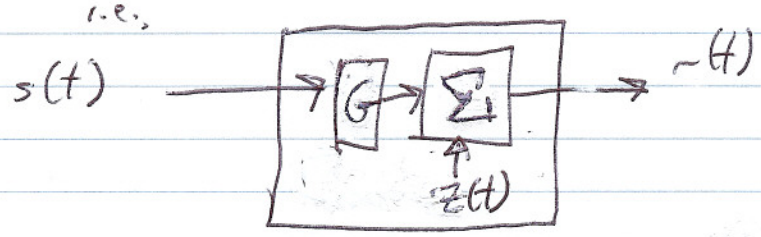
(16) Fourier Analysis – Application. Noise & Variability II.

Mixture of truly periodic component + noise, e.g.

but G is realistic

i.e.,



(t) characterized by $P_Z(\omega)$.

Our recipe for characterizing G was to measure the response at the frequency ω_0, i.e.,

$$\hat{r}(\omega_0) = \frac{1}{P}\int_0^P e^{-i\omega_0 t} r(t)\,dt$$

for $P = \frac{2\pi}{\omega_0}$.

But estimate of $F(\omega)$ will differ, because of the noise $Z(t)$

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

$$F(r, \omega_0, L, T) = \int_T^{T+L} e^{-i\omega_0 t} r(t)\,dt$$

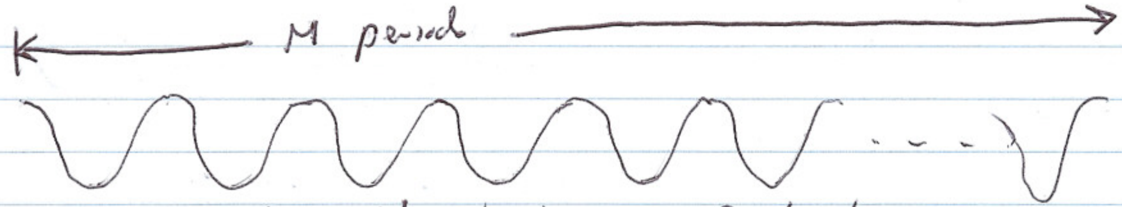
Easiest to keep L a multiple of $P = \frac{2\pi}{\omega_0}$. (we know ω_0)

Also, put $T = 0$: we know when we start the stimulus,
AND, our stimulus has a "clock" – so F may depend on T.

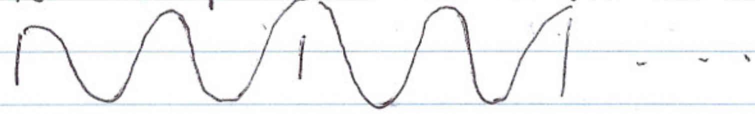
$$F(r, \omega, MP, 0) = \int_0^{MP} e^{-i\omega_0 t} r(t)\,dt = \sum_{m=1}^{M} \int_{(m-1)P}^{mP} e^{-i\omega_0 t} r(t)\,dt$$

(17)

$s(t)$



break this up into N chunks, each of length n.



If n is long enough, the noise terms in each chunk are independent. That is, we can decompose:

$r(t) = g(t) + z(t)$, i.e., $\int r(t) e^{-i\omega t} dt = \int g(t) e^{-i\omega t} dt + \int z(t) e^{i\omega t} dt$

$g(t)$ is the noise-free part of $r(t)$

$e^{i\omega_0 t}\ \hat{g}(\omega_0)$, so $\int_0^{nP} g(t) e^{-i\omega_0 t} dt = nP\,\hat{g}(\omega_0)$.

$z(t)$ is $\sim$ independent in each chunk of length nP.

$P_z(\omega) = \lim_{L\to\infty} \frac{1}{L} \langle | \int_0^L e^{i\omega t} z(t) dt |^2 \rangle$

and $\langle \int_0^L e^{i\omega t} z(t) dt \rangle = 0$.

So a Fourier estimate $\int_0^{nP} e^{-i\omega_0 t} g(t) dt$ has two parts:

one from $\hat{g}(\omega)$, equal to $nP\hat{g}(\omega_0)$

one from the noise term, of mean 0 & variance $(nP) P_z(\omega_0)$.

Or, $\frac{1}{T} \int_0^T e^{-i\omega_0 t} g(t) dt$ has mean $\hat{g}(\omega_0)$

and variance $\frac{1}{T} P_z(\omega_0)$.

If input is $\hat{s}(\omega_0) e^{i\omega t_0}$,
then output at ω_0 has mean $\hat{g}(\omega_0)\hat{s}(\omega_0)$

(8)

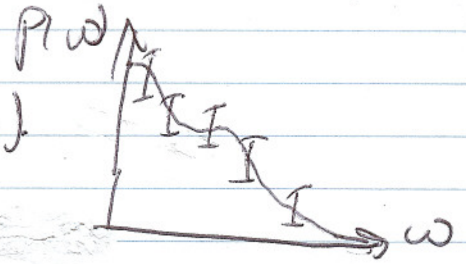
Summary: (slightly different view)

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)

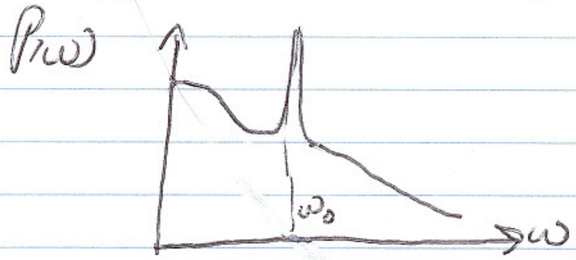
When a periodic component is also present, it adds a bias to each step, so expected magnitude grows in proportion to T.

Lets say you don't know that a deterministic component is present, & you measure the power spectrum with a data length T.

(Later - how to do this)



Now, re-measure with a data length $K_1 T$.



The growth of the peak suggests a periodic component at ω_0.

But maybe not. $K_2 T > K_1 T$ might not continue to show growth.

What's happening? In a time period T, you can't tell whether there is an oscillator of period ω_0, or merely a noise band confined to $\omega_0 \pm \frac{\pi}{T}$.

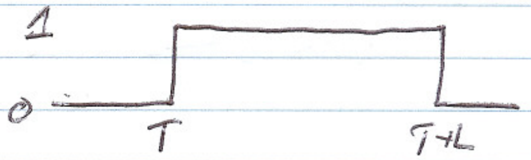
To get a better understanding of this (& also power spectra in general) we need to look more closely at the estimate

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

View this as

$$\int_{-\infty}^{\infty} g(t)\, W(t)\, e^{-i\omega_0 t}\, dt \qquad (*)$$

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



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

∴ But this is "just like" looking at the power spectrum of

$$Y(t) = g(t)\, W(t).$$

Familiar: Multiplication in freq domain $\longleftrightarrow$ convolve in time domain

⇓

Multiplication in time domain $\longleftrightarrow$ convolve in frequency domain.

So the power spectrum of Y, i.e, the variance of (*), is determined by Fourier estimates of Y, each of which is a convolution of Fourier estimates of g and W.

$$\tilde{y}(\omega_0) = \int \hat{g}(\omega)\, \hat{W}(\omega_0 - \omega)\, d\omega$$

$$\left\langle |\tilde{y}(\omega_0)|^2 \right\rangle = \left\langle \int \tilde{g}(\omega)\, \hat{W}(\omega_0 - \omega)\, d\omega \int \overline{\hat{g}(\omega')\, \hat{W}(\omega_0 - \omega')\, d\omega'} \right\rangle$$

will have terms like $\tilde{g}(\omega)\overline{\tilde{g}(\omega')}$.
which $\to 0$ unless $\omega = \omega'$. $\quad (e^{i\omega T} e^{-i\omega' T} = 1)$

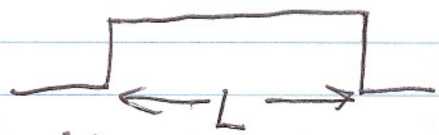
So the only terms that contribute to

$\langle|\tilde{y}(\omega_0)|^2\rangle$ are $\int|\hat{g}(\omega)|^2|\tilde{W}(\omega_0-\omega)|^2\,d\omega$

i.e., the power spectrum of Y is the convolution of the power spectrum of g, and $|\tilde{W}|^2$.

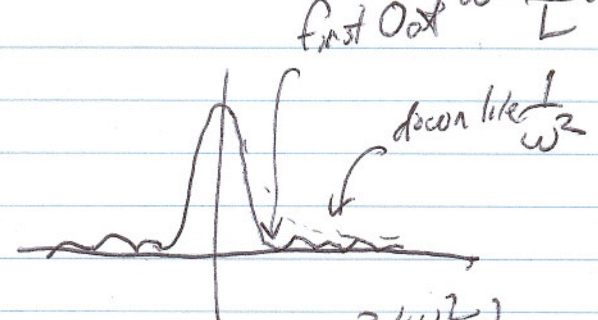
[see also PSPC 28-30 of 2003-4 notes]

Recall what is $\tilde{W}(\omega)$ for



Then $\text{sinc}\,\frac{\omega L}{2} = \frac{\sin(\frac{\omega L}{2})}{\omega L/2}$.

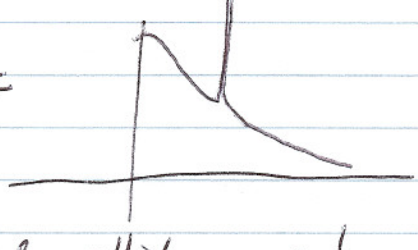
True power spectrum is blurred by



$= \frac{\sin^2(\frac{\omega L}{2})}{(\frac{\omega L}{2})^2}$

Now what happens if true spectrum looks like

$g(\omega) =$

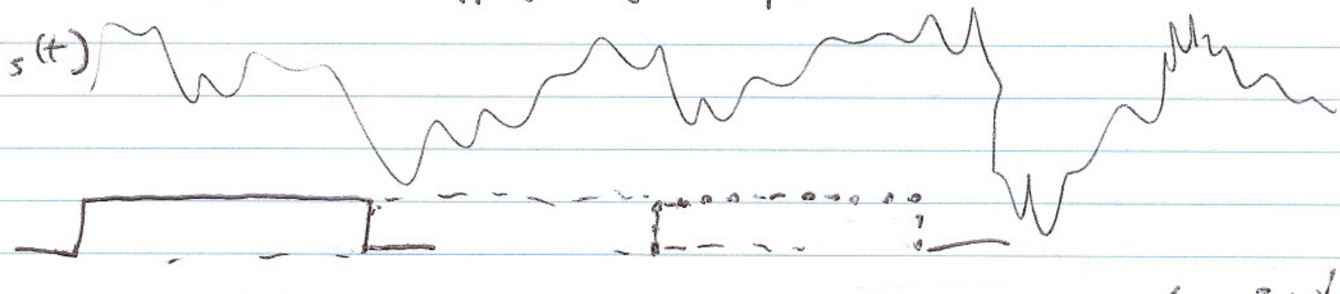


Estimate will be severely corrupted at nearby frequencies.

Lots of nearly-periodic signals in the real world.

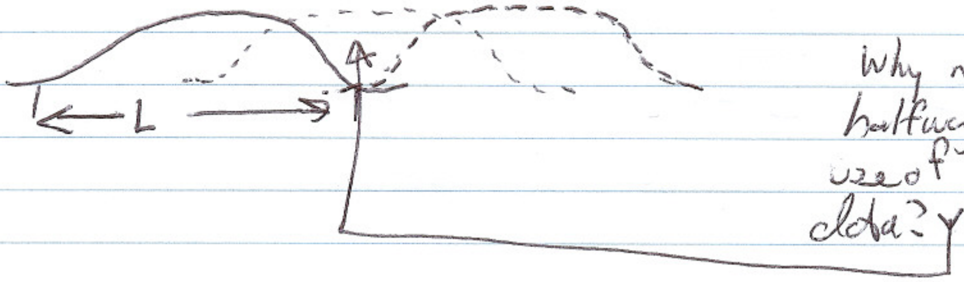
How to choose L? Or, how to choose W?

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



But the hard edges lead to out-of-band spread $\left(\operatorname{sinc}^2 \frac{\omega L}{2}\right)$

∴ "Welch" strategy, using cosine bells



Why not place cosine bells halfway in between to make use of the minimally-used data?

Better idea: consider the properties of overlapping windows:

$$\int g(t) W_1(t)\, e^{i\omega t}\, dt \quad \text{and} \quad \int g(t) W_2(t)\, e^{i\omega t}\, dt.$$

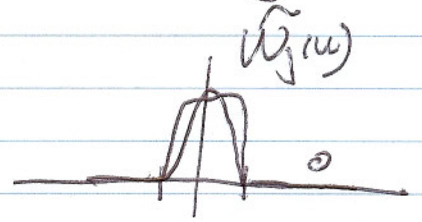
One is $\tilde{g}(\omega) * \tilde{W}_1(\omega)$, the other is $\tilde{g}(\omega) * \tilde{W}_2(\omega)$.

The "surprise" is that these estimates are <u>uncorrelated</u> if $\hat{g}(\omega)$ is locally flat <u>AND</u> $\int W_1(t) \cdot W_2(t)\, dt = 0$.

The implication of this observation (Thompson's Multitaper Estimate) is that we can get multiple estimates of $\int g(t) W_k(t) e^{i\omega t} dt$ 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



To see that these estimates are approximately independent:

Instead of considering $\int g(t) W_j(t) e^{i\omega t} dt$, look at $h(t) = g(t) e^{i\omega t}$ & consider $\int h(t) W_j(t) dt = \int h(t) W_j(-t) dt$

This is a convolution, of $h * W_j(0)$, so it's also $\int \hat{h}(\omega) \hat{W}_j(\omega) d\omega$

But $\hat{W}_j$ is 0 outside of a narrow range $\Delta\omega$, and $\hat{h}(\omega)$ is assumed to be flat near $\omega = 0$ (since $\tilde{g}(\omega)$ was assumed flat)

So $\hat{h}(\omega)$ might as well be flat everywhere.

So we might as well have considered the behavior of $\int h(t) W_j(t) dt$ for $h(t)$ being WHITE NOISE, $\langle h(t) h(t') \rangle = K \delta(t-t')$

Now $\langle \int h(t) W_j(t) dt \int h(t') W_k(-t') dt' \rangle = \langle \iint h(t) h(t') W_j(t) W_k(-t') dt dt' \rangle$

only $t = t'$ survives (since $\langle h(t) h(t') \rangle = 0$ otherwise)

$= K \int W_j(t) W_k(t) dt.$

(23)

So the multitaper estimate is

$$P_x(\omega) = \frac{1}{K} \sum_{j=1}^{K} \left| \int_0^T x(t)\, W_j(t,T)\, e^{-i\omega t}\, dt \right|^2$$

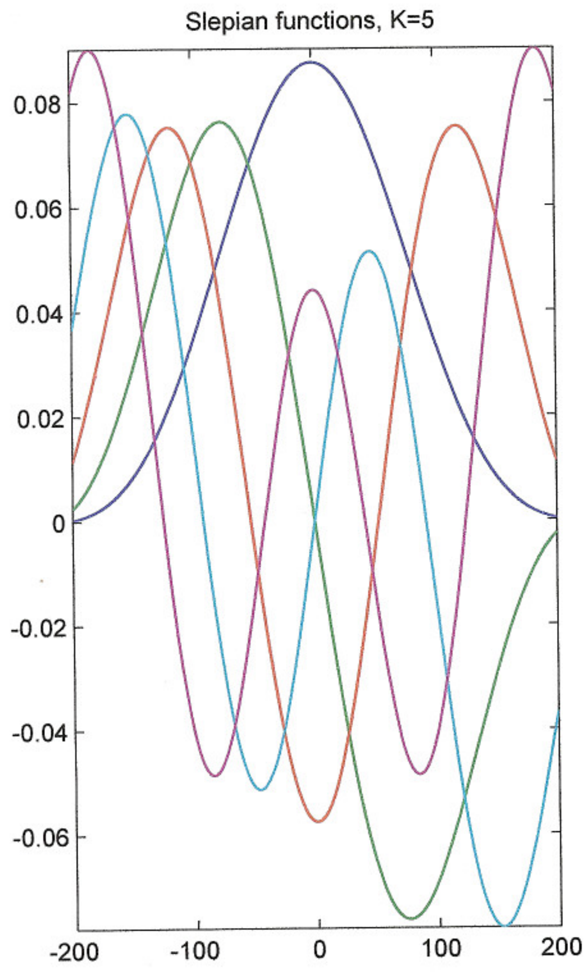
$K = 2N+1$ is the number of orthogonal "tapers" &

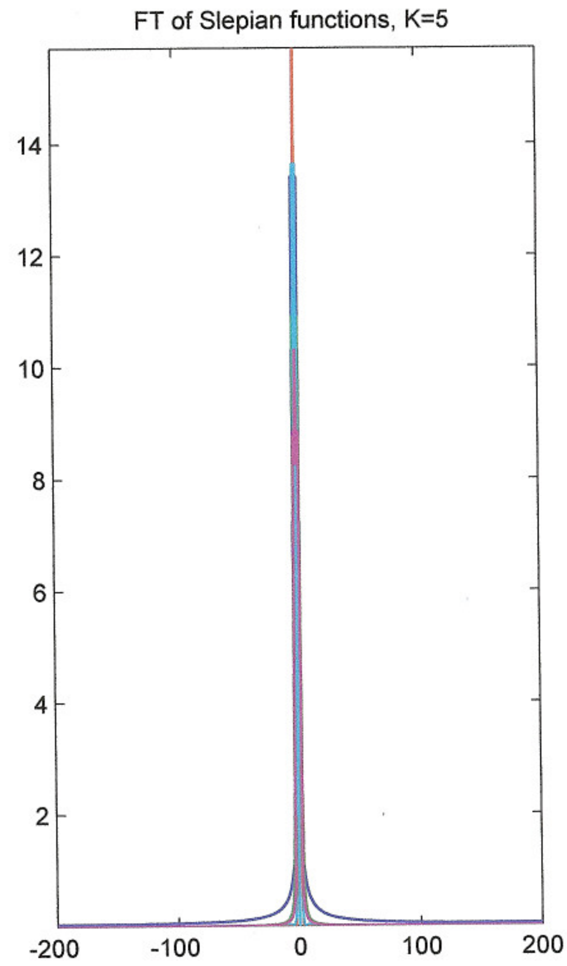
$W_j(t,T)$ are known-in-advance functions that

① are orthogonal [0, T]

② have F.T.'s ~ confined to $\frac{2\pi}{T} \cdot N$

$\boxed{\text{"optimal" estimator if } P_x(\omega) \text{ is locally flat}}$





(21)

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: $\tilde{x}(\omega)$ is constant.

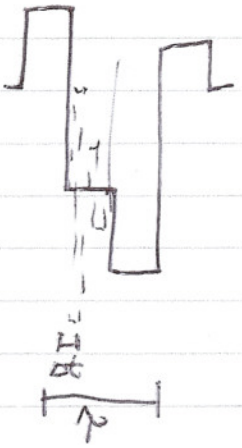
Problem with time-domain definition: even within a small window Δt, there are infinitely many samples, & some are very large.

Problem with frequency-domain definition: $\tilde{x}(\omega)$ constant even as $\omega \to \infty$!

But it's still a useful concept.

Resolution is that generally, we only care about how a system responds to white noise, not the white noise itself -- and any physical system has a finite bandwidth (or time resolution).

If we refine the white noise, eventually the system shouldn't care



etc. So when one simulates a white noise, the variance of the distribution at each time point needs to depend on the $1/\Delta t$ of the simulation - to ensure that the behavior of a system that averages over T is invariant.

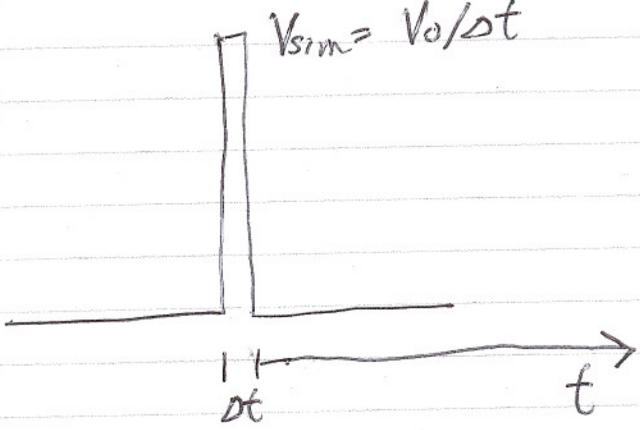
Reduce Δt by a factor N $\rightarrow$ increase variance by N.

($V_{an} \cdot \Delta t$) specifies the noise; Variance of simulation $= V_0/\Delta t$

(25)

This makes sense because the autocovariance is

$\langle x(t')x(t'+t)\rangle = c(t)$



Area $= V_0$, constant.

Power spectrum = FT of autocovariance $= \int_{-\infty}^{\infty} e^{i\omega t} c(t)\,dt$

$= V_0$ if $\omega \ll \frac{1}{\Delta t}$