Class 3 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 Fourier Convolution Theorem

Amazingly, a beautiful formula for the convolution of two functions can be given in terms of their Fourier transforms. Specifically, if $h(s) = (f * g)(s)$, then we actually find that $\tilde{h}(\omega) = \sqrt{2\pi}\tilde{f}(\omega)\tilde{h}(\omega)$. This is known as the *Fourier Convolution Theorem*, and we will prove it now.

In order to complete the proof, first we need to know about Fubini's theorem. A special case of Fubini's theorem states that if f and g are integrable, and absolutely convergent, then the following equality holds:

$$\int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

This theorem is almost painfully obvious, if you think about it for a while. It's simply stating that the order in which we perform the integrals doesn't matter, so long as they both exist, and what not. If you seek to amuse yourself, try showing the Wikipedia article on Fubini's theorem to a physicist. He or she will doubtless respond: "Wait, there's a name for that? I always just assumed you could safely swap around integrals." Well, it turns out to take a little machinery to prove formally, but it makes enough intuitive sense, so we'll proceed simply assuming Fubini's theorem.

Now, we proceed on definitions:

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

Now we will expand $g(s-x)$ in terms of its Fourier transform, as per the notes of week 1. Recall that any function $\psi : \mathbb{R} \to \mathbb{C}$ can be expanded as:

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

Thus, we can rewrite the above as:

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

Now we can apply Fubini's theorem to this last line, to exchange the integral over ω with the integral over x:

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

On the second line we split up the $e^{i\omega(s-x)}$ and pulled $\tilde{g}(\omega)$ out of the inner integral, (which is legal because it doesn't depend on x). With a little more rearrangement, we get the last line, wherein the inner integral is simply the Fourier transform of f! Thus:

$$h(s) = \int_{-\infty}^{\infty} e^{i\omega s} \tilde{f}(\omega) \tilde{g}(\omega)\,\mathrm{d}\omega$$

Now, observe that this integral is simply the inverse Fourier transform of $\tilde{f} \cdot \tilde{g}$ using s as the variable, instead of x, and missing the $1/\sqrt{2\pi}$ out front. But if $h(s) = \mathcal{F}^{-1}\{\sqrt{2\pi}\tilde{f}(\omega)\tilde{g}(\omega)\}(s)$, then assuredly we can conclude that:

$$\tilde{h}(\omega) = \sqrt{2\pi}\tilde{f}(\omega)\tilde{g}(\omega)$$

The proof is complete.

Another way of saying this is that you can always compute the convolution of two functions by taking their Fourier tansforms, multiplying them together (and multiplying by a factor of $\sqrt{2\pi}$), then taking the inverse Fourier transform. This is actually an immensely practical trick, and one that computers use to perform convolutions all the time. It turns out that convolutions are incredibly useful all across engineering, be it for designing linear filters, blurring/deblurring images, creating special effects in graphics, simulating PDEs, or even just multiplying large numbers! Later when we talk about impulse responses we'll see why this computational trick for convolution is so useful for writing programs to simulate things like the flow of waves in water, or how heat spreads in a metal object.

If you can't wait to see how this trick gets used, try looking up:

http://en.wikipedia.org/wiki/Sch%C3%B6nhage%E2%80%93Strassen_algorithm

Now, it's worth mentioning a dual rule to this. When $h(x) = f(x)g(x)$, we get that $\tilde{h}(\omega) = \tilde{f}(\omega) * \tilde{g}(\omega)/\sqrt{2\pi}$. By now a clear pattern of duality should be clear: If something holds for the Fourier transform, almost exactly the same thing must hold for the inverse transform, because they're almost the same. I leave proving this dual rule as an exercise to the reader. **Hint:** Do the above proof over again, only this time using inverse Fourier transforms in the place of Fourier transforms, and vice versa.

1.1 Associativity, at Long Last!

This result is actually quite profound. When you convolve two functions, you're actually just multiplying their Fourier transforms. Note that this immediately implies that convolution is associative and commutative, simply because plain old multiplication is! Proof:

$$f * (g * h) = \mathcal{F}^{-1}\{2\pi \tilde{f}\tilde{g}\tilde{h}\} = (f * g) * h$$

Thus, our long awaited proof of associativity is now just a single line. Gorgeous.

2 Borwein Integrals

Alright, let's finally finish off those pesky Borwein integrals. Here's going to be our strategy. Let's say we have a function $f : \mathbb{R} \to \mathbb{C}$ that we want to know the integral of, from $-\infty$ to ∞, that is:

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$$

Well, if we could only compute $\tilde{f}$, then suddenly $\tilde{f}(0)$ is almost what we want:

$$\tilde{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i0x} f(x)\,\mathrm{d}x$$
$$\sqrt{2\pi}\tilde{f}(0) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$$

Alright, so if we can figure out the Fourier transform of the function to be integrated, then evaluating it at the origin and multiplying by $\sqrt{2\pi}$ gives us the desired integral. This is excellent.

Recall the first Borwein integral from the first week:[1]

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\,\mathrm{d}x$$

How are we to evaluate this? We can use our above trick, but we'll need to be able to take the Fourier transform of that sinc function. Let's do that first.

[1]Technically, in class and in the notes I defined the Borwein integrals as $\int_0^\infty$ instead of $\int_{-\infty}^\infty$, but that's okay. The sinc function is even, so the integral is just twice as large if you integrate in both directions. Very minor correction.

2.1 Fourier Transform of sinc

Let's examine the Fourier transform of $f(x) = \sin(x)$. Again, we can decompose:

$$f(x) = \frac{e^{ix} - e^{-ix}}{2i}$$

Splitting each term up, and recalling the Fourier transform of a complex exponential (from the first week), we get:

$$\tilde{f}(\omega) = i\sqrt{\frac{\pi}{2}}\delta(\omega + 1) - i\sqrt{\frac{\pi}{2}}\delta(\omega - 1)$$

Warning! In class I accidentally swapped these two terms, because I did my prep in for the class in Mathematica, which uses a subtly different Fourier transform convention. This was why I kept getting obnoxious sign errors later.

Now, in the second week's lecture (and in its notes) we showed that if $g(x) = f(x)/x$, then:

$$\tilde{g}(\omega) = -i \int \tilde{f}(\omega)\, \mathrm{d}\omega$$

To jog your recollection, this is simply applying the rule that "$g = f'$ implies that $\tilde{g} = i\omega\tilde{f}$", with a little rearrangement.

Thus, we actually have a way of computing this Fourier transform! We know $\mathcal{F}\{\sin(x)\}$, and we know that the Fourier transform of $\sin(x)/x$ is simply $-i$ times the integral of $\mathcal{F}\{\sin(x)\}$. Alright, let's evaluate this integral, using a particular anti-derivative of our choice for the indefinite integral:

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

Now, this last integral is easy to perform. We are computing the indefinite integral of the sum of two Dirac delta functions, one negative, and one positive. Observe that for $\omega < -1$, neither Dirac delta's spike gets included in the integral, and thus the integrand is zero. When $-1 < \omega < 1$, only the left term's spike gets swept up in the integral, and the resultant answer is $\sqrt{\pi/2}$. When $\omega > 1$, both the left and right terms' spikes get swept up in the integral, but they are of opposite signs, and thus cancel.

We could define this in terms of our previous rect function, but we're going to end up needing to manipulate so many of these later that it will be easier to just define a new function called box_w. We will define box_w to be:

$$\mathrm{box}_w(x) = \begin{cases} \frac{1}{w} & \text{If } -w < x < w, \\ 0 & \text{otherwise.} \end{cases}$$

That is to say, $\text{box}_w(x)$ is the box whose total area is 2, and which goes from $-w$ to w.

Therefore, the overall answer is:

$$\mathcal{F}\left\{\frac{\sin(x)}{x}\right\}(\omega) = \sqrt{\frac{\pi}{2}}\,\text{box}_1(\omega)$$

That is to say, the Fourier transform of a sinc is a box!

DSP Application: This property alone has *insanely* many implications in Digital Signal Processing (DSP) theory. It turns out that if you want to implement a *perfect* low-pass filter, you convolve your singal with a sinc, because it perfectly multiplies the function by a box in frequency domain, killing all the high frequency components. This is sometimes called a brick-wall filter. I encourage those who are interested to read the Wikipedia articles on the sinc function, rectangular function, and triangular function.

Now, returning to the topic at hand. We have figured out the Fourier transform of $\sin(x)/x$. Thus, we can now compute:

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\sin x}{x}\,\mathrm{d}x &= \sqrt{2\pi}\tilde{f}(0) \\
&= \sqrt{2\pi}\cdot\sqrt{\frac{\pi}{2}}\,\text{box}_1(0) \\
&= \pi
\end{aligned}$$

Huzzah! We have completed our first tricky integral.

2.2 The General Borwein Integral

Recall from the second week's notes when we proved that if $g(x) = f(\lambda x)$, then $\tilde{g}(\omega) = \frac{1}{\lambda}\tilde{f}(\omega/\lambda)$. An immediate conclusion of this is that the Fourier transform of $f(x) = \text{sinc}\left(\frac{x}{\lambda}\right)$ is given by:

$$\tilde{f}(\omega) = \sqrt{\frac{\pi}{2}}\,\text{box}_{1/\lambda}(\omega)$$

This is exactly as we wanted: When you spread the sinc out, you squish in the box that is its Fourier transform. We defined box perfectly to cancel out all the obnoxious constant factors.

Let's now handle the general case, with the product of k different sincs, using a closely related strategy:

$$\int_{-\infty}^{\infty} \underbrace{\text{sinc}(x)\cdot\text{sinc}\left(\frac{x}{3}\right)\cdot\text{sinc}\left(\frac{x}{5}\right)\cdots\text{sinc}\left(\frac{x}{n}\right)}_{k\text{ factors}}\,\mathrm{d}x$$

Again, we will define $f(x)$ to be the integrand, and attempt to evaluate $\sqrt{2\pi}\tilde{f}(0)$.

Because f is the product of a bunch of sincs, we can apply the dual Fourier Convolution Theorem we derived above, and see that $\tilde{f}$ is the convolution of a bunch of box functions.

Now, there are going to be a *lot* of obnoxious little constants to ferry around. Let's make sure we do it out formally, because I unfortunately ran out of time to do it properly in class. Firstly, we're convolving k things together, so we get $k-1$ factors of $1/\sqrt{2\pi}$ from applying the dual convolution theorem to each of the $k-1$ time-domain products we'd like to convert into a Fourier domain convolution. Next, each sinc becomes a box function with a $\sqrt{\frac{\pi}{2}}$ factor out front, so we can merge all of these immediately. Putting this all together, we get:

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}^{k-1}} \sqrt{\frac{\pi}{2}}^{k} \operatorname{box}_1 * \left(\operatorname{box}_{1/3} * \cdots * \operatorname{box}_{1/n}\right)$$

Here I carefully subdivided the convolutions into two parts: A single "outermost" convolution, and a bunch of smaller convolutions. Now, when we evaluate at $\omega = 0$, we're asking for this outermost convolution evaluated at zero. But evaluating a convolution at zero is simply the same as asking for a single inner product, where the relative shift between the functions is zero. Namely:

$$\tilde{f}(0) = \frac{1}{\sqrt{2\pi}^{k-1}} \sqrt{\frac{\pi}{2}}^{k} \int_{-\infty}^{\infty} \operatorname{box}_1(x) \underbrace{\left(\operatorname{box}_{1/3} * \cdots * \operatorname{box}_{1/n}\right)}_{\psi}(x)\, \mathrm{d}x$$

(Here I took advantage of the fact that box_1 is even, so we can ignore the reflection of one function implicit in convolution.)

Alright. Now, we're evaluating some sort of integral of $\operatorname{box}_1(x)$ times all the remaining convolutions, the result of which we'll call the function ψ for convenience. Observe that if ψ that has all of its support (read: area) entirely within the interval $[-1, 1]$, then multiplication by box_1 does nothing to its integral over $-\infty$ to ∞. Thus, there are two cases. Either this inner convolution yields a function ψ narrower than $[-1, 1]$, or it yields something wider.

Now in order to manipulate further, we're going to need some lemmas, that I forgot to prove in class, but realize now that I need. Let's examine how the width of functions behaves under convolution.

2.2.1 Width of the Convolution of Functions

The support of a function is the region in which it is non-zero. For example, box_w has support on the interval $[-w, w]$. As I argued in lecture, if I convolve two functions, one of whom has support in the interval $[-a, a]$, and the other of which has support in the interval $[-b, b]$, then the their convolution has support in the interval $[-a-b, a+b]$. This is because blurring out one function with the other effectively adds their two widths. I leave convincing yourself of this little fact as an exercise to the reader. To convince yourself, examine what happens as you slide one function along the other, and think about when the edges of the functions pass each other. It's pretty easy to visually convince yourself that the result of the convolution has a width that is the sum of the widths of the functions being convolved.

2.2.2 Area of the Convolution of Functinos

Further observe that convolving two functions yields a function whose total area under its curve is the product of the two areas. That is, given $f, g : \mathbb{R} \to \mathbb{C}$, we get that:

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

First we apply Fubini's theorem to swap the order of the integrals. Then we pull out $f(x)$ from the inner integral. Next, we apply the u substitution $u = s - x$. This makes the inner integral independent of x, so we can pull it out entirely, yielding a product of the two integrals.

This theorem is powerful! It says that when you convolve two functions, you multiply up their areas. This sort of makes sense: If we use one function to as a template to make echoes of the other, that we get the product of the areas is a little unsurprising.

2.2.3 Back to Borwein

Okay, now that we have these two lemmas, let's continue our investigation. Because $\psi = \mathrm{box}_{1/3} * \mathrm{box}_{1/5} * \cdots * \mathrm{box}_{1/n}$, from our first lemma we see that ψ must have support on the interval:

$$\left[-\frac{1}{3} - \frac{1}{5} - \cdots - \frac{1}{n}, \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{n}\right]$$

So, whenever this sum of reciprocals is less than 1, then ψ "fits" entirely within box_1, and the integral is simply the integral over ψ.

Now, by the second lemma, the integral over ψ is simply the product of the integrals over each of the boxes that ψ is a convolution of. We defined box_w to always have an area of two, regardless of w, so $\int_{-\infty}^{\infty} \psi(x)\,\mathrm{d}x = 2^{k-1}$. Putting all of this together, we get that when the sum of reciprocals of the denominators in the sincs (excluding the first) is less than 1, we get an overall answer of:

$$\begin{aligned}
\sqrt{2\pi}\tilde{f}(0) &= \frac{1}{\sqrt{2\pi}^{k-2}}\sqrt{\frac{\pi}{2}}^{k}\int_{-\infty}^{\infty} \mathrm{box}_1(x)\cdot\psi(x)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}^{k-2}}\sqrt{\frac{\pi}{2}}^{k} 2^{k-1} \\
&= \pi
\end{aligned}$$

However, when ψ is wider than box_1, then the integral on the right comes out a little bit deficient, because box_1 "cuts off" ψ. Thus, the Borwein integral going all the way up to $\operatorname{sinc}\left(\frac{x}{13}\right)$ is still equal to π, because:

$$\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{9}+\frac{1}{11}+\frac{1}{13}\approx 0.9551<1$$

But, the Borwein integral going all the way up to $\operatorname{sinc}\left(\frac{x}{15}\right)$ is slightly deficient, because:

$$\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{9}+\frac{1}{11}+\frac{1}{13}+\frac{1}{15}\approx 1.0218>1$$

2.3 Final Remarks

Here we used odd integers as the denominators on the sincs, but that is in no way fundamental. We could also have any other sequence whose sum of reciprocals eventually flips past one, in which case it might have have taken *waaaay* more terms! For example, examine the following Borwein-style integral:

$$\int_{-\infty}^{\infty}\operatorname{sinc}(x)\operatorname{sinc}\left(\frac{x}{31}\right)\operatorname{sinc}\left(\frac{x}{37}\right)\operatorname{sinc}\left(\frac{x}{41}\right)\cdots\operatorname{sinc}\left(\frac{x}{16103}\right)\,\mathrm{d}x=\pi$$

Here I choose the denominators to be primes, starting at 31, counting upwards. (If you count up the number of primes I used, this has a total of 1,867 sincs in it.)

However, if we include just *one* more prime, then suddenly:

$$\int_{-\infty}^{\infty}\operatorname{sinc}(x)\operatorname{sinc}\left(\frac{x}{31}\right)\operatorname{sinc}\left(\frac{x}{37}\right)\operatorname{sinc}\left(\frac{x}{41}\right)\cdots\operatorname{sinc}\left(\frac{x}{16103}\right)\cdot\operatorname{sinc}\left(\frac{x}{16111}\right)\,\mathrm{d}x<\pi$$

In fact, because the sum of the reciprocals of the primes diverge, you can always construct arbitrarily pathological Borwein-style integrals like this.

3 Bonus Section: Proof that the Sum of the Reciprocals of the Primes Diverges

I proved this after class. This proof is due to Erdős, and is stunningly gorgeous.

Let p_i be the ith prime number. Assume for contradiction that the sum of the reciprocals of the primes converges, namely:

$$\sum_{i=1}^{\infty}\frac{1}{p_i}<\infty$$

In this case, there must exist some k such that summing up the reciprocals of the first k primes gets you within $1/2$ of the final convergent value. In other words, there must exist a k such that:

$$\sum_{i=k+1}^{\infty}\frac{1}{p_i}<\frac{1}{2}$$

Now define $S_{j,n}$ to be the set of all positive integers at most equal to n that can be written as a product of the first j primes. For example, $S_{2,10} = \{1, 2, 3, 4, 6, 8, 9\}$. Here 5, 7, and 10 are excluded because they use primes other than 2 and 3. Now we will find contradictory estimates for $|S_{k,n}|$, as n is allowed to grow large.

An upper bound:
For any given $r \in S_{k,n}$, we can write r uniquely as the product of a square-free part, and a perfect square. That is, $r = ab^2$, with a having no more than two of any prime factor, and b being arbitrary. There are at most 2^k ways of choose a, because a either has or fails to have each of the k allowable prime factors. There are at most $\sqrt{n}$ ways to choose b, because b^2 is not allowed to exceed n, or r couldn't possibly be in $S_{k,n}$. Therefore, $|S_{k,n}| \leq 2^k\sqrt{n}$.

A lower bound:
The only way a number r satisfying $1 < r < n$ could fail to be in $S_{k,n}$ would be if r is divisible by some p_i with $i > k$. Therefore, we can lower bound the size of $S_{k,n}$ by subtracting off over-estimates of the number of multiples of p_i in the interval $[1, n]$ for each i. In other words, there are at most n/p_{k+1} integers in the interval $[1, n]$ that are divisible by the first "banned prime", p_{k+1}. Likewise, there are at most n/p_{k+2} numbers disqualified for being divisible by the second "banned prime." Thus, we can lower bound:

$$
\begin{aligned}
|S_{n,k}| &\geq n - \frac{n}{p_{k+1}} - \frac{n}{p_{k+2}} - \cdots \\
&\geq n - n\sum_{i=k+1}^{\infty}\frac{1}{p_i} \\
&\geq n - n\frac{1}{2} \\
&\geq \frac{1}{2}n
\end{aligned}
$$

We have now produced two bounds on the size of $S_{k,n}$. They contradict at, for example, $n = 4^{k+2}$. By contradiction, no such k can exist, and the sum of the reciprocals of the primes diverges.