Lecture 7:

Reading: Bendat and Piersol, Ch. 8.5.4, Ch. 5.2.3

Recap

Last time we looked at the Fourier transform, which lets us re-represent data in the time (or space) domain in terms of coefficients of sines and cosines.

Today we're going to look at 3 really key concepts for Fourier transforms.

Three great traits of the Fourier transform

We've talked about the effectiveness of the Fourier transform for identifying frequencies that are particularly energetic without having to know a priori what frequencies might have resonant peaks, and we've noted that the Fourier transform is useful for evaluating the size of one peak relative to another.

1. Derivatives in time become multiplication in the frequency domain. Fourier coefficients have some additional mathematical power. For example, suppose I want to take the time derivative of my data. If I start with

$$A(t) = \sum_{n=-\infty}^{\infty} a_n e^{-i2\pi f_n t} \tag{1}$$

then

$$\frac{\partial A(t)}{\partial t} = \sum_{n=-\infty}^{\infty} a_n \frac{\partial e^{-i2\pi f_n t}}{\partial t} = \sum_{n=-\infty}^{\infty} -i2\pi f_n a_n e^{-i2\pi f_n t} \tag{2}$$

So the first derivative become a multiplication by frequency. Higher derivatives are similarly simple

$$\frac{\partial^q A(t)}{\partial t^q} = \sum_{n=-\infty}^{\infty} (-i2\pi f_n)^q a_n e^{-i2\pi f_n t}. \tag{3}$$

Integration can be represented as a division operation:

$$\int A(t)\, dt = \sum_{n=-\infty}^{\infty} (i2\pi f_n)^{-1} a_n e^{-i2\pi f_n t} \tag{4}$$

though we'll run into a bit of trouble if $f_0 \neq 0$, that is if the record has a non-zero mean. That can mean that we might want to remove the mean before we start doing anything more complicated.

In class we illustrated this by looking at the time series of the Southern Annular Mode from http://www.nerc-bas.ac.uk/icd/gjma/sam.html. I had done a bit of pre-editing of the ASCII data file to remove the header and make it a full matrix. Then we did the following

```
% read the data
data=load('sam_nohead.txt');
data=data(:,2:13);  % remove the first column with the years
data=data'; % rotate the data so that months run down.  By doing
          % this, we can obtain a time series by using data(:);
%
% compute FFTs
fft_data = fft(data(1:717)); % eliminate the last 3 points which
                             % were NaN
```

```
fft_ddata = fft(diff(data(1:717))); % compute the fft of the first
                                    % derivative
%
% plot comparisons
semilogy(0:716,abs(fft_data).^2);  % plot the squared amplitude of
            % the fft note the symmetry, since we haven't truncated
            % at the Nyquist frequency
hold on
semilogy(0:715,abs(fft_ddata).^2,'r'); % plot the squared amplitude
                                       % of the first derivative
semilogy(0:357,abs(fft_data(1:358)).^2 .*((0:357)).^2 /715/25,'m')
    % plot the amplitude scaled by frequency, with an arbitrary
    % multiplicative factor to help us get the amplitudes to match
legend('squared fft of data','squared fft of data derivative',...
      'frequency squared times squared fft of data')
```

The results are shown in Figure 1. Here we've done this is the sloppiest way possible, but it still gives us a demonstration that the fft of the first derivative has the same spectral structure as the fft multiplied by frequency

2. Fourier transforms simplify convolution.

Suppose you plot some noisy data—the data features crazy amplitude swings, and no one can make any sense of it, but you think that hiding behind all this noise, there might be a slowly varying signal. You might be told, just do a running mean to smooth it out. That running mean is a convolution.

Convolution plays an important role in thinking about the Fourier transform, so we need to spend a little time on the concept. Here's the basic convolution integral:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau. \tag{5}$$

You can think of x as the data, and h as a filtering operator (such as a "boxcar" filter, or a triangle filter, or a roughly Gaussian-shaped window, or anything else that suits you.

In Matlab you can do this as:

```
y=conv(data(:),boxcar(12)/12);
```

which produces the same results as:

```
y=filter(boxcar(12)/12,1,data(:));
```

In both cases these will be shifted by half the width of the filter, so we can plot:

```
plot(data(:))
hold on
t(-6:731-7,conv(data(:),boxcar(12)/12),'r','LineWidth',2)
xlabel('time (months)','FontSize',14)
ylabel('SAM','FontSize',14)
legend('monthly SAM','one-year running mean of SAM')
```

See Figure 2

Formally the notation for a convolution of two records h and x is written

$$h * x = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau. \tag{6}$$

What happens if we Fourier transform this?

$$\begin{aligned}
\mathcal{F}(h * x) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau\right] e^{-it2\pi f}\, dt & (7)\\
&= \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} \left[x(t-\tau)e^{-it2\pi f}\, dt\right] d\tau & (8)\\
&= \int_{-\infty}^{\infty} h(\tau)e^{-i\tau 2\pi f}\mathcal{F}(x(f))\, d\tau & (9)\\
&= \mathcal{F}(h)\mathcal{F}(x) & (10)
\end{aligned}$$

where here I've represented the Fourier transform with a script $\mathcal{F}$.

This has profound consequences. It means that anything that required a convolution in the time domain I can handle trivially in the Fourier domain. Suppose I want to filter my data. If I don't like the hassle of convolving, I can just Fourier transform, multiply by the Fourier transform of my filter, and inverse Fourier transform. This will prove to be amazingly powerful.

3. Parseval's theorem: Total variance in the time domain equals total variance in the frequency domain

The third trait of the Fourier transform is that it conserves energy (or variance). Formally, we refer to this as Parseval's theorem, and we'll take a closer look later.

Parseval's theorem provides a critical link between total energy in the time domain and total energy in the Fourier transform domain. There are a couple of ways to thing about this.

3.1 Parseval's theorem via convolution. Let's start with the convolution of a data record with itself:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)x(t-\tau)d\tau. \tag{11}$$

What happens if I convolve my data ($x(t)$) with the time reversal of itself ($x(-t)$)?

$$y(t) = \int_{-\infty}^{\infty} x(\tau)x(t+\tau)d\tau. \tag{12}$$

More conventionally we might write:

$$y(\Delta t) = \int_{-\infty}^{\infty} x(t)x(\Delta t + t)dt. \tag{13}$$

So we're looking at the data multiplied by itself for a time lag Δt. At zero lag, this is the variance, and as we vary Δt we're looking at the lagged covariance for different time lags.

What about the Fourier transform? The Fourier transform of a convolution is simply the product of the Fourier transforms of each variable so the Fourier transform of $y(\Delta t)$ shoud be X^*X. (We use the complex conjugate of X since we convolved x with its time reversal.) We could inverse transform this back to produce $y(\Delta t)$:

$$y(\Delta t) = \int_{-\infty}^{\infty} X^*Xe^{i2\pi f\Delta t}\, df \tag{14}$$

Now, focus on the case when $\Delta t = 0$. This implies that

$$y(0) = \int_{-\infty}^{\infty} x(t)^2 dt = \int_{-\infty}^{\infty} X^* X \, df, \tag{15}$$

which tells us that the total variability in x is equivalent to the total variability in its Fourier transform X.

3.2 Parseval's theorem from the definition of the Fourier transform. Maybe a clearer way to understand Parseval's theorem is to think about the product of two variables, x_1 and x_2. We can rewrite the product, substituting the inverse Fourier transform of the Fourier transform of $x_2(t)$:

$$x_1(t)x_2(t) = x_1(t) \int_{-\infty}^{\infty} X_2(f) e^{i2\pi ft} \, df \tag{16}$$

so we can integrate this in time:

$$\int_{-\infty}^{\infty} x_1(t)x_2(t)dt = \int_{-\infty}^{\infty} \left[x_1(t) \int_{-\infty}^{\infty} X_2(f) e^{i2\pi ft} \, df \right] dt \tag{17}$$

$$= \int_{-\infty}^{\infty} X_2(f) \left[\int_{-\infty}^{\infty} x_1(t) e^{i2\pi ft} dt \right] df \tag{18}$$

$$= \int_{-\infty}^{\infty} {X_1}^*(f) X_2(f) \, df. \tag{19}$$

(My edition of Bendat and Piersol has a typo in this derivation, which appears just prior to equation 5.83, and this has caused no end of confusion.) Here we use the complex conjugate of the Fourier transform of x_1, because we computed the Fourier transform with $e^{+i2\pi ft}$ instead of the standard $e^{-i2\pi ft}$.

Put succinctly, if $x_1 = x_2$:

$$\int_{-\infty}^{\infty} x^2(t)dt = \int_{-\infty}^{\infty} |X(f)|^2 \, df \tag{20}$$

This is Parseval's relationship.

It's worth noting that if we worked with $\sigma = 2\pi f$ rather than f, we'd have to normalize by 2π:

$$\int_{-\infty}^{\infty} x^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\sigma)|^2 \, d\sigma \tag{21}$$

In thinking about the time domain vs the frequency domain, one thing to keep in mind is the distinction between integrating over all time (on the left in the above equation) and integrating over all space (on the right). This implies that we're going to need to keep track of our frequency information carefully. In essence the Fourier coefficients in X (e.g. $|a_m|^2$) do not have the same units as the time domain values in x^2, because x is integrated in time and $|a_m|$ is integrated in frequency. If the total integral of x^2 is equal to the total integral of $|X|^2$, then we're going to need to adjust by factors of δf, and this will influence how we label our axes.

More on the formalism of the Fourier transform

Now we can use this to verify that our Fourier coefficients are consistent. If I have a data set $x(t)$ that can be expressed as

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \exp(i2\pi f_n t) \tag{22}$$

and

$$a_m = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\exp(-i2\pi f_m t)\,dt \tag{23}$$

then let's check that our coefficients work out. We can substitute in $x(t)$ to obtain

$$\begin{aligned} a_m &= \frac{1}{T}\int_{-T/2}^{T/2}\sum_{n=-\infty}^{\infty} a_n \exp(i2\pi f_n t)\exp(-i2\pi f_m t)\,dt & (24)\\ &= \frac{1}{T}\int_{-T/2}^{T/2}\sum_{n=-\infty}^{\infty} a_n \exp(i2\pi(f_n - f_m)t)\,dt & (25)\\ &= \sum_{n=-\infty}^{\infty}\frac{a_n}{T}\int_{-T/2}^{T/2}\exp(i(f_n - f_m)t)\,dt. & (26)\end{aligned}$$

When $n = m$, the integral goes to T, and the summed expression becomes a_n. When $n \neq m$, we're dealing with orthogonal cosines and sines, and the integral goes to zero. Thus the net result is that

$$\begin{aligned} a_m &= \sum_{n=-\infty}^{\infty} a_n \delta_{nm} & (27)\\ &= a_m & (28)\end{aligned}$$

where δ_{nm} is called the Kronecker delta function, with $\delta_{nm} = 1$ if $n = m$ and $\delta_{nm} = 0$ otherwise. (Formally in continuous form the δ function can be thought of as a distribution, like a pdf, that has shrunk to be infinitely high and infinitesimally narrow, so that the area under the distribution is exactly 1.)

Red, white, and blue spectra

Now let's look at a few spectra. We use words associated with light to talk about spectra. Red colors have long wavelengths (e.g. infrared), while blues and purples have short wavelengths (e.g. ultraviolet). If a spectrum is dominated by low frequencies or long wavelengths, we refer to it as "red". If it is dominated by short wavelengths or high frequencies, is is "blue". If it has nearly the same energy levels at all frequencies or wavelengths, then it is "white", like the white broad-spectrum lights that we use for electric lighting.

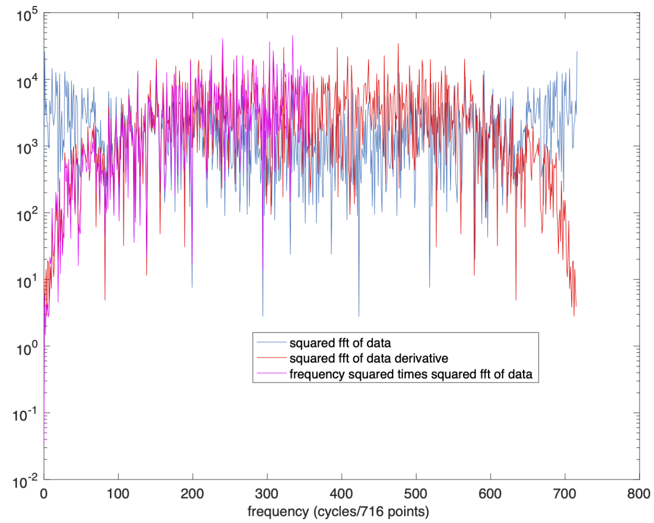

Figure 1: Squared Fourier amplitudes computed from the time series of the Southern Annular mode, as discussed in the text.

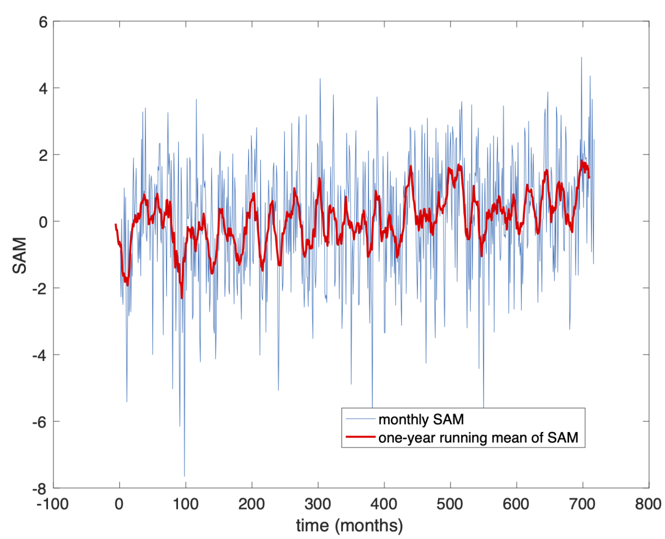

Figure 2: Time series of the Southern Annular Mode (SAM) and a one-year running mean.