Lecture 3

Power spectrum issues
- Frequentist approach
- Bayesian approach (some comments)

On web page
- Assignment 1
- Notes: Linear Shift Invariant Systems and Fourier transforms

Other reading: Gregory, Appendix B (DFTs)
LSI Systems

In the notes Linear, Shift-invariant Systems and Fourier Transforms on the course website it is shown that exponentials are an appropriate basis for LSI systems. LSI systems involve an impulse response (equivalent to a Green’s function): \( h(t) \) that is convolved with the input of the system to obtain the output. There are then two basic kinds of systems if we consider them to involve time-domain quantities:

Causal

\( h(t) = 0 \) for \( t < 0 \)

The output depends only on past values of the input.

\[
H(s) = \int_{-\infty}^{0} dt e^{-st} h(t)
\]

Laplace transform

Acausal

\( h(t) \) not necessarily \( 0 \) for \( t < 0 \)

\[
H(s) = \int_{-\infty}^{\infty} dt e^{-st} h(t) |_{s = i\omega}
\]

Fourier transform

Exponentials are useful for describing the action of a linear system because they are eigenfunctions of the system. If we can describe the actual input function in terms of exponential functions then determining the resultant output becomes trivial. This is, of course, the essence of Fourier transform treatments of linear systems and their underlying differential equations.

Focus on Fourier Transforms

We will use only FTs because we will use them for spatial (and other) domain analyses. Also, with digital data sets the notion of past and future becomes blurred unless computations are being done in real time.

There are three classes of Fourier Transforms available whose utility depends on the nature of the quantity being considered:

Fourier Transform (FT): applies to continuous, aperiodic functions:

\[
s(t) = \int_{-\infty}^{\infty} dt e^{i\omega t} \tilde{X}(\omega)
\]

The basis functions \( e^{i\omega t} \) are orthonormal on \([-\infty, \infty]\):

\[
\int_{-\infty}^{\infty} dt e^{i\omega_1 t} e^{-i\omega_2 t} = \delta(\omega_1 - \omega_2)
\]

Fourier Series: applies to continuous, periodic functions with period \( P \), discrete frequencies:

\[
s(t) = \sum_{n=-\infty}^{\infty} \tilde{X}_n \cos(2\pi nt/P)
\]

\[
\tilde{X}_n = \frac{1}{P} \int_{0}^{P} dt e^{-i2\pi nt/P} s(t)
\]

\( s(t) \) periodic with period \( P \), orthonormality on \([0, P]\):

\[
\int_{0}^{P} \cos(2\pi nt/P) \cos(2\pi nt'/P) = \delta_{n,n'}
\]

Discrete Fourier Transform (DFT): applies to discrete time and discrete frequency functions:

\[
x_k = \sum_{n=0}^{N-1} \tilde{X}_n e^{-i2\pi nk/N}
\]

\[
\tilde{X}_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{i2\pi nk/N}
\]

\( x_k, \tilde{X}_n \) periodic with period \( N \), orthonormality on \([0, N]\):

\[
\frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi nk} = \delta_{n,0}
\]

The Fourier transform is the most general of these because aperiodic functions are the most general and therefore the other two can be derived from it.

Note that the DFT is not “just” a sampled version of the FT. Nontrivial consequences take place upon digitization, as we shall see.
Attached is a table of Fourier transform theorems for continuous functions. These are enormously useful because they form the basis for particular algorithms and they also allow you to do complicated manipulations without explicitly doing any integrals.

1. Demonstrate the first two entries in the table by using a Gaussian function, taking its limit as its width $\rightarrow 0$. Start with the classic FT pair for a Gaussian function:

$$\tilde{\delta}(t) \overset{DFT}{\Rightarrow} e^{-\pi \omega^2}$$

Construct a Gaussian in time $g(t)$, using only the scale and shift theorems in the table (no integrals), which is the limit of an infinitesimal width, acts as a Dirac $\delta$-function, namely

$$f(t_0) = \int dt f(t)\delta(t - t_0),$$

where $f(t)$ is an arbitrary function. You need to pay attention to normalization of the function (e.g. unit area). Without doing any integrals investigate the Fourier transform of your Gaussian representation of the $\delta$-function. What is its shape and amplitude? A sketch is always helpful. What is the phase of the FT of $b(t) - t_0$ for an arbitrary time shift, $b_0$?

2. Prove the remaining theorems in the table, down to (and including) the integration theorem.

3. What does the integration theorem imply for the high-frequency components of the FT of the original function? Same question for the derivative theorem and the low-frequency components.

4. Derive the convolution and Parseval’s theorem for the DFT.

5. Analyze a time series by using discrete samples of the continuous-time model

$$x(t) = k \text{ for } x(t).$$

Let the sample interval be $\Delta t = 1$. Separately consider frequencies corresponding to periods $T = 1/\Delta t = 10, 30,$ and $0.3333$ s and use a total time span of $T = 256$ s. The instructor will put example python code on the course web site that you can use or refer to if you like.

(a) Before doing anything numerically, calculate the discrete frequency bin $\omega$ in which you expect each omnidirectional to be in or near.

(b) Numerically simulate the time series with your choice of amplitude $A$ and type of noise $\sigma$. That is, generate noise samples that are statistically independent but with an amplitude distribution that is non-Gaussian of any type you like.

(c) For each time series calculate the power spectrum from the DFT as given in class. You will want to plot each spectrum. Explain your results.

(d) Calculate a histogram of the spectrum values and explain the results.

(e) Test Parseval’s theorem using your dual-domain results.

### 1D Fourier Transform Theorems

<table>
<thead>
<tr>
<th>Function</th>
<th>Fourier transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1$</td>
<td>$\delta(f)$</td>
</tr>
<tr>
<td>$\delta(t)$</td>
<td>$\delta(f)$</td>
</tr>
<tr>
<td>$\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta)$</td>
<td>$\tilde{\delta}(f) = \Delta \sum_{n=-\infty}^{\infty} \delta(f - n/\Delta)$</td>
</tr>
<tr>
<td>$\delta(t) \ast \delta(t)$</td>
<td>$\tilde{\tilde{X}}(f) \tilde{\tilde{Y}}(f)$</td>
</tr>
<tr>
<td>$\delta(t) \ast \delta(t)$</td>
<td>$\tilde{\tilde{X}}(f) \tilde{\tilde{Y}}(f)$</td>
</tr>
<tr>
<td>$\delta(t) \ast \delta(t)$</td>
<td>$\tilde{\tilde{X}}(f) \times \tilde{\tilde{Y}}(f)$</td>
</tr>
<tr>
<td>$\delta(t) \ast \delta(t)$</td>
<td>$\tilde{\tilde{X}}(f) \times \tilde{\tilde{Y}}(f)$</td>
</tr>
<tr>
<td>$\int_{-\infty}^{\infty}</td>
<td>\tilde{f}(f)</td>
</tr>
<tr>
<td>$\frac{df}{df}$</td>
<td>$= 2\pi f \tilde{X}(f)$</td>
</tr>
<tr>
<td>$f \ast \delta(t)$</td>
<td>$\tilde{\int_{t}} \tilde{X}(f) \tilde{\delta}(f)$</td>
</tr>
<tr>
<td>$\int_{-\infty}^{\infty}</td>
<td>\tilde{f}(f)</td>
</tr>
<tr>
<td>$f \ast \delta(t)$</td>
<td>$\tilde{\int_{t}} \tilde{X}(f) \tilde{\delta}(f)$</td>
</tr>
<tr>
<td>$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$</td>
<td>$\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$</td>
</tr>
<tr>
<td>$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$</td>
<td>$\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$</td>
</tr>
<tr>
<td>$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$</td>
<td>$\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$</td>
</tr>
<tr>
<td>$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$</td>
<td>$\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$</td>
</tr>
</tbody>
</table>

$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Sampling theorem |
| $\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Bandlimited $\Delta f = \text{half BW.}$ |

$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Sampling theorem |
| $\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Bandlimited $\Delta f = \text{half BW.}$ |

$\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Sampling theorem |
| $\delta(t) \equiv \frac{2\pi \delta(f)}{2\pi (1 - \epsilon \epsilon^*)}$ | $\equiv \sum_{n=-\infty}^{\infty} e^{-2\pi i n\epsilon \epsilon^*} \frac{1}{\sqrt{2\pi}}$ | Bandlimited $\Delta f = \text{half BW.}$ |
Utility

- Conceptual, making guesses, actually obtaining FTs of complicated functions
- Basis for algorithms:
  - E.g. Shift theorem:
    \[ x(t - t_0) \iff \tilde{X}(f) e^{-2\pi if t_0} \]

- Can use matched filtering in the t domain or estimate the phase slope of the Fourier transform.
- There is a theoretical limit to how precisely an object can be located \( \sim 1 / (\text{signal to noise ratio}) \)
- In pulse localization, frequency domain approaches are easier to implement and achieve the theoretical limit

----

Hi, Dr. Elizabeth?
Yeah, uh... I accidentally took the Fourier transform of my cat...

Meow!
Mapping Frequency to Frequency Bin

Consider

\[ x(t) = e^{i\omega_0 t} = e^{i2\pi ft} \]

In an N-point DFT, in which bin does the signal fall (mostly)?

As before we have \( T = N\delta t \) and \( \delta f = 1/T \)

The frequency mapping is

\[ f_j = j\delta f \quad \text{for} \quad j = 0, \ldots, N/2 \]
\[ f_j = (N - j)\delta f \quad \text{for} \quad j = N/2 + 1, \ldots, N - 1 \]

The Nyquist frequency is the maximum frequency in the spectrum

\[ f_N = \frac{1}{2\delta t} \]

So for \( f_0 \leq f_N \) we the corresponding frequency bin \( k \) in the DFT is

\[ k_0 = f_0N\delta t \]

Negative frequencies correspond to \( k \) above the Nyquist frequency.

Frequencies \( |f_0| > f_N \) are still represented in the DFT (remember ... it is lossless) but they appear at aliased frequencies.

The frequency bin \( k_0 \) varies with \( N \) for fixed \( f_0 \) and \( \delta t \) and varies with \( \delta t \) for fixed \( f_0 \) and \( N \).

How do we fill an array to get a real signal in the other domain?

• \( k = 0, 1, \ldots, N-1 \)
• Need to know \( N/2+1 \) unique values

\[ \tilde{X}_0 \equiv \tilde{X}_{N-1} \]
\[ \tilde{X}_1 \equiv \tilde{X}_{N-2} \]

unique values
Fourier Comments

• Recall that Fourier series/transforms involve exponential basis functions that are orthogonal over a relevant interval (e.g. [0, N-1] for the DFT)

• If there are unknown values of the function in either domain, then orthonormality is broken
  • gaps
  • nonuniform sampling
  • misestimated values before FT

• The three FT forms are similar but not always interchangeable
  • sampling and transforming do not commute

• Issues:
  • aliasing
  • periodicity and convolution

Vector form of DFT

A time series can be written as a data vector. So can its Fourier transform. An N-point FT can be written as the product of a matrix and the data vector.

What does this matrix look like?
DFT of a Complex Exponential + Noise
Consider a time series

\[ x_n = A e^{i\omega_0 n \delta t} + n_n, \quad n = 0, \ldots, N - 1 \]

where \( n_n \) is complex white noise.

What are the properties of white noise? By definition, white noise has a flat spectrum. But this means flat in the mean. Mean over what? Over a statistical ensemble. What is an ensemble?

We will define these later. But it needs to be clear that we have one realization of data that is conceptually part of an ensemble of all possible realizations.
Fourier-based Power Spectrum

\[ S_k = \frac{|\tilde{X}_k|^2}{N} \]

Often called the periodogram [Shuster, 1898]


In Python speak: \( S_k = \text{abs(fft(x))}^2/N \)
1st and 2nd-order Moments

We characterize the noise using the statistical moments. These are the first moment (the mean) and the second moment, the autocorrelation function (ACF).

Ensemble average moments are designated with angular brackets \( \langle \cdots \rangle \):

The mean of the white noise is
\[
\langle n_n \rangle = 0 \quad \text{zero mean}
\]
and the ACF is written in terms of a Kronecker delta,
\[
\langle n_n n_m^* \rangle = \sigma_n^2 \delta_{nm} \quad \text{white noise}.
\]

The \( \delta \)-form of the ACF is consistent with any pair of values \( n_n \) and \( n_m \) being statistically independent. Note that we take a conjugate \( * \) because the noise is complex. One gets a different answer without the conjugate!

By definition (as we will see formally later on), the power spectrum is the Fourier transform of the ensemble-average ACF where the ACF is assumed to depend only on the difference between the two times \( n \) and \( m \). This is a property of stochastic processes that have stationary statistics.

For white noise we could write the ACF as
\[
R(n, m) \rightarrow R(n - m) = \sigma_n^2 \delta_{n,m}.
\]
The DFT of a delta function is a constant so we have shown that our definitions are consistent.

\[\int \sqrt{N} = \frac{\int}{\int} \]

We can calculate the DFT of the complex exponential because it is simply a geometric series. Usually it is not so simple!

The DFT of \( X_n \) is
\[
\tilde{X}_k = N^{-1} \sum_{n=0}^{N-1} X_n e^{-2\pi i nk/N}
\]
\[
= A \sum_{n=0}^{N-1} e^{i\phi n} e^{-2\pi i nk/N} + \sum_{n=0}^{N-1} n_n e^{-2\pi i nk/N}
\]
\[
= A e^{i\phi} \frac{\sin \frac{\pi}{N}(\omega_0 \delta t - 2\pi k/N)}{\sin \frac{\pi}{N}(\omega_0 \delta t - 2\pi k/N)} + \tilde{N}_k
\]
where \( \phi_n \) is an uninteresting phase factor and \( \tilde{N}_k \) is the DFT of the white noise.

The amplitude of the spectral line term is \( A \) (the limit where the arguments of the \( \sin \)’s \( \rightarrow 0 \)).

The noise term \( \tilde{N}_k \) is a zero mean random process with second moment
\[
\langle \tilde{N}_k \tilde{N}_k^* \rangle = N^{-2} \sum_{n=0}^{N-1} \sum_{n=0}^{N-1} \langle n_n n_m^* \rangle e^{-2\pi i (nk - n'm)/N}
\]
\[
= N^{-2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \sigma_n^2 \delta_{nm} e^{-2\pi i (nk - n'm)/N}
\]
\[
= \langle \sigma_n^2 \rangle \sum_{n=0}^{N-1} e^{-2\pi i (k-k')/N}
\]
\[
= \langle \sigma_n^2 / N \rangle \delta_{kk'}
\]
The second moment of the noise has the same form in both the time and frequency domains.
Signal to noise ratio:

The rms amplitude of the noise term (in the frequency domain) is therefore $\sigma_{\tilde{N}} = \sigma_n/\sqrt{N}$ and the signal-to-noise ratio is

$$(S/N)_{DFT} = \frac{\text{line peak}}{\text{rms noise}} = \sqrt{N} \frac{A}{\sigma_n}.$$ 

Thus, the $S/N$ of the line is $\sqrt{N}$ larger than the $S/N$ of the time series

$$(S/N)_{\text{time series}} = \frac{\text{amplitude of exponential}}{\text{rms noise}} = \frac{A}{\sigma_n}.$$ 

In practice, we must investigate the $S/N$ of the squared magnitude of the DFT. Let $\omega_0 = 2\pi f_c \delta t = 2\pi k_0/N$ so that the frequency is commensurate with the sampling in frequency space. Then $\tilde{X}_k = A \delta_{kk_0} + \tilde{N}_k$ and the spectral estimate becomes

$$S_k = |\tilde{X}_k|^2 = |A \delta_{kk_0} + \tilde{N}_k|^2$$

$$= A^2 \delta_{kk_0} + A \delta_{kk_0} (\tilde{N}_k + \tilde{N}_k^*) + |\tilde{N}_k|^2. \quad (1)$$

The ensemble average of the estimator is

$$\langle S_k \rangle = \langle |\tilde{X}_k|^2 \rangle = A^2 \delta_{kk_0} + \langle |\tilde{N}_k|^2 \rangle$$

$$= A^2 \delta_{kk_0} + \sigma_n^2/N. \quad (2)$$

The ratio of the peak to the off line mean is $NA^2/\sigma_n^2$, consistent with $(S/N)_{DFT}$ calculated before.

Detection . . . or not?

Suppose you have a data set that you think may have the form of the model given above. To answer the question “is there a signal in the data” we have to assess what are the fluctuations in the DFT (or, more usefully, the squared magnitude of the DFT = an estimate for the power spectrum) due to the additive noise. We would like to have confidence that a feature in the DFT or the spectrum is “real” as opposed to being a noise fluctuation that is spurious. To quantify our confidence, we need to know the properties of our test statistic. The following develops an approach that is applicable to the particular problem and illustrates generally how we go about assessing test statistics.
No signal case \( S/N = 0.01 \)

Why is the spectrum noisy?
What is the probability density function for the fluctuations?
What is a reasonable detection threshold?

PDF & CDF

- \( S = \) spectral amplitude
- Consider \( S \) to be a *random variable* --- a mapping from event space to real numbers (in this case)
- PDF = probability density function
  \[ f_S(S)dS = P\{\text{amplitude is in } [S, S + dS]\} \]
- CDF = cumulative distribution function
  \[ F_S(S) = P\{< S\} = \int_{-\infty}^{\infty} dS' f_S(S') \]
No signal case ($S/N = 0.01$)

Why is the spectrum noisy?
What is the probability density function for the fluctuations?
What is a reasonable detection threshold?
Figure 1: Spectrum of a signal with noise. The spectrum is shown for two cases: (SNR) = 0.500 and (SNR) = 1.000. The figure illustrates the effect of increasing signal-to-noise ratio on the spectral content.

Key points:
- N = 1024 samples
- P = 2.6 samples
- (SNR) = 0.500
- (SNR) = 1.000

The inset shows a close-up of the spectrum, highlighting the peak at the desired frequency index.
Python Code

```python
from scipy import *
from matplotlib.pyplot import *

N = 1024
np = 4.1

# dimensional quantities based on 1 sec per sample
dt = 1.
P = np*dt
tvec = arange(N)*dt

# Generate time series:
rms_phase = 0.
arg = 2.*pi*tvec/P + rms_phase * randn(N)
signal = SNR * (cos(arg) + 1j*sin(arg))
oise = (randn(N) + 1j*randn(N))/sqrt(2.)
x = signal + noise

# Calculate spectrum:
xfft = fft(x)
Sx = abs(xfft)**2/N
```

Full code: sinusoid.py
Full code: sinusoid.py
On web page

The Probability of False Alarm:

Suppose we want to test whether a feature in a spectrum is signal or noise. Let’s suppose that there is no signal (a ‘null’ hypothesis) in which case we can calculate the probability that a given amplitude is just a noise fluctuation.

If there is only noise, the probability density function of \( S_k \) for any given \( k \) is a one-sided exponential because \( S_k \) is \( \chi^2_2 \):

\[
\tilde{f}_{S_k}(S) = \frac{1}{(S_k)^{\frac{1}{2}}} e^{-S/(2S_k)} U(S)
\]

Why is the spectrum distributed as \( \chi^2_2 \) with two degrees of freedom?
Suppose there is a spike in the spectrum of amplitude $\eta(S_k)$

The noise-like aspect of $S_k$ implies that there can be spikes above a specified detection threshold that are spurious ("false alarms"). The probability that a spike has an amplitude $\geq \eta(S_k)$ is

$$P(S \geq \eta(S_k)) = \int_{\eta(S_k)}^\infty ds f_S(s) = e^{-\eta}$$

If the DFT length is $N_{\text{DFT}}$, there are $N_{\text{DFT}}$ unique values of the spectrum. Note this is true for a complex process but not for a real one. Why?

The expected number of spurious (i.e. false-alarm) spikes that equal or exceed $\eta(S_k)$ is

$$N_{\text{spurious}} = N_{\text{DFT}} e^{-\eta}$$

To have $N_{\text{spurious}} \leq 1$ we must have

$$N_{\text{DFT}} e^{-\eta} \leq 1$$

we need

$$\eta \geq \ln N_{\text{DFT}}$$

<table>
<thead>
<tr>
<th>$N_{\text{DFT}}$</th>
<th>$\eta$ to have $N_{\text{spurious}} \leq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>4.9</td>
</tr>
<tr>
<td>1k</td>
<td>6.9</td>
</tr>
<tr>
<td>16k</td>
<td>9.7</td>
</tr>
<tr>
<td>1M</td>
<td>13.9</td>
</tr>
<tr>
<td>1G</td>
<td>20.8</td>
</tr>
<tr>
<td>1T</td>
<td>27.7</td>
</tr>
</tbody>
</table>

The larger the number of trials, the higher the threshold that is needed to have a specified number of false positives.

There are never zero false positives!