\alpha(\{a\}) = \lim_{\epsilon \downarrow 0} \mu_\alpha((a-\epsilon,a+\epsilon)) = \alpha(a+0) - \alpha(a-0),$$

and similarly for $\mu_\alpha(\{b\})$. Thus by (2) we have

$$\begin{aligned}\mu_\alpha([a,b]) &= \mu_\alpha((a,b)) + \mu_\alpha(\{a\}) + \mu_\alpha(\{b\}) \\ &= (\alpha(b-0) - \alpha(a+0)) + (\alpha(a+0) - \alpha(a-0)) + (\alpha(b+0)-\alpha(b-0)) \\ &= \alpha(b+0) - \alpha(a-0).\end{aligned}$$

Notice that $\mu(\{a\})$ is nonzero if and only if $\alpha(a+0) - \alpha(a-0) \neq 0$, i.e., if α is discontinuous at a.

Page 20, Example 1:

The easiest way to prove the assertion $\int f\, dx = \int f \left(\frac{d\alpha}{dx}\right) dx$ is to prove it first for step functions

$$f = \sum_k c_k \chi_{[a_k,a_{k+1}]}(x),$$

where $\{a_k\}$ is a sequence of points in the interval [a,b]. In that case the integral

$$\int f\, dx = \sum_k c_k(\alpha(a_{k+1}) - \alpha(a_k)),$$

while

$$\int f \left(\frac{d\alpha}{dx}\right) dx = \sum_k c_k \int_{a_k}^{a_{k+1}} \frac{d\alpha}{dx}\, dx = \sum_k c_k(\alpha(a_{k+1}) - \alpha(a_k)),$$

where the last equality follows from the fact that the last integral is just a Riemann integral, whose value we can calculate using the fundamental theorem of calculus.

Page 21, top: That is, the numbers in the cantor set are those whose decimal expansions in base 3 have no 1's in them, e.g., the number .022020020222… would be in the Cantor set. The point here is that to show that this set is uncountable, we can identify each number in the Cantor set with the number which has exactly the same digits, except that every 2 is replaced by a 1. Thus the above number would correspond to .011010010111. Now this can be viewed as a number in base 2, and notice that we can get every number between 0 and 1 in base 2 this way. Thus we have found a one-to-one correspondence between numbers in the Cantor set and all numbers in the interval [0,1]. This shows that the Cantor set is uncountable, since the set [0,1] is.

Page 21, middle: "Extend α by making it continuous."

Notice that at this point α is defined on the entire set S, which is easily shown to be dense. There is a basic theorem of analysis (the Tietze extension theorem) which states that any function which is continuous on a dense subset of a metric space X can be extended to a continuous function on X. In this case the dense subset is S, and to see that α as defined on S is indeed continuous, it suffices to show that for any $x \in S$, $\lim_{y \uparrow x} \alpha(y) = \alpha(x)$, and $\lim_{y \downarrow x} \alpha(y) = \alpha(x)$, where here y is of course restricted to be in S. Let us verify the first statement, since the second one follows similarly.

Notice that since α is monotonically increasing on S, as $y \uparrow x$, $\alpha(y)$ increases to some limiting value. We need to show that this limiting value is $\alpha(x)$. Suppose that this were not true. Then as $y \uparrow x$, $\alpha(y)$ would increase to a limiting value L which was less than $\alpha(x)$, and thus there would be no $y \in S$ such that $L<\alpha(y)<\alpha(x)$. Notice now that the range of the function α (i.e., the possible values α may take) consists of all numbers of the form $k/2^n$, where n is a positive integer, and $0<k<2^n$ is also an integer. This forms a dense set in the interval (0,1). Thus there is a $y \in S$ such that $\alpha(y)$ lies in any given interval contained in [0,1], which contradicts the assertion that there is no y such that $L<\alpha(y)<\alpha(x)$. This contradiction shows that indeed $\lim_{y \uparrow x} \alpha(y) = \alpha(x)$. This together with $\lim_{y \downarrow x} \alpha(y) = \alpha(x)$, which can be proved similarly, shows that α is continuous on S and thus can be extended continuously to all of [0,1].

Notice also that almost everywhere, $\alpha(x)$ is just a constant function on some interval I which is part of S, so that its derivative α' is well-defined, and equal to 0 almost everywhere (see the illustration). Finally note that for any interval I = (a,b) making up S, $\mu_\alpha(I) = \alpha(b) - \alpha(a) = 0$, since α is constant on I. From this we conclude that $\mu_\alpha(S) = 0$, since S is a countable union of such intervals.

Page 22, top:

Note that a compact set on the real line is one which is closed and bounded.

In the Definition, it is not difficult to check that a pure point measure is one which has the property that there is a countable or finite collection of points $\{x_k\}$ such that $\mu(\{x_k\}) > 0$ for each x_k, and such that μ is 0 on any set not containing any x_k. For if there were an uncountable set of such x_k, then the set S consisting of all x_k would have infinite measure, since the sum of any uncountable collection of positive numbers is infinite. Clearly also it must be that a pure point measure μ is 0 on any set not containing any x_k, just by the definition of the pure point measure.

Page 22, Theorem I.13:
The fact that the decomposition is unique follows from the fact that in any such decomposition of μ, μ_{pp} is non-zero only on all pure points of μ, which form a unique set.

Page 23, top:
The integral in this abstract setting can be defined in exactly the same way as it was defined on the real line earlier, and all of the theorems still hold in this more general setting.

Page 25, top:
Here $\chi_A(x)$ (characteristic function) denotes the function which is 1 if $x \in A$, and 0 if $x \notin A$.

Page 25, bottom:
The material on product measures (from here through the end of section 1.4) is optional, though interesting.

Page 25, middle: The material from here through the end of section I.4 (on product measures) is optional. It is nevertheless interesting and not very difficult, and so worth reading if you have the time.

Section I.5: Two convergence arguments

Page 27, middle: The inequality in the displayed equation follows from the fact that for any sequence of numbers a_n which has a limit, for any $N > 0$, $\sup_{n \geq N} a_n \geq \lim_{n \to \infty} a_n$.

Page 28: The material from the middle of this page through the end of the chapter is optional. Again the material is not difficult, and is worth reading if you have the time.

Reed and and Simon, Chapter 2
Hilbert Spaces

Section II.1: The geometry of Hilbert space

Page 37, top: If α = a + bi is a complex number, then its conjugate is defined as $\overline{\alpha} \equiv$ a - bi. A function f defined on a vector space is *conjugate linear* if f(x + y) = f(x) + f(y), and $f(\alpha x) = \overline{\alpha} f(x)$.

Note also that if α = a + bi is complex, we define $|\alpha|^2 = a^2 + b^2$, and that we can also write $\alpha\overline{\alpha} = |\alpha|^2$.

Page 37, bottom: To show that the vectors $\sum_{n=1}^{N}(x_n,x)\, x_n$ and $x - \sum_{n=1}^{N}(x_n,x)\, x_n$ are orthogonal, take their inner product:

$$\left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \ x - \sum_{n=1}^{N}(x_n,x)\, x_n\right) = \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \ x - \sum_{n=1}^{N}(x_n,x)\, x_n\right)$$

$$= \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \ x\right) - \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \sum_{n=1}^{N}(x_n,x)\, x_n\right)$$

$$= \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \ x\right) - \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \sum_{m=1}^{N}(x_m,x)\, x_m\right)$$

$$= \sum_{n=1}^{N}\overline{(x_n,x)}\,(x_n, x) - \left(\sum_{n=1}^{N}(x_n,x)\, x_n,\ \sum_{m=1}^{N}(x_m,x)\, x_m\right)$$

$$= \sum_{n=1}^{N}|(x_n,x)|^2 - \sum_{n=1}^{N}\sum_{m=1}^{N}\left((x_n,x)\, x_n,\ (x_m,x)\, x_m\right)$$

$$= \sum_{n=1}^{N}|(x_n,x)|^2 - \sum_{n=1}^{N}\sum_{m=1}^{N}\overline{(x_n,x)}\,(x_m,x)\,(x_n,x_m)$$

$$= \sum_{n=1}^{N}|(x_n,x)|^2 - \sum_{n=1}^{N}\sum_{m=1}^{N}\overline{(x_n,x)}\,(x_m,x)\,\delta_{nm}$$

$$\left(\text{where } \delta_{nm} = \begin{cases} 1 \text{ if } n = m \\ 0 \text{ if } n \neq m \end{cases} \text{ is the Kronecker delta function}\right)$$

$$= \sum_{n=1}^{N} |(x_n,x)|^2 - \sum_{n=1}^{N} \overline{(x_n,x)}\,(x_n,x)$$

$$= 0$$

Note that in the third equality we had to change the index of the second sum in order to combine the two sums into a single one.

Page 37, top: Third line: Note that since the x_n are orthonormal, we have

$$\| \sum_{n=1}^{N}(x_n,x)\, x_n \|^2 = \left(\sum_{n=1}^{N}(x_n,x)\, x_n \,,\, \sum_{n=1}^{N}(x_n,x)\, x_n \right)$$

$$= \left(\sum_{n=1}^{N}(x_n,x)\, x_n \,,\, \sum_{m=1}^{N}(x_m,x)\, x_m \right)$$

$$= \sum_{n=1}^{N} \sum_{m=1}^{N} \overline{(x_n,x)}\,(x_m,x)\,(x_n,x_m)$$

$$= \sum_{n=1}^{N} \sum_{m=1}^{N} |(x_n,x)|^2$$

Page 40, top: finite dimensional Hilbert spaces
A Hilbert space is finite dimensional if the largest orthonormal set of vectors in the space is finite.

Page 42, top: The second equality in the sequence of equations follows from the parallelogram law, namely:

$$\| (y_n - x) - (y_m - x) \|^2 + \| (y_n - x) + (y_m - x) \|^2 = 2 \| y_n - x \|^2 + 2 \| y_m - x \|^2.$$

Page 42, bottom: "… which implies Re(w,y) = 0."

This follows from the simple fact that for the parabola $at + bt^2$, the parabola crosses the t axis if a is not 0.

Page 43, middle: "By the continuity of T, $\mathcal{N}$ is a closed subspace."

This follows from the fact that if $\{x_n\}$ is a Cauchy sequence in T, then since T is continuous we have $T\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} T(x_n) = 0$, so that $\lim_{n\to\infty} x_n$ is also in $\mathcal{N}$, so that $\mathcal{N}$ is indeed closed.

Page 43, bottom: Note that in the equation $y = \left(y - \frac{T(y)}{T(x_0)} x_0\right) + \frac{T(y)}{T(x_0)} x_0$, it is easy to

check that the first group on the right side is in $\mathcal{N}$, by applying T to it.

"If T(x) = (y′,x) also, then. . . "

Note we have

$$\begin{aligned} \| y' - y_T \|^2 &= (y' - y_T, y' - y_T) \\ &= (y',y') - (y',y_T) - (y_T,y') - (y_T,y_T) \\ &= T(y') - T(y_T) - T(y') - T(y_T) \\ &= 0. \end{aligned}$$

TOPIC : A NOTE ON COUNTABLE AND UNCOUNTABLE SETS AND SUMS

Countable and uncountable numbers

An infinite set A is *countable* if it can be put in one to one correspondence with the natural numbers $\mathbb{N} = \{1,2,3,4\ldots\}$. Notice therefore that $\mathbb{N}$ itself is countable, and that the set $\mathbb{Z}$ of all integers is countable, since we can make the one-to-one correspondence:

$1 \rightarrow 0$
$2 \rightarrow 1$
$3 \rightarrow -1$
$4 \rightarrow 2$
$5 \rightarrow -2$
$6 \rightarrow 3$
$7 \rightarrow -3$
$8 \rightarrow 4$
$9 \rightarrow -4$

and so on. Further, through a slightly more complicated argument, we can show that the set $\mathbb{Q}$ of all rational numbers (i.e., all numbers expressible as a ratio of two integers) is also countable. However, as Cantor showed in the late 19th century, the set $\mathbb{R}$ of all real numbers is not countable. That is, there exists no one to one correspondence between $\mathbb{R}$ and $\mathbb{N}$.

Uncountable sums

Most infinite sums which are considered are of the form $\sum_{i=1}^{\infty} a_i$. By definition this means

$$\sum_{i=1}^{\infty} a_i \equiv \lim_{N\to\infty} \sum_{i=1}^{N} a_i .$$

Since these are sums over all natural numbers $\mathbb{N}$, such a sum can be written in the form $\sum_{i\in\mathbb{N}} a_i$. This is an example of a countable sum, since it is over the set $\mathbb{N}$ which is

countable. Consider now an example of an uncountable sum, say where the index i ranges not over $\mathbb{N}$ but over $\mathbb{R}$ or some other uncountable set. We then have $\sum_{i\in\mathbb{R}} a_i$.

What does such a sum mean? First, any reasonable definition of an *uncountable* sum where all of the numbers a_i are all positive would give an answer of ∞. Therefore, we say that an uncountable sum converges *only if* the set of a_i which are non-zero is much smaller than the whole set of a_i. Namely, we require that the set of a_i which do not vanish is *countable.* Otherwise the sum is said automatically to diverge.

Thus, if we have a sum $\sum_{i\in A} a_i$ where A is an uncountable set, we will always assume that all but a countable number of the a_i are zero if the sum is to converge.

Note that in Theorem II.6, the sum (II.1) might be over an uncountable set if A is uncountable, and in this case all but a countable number of the terms are assumed to be 0.

Page 45, bottom: $\| \sum_{j=m+1}^{n} (x_{\alpha_j}, y)\, x_{\alpha_j} \|^2 = \sum_{j=m+1}^{n} |(x_{\alpha_j}, y)|^2$

This follows from Theorem II.1, where in the statement of the Pythagorean theorem, the second term on the right side vanishes in this case, since x in this case is given by a finite sum $\sum_{j=m+1}^{n} (x_{\alpha_j}, y)\, x_{\alpha_j}$.

"Therefore $\{y_n\}$ is a Cauchy sequence . . . "

Indeed, the sum $\sum_{j=m+1}^{n} |(x_{\alpha_j}, y)|^2$ can easily be shown to go to 0 as $n,m \to \infty$,

since the full sum converges. Since this defines $y_n - y_m$, we have that $\{y_n\}$ forms a Cauchy sequence.

"$(y - y', x_{\alpha_\ell}) = \lim_{n\to\infty} \left(y - \sum_{j=1}^{n} (x_{\alpha_j}, y) x_{\alpha_j}, x_{\alpha_\ell} \right)$ "

Note that since the inner product is continuous, we have

$$\lim_{n\to\infty}\left(y - \sum_{j=1}^{n}(x_{\alpha_j},y)x_{\alpha_j},\ x_{\alpha_\ell}\right) = \left(y - \lim_{n\to\infty}\sum_{j=1}^{n}(x_{\alpha_j},y)x_{\alpha_j},\ x_{\alpha_\ell}\right) = (y - y', x_{\alpha_\ell})\ .$$

Page 46, top: The second equality in this series follows directly from the Pythagorean theorem on page 37.

Page 46, middle: There is a typographical error in this sequence of equalities. The second line should read $w_2 = u_2 - (v_1,u_2)\, v_1$.

Page 47, middle: "By throwing out some of the x_n's. . . "

More precisely, this means that we start with x_1, x_2, and throw out x_2 if it depends on x_1. We then look at x_3, and keep it if it is independent of the previous vectors we have kept, and throw it out otherwise. We thus successively look at each vector in our sequence x_n, and throw it out if it is a linear combination of the previous vectors we have kept, and keep it otherwise. At the end of this infinite process, we have a collection of vectors in which each vector is linearly independent of all the previous ones.

". . . the set of finite linear combinations of the y_n with rational coefficients is dense in $\mathcal{H}$. . . "

This follows from the fact that by theorem (II.6) every $y \in \mathcal{H}$ can be written in the form (II.1) as a countable sum. Thus y can be arbitrarily well approximated by finite sums of this form. Since a finite sum of this form has complex coefficients (x_α,y), these coefficients can in turn be arbitrarily well approximated by rational complex numbers (complex numbers whose real and imaginary parts are rational). Thus any $y \in \mathcal{H}$ can be arbitrarily well approximated by finite sums of the form (II.1), with rational (complex) coefficients. This says that the set of such sums is dense in $\mathcal{H}$.

"It is easy to show that it is unitary."

To see that the map is unitary, assume that we have x_1, x_2, so that

$$\mathcal{U}x_1 = \{(y_n,x_1)\}_n\ , \qquad \mathcal{U}x_2 = \{(y_n,x_2)\}_n\ .$$

Then we have that the inner product of the sequences, as defined in Example 3 of Section II.1, is

$$(\{(y_n,x_1)\}_n\ ,\ \{(y_n,x_2)\}_n) = \sum_{n=1}^{\infty}\overline{(y_n,x_1)}\,(y_n,x_2)$$

$$= \sum_{n=1}^{\infty}(x_1,y_n)\,(y_n,x_2)$$

$$= \sum_{n,m=1}^{\infty} (x_1,y_n)\, \delta_{nm}\, (y_m,x_2)$$

$$= \sum_{n,m=1}^{\infty} (x_1,y_n)\, (y_n,y_m)\, (y_m,x_2)$$

$$= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (x_1,y_n)\, (y_n,y_m)\, (y_m,x_2)$$

$$= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} ((x_1,y_n)\, y_n,y_m))\, (y_m,x_2)$$

$$= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} ((y_n,x_1)\, y_n,(y_m,x_2)y_m))$$

$$= \left(\sum_{n=1}^{\infty} (y_n,x_1)\, y_n, \sum_{m=1}^{\infty} (y_m,x_2)y_m) \right)$$

$$= (x_1,x_2)$$

Thus our $\mathcal{U}$ map is indeed unitary.

Page 48, bottom:

Theorem II.9:

To show that the sum given in the statement of the theorem indeed converges in L^2 to f(x), we need to show that the system $\{(2\pi)^{-1} e^{inx}\}_{n\in\mathbb{Z}}$ is a complete orthonormal set (i.e., a basis) for $\mathcal{H} = L^2$. We know that the set is orthonormal, so we need only show that it is complete, i.e., that finite linear combinations of functions in the set $(2\pi)^{-1} e^{inx}\}_{n\in\mathbb{Z}}$ are dense in $\mathcal{H}$. The proof that this set is complete starts here with the proof of the claim that the set $C^1_p[0,2\pi]$ is dense in $\mathcal{H} = L^2[0,2\pi]$. This fact is then used in the second paragraph of the proof to complete it.

Note that we are working in $[0,2\pi]$, so that by a periodic function (of period 2π) what we mean here is just a function $\phi(x)$ such that $\phi(0) = \phi(2\pi)$.

“But a step function can be approximated (in L^2) by a $C^1_p[0,2\pi]$ function by . . . "

If you draw a step function defined somewhere in the interval $[0,2\pi]$, then you can see that if you draw the ends of the function so that they don't drop abruptly to 0 but instead slant down at a steep angle, then you have a continuous function which approximates the original step function. This function is still not periodic (that is, its values at the endpoints 0 and 2π still don't match). To make it periodic, erase the graph of function very near to the point x = 2π, and from the point on the graph of the function where the erasure has started, draw a straight line to the point (2π,a), where a is the height of the original function at 0, i.e., a = f(0). This forces the new function to have the same value at 0 and 2π, without making much of a change in the original function. In this way it can be shown that for any original step function f(x), there is a periodic continuous function which approximates it arbitrarily well (in this case in the L^2 metric).

"To show that $\{(2\pi)^{-1/2}e^{inx}\}_{-\infty}^{\infty}$ is a complete set we need only show that..."

Indeed if $(e^{inx},g) = 0$ implies that g = 0, then this means that the set $\{e^{inx}\}_n$ cannot be extended to a larger orthonormal set, since otherwise we could make g a member of this larger orthonormal set, and still have $(e^{inx},g) = 0$ without having g = 0. If the set cannot be extended to a larger orthonormal set, it must already be a basis. Thus if we can prove that $(e^{inx},g) = 0$ implies that g = 0, then we will have shown that $\{e^{inx}\}_n$ is a basis and we will be finished.

Page 49, top: "which implies g = 0."

Indeed, if an element g is orthogonal to every element in a dense set in $\mathcal{H}$, it is not difficult to show by approximation that it is orthogonal to every element in $\mathcal{H}$.

Note: Section II.4 is optional reading, and will not be covered in class.

Note: this material will be covered after the material in chapter 4; this section gives applications of the Riesz representation theorem.

Section II.5: Ergodic Theory

Page 55, top: The phase space Γ is usually a vector space such that each vector corresponds to a single physical state of the system being described. For example, if the system consists of a mass moving back and forth on a spring (along the real line $\mathbb{R}$ say), we define x_1 to be the position of the mass along $\mathbb{R}$, and x_2 to be the velocity, so that at any moment the position and velocity of the mass would correspond to the point (x_1,x_2) in the plane. There is now a unique correspondence between points in the plane (which is Γ in this case) and states of the system. If the system consisted of *two* masses in motion (again on the line say), then we would define x_1 and x_2 to be the position and velocity of the first mass, and x_3 and x_4 to be the position and velocity of the second mass, so that the entire state of the system would be described by the points (x_1,x_2,x_3,x_4), i.e., a point in a four dimensional space which in this case makes up Γ. Or we might have two particles in three dimensions, in which case we would have the coordinates x_1,x_2,x_3 and x_4,x_5,x_6 to describe the components of position and velocity of the first particle, and coordinates x_7,x_8,x_9 for position of the second particle, and x_{10},x_{11},x_{12} for the velocity of the second particle. Thus such a system would be described by coordinates $(x_1,x_2,\ldots,x_{12})$, i.e., a point in a 12 dimensional space Γ. Thus in general Γ is a vector space whose points are in one-to-one correspondence with the possible states of the physical system.

Note an unfortunate coincidence of notation: The letter T denotes both a time ($T \in \mathbb{R}$), and an operator T_t - don't confuse these.

"Obviously, $T_{t+s} = T_t T_s$."

Note that T_{t+s} is the operation which takes a point x in Γ corresponding to a physical state at time 0 and maps this point to the resulting physical state at time t + s. However, taking a physical state x at time 0 to the resulting state at a later time t + s is equivalent to first taking the state to the resulting state at time t (i.e., by applying T_t), and then taking that to the resulting state s time units later (by applying T_s). Thus we have that for a physical state x at time 0, we have

$$T_{t+s}x = T_{t+s}(x) = T_t(T_s(x)) = T_t T_s x$$

for each $x \in \Gamma$, so that $T_{t+s} = T_t T_s$.

"observables of the system . . . are functions on phase space"

Thus for example, if we want to consider the energy E of the physical system, since E depends on the physical state of the system, it depends on the numbers $(x_1,x_2,x_3, \dots)$ defined above, which define the state of the physical system. Thus in fact the energy E is given by $E = E(x_1,x_2,x_3,\dots)$ as a function of the variables x_i describing the system. More generally, for any other physical observable $\mathcal{O}$, we can write $\mathcal{O}$ as a function of the physical state of the system, so that $\mathcal{O} = f(x_1,x_2,\dots,x_n)$ is a function defined on the space Γ. This is why functions f on Γ are important here.

Page 55, middle: The integral

$$(1) \qquad \lim_{T\to\infty} (1/T) \int_0^T f(T_t x)\, dt$$

represents the following. For concreteness, let $x \in \Gamma$ be a physical state of the system, and assume that f(x) represents the energy of the system in state x. Then first note that if x is a state of the system at time t = 0, $T_t x$ is the corresponding physical state t units of time later. Thus $f(T_t x)$ represents the energy of the system t units of time later. Therefore $(1/T)\int_0^T f(T_t x)\, dt$ is the *average value of* $f(T_t x)$ over all t from 0 to T (note that T is being used in two different senses in this equation).

In ergodic theory, one goal is to prove things like the assertion that the energy of a physical system approaches a limit as time t goes to ∞. That is, it is desired to prove the assertion that $f(T_t x)$ approaches some limit as $t \to \infty$. Sometimes this is either very difficult to prove or not true, and so in those cases a weaker assertion is studied. Instead of trying to prove that the energy of the system approaches a limit as $t \to \infty$, one instead tries to show that the *average energy* of the system from time 0 to t approaches a limit as $t \to \infty$. That is, one tries to prove exactly that the limit $(1/T)\int_0^T f(T_t x)\, dt$ exists (here the upper limit T of integration represents time t, and we let it approach ∞). This is why such limits are interesting.

Note that the word observable here denotes a quantity like the energy, i.e., a function f on the phase space Γ.

"energy is a conserved quantity. . . "

That is, the energy $f(T_t x)$ at time t is the same for all t, i.e., it is independent of time.

Page 55, bottom:

The constant energy surface Ω_E is simply the set of physical states $x = (x_1, x_2, \ldots, x_n)$ for which the energy $E(x_1, x_2, \ldots)$ takes on a fixed numerical value E, i.e., such that $E(x_1, x_2, \ldots, x_n) = E$. The idea here is that even if the limit (1) above is not the same for all initial conditions x (here f(x) represents any physical observable which depends on the configuration x), it *is* the same for all initial conditions x which have the same energy E. That is, that it is true for all x in the set Ω_E. This limit (1), which depends on f, is denoted by $\mu(f)$.

To verify the three stated properties of μ, note that (a) states that if the function f is 1 (i.e., takes the value 1 for all x), then the average value μ must of course also be 1. The second condition (b) of linearity is easy to verify directly, as is (c).

Finally, a version of the Riesz representation theorem in Chapter IV shows that since μ satisfies properties (a) through (c), it is a *positive linear functional*, and so there exists a measure $\hat{\mu}$ on Ω_E such that

$$\mu(f) = \int_{\Omega_E} f(w)\, d\hat{\mu}(w)$$

holds for all continuous functions f defined on Ω_E.

Page 56, top:

Note that χ_F being the characteristic function of F means that

$\chi_F(x) = \begin{cases} 1 \text{ if } x \in F \\ 0 \text{ if } x \notin F \end{cases}$. Note also that in this section we are assuming that Ω is just a measure space, and hence so is the subset Ω_E. Thus we need to define what we mean by a measurable set in Ω_E. In this case, we will define the measurable sets in the same way the Borel sets are defined on the real line, namely as the smallest σ-field of sets which contain the open sets in Ω_E (here an open set in Ω_E can be defined independently of all of Ω, since Ω_E is itself a metric space).

Page 56, middle:

Note that one can verify directly that $\chi_{T_s^{-1}F}(x) = \chi_F(T_s x)$, using the above definition of a characteristic function. This implies therefore that

$$\chi_{T_s^{-1}F}(T_t w) = \chi_F(T_t T_s w).$$

As stated, we assume the limit $\lim_{T\to\infty} \frac{1}{T}\int_0^T f(T_t w)dt$ exists for all continuous f. In addition, if the limit exists when $f = \chi_F(x)$ is a characteristic function, then we have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{T_s^{-1}F}(T_t w)dt = \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_F(T_s T_t w)dt$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_F(T_{t+s} w)dt \tag{1}$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_s^{T+s} \chi_F(T_{t+s} w)dt - \lim_{T\to\infty}\frac{1}{T}\int_T^{T+s} \chi_F(T_{t+s} w)dt + \lim_{T\to\infty}\frac{1}{T}\int_0^{s} \chi_F(T_{t+s} w)dt$$

Now note that the maximum of the last integrand is 1, since the integrand is a characteristic function. Therefore we can bound the middle integral on the last line of (1) above by multiplying 1 by the length of the interval of integration s, i.e.

$$\int_T^{T+s} \chi_F(T_{s+t} w)dt \le s\cdot 1,$$

so that $\lim\limits_{T\to\infty}\frac{1}{T}\int_T^{T+s}\chi_F(T_{t+s}w)dt = 0$ since s is fixed. The last integral in (1) is also 0, and we get, letting $T\to\infty$ (using a change of variables for t in the second equality)

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{T_s^{-1}F}(T_t w)dt = \lim_{T\to\infty}\frac{1}{T}\int_{\mathrm{s}}^{T+\mathrm{s}} \chi_F(T_{t+s} w)dt = \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_F(T_t w)dt.$$

By what has been stated (assuming that what holds above for continuous $f(x)$ also holds for the discontinuous function $\chi_F(x)$) we then have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_F(T_t w)dt = \int_{\Omega_E}\chi_F(w)d\widehat{\mu}(w) = \widehat{\mu}(F),$$

since the integral of a characteristic function of F gives the measure of F. On the other hand we similarly have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{T_s^{-1}F}(T_t w)dt = \widehat{\mu}(T_s^{-1}F),$$

from which we conclude that $\widehat{\mu}(F) = \widehat{\mu}(T_s^{-1}F)$, i.e, that μ is invariant under the transformation T_s for all s.

Reed and and Simon, Chapter 3
Banach Spaces

Section 3.1: Definition and examples

Page 68, top:

"... the uniform limit of continuous functions is continuous":

The uniform limit of a sequence of functions f_n is a limiting function f such that $f_n \to f$ uniformly, i.e., $\sup_x |f(x) - f_n(x)| \underset{n \to \infty}{\longrightarrow} 0$.

Page 68, middle:

Remember that when we consider classes of functions like L^p and L^∞, we are not working with individual functions $f(x)$ on a measure space, but rather equivalence classes. Two functions are equivalent, or belong to the same equivalence class, if they differ only on a set of measure 0.

p. 68, bottom:

The Riesz-Fisher theorem is not proved in the supplemental section, but a proof is in the exercises.

p. 70, bottom:

"the norm of A is actually equal to C":

We have by the triangle inequality that $A_n = A_n - A + A \Rightarrow$

$$\|A_n\| \le \|A_n - A\| + \|A\|.$$

Similarly,

$$\|A\| \le \|A_n - A\| + \|A_n\|.$$

Combining the two,

$$|\,\|A_n\| - \|A\|\,| \le \|A_n - A\| \underset{n \to \infty}{\longrightarrow} 0,$$

so

$$\|A\| = \lim_{n \to \infty} \|A_n\| = C.$$

p. 73, middle:

"the map which identifies $\mathcal{H}$ with its dual $\mathcal{H}^*$ is conjugate linear"

Note that if G is a linear functional, the correspondence $G \longleftrightarrow g$ given by

$$\langle g, f\rangle = G(f)$$

is such that if we multiply g by a complex number α, we have a new operation on f given by

$$\langle \alpha g, f\rangle = \overline{\alpha}\langle g, f\rangle = \overline{\alpha}G(f) = (\overline{\alpha}G)(f),$$

where the new linear functional (αG) is defined by the above equation. Thus multiplication of g by α corresponds to multiplying G by $\overline{\alpha}$. The correspondence is thus conjugate linear, in that constants α on one side of the relationship are replaced by their complex conjugates $\overline{\alpha}$ on the other side.

"... with norm equal to $\sum_{k=1}^{\infty}|\lambda_k|$..."

Note that

$$\|\Lambda\{a_k\}\| = \|\sum_{k=1}^{\infty}\lambda_k a_k\| \le \sup_k|a_k|\sum_{k=1}^{\infty}|\lambda_k| = \left(\sum_{k=1}^{\infty}|\lambda_k|\right)\|\{a_k\}\|. \tag{1}$$

By the definition of the norm of an operator, it follows that $\|\Lambda\|$, which is the smallest constant C such that

$$\|\Lambda\{a_k\}\| \le C\|\{a_k\}\|.$$

Why is this equivalent to the definition of the norm of a linear functional give on the top of p. 72?

In any case, for all sequences $\{a_k\} \in c_0$, the norm $\|\Lambda\|$ satisfies

$$\|\Lambda\| \le \sum_{k=1}^{\infty}|\lambda_k|. \tag{2}$$

Assume without loss that Λ is not the 0 functional. Then, if we consider the sequence

$$\{a_k\} = \frac{|\lambda_1|}{\lambda_1}, \frac{|\lambda_2|}{\lambda_2}, \frac{|\lambda_3|}{\lambda_3}, \ldots, \frac{|\lambda_N|}{\lambda_N}, 0, 0, 0, 0\ldots$$

with N non-zero terms followed by all zeroes (define $\frac{0}{0} = 1$), then we see that

$$\|\Lambda\{a_k\}\| = |\sum_{k=1}^{N}\lambda_k\frac{|\lambda_k|}{\lambda_k}| = \sum_{k=1}^{N}|\lambda_k|\,\|\{a_k\}\|.$$

Thus again by definition of $\|\Lambda\|$, we see that

$$\|\Lambda\| \ge \sum_{k=1}^{N}|\lambda_k|.$$

Letting $N \to \infty$ and combining with (2), we see

$$\|\Lambda\| = \sum_{k=1}^{\infty}|\lambda_k|,$$

as desired.

p. 74, middle

“It follows from Theorems III.5 and III.6 that, given $\boldsymbol{x}$, we can find a $\boldsymbol{\lambda \in X^*}$ so that $\boldsymbol{\|\lambda\|_{X^*} = 1}$ and $\boldsymbol{\lambda(x) = \|x\|_X}$.”

This actually follows from Corollary 2 on page 77. Namely, we are given an element $x \in X$ and Corollary 2 states that there exists a functional $\Lambda \in X^*$ such that

$$\Lambda(x) = \|\Lambda\|_{X^*}\|x\|_X \tag{3}$$

. By defining $\lambda(x) = \Lambda(x)/\|\Lambda\|_{X^*}$, we get that $\|\lambda\|_{X^*} = 1$, and by (4) we see

$$\lambda(x) = \|x\|_X,$$

as desired.

“This shows that

$$\boldsymbol{\|\tilde{x}\|_{X^{**}} = \sup_{\lambda \in X^*,\, \|\lambda\| \leq 1} |\tilde{x}(\lambda)| \geq \|x\|_X} \tag{4}$$

...

The last inequality follows from the fact that $\tilde{x}(\lambda) = \lambda(x)$. Thus we have

$$\sup_{\lambda \in X^*, \|\lambda\| \leq 1} |\tilde{x}(\lambda)| = \sup_{\lambda \in X^*, \|\lambda\| \leq 1} |\lambda(x)| \geq |\lambda_1(x)| = \|x\|_X,$$

where $\lambda_1 \in X^*$ is referred to on the previous displayed equation as having the properties

$$\|\lambda_1\|_{X^{**}} = 1, \quad \lambda_1(x) = \|x\|_X.$$

On the other hand, the previous equation in the text, which states that for all $\lambda \in X^*$

$$|\tilde{x}(\lambda)| \leq \|x\|_X\|\lambda\|_{X^*},$$

which states that, by definition of the norm of a linear functional,

$$\|\tilde{x}\|_{X^{**}} \leq \|x\|_X.$$

Combining this with (4), we obtain the conclusion

$$\|\tilde{x}\|_{X^{**}} = \|x\|_X,$$

as desired.

p. 75, middle
“the extension of $\boldsymbol{\lambda}$ to $\boldsymbol{\tilde{Y}}$, call it $\boldsymbol{\tilde{\lambda}}$...”

There is a typo here; note that Y should indeed be $\tilde{Y}$.

p. 76, top
“We can therefore find a real number $\boldsymbol{a}$ such that ..."

We can take the supremum of the left side of the previous displayed equation over all α, y_1 and still have the inequality hold. Then we can take the infimum of the right side of the equation over all β, y_2 without changing the inequality (note the infimum on the right is written with respect to α, y instead of β, y, which is fine as these are dummy variables. Then we can find an a between the infimum and the supremum.

"It may easily be verified that the resulting extension satisfies $\tilde{\lambda}(x) \leq p(x)$ for all $x \in \tilde{Y}$."

This follows from the fact that for any vector $y + \alpha z \in \tilde{Y}$, we now have for the new extended functional $\tilde{\lambda}$ (defined on this extended space as above), first assuming $\alpha > 0$:

$$\tilde{\lambda}(y + \alpha z) = \tilde{\lambda}(y) + \alpha\tilde{\lambda}(z) = \lambda(y) + \alpha a \leq \lambda(y) + \alpha\left[\frac{1}{\alpha}(p(y + \alpha z) - \lambda(y))\right] \qquad (1)$$

$$= \lambda(y) + p(y + \alpha z) - \lambda(y) = p(y + \alpha z)$$

Now assuming $\alpha < 0$, we have similarly

$$\tilde{\lambda}(y + \alpha z) = \lambda(y) + \alpha a \leq \lambda(y) + \alpha\left[\frac{1}{-\alpha}(-p(y - (-\alpha)z)) + \lambda(y))\right] \qquad (2)$$

$$= p(y + \alpha z).$$

Note that above we must replace α by $-\alpha$ since the sup in the first equation on p. 76 is over positive α only. Nevertheless, since the sup is over *all* α (as long as they're positive), it is OK to replace α by $-\alpha$ everywhere (in the brackets), making it positive. Note also that in (2) we use the sup part of the first equation on p. 76, while in (1) we use the inf part. That's because α is positive in (1) and negative in (2).