Mixture of truly periodic component + noise, etc.

\[ s(t) = e^{i\omega t} \xrightarrow{G} r(t) \text{ but } G \text{ is real}. \]

\[ s(t) \xrightarrow{G} z(t) \xrightarrow{\Xi} r(t) \]

Our recipe for characterizing \( G \) was to measure the response at the frequency \( \omega_0 \), i.e.,

\[
\hat{F}(\omega) = \frac{1}{P} \int_{-P}^{P} e^{-i\omega t} r(t) dt
\]

\[
P = \frac{2\pi}{\omega_0}.
\]

But estimate of \( F(\omega) \) will differ, because of the noise \( z(t) \).

Just like we did for the pure-noise study, we can look at Fourier estimates

\[
F(\tau, \omega_0, L, T) = \int_{-T}^{T} e^{-i\omega_0 t} r(t) dt
\]

Easiest to keep \( L \) a multiple of \( P = \frac{2\pi}{\omega_0} \). (We knew \( \omega_0 \))

\( \forall \omega_0, \text{ put } T = 0 \) : we know when we start the stimulus. Also, our stimulus has a "clock" - so \( F \) may depend on \( T \).

\[
F(\tau, \omega_0, M \cdot P, 0) = \sum_{m=1}^{M} \int_{(m-1)P}^{MP} e^{-i\omega_0 t} r(t) dt = \sum_{m=1}^{M} \int_{(m-1)P}^{MP} e^{-i\omega_0 t} r(t) dt
\]
If $n$ is long enough, the noise terms in each chunk are independent. Thus, we can decompose:

$$ r(t) = q(t) + z(t), \quad r(t) \sim N(0, \sigma^2) $$

$$ e^{-i\omega t} \hat{g}(\omega), \quad \mathbb{E}\left\{ e^{-i\omega t} \hat{z}(\omega) \right\} = 0 $$

$z(t)$ is independent in each chunk of length $n\sigma$.

$$ P_z(\omega) = \lim_{n \to \infty} \frac{1}{T} \mathbb{E}\left\{ \int_0^T e^{-i\omega t} z(t) dt \right\} $$

and $\mathbb{E}\left\{ \int_0^T e^{-i\omega t} z(t) dt \right\} = 0$.

So a Fourier estimate:

$$ \int_0^n e^{-i\omega t} q(t) dt \text{ has two parts:} $$

one from $\hat{g}(\omega)$, equal to $n\sigma \hat{g}(\omega)$

one from the noise term, of mean 0 and variance $(n\sigma)^2 P_z(\omega)$.

Or,

$$ \frac{1}{T} \int_0^T e^{-i\omega t} q(t) dt \text{ has mean } \hat{g}(\omega) $$

If input is $\sin(\omega_0 t)$

then output at $\omega_0$ has mean $\hat{g}(\omega_0)$.
Summary: (slightly different view)

For a pure noise (no deterministic component) Fourier estimates have a mean of 0 and an expected magnitude that grows in proportion to $\sqrt{T}$ (variance proportional to $T$).

When a periodic component is also present, it adds a bias to each step, so expected magnitude grows in proportion to $T$.

Let's say you don't know if a deterministic component is present. You measure the power spectrum with a data length $T$.

Now, re-measure with a data length $KT$.

The growth of the peak suggests a periodic component at $\omega_0$.

But maybe not. $K_T > K$ might not continue to show growth.

What's happening? In a time period $T$, you can't tell whether there's an oscillation at period $\omega_0$, or merely a noisy band centered to $\omega_0 \pm \frac{\pi}{T}$. 

\[ \text{Power Spectrum} \]

\[ \text{Periodic Component} \]

\[ \text{Noise Band} \]
To get a better understanding of this (and how power spectra in general), we need to look more closely at the estimate

$$F(g, \omega_0, L, T) = \sum_{T+L}^{T} g(t) e^{-i\omega t} dt$$

View this as

$$\int_{-\infty}^{\infty} g(t) W(t) e^{-i\omega t} dt$$

where $W(t)$ is a "window" function.

We'd like to understand how the variance of $F(g, \omega_0, W)$ behaves, as $\omega_1 \rightarrow 1W$.

But this is "just like" looking at the power spectrum of $Y(t) = g(t) W(t)$.

**Fourier Multipliers in freq domain $\leftrightarrow$ convolve in time domain**

$\downarrow$

**Multiply in time domain $\leftrightarrow$ convolve in frequency domain**

So the power spectrum of $Y$ is the variance of $Y$

$\hat{g}(\omega) \widehat{W}(\omega_0-\omega_0)$

Each of which is a convolution of Fourier estimates $g$ and $W$.

$$\gamma(\omega) = \mathcal{F}\{g(\cdot) \cdot W(\cdot)\}$$

$$\langle |\gamma(\omega)|^2 \rangle = \langle \mathcal{F}\{g(\cdot) \cdot W(\cdot)\} \rangle$$

will have terms like $\hat{g}(\omega) \hat{g}(\omega')$, which $\Rightarrow 0$ unless $\omega = \omega'$

$$(e^{i\omega T} e^{-i\omega' T} = 1)$$

must be modified
So the only terms that contribute to
\[ \langle \hat{\gamma}(\omega) \rangle = \int \langle \hat{\gamma}(\omega) \rangle \hat{W}(\omega, -\omega) d\omega \]
i.e., the power spectrum of \( \gamma \) is the convolution of the
power spectrum of \( \hat{\gamma} \), and \( \hat{W} \).

[see p. 89 of 2003-4 notes]

Recall Whitney \( \hat{W}(\omega) \) for

\[ \text{sinc} \left( \frac{\omega L}{2} \right) = \frac{\sin (\frac{\omega L}{2})}{\omega L/2} \]

True power spectrum is blurred by

\[ \frac{\sin^2 \left( \frac{\omega L}{2} \right)}{(\omega L/2)^2} \]

Now what happens if true spectrum looks like

\[ \hat{\gamma}(\omega) = \]

Estimates will be severely corrupted at noisy frequencies.
Lots of nearly-periodic signals in the real world.
How to choose \( L \)? Or how to choose \( W \)?

"Obvious" approach to spectral estimation is to choose windows that are non-overlapping, to get independent estimates:

\[
\hat{s}(t)
\]

But the hard edges lead to out-of-band spread \( \sin \frac{\omega}{2} \):

"Walsh" strategies, using cosine bells

Why not place cosine bells butting in between to make use of the minimally-redundant data? Y

Better idea: consider the properties of non-overlapping windows:

\[
\int g(t) W_1(t) e^{j\omega t} dt \quad \text{and} \quad \int g(t) W_2(t) e^{j\omega t} dt.
\]

One is \( \hat{g}(\omega) \ast W_1(\omega) \), the other, \( \hat{g}(\omega) \ast \hat{W}_2(\omega) \).

The "surprise" is that these estimates are uncorrelated if \( g(t) \) is locally flat over \( \int W_1(t) W_2(t) dt = 0 \).
The implication of this observation (Thompson's Multiplier Estimator) is that we can get multiple estimates of
\[ \int g(t)W_k(t) dt \text{ from the entire data length} \]
we should choose \( W_k(t) \) to be orthogonal
\[ \int W_j(t)W_k(t)dt = 0 \]
and to have spectra like \[ \text{[Diagram]} \]

To see that these estimates are approximately independent:
Instead of considering \( \int g(t)W_j(t) dt \text{ which leads to} \)
\[ h(t) = g(t) e^{jyt} \]
consider \( \int h(t)W_j(t)dt = \int h(t)W_j(-t)dt \)
This is a convolution of \( h \ast W_j (0) \), so it is also \( \int \text{H} \ast \text{W} \int \text{W} \) with \( R + \text{W} \) is 0 outside a narrow range \( \text{W} \) and \( h(t) \)
is assumed to be flat near \( w=0 \) (since \( \text{W} \) was assumed flat)
so \( h(t) \) might as well be flat everywhere.

So we might as well have considered the behavior of \( \int h(t)W_j(t)dt \) when \( h(t) \) being white noise < \( h(t)h(t') = K \delta(t-t') \)
\[ \int \int h(t)W_j(t)W_k(t)dt = \int \int h(t)W_j(t)W_k(t')dt \]
for \( t - t' \) survive (since \( \int h(t)h(t') = 0 \), otherwise)
\[ \Rightarrow = K \int W_j(t)W_k(t)dt. \]
So the multitaper estimate is

\[ P_x(w) = \frac{1}{K} \sum_{j=1}^{K} \left| \int_0^T x(t) W_j(t, \tau) e^{-j \omega \tau} d\tau \right|^2 \]

\[ K = 2N+1 \] is the number of orthogonal "tapers" or

\[ W_j(t, \tau) \] are known-in-advance functions that

1. are orthogonal \([-\pi, \pi]\]

2. have F.T.'s non confined to \(\frac{\pi}{T} \cdot N\)

Optimal estimate if \(P_x(w)\) is locally flat.
What is "white noise"?

Informal time-domain def: Each value of $x(t)$ is independently chosen from the same distribution (often specified as a Gaussian).

Informal frequency-domain def: $X(w)$ is constant.

Problem with time-domain definition: even within a small window $dt$, there are infinitely many samples some are very large.

Problem with frequency-domain definition: $X(w)$ constant even as $w \to \infty$.

But it's still a useful concept.

Resolution is that generally, we only care about how a system respond to white noise, not the white noise itself — only physical system has a finite bandwidth (or true resolution).

$x(t) \quad L(X(f))$

If we refine the white noise, eventually the system should care.

$\text{etc. So when one simulates a white noise,}$

$\text{the variance of the distribution}$

$\text{at each time point needs to depend}$

$\text{on the width of the simulation}$

$\text{to ensure that the behavior of a}$

$\text{system that averages over } T$

$\text{is meaningful.}$

Reduce $\Delta t$ by a factor $N \rightarrow$ increase variance by $N$.

$(\text{Var. } \Delta t)$ specifies the noise, $\text{Variance of simulation } = \frac{V}{\Delta t}$
This makes sense because the autocovariance is
\[ V_{ss}(t) = \frac{V_0}{\Delta t} \]
\[ \langle x(t)x(t+\Delta t) \rangle = c(t) \]

\[ \text{Area} = V_0, \text{ constant.} \]

\[ \text{Power spectrum} = FT \text{ of autocovariance} = \int_{-\infty}^{\infty} e^{-i\omega t} c(t) dt \]
\[ = V_0 \text{ if } \omega \ll \frac{1}{\Delta t} \]