Class 2 Notes

Peter Schmidt-Nielsen

March 20, 2015

These notes follow approximately what was lectured about.

There is some extra material in here that I didn't cover in lecture – it's marked with red warning text. Feel free to skip it.

1 The Fourier Transform

As a brief reminder, given $f, g : \mathbb{R} \to \mathbb{C}$, we define their inner product (or *Hermitian form*) to be:

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

Now, in infinite dimensions we call a basis e_x (parameterized by $x \in \mathbb{R}$) orthonormal if $\langle e_x, e_{x'} \rangle = \delta(x - x')$. This is analogous to us calling a finite dimensional basis orthonormal if $\langle e_i, e_j \rangle = \delta_{ij}$. In other words, the Dirac delta is the correct infinite-dimensional generalization of the Kronecker delta.

As a brief reminder from last time, we defined the Fourier Transform as:

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

We further observed that it corresponds to changing into the $e^{i\omega x}/\sqrt{2\pi}$ basis. That is to say, $\tilde{f}(\omega)$ is the expansion coefficient of some particular $e^{i\omega x}/\sqrt{2\pi}$ when we write f as a sum of such basis functions.

Let's examine the function $f(x) = \sin(100x) + \sin(3x)$, and examine its Fourier transform $\tilde{f}(\omega)$. We can decompose each sine as:

$$\sin(\omega x) = \frac{e^{i\omega x} - e^{-i\omega x}}{2i}$$

Further, recall the very end of notes 1, in which we show that $\langle e^{i\omega x}, e^{i\omega' x} \rangle = 2\pi\delta(\omega - \omega')$, a very useful identity.[1]

[1] I briefly reviewed the argument in class, but I won't repeat it here – simply refer to the derivation of $\langle e_\omega, e_{\omega'} \rangle$ at the end of notes 1.

The result is that when we're evaluating $\tilde{f}(\omega)$, it becomes infinite at $\omega = \pm 100$ and $\omega = \pm 3$, and is otherwise zero. In other words, our function $\tilde{f}(\omega)$ screams out whenever you plug in an ω such that the input function contains a sinusoid of frequency $\pm\omega$.

2 Properties

Now, we prove some important properties of the Fourier transform.

2.1 Linearity

Given two functions, $f, g : \mathbb{R} \to \mathbb{C}$, we find that:

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

That is to say, the Fourier transform of a sum of functions is equal to the sum of the Fourier transforms of the functions.

Closely relatedly, we can obviously pull out constant factors. That is to say, $\mathcal{F}\{cf(x)\} = c\mathcal{F}\{f(x)\}$. In words, the Fourier transform of c times a function is c times the Fourier transform of the function.

For those who are familiar with the terminology, an operator with the above two properties (splitting over sums, and pulling out constants) is known as *linear*. In a single condition, for $c, d \in \mathbb{C}$, and $f, g : \mathbb{R} \to \mathbb{C}$, we get that $\mathcal{F}\{cf(x) + dg(x)\}(\omega) = c\tilde{f}(\omega) + d\tilde{g}(\omega)$.

2.2 Modulation, or "Shift" Property

This property is actually how heterodyning works in modern AM and FM radios. It is one of *the* most used properties of the Fourier transform for radio engineering. Let g be a shifted copy of f, shifted to the right by an amount λ, namely $g(x) = f(x - \lambda)$. Observe:

$$\begin{aligned}
\tilde{g}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x} f(x-\lambda)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega(u+\lambda)} f(u)\,\mathrm{d}u \\
&= e^{-i\omega\lambda} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega u} f(u)\,\mathrm{d}u \\
&= e^{-i\omega\lambda} \tilde{f}(\omega)
\end{aligned}$$

That is to say, shifting a function over in time domain multiplies it through by a complex exponential in frequency domain.

An Application: The dual of this property is actually how radios recieve AM via heterodyning! If I multiply the signal $f(t)$ coming in over the antenna by $e^{-i\omega t}$, then I'm effectively shifting all the frequencies present in f down by ω. Thus, if $\omega = 1420$ kHz, then I'll end up shifting the signal being transmitted by the AM radio station WBEC at 1420 kHz down to being in the audible frequency range. For more information, read the Wikipedia heterodyning article.

2.3 Scaling

The Fourier transform has an interesting property scaling reciprocity property. Given a function $f : \mathbb{R} \to \mathbb{C}$, we can make a new function g that "goes λ times faster" by setting $g(x) = f(\lambda x)$. As you'll recall (or realize from briefly pondering it), if we set $\lambda = 2$, then g is a version of f that is "squished in" by a factor of two.

How does this squishing affect $\tilde{g}$? We can evaluate this:

$$\tilde{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x} f(\lambda x)\, \mathrm{d}x$$

Apply the u-substitution $u = \lambda x$, $\mathrm{d}u = \lambda \mathrm{d}x$:

$$\begin{aligned} &= \frac{1}{\lambda\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega u/\lambda} f(u)\, \mathrm{d}u \\ &= \frac{1}{\lambda} \tilde{f}(\omega/\lambda) \end{aligned}$$

Very interesting! We just showed that if you squish a function inwards by a factor of λ in time domain, then it gets stretched out by a factor of λ in the frequency domain (also the amplitude falls by a factor of λ).

Does this result make sense? Let's convince ourselves that it really does. If we squish a function f in by a factor of λ, we've effectively made all the sinusoids in f go at a higher frequency. Thus, the Fourier transform has to be stretched out to represent this fact! That is to say, if $\tilde{f}$ had a peak at $\omega = 7$, and g goes twice as fast as f (that is, $\lambda = 2$), then $\tilde{g}$ better have a peak at $\omega = 14$. The scaling by a factor of $1/\lambda$ just has to do with the chain rule.[2]

This property has a profound connection to the idea of *covariant* and *contravariant* coordinates, from modern physics. I strongly encourage the interested student to look up covariance and contravariance on Wikipedia. The connection should be clear. (Hint: x is contravariant, ω is covariant.)

2.4 Hermitian Symmetry

The Fourier transform has some interesting symmetries. Given a pure real function $f : \mathbb{R} \to \mathbb{R}$ as input, we find that $\tilde{f}(-\omega) = \overline{\tilde{f}(\omega)}$, where $\bar{c}$ is the complex conjugate of a

[2] Well, it's actually a little bit more profound than that. It also has to do with the fact that energy is proportional to frequency, but we'll get to that later.

complex number c. To prove this:

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

We call a function with this property that reversing it conjugates it a *Hermitian function*. That the Fourier transform takes pure real functions to Hermitian functions is known as *Hermitian symmetry*.

New Material: Interestingly, the Fourier transform has many more closely related symmetries, all of which can be shown just as readily as the above argument. Here is a brief list of them; I encourage the interested student to try to prove them all.

1. Given a pure real f, $\tilde{f}$ has even real part, and odd imaginary part. (Equivalent to being a Hermitian function.)
2. Given a pure imaginary f, $\tilde{f}$ has odd real part, and even imaginary part. (The opposite of being a Hermitian function.)
3. Given an even f, we get that $\tilde{f}$ is pure real.
4. Given an odd f, we get that $\tilde{f}$ is pure imaginary.

2.5 Derivative Rule

Given a function $f : \mathbb{R} \to \mathbb{C}$ that tends to zero as $x \to \pm\infty$, let g be its derivative, $g = f'$. We will now evaluate $\tilde{g}$, by integrating by parts.

$$\begin{aligned}
\tilde{g}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \overbrace{e^{-i\omega x}}^{u} \overbrace{f'(x)}^{v'}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \left(\overbrace{e^{-i\omega x}}^{u} \overbrace{f(x)}^{v} \Bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \overbrace{\left(\frac{\mathrm{d}}{\mathrm{d}x} e^{-i\omega x}\right)}^{u'} \overbrace{f(x)}^{v}\,\mathrm{d}x \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\cancelto{0}{e^{-i\omega x} f(x) \Bigg|_{-\infty}^{\infty}} + i\omega \int_{-\infty}^{\infty} e^{-i\omega x} f(x)\,\mathrm{d}x \right) \\
&= i\omega \tilde{f}(\omega)
\end{aligned}$$

Here we can ignore the boundary term because of the assumption that f tends towards zero as x goes to infinity.

This is a *very* cool result. It says that given a function f, we can differentiate it in Fourier domain simply by multiplying $\tilde{f}$ by $i\omega$. Naturally, by the uniqueness of derivatives, this immediately implies that multiplying by $(i\omega)^2 = -\omega^2$ acheives two derivatives. Further, by the (almost) uniqueness of integrals, we can immediately conclude that $\tilde{f}(\omega)/(i\omega)$ must be the Fourier transform of the integral of f!

Subtle Aside: But wait, you object, when you integrate a function, the result is defined up to the addition of a constant, namely the infamous $+C$. So too should our answer be. Observe that adding $+C$ in the time domain corresponds to adding the Fourier transform of C in the Fourier domain. Let's rewrite C via:

$$C = \sqrt{2\pi}Ce_0(x) = \sqrt{2\pi}C \cdot \frac{1}{\sqrt{2\pi}}e^{i0x}$$

From this we see that the Fourier transform C must be equal to $\sqrt{2\pi}C\delta(\omega)$, as per the calculations in the last notes. This is an infinite spike at the origin. The dual way of saying this is that the value of the Fourier transform at the origin gives the constant offset of the function. This is why the Fourier transform evaluated at zero is sometimes called the "DC offset," because the zero frequency corresponds to the absolute offset of the function, or what it integrates up to.

Thus, if g is the integral of f, we get that:

$$\tilde{g}(\omega) = \frac{1}{i\omega}\tilde{f}(\omega) + \sqrt{2\pi}C\delta(\omega) \qquad \text{For any } C.$$

Sure enough, this actually makes perfect sense! When we differentiate, we multiply by $i\omega$, which has a root at $\omega = 0$. Thus, we "forget" the DC offset of $\tilde{f}$ by differentiating. Now, when we integrate, we can choose any DC offset we want again, corresponding to the freedom to add any multiple of a Dirac delta at the origin.

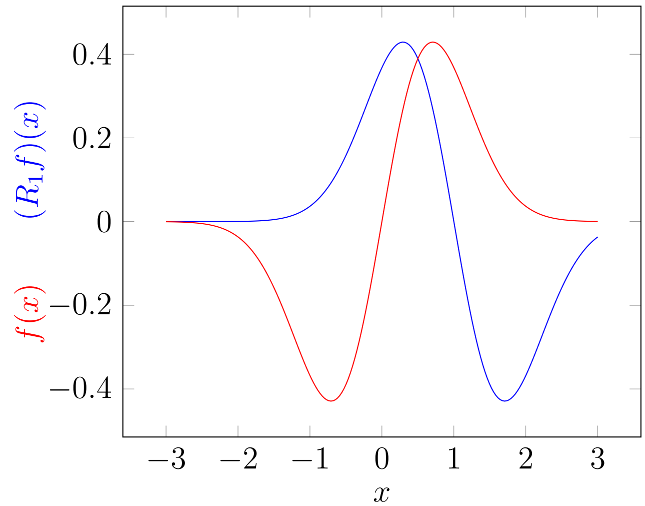
Don't worry if the above aside doesn't make sense to you, it's just an interesting side oservation.

Unimportant, But Fun Material: In fact, this rule is *so* powerful that it even lets us define the fabled "half derivative." So, if multiplying by $i\omega$ in the frequency domain achieves differentiation, then clearly multiplying by $\sqrt{i\omega}$ should acheive half a derivative. Sure enough, it does. That is to say, $\mathcal{F}^{-1}\{\sqrt{i\omega}\tilde{f}(\omega)\} = \sqrt{\partial}f$. Don't worry about it if you don't follow this, it's merely a mathematical curiosity.

3 Convolution

We're about to get into the serious meat of the Fourier transform. This is why people care about it at all.

Given a function f, let R_s be the operator that reverses its input function, then shifts it over to the right by an amount s. So, for example, given $f(x) = xe^{-x^2}$, we get that $(R_1f)(x) = f(1-x) = (1-x)e^{-(1-x)^2}$. Sure enough, if we plot these functions, we visually see that R_1 flipped the function around, and shifted it over a bit:



Here the original function is in red, and the reversed shifted copy is in blue.

In general, we can define $(R_s f)(x)$ according to:

$$(R_s f)(x) = f(s - x)$$

Thinking about this for a few seconds reveals that this does what we want.

Now, given two functions f and g, we will speak of their *convolution*, a new function that we will denote by $f * g$. We'll see in a second what the convolution of two functions looks like, but for now, a definition:

$$(f * g)(s) = \langle \overline{f}, R_s g \rangle$$

In other words to evaluate the convolution of f and g at some input s, you reverse and shift the second function by s, then take the inner product against the (conjugate of the) first function.

This definition is complex, but we can also write it out more simply. Given functions $f, g : \mathbb{R} \to \mathbb{C}$, we can expand this definition as:

$$\begin{aligned}
(f * g)(s) &= \langle \overline{f(x)}, g(s - x) \rangle \\
&= \int_{-\infty}^{\infty} \overline{\overline{f(x)}} g(s - x) \, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x) g(s - x) \, \mathrm{d}x
\end{aligned}$$

Observe that this definition immediately has symmetry because we can apply the u-

substitution $u = s - x$:

$$\begin{aligned}(f * g)(s) &= \int_{-\infty}^{\infty} f(x) g(s-x)\,\mathrm{d}x \\ &= -\int_{\infty}^{-\infty} f(s-u) g(u)\,\mathrm{d}u \\ &= \int_{-\infty}^{\infty} g(u) f(s-u)\,\mathrm{d}u \\ &= (g * f)(s)\end{aligned}$$

Thus, the operator is commutative. Further, we can also immediately see that $f*(g+h) = f*g + f*h$, simply from the linearity of integration. Further, $(cf)*(dg) = (cd)(f*g)$ for any complex constants c, d. I leave these last two properties as trivial exercises to the reader.

Now, it's much harder to show, but it actually ends up being the case that convolution is associative, namely $f*(g*h) = (f*g)*h$. We will prove this in a moment using Fourier transforms.

3.1 Intuition Please

Okay, what's going on here? We have a definition for convolution, and we even know a handful of properties of it, but absolutely no intuition. To build up intuition, let's first try the case where g is a spike to infinity at some $x = \lambda$, namely $g(x) = \delta(\lambda - x)$. Let's evaluate this on our definition:

$$\begin{aligned}\left(f(x) * \delta(\lambda - x)\right)(s) &= \int_{-\infty}^{\infty} f(x) * \delta(s + x - \lambda)\,\mathrm{d}x \\ &= f(s - \lambda)\end{aligned}$$

Okay, so the result was pretty clear. If g is a spike to infinity at some position $x = \lambda$, then f simply gets shifted to the right by x.

Now here's the really clever observation. Recall that the first notes ended in the subtle observation that any function can be split up into an infinite sum of Dirac deltas:

$$g = \int_{-\infty}^{\infty} g(x')\delta(x - x')\,\mathrm{d}x'$$

Further, observe that $f*(g+h) = f*g + f*h$, meaning that convolution distributes over the integral. That is to say, each little infinitesimal delta function that builds up g produces its own shifted copy of f: an echo, if you will. When you integrate up all these infinitesimal shifted echoes of f, you've computed the convolution of f with g![3]

This intuition is powerful and important, so please try to ponder it for a little while. I would also recommend looking at the Wikipedia article on convolution, here:

[3] What's even more mind boggling is that using g as a template to produce a bunch of shifted echoes of f ends up being the same as using f as a template to produce a bunch of shifted echoes of g.

`http://en.wikipedia.org/wiki/Convolution`

Observe that the nice animated gifs on Wikipedia that show convolution use the intuition of computing the inner product between the two functions, while one function is reversed, then slid past the other. This is completely analogous to the $\langle \overline{f}, R_s g \rangle$ definition from above, where you can view varying the argument to the convolution as varying the shift value between the functions.

Let's define:

$$\mathrm{rect}(x) = \begin{cases} 1 & \text{If } -1/2 < x < 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

This is the function that is a little box of width 1, and height 1, cenetered at the origin.

In lecture I argued that:

$$(\mathrm{rect}(x) * \mathrm{rect}(x))(s) = \begin{cases} 0 & \text{If } s \leq -1, \\ 1+s & \text{if } -1 \leq s \leq 0, \\ 1-s & \text{if } 0 \leq s \leq 1, \\ 0 & \text{if } s > 1. \end{cases}$$

If you plot it, this function is actually a little triangular hat from -1 to 1, reaching a height of 1, centered about the origin. This convolution is depicted really nicely on Wikipedia here:

`http://tinyurl.com/mh7gxea`

Again, try to build up the intuition that bluring a box out using another box produces a little triangular hat. Another way of thinking of this is that we're superimposing a bunch of echoes of the box, each shifted by an amount between $-1/2$ and $1/2$.

Yet *another* intuition for this convolution can be gotten from the $\langle \overline{f}, R_s g \rangle$ definition. If we choose $s = 0$, then $f = R_s g = \mathrm{rect}$, and we're taking the inner product of a box with itself. This gives a value of 1. However, as we increase s, we're shifting the right box, misaligning it with the first one. This decreases their inner product linearly. Finally, if $|s| > 1$, then the boxes don't overlap at all, and their inner product is zero.

Again, convolution is a really important concept, so I encourage you to think hard about the above arguments. Next week we'll prove the Fourier Convolution Theorem, which will show why this is such an important operation.