Class 1 Notes

Peter Schmidt-Nielsen

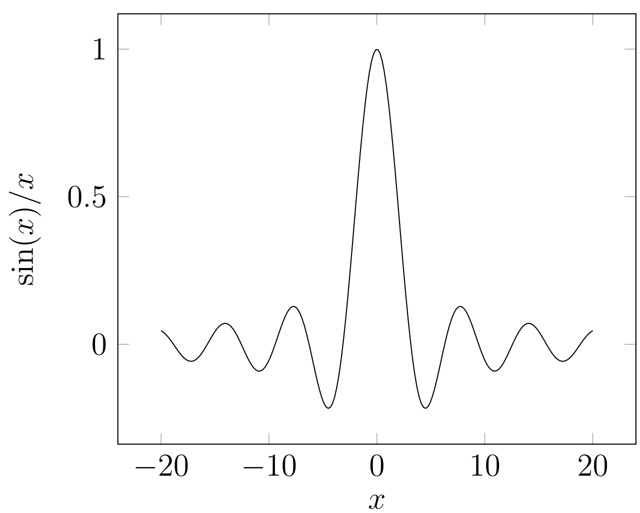
February 26, 2015

These notes follow approximately what was lectured about. There will be a lot of very technical footnotes for those who are interested! Don't worry if you don't understand all the footnotes.

1 Borwein Integrals

We are going to explore an integral involving trig functions as our first motivation for this class. First, let's define the sinc function according to:

$$\operatorname{sinc}(x) = \frac{\sin x}{x}$$



What about sinc(0), which seems poorly defined, because of a division by zero? Well, both the numerator and denominator go to zero as $x \to 0$, so by applying l'Hôpital's rule we can find that $\lim_{x\to 0} \operatorname{sinc}(x) = 1$. Thus, we choose to fill in $\operatorname{sinc}(0) = 1$ by definition, maintaining continuity.[1]

As a motivational example, let's examine an integral in terms of this function:

$$\int_0^\infty \operatorname{sinc}(x)\,\mathrm{d}x = \frac{\pi}{2}$$

Finding this integral is a little tricky, but for now we're not interested in the details of how it's done, just in the big picture.

So, we get out $\pi/2$, big whoop – integrals of trig functions give answers in terms of π all the time. Well, let's examine another integral. Again, I won't show you how to actually do the integral (yet), but I'll just fill in the correct answer:

$$\int_0^\infty \operatorname{sinc}(x)\operatorname{sinc}(x/3)\,\mathrm{d}x = \frac{\pi}{2}$$

That's interesting, we got the same answer out. Can we extend this pattern somehow? Sure enough, we can:

$$\int_0^\infty \operatorname{sinc}(x)\operatorname{sinc}\left(\frac{x}{3}\right)\operatorname{sinc}\left(\frac{x}{5}\right)\,\mathrm{d}x = \frac{\pi}{2}$$

And, sure enough, we can keep going like this, adding new factors of sinc, and the integral remains constant:

$$\int_0^\infty \operatorname{sinc}(x)\operatorname{sinc}\left(\frac{x}{3}\right)\operatorname{sinc}\left(\frac{x}{5}\right)\operatorname{sinc}\left(\frac{x}{7}\right)\operatorname{sinc}\left(\frac{x}{9}\right)\operatorname{sinc}\left(\frac{x}{11}\right)\operatorname{sinc}\left(\frac{x}{13}\right)\,\mathrm{d}x = \frac{\pi}{2}$$

But, if we go one further, and add a factor $\operatorname{sinc}(x/15)$, then everything breaks down!

$$\begin{aligned}\int_0^\infty \operatorname{sinc}(x)\operatorname{sinc}\left(\frac{x}{3}\right)\cdots\operatorname{sinc}\left(\frac{x}{15}\right)\,\mathrm{d}x &= \frac{\pi}{2} - \frac{6879714958723010531}{935615849440640907310521750000}\pi \\ &\approx \pi - 2\times 10^{-11}\end{aligned}$$

The integral comes out about twenty parts per *trillion* defficient! What happened when we reached fifteen, and what's so special about that number? Something seriously weird is going on here.

Consider these integrals to be a hint, perhaps a crack in the wall. We will peek through this crack, and find that a gorgeous theory lies behind their bizarre behaviour. The purpose of the first few weeks will be to provide a theoretical framework involving Fourier transforms that explains the above integrals (known as *Borwein integrals*), and allows us to evaluate them with ease.

[1]For those who are interested, the technical term for what we just did is filling in a *removable discontinuity*. To be extremely nitpicky, this means we've actually defined $\operatorname{sinc}(x)$ as:

$$\operatorname{sinc}(x) = \begin{cases} 1 & \text{If } x = 0, \\ \frac{\sin x}{x} & \text{otherwise.}\end{cases}$$

A closely related theory (of Fourier series) will allow us to evaluate various trig integrals over finite intervals, such as:

$$\int_0^{2\pi} \cos^n(x)\,\mathrm{d}x = \begin{cases} 2^{-n+1}\binom{n}{n/2}\pi & \text{For even } n, \\ 0 & \text{otherwise.} \end{cases}$$

This integral, crazy as it seems, came up as part of a problem in an assignment in MIT's elliptic curves class last week.[2]

2 Vector Spaces

2.1 Definition

What is a vector? The basic example most are familiar with is a list of numbers. That is, we think of $(3, 2)$ as the vector whose x-coordinate is three, and whose y-coordinate is two. Given two vectors, (x_1, y_1) and (x_2, y_2) they have a sum:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

Further, a vector can be multiplied by a scalar (a.k.a. number), so, for example, $3(x, y) = (3x, 3y)$.

Mathematically, we generalize this example by defining a vector space as a set V of vectors subject to four axioms:

1. Given two vectors $v_1 \in V$ and $v_2 \in V$, they have a unique sum $(v_1 + v_2)$ which is also a vector. We want this addition to be commutative and associative, just like regular addition of numbers.

   > *Hint:* This axiom says nothing profound! I just want to exclude the possibility that the sum of two vectors is an entirely different object, or fails to be defined in some cases, or that there might be multiple valid answers for $v_1 + v_2$. These would all be bad.

2. Given a number c and a vector $v \in V$, the scalar multiplication of the vector by the number, denoted cV, is also a vector. Scalar multiplication must distribute over vector addition. That is, $c(v_1 + v_2) = cv_1 + cv_2$. Further, we want obvious things to hold, like $c_1(c_2 v) = (c_1 c_2)v$, and that $1v = v$. Again, I'm not saying anything particularly profound, we just want operation in vector spaces to behave as we algebraically expect.[3]

[2] For those interested in the details, the problem involved computing moments of the distribution of traces of Frobenius on elliptic curves. For the extremely interested, try examing: `http://math.mit.edu/~drew/Brandeis.pdf` and reading about the Sato-Tate conjecture.

[3] Technical detail for the interested: the number c must come from a field. Thus, any vector space V has a specific field that it is defined with respect to. This is known as its *ground field*, and if the ground field of a vector space is $\mathbb{R}$, then it is said to be a real vector space, and likewise a ground field of $\mathbb{C}$ makes a vector space complex.

3. There exists a special zero vector $\vec{0} \in V$, satisfying $\vec{0} + v = v$ for every $v \in V$. To distinguish the vector zero from the scalar zero we will use the little hat. I may sometimes get lazy, and just use 0 for this vector. We'll see.

4. Every vector $v \in V$ has a "negative," or "additive inverse" denoted by $-v$ that sums to zero with v. That is, $v + (-v) = (-v) + v = \vec{0}$.

2.2 Some Examples

Let's quickly check that our prototypical vector space of "lists of numbers" satisfies all of these axioms, at least in the 2-dimensional case.

1. We defined the sum of $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$ as $v_1 + v_2 = (x_1 + x_2, y_1 + y_2)$. This clearly inherits commutativity and associativity from plain old addition of numbers. The answer is unique and always exists for any v_1 and v_2.

2. Again, it's almost too obvious. Given a number c and a vector $v = (x, y)$, we get that $cv = (cx, cy)$. This obeys all the properties we need.

3. The zero vector $\vec{0} = (0, 0)$, and obeys the sum law we need, because:

$$\vec{0} + v = (0 + x, 0 + y) = (x, y) = v$$

4. Given a vector $v = (x, y)$, we define its negative as $-v = (-x, -y)$, which clearly obeys the law we need:

$$v + (-v) = (x, y) + (-x, -y) = (x - x, y - y) = (0, 0) = \vec{0}$$

Okay, so we've shown that our traditional idea of vectors as lists of numbers define a vector space.

Now let's examine a more complicated example. We can consider the space of functions $f : \mathbb{R} \to \mathbb{R}$ (or $\mathbb{R} \to \mathbb{C}$) to be a vector space, with addition and scalar multiplication being defined in the usual ways, namely $(f + g)(x) = f(x) + g(x)$ and that $(cf)(x) = cf(x)$. It's pretty straight forward, but the reader can verify that this space satisfies the axioms of a vector space.

3 Finite Dimensional Vector Spaces

We will now explore some basic results in finite dimensional vector spaces.

3.1 Expansion in a Basis

It's often of interest to represent one vector as a sum (read: linear combination) of others. In the finite dimensional case of n-dimensional vectors we can pick a basis $\{\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n\}$, with each $\hat{e}_i$ being the vector that is all zeroes, except with a one in position i. For example, with three dimensions:

$$\hat{e}_1 = (1, 0, 0) \qquad \hat{e}_2 = (0, 1, 0) \qquad \hat{e}_3 = (0, 0, 1)$$

Given such a basis, we can expand a vector in this basis, by writing it as a linear combination. For example, for $v = (2, -3, 4)$, we get:

$$v = 2\hat{e}_1 - 3\hat{e}_2 + 4\hat{e}_3$$

In general, for any vector v there will be a set of *expansion coefficients* $c_1, c_2, \ldots c_n$ such that:

$$v = c_1\hat{e}_1 + c_2\hat{e}_2 + \cdots c_n\hat{e}_n = \sum_{i=1}^{n} c_i\hat{e}_i$$

Naturally, these expansion coefficients will be simply be the coordinates of v.[4] That is, in our previous example, $c_1 = 2$, $c_2 = -3$, $c_3 = 4$.

3.2 Inner Product

In finite dimensions, we're familiar with the *dot product* of two vectors v_1 and v_2 denoted by $v_1 \cdot v_2$, defined as:

$$v_1 \cdot v_2 = (x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = x_1x_2 + y_1y_2 + z_1z_2$$

Hopefully you already have the intuition that the dot product of two vectors says how much they point in the same direction. As a brief reminder, there is a great formula for the dot product in terms of the lengths of v_1 and v_2 and the angle between them. Specifically, where $|v|$ is the Euclidean (read: plain old) length of the vector v, and θ is the angle between v_1 and v_2, we get that:

$$v_1 \cdot v_2 = |v_1||v_2| \cos\theta$$

This has some interesting implications. If the two vectors are perfectly aligned (point in the same direction) then $\theta = 0$ and $\cos 0 = 1$, and the dot product is just the product of the lengths. On the other hand, if the two vectors point in exact opposite directions then $\theta = \pi$ and $\cos\pi = -1$, and the dot product is minus the product of the lengths. In between these two extreme cases we find that if the two vectors are orthogonal to each other, then $\theta = \pi/2$, and $\cos(\pi/2) = 0$, and the dot product is zero. Thus, the dot

[4] Technical note: Although this all seems trivial, we're basically saying that every coordinate-free finite dimensional vector space is isomorphic to some $\mathbb{F}^n$, where $\mathbb{F}$ is the ground field of our vector space. This is an important point.

product varies continuously, telling us approximately how much the vectors point in the same direction, but scaled up by the magnitudes of the vectors.

As another crucial special case, examine taking the dot product of a vector with itself. See that $v \cdot v = |v||v| \cos 0 = |v|^2$. Thus, taking the dot product of a vector with itself gives the square of its length. This will come up again later.

In preparation to move into the general infinite dimensional setting, we will start shifting terminology. Specifically, we will start using the term *inner product* instead of dot product. The inner product is a specific case of the more general concept of a *bilinear form*, so you may hear me use this term instead by accident.[5] We will denote the inner product of two vectors $v_1, v_2 \in V$ by $\langle v_1, v_2 \rangle$.

Observe that taking the dot product against one of the basis vectors $\hat{e}_i$ from before pulls out a single coordinate from a vector. That is, for $v = (x, y, z)$, we get, for example:

$$\langle e_2, v \rangle = \langle (0, 1, 0), (x, y, z) \rangle = 0x + 1y + 0z = y$$

Armed with this observation, we can now write out a simple formula for writing v as a sum of our basis vectors:

$$v = \sum_{i=1}^{n} \langle \hat{e}_i, v \rangle \hat{e}_i$$

Another way of saying this is that the expansion coefficient c_i is given by $c_i = \langle \hat{e}_i, v \rangle$.

This observation is actually more general, and leads to our first main result, which we will not prove. A basis of vectors $\{e_1, e_2, \ldots e_n\}$ is said to be *orthogonal* if all pairs of vectors are at right angles to each other. This occurs when their dot product is zero, and thus we require that $\langle e_i, e_j \rangle = 0$ when $i \neq j$.

A vector is said to be *normalized* if it is of unit length. We can write this down as $\langle e_i, e_i \rangle = 1$, recalling that $\langle e_i, e_i \rangle$ is the square of the length of e_i.

A basis is said to be *orthonormal* if it is orthogonal, and contains all normalized vectors.[6] We can actually combine these two conditions into one as:

$$\langle e_i, e_j \rangle = \begin{cases} 1 & \text{If } i = j \\ 0 & \text{otherwise.} \end{cases}$$

To compactly denote this, we will introduce the *Kronecker delta* symbol denoted by δ_{ij}, defined according to:

$$\delta_{ij} = \begin{cases} 1 & \text{If } i = j \\ 0 & \text{otherwise.} \end{cases}$$

[5]Technical note: If the ground field is $\mathbb{C}$ then we will instead concern ourselves with *sesquilinear forms* (a.k.a. *antilinear form* or *conjugate-linear form*) or *Hermitian forms*. Those who are interested are encouraged to read the definitions of these various objects on Wikipedia. They all differ slightly. In this case we will mostly use Hermitian forms.

[6]Technical note: Nomenclature is actually a mess here. In pure math, what I call "orthonormal" will sometimes just be called "orthogonal," specifically in the context of matrices. Engineering practice agrees with my convention, and I prefer it. However, as a concession to avoid confusion for students who are familiar with some linear algebra, I will refrain from talking about "orthogonal matrices" to side-step the ambiguity.

Now we can write that a basis is orthonormal iff $\langle e_i, e_j \rangle = \delta_{ij}$. This symbol is also useful because we can say that $\hat{e}_i$'s jth coordinate is δ_{ij}. This brings us to our first main result.

Theorem: *Given a real vector space V, and an orthonormal basis $\{e_1, e_2, \ldots e_n\}$ of that vector space, any vector v can be decomposed into a sum of the basis vectors as:*

$$v = \sum_{i=1}^{n} \langle e_i, v \rangle e_i$$

For a proof of this theorem, I refer the interested reader to *Linear Algebra Done Right* by Axler.

Let's do an example out, just to verify the power of this theorem. We will examine $\mathbb{R}^2$ as our example vector space. Our orthonormal basis will be:

$$e_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \qquad e_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

Observe that $\langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = 1$, and $\langle e_1, e_2 \rangle = 0$, verifying orthonormality. Now, given a vector $v = (6, 7)$, we will test the decomposition. According to our main theorem, we should get that:

$$v = \langle e_1, v \rangle e_1 + \langle e_2, v \rangle e_2$$

We get that $\langle e_1, v \rangle = 13/\sqrt{2}$, and $\langle e_2, v \rangle = -1/\sqrt{2}$. Sure enough:

$$\begin{aligned}
(6,7) &= \frac{13}{\sqrt{2}} \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) - \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \\
&= \left(\frac{13}{2}, \frac{13}{2} \right) - \left(\frac{1}{2}, -\frac{1}{2} \right) \\
&= (6,7)
\end{aligned}$$

Huzzah!

3.3 Complex Vector Spaces

In complex vector spaces, almost exactly the same notion holds, only now we need to modify our sense of inner product. First, observe a deficiency in our current definition. Given the vector $v = (i, i)$, we get that $\langle v, v \rangle = i^2 + i^2 = -2$. What? But we said that $\langle v, v \rangle$ should be the square of the length of v. We really don't want complex lengths – the length of a vector really should be a non-negative real quantity. To fix this, we modify our definition of the inner product. Where v is the vector whose coordinates are v_1 through v_n (that is $v = (v_1, v_2, \ldots v_n)$) and likewise for w, we redefine $\langle v, w \rangle$ as:

$$\langle v, w \rangle = \sum_{i=1}^{n} \overline{v_i} w_i$$

(Where $\overline{x}$ is the complex conjugate of x.)

Now observe that $\langle (i,i),(i,i)\rangle = (-i)\cdot i + (-i)\cdot i = 2$. Because the product of a number with its conjugate is always non-negative we get that this new inner product is always non-negative as well.

Observe that this new definition entails the old inner product definition when the vectors are real, as the complex conjugate of a real number is itself. It is called the *standard Hermitian form*, and is actually in some sense the *correct* inner product for complex vector spaces. Sure enough, our main theorem above holds in complex vector spaces (specifically C^n) when using this inner product.

4 Infinite Dimensional Vector Spaces

Let us now move into the infinite dimensional case, where functions take the place of vectors. Given a function f, we'd like to write it as the sum of basis functions e_i. Unfortunately, no finite basis can do it in general! The intuition is that a function $f : \mathbb{R} \to \mathbb{R}$ varies in infinitely many places, and thus has infinitely many dimensions.

However, not to worry – we can simply instead use *infinitely many* basis vectors, and integrate instead of summing. Previously we used basis vectors $\hat{e}_i$. To unify the notation we use in the infinite and finite dimensional settings, we will momentarily treat finite dimension vectors as functions, mapping indices to coordinates. That is, where $v = (2, \pi, 8)$, we will let $v(1) = 2$, $v(2) = \pi$, $v(3) = 8$.

Given this notation, see that $\hat{e}_i(j) = \delta_{ij}$. Why not try this same trick in the infinite dimension case? We will define *basis functions* $\hat{e}_x : \mathbb{R} \to \mathbb{R}$ parameterized by $x \in \mathbb{R}$, such that each $\hat{e}_x$ satisfies $\hat{e}_x(x') = \delta_{xx'}$.

Unfortunately, it is impossible to expand anything whose integral doesn't vanish in this basis! For example, take $f(x) = 1$. It seems intuitively like we want to integrate up $\hat{e}_x$ for every single possible x, to set the value of the function to 1 everywhere. Thus, it seems like $g = \int_{-\infty}^{\infty} \hat{e}_{x'}\, \mathrm{d}x'$ should work to make $g = f$. Let's try evaluating this candidate function g somewhere to see if it is actually equal to f.

$$\begin{aligned} g(3) &= \left(\int_{-\infty}^{\infty} \hat{e}_{x'}\, \mathrm{d}x' \right)(3) \\ &= \int_{-\infty}^{\infty} \hat{e}_{x'}(3)\, \mathrm{d}x' \\ &= \int_{-\infty}^{\infty} \delta_{x'3}\, \mathrm{d}x' \end{aligned}$$

Note that the last line is the integral of a function of x' that is equal to 1 when x' is equal to 3, and zero otherwise. This function reaches the finite value of 1, but only at the single point $x' = 3$, so the area under the curve is clearly zero. Thus, the integral vanishes.

Uh oh! We just showed that $g(3) = 0$, and in fact, g is identically zero everywhere. This is deeply disappointing.

It turns out that there is a fix. If only we could make the basis function get infinitely tall, then perhaps it could accumulate some real area in an infinitely narrow spike. Thus, we define the *Dirac delta function*[7] $\delta(x)$ as:

$$\delta(x) = \begin{cases} \infty & \text{If } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

This function is zero everywhere except with a single spike at the origin. However, this spike is so infinitely tall that it manages to have an area of one under it, despite having no width. In some sense we are declaring that $\infty \cdot 0 = 1$.

This definition is, unfortunately, a little informal, and shouldn't be relied upon other than for intuition. The *real* definition is that the integral of the Dirac delta function across an interval is one if the interval contains 0, and zero otherwise. For example:

$$\int_{-1}^{1} \delta(x)\,\mathrm{d}x = 1 \qquad \int_{1}^{2} \delta(x)\,\mathrm{d}x = 0$$

Let's not concern ourselves too much with what happens when one of the bounds of the integral is *exactly* zero. Again, the treatment is, so far, relatively informal.

Now let's try our experiment again with the new basis $\hat{e}_x(x') = \delta(x - x')$. Thus, we now get $\hat{e}_x(x') = \infty$ when $x = x'$ and zero otherwise. Now the integral proceed as:

$$\begin{aligned} g(3) &= \left(\int_{-\infty}^{\infty} \hat{e}_{x'}\,\mathrm{d}x'\right)(3) \\ &= \int_{-\infty}^{\infty} \hat{e}_{x'}(3)\,\mathrm{d}x' \\ &= \int_{-\infty}^{\infty} \delta(x' - 3)\,\mathrm{d}x' \\ &= 1 \end{aligned}$$

Huzzah!

Sure enough, we can now use this basis to expand any function, by putting the right expansion coefficients in front of the δs. Now we will let f be any old function $\mathbb{R} \to \mathbb{R}$.

$$\begin{aligned} g(x'') &= \left(\int_{-\infty}^{\infty} f(x')\hat{e}_{x'}\,\mathrm{d}x'\right)(x'') \\ &= \int_{-\infty}^{\infty} f(x')\hat{e}_{x'}(x'')\,\mathrm{d}x' \\ &= \int_{-\infty}^{\infty} f(x')\delta(x'' - x')\,\mathrm{d}x' \end{aligned}$$

[7] Which isn't actually a function. We'll get to this later, when we formalize distributions and measures.

Observe that $\delta(x'' - x')$ is zero whenever $x'' \neq x'$, so the integrand vanishes whenever x' is not simply equal to x''i. Thus, we can replace $f(x')$ with $f(x'')$ without affecting the integral:

$$\begin{aligned} g(x'') &= \int_{-\infty}^{\infty} f(x'')\delta(x'' - x')\,\mathrm{d}x' \\ &= f(x'') \int_{-\infty}^{\infty} \delta(x'' - x')\,\mathrm{d}x' \\ &= f(x'') \cdot 1 \end{aligned}$$

Thus, $g = f$. Sure enough, we've expanded f as an infinite linear combination of Dirac deltas.

But wait, you protest, what does $f(x'')\delta(x'' - x')$ even mean? If $f(x'')$ happens to be equal to 7, then is $7\delta(x'' - x')$ the function that is seven times infinity when $x'' = x'$, and zero otherwise? Is seven times infinity even really different than infinity in the first place? This is a valid objection, but for now we'll ignore it in a haze of informality. Later, when we speak of measure theory, we will explain how to make more sense of such expressions.

4.1 Hermitian Forms in Infinite Dimensions

We will generalize our inner product to infinite dimensions be defining, for $f, g : \mathbb{R} \to \mathbb{C}$ the inner product $\langle f, g \rangle$ as:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} \overline{f(x)} g(x)\,\mathrm{d}x$$

It turns out that this definition does basically everything we want, and is really the correct generalization of the inner product.

Let's do one brief example, namely, $\langle \hat{e}_0, f \rangle$.

$$\begin{aligned} \langle \hat{e}_0, f \rangle &= \int_{-\infty}^{\infty} \overline{\delta(x)} f(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \overline{\delta(x)} f(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \overline{\delta(x)} f(0)\,\mathrm{d}x \\ &= f(0) \end{aligned}$$

We just showed that taking the inner product against $\hat{e}_0$ pulls out the value of the function at $x = 0$. Note that this is completely analogous to the finite dimensional case where taking the dot product against $\hat{e}_i$ pulled out the ith coordinate! This is a *good* generalization! I leave it as an exercise to show that $\langle \hat{e}_x, f \rangle = f(x)$ in general.

Let's apply this identity case where $f = \hat{e}_y$. This tells us that $\langle \hat{e}_x, \hat{e}_y \rangle = \hat{e}_y(x) = \delta(x-y)$. In other words, the inner product of two of our Dirac delta basis functions is zero, unless the peaks perfectly line up ($x = y$), in which case it is infinite.

This actually turns out the be the correct infinite dimensional generalization of orthonormality. Where previously we wanted that $\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$, now we demand that $\langle \hat{e}_x, \hat{e}_y \rangle = \delta(x-y)$. By now the pattern should be clear: To go from finite to infinite dimensional, simply replace the Kronecker delta with the Dirac delta.

Sure enough, our Dirac delta basis satisfies this notion of orthonormality.

4.2 Infinite Dimensional Theorem

We'd like to generalize our main finite dimensional theorem into infinite dimensions. Unfortunately, it's really *really* hard, in part because you get constantly bollocksed up by convergence issues. Also, $\mathbb{R} \to \mathbb{R}$ is a *really* big space! It contains nasty functions that are discontinuous everywhere, and unbounded on any interval.

Thus, the following result comes with pages of caveats that we won't address formally until much later in the course. We will now switch to "physicist mode," where we pretend expressions like $\infty \cdot 0 = 1$ make sense, and proceed with reckless abandon. Here's the statement informally.

"Theorem": *Given the complex infinite dimensional vector space of functions $\mathbb{R} \to \mathbb{C}$, and a suitable orthonormal basis of functions $e_x : \mathbb{R} \to \mathbb{C}$, (that is, one basis function for each $x \in \mathbb{R}$), pretty much any reasonable function $f : \mathbb{R} \to \mathbb{C}$ can be decomposed into a linear combination of the basis vectors as:*

$$f = \int_{-\infty}^{\infty} \langle e_x, f \rangle e_x \, \mathrm{d}x$$

Sure enough, our previous example of decomposing a function using Dirac deltas as our basis verifies this idea.[8]

5 The Fourier Transform

We only skimmed this material in the last few minutes of the class. I'll go over it again more carefully in session 2, but here's approximately what I talked about, to reference.

Let's examine the basis $e_\omega(x) = \frac{1}{\sqrt{2\pi}} e^{i\omega x}$. It has many interesting properties. It can

[8] If you really want to see a formal proof of this, with all the caveats defined appropriately, try reading about real analysis and functional analysis. You will probably need to know a fair amount of topology, and be comfortable with the idea of Cauchy sequences in infinite dimensional spaces. Often the theorem is proven in a restricted setting, such as Schwartz space.

be observed that it is orthonormal because:

$$\begin{aligned}\langle e_\omega, e_{\omega'}\rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{e^{i\omega x}} e^{i\omega' x}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x} e^{i\omega' x}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(\omega'-\omega)x}\,\mathrm{d}x\end{aligned}$$

Observe that if $\omega' = \omega$, then the integrand becomes $e^0 = 1$, and thus the integral is infinite. However, when $\omega' \neq \omega$, then the integrand is oscillatory, and the integral neither diverges nor converges. We are still in informal physicist mode, so we get to say that this oscillatory integral is simply zero. Thus, $\langle e_\omega, e_{\omega'}\rangle = \delta(\omega - \omega')$.

Wait, but what about the $1/\sqrt{2\pi}$ factor? It's a really complicated story, and relates to the observation that 3δ really doesn't make much sense, because 3∞ isn't a sensible quantity. Just like before, I promise that it will become clear in time! For now we just accept it as being necessary.

Now, the Fourier transform is easy to define. Given a function $f : \mathbb{R} \to \mathbb{C}$, we will rewrite it as a sum of our oscillatory basis functions e_ω. We will define the *Fourier transform* of f to be the expansion coefficients in the linear combination, and denote it by $\tilde{f}(\omega)$. According to our "theorem" we can do this as:

$$f(x) = \int_{-\infty}^{\infty} \underbrace{\left\langle \frac{1}{\sqrt{2\pi}} e^{i\omega x}, f \right\rangle}_{\tilde{f}(\omega)} \frac{1}{\sqrt{2\pi}} e^{i\omega x}\,\mathrm{d}\omega$$

Thus, we get that:

$$\begin{aligned}\tilde{f}(\omega) &= \left\langle \frac{1}{\sqrt{2\pi}} e^{i\omega x}, f \right\rangle \\ &= \frac{1}{\sqrt{2\pi}}\langle e^{i\omega x}, f\rangle \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \overline{e^{i\omega x}} f(x)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega x} f(x)\,\mathrm{d}x\end{aligned}$$

And there you have it, this is the Fourier transform! It takes a function $f : \mathbb{R} \to \mathbb{C}$ and converts it to a function $\tilde{f} : \mathbb{R} \to \mathbb{C}$ according to the above formula. What's more, this function $\tilde{f}$ tells you the expansion coefficients to write f as a sum of complex exponentials of the form $\frac{1}{\sqrt{2\pi}} e^{i\omega x}$.

I've now defined the Fourier transform, but given absolutely zero motivation. This will come next week.

Now, in class, I rushed through a derivation of the half derivative. I really shouldn't have, because I didn't do it proper justice. Rather than putting this in the notes, we'll just pick up next week starting here, with the definition of the Fourier transform in hand.