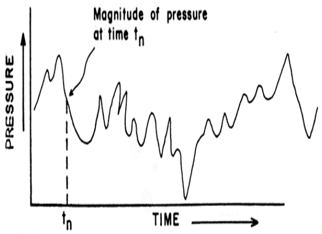
Part 3: Time Series I



(Figure from Panofsky and Brier 1968)

- Autocorrelation Function
- Harmonic Analysis
- Spectrum Analysis
- Data Window
- Significance Tests

Purpose of Time Series Analysis

Some major purposes of the statistical analysis of time series are:

- To understand the variability of the time series.
- To identify the regular and irregular oscillations of the time series.
- To describe the characteristics of these oscillations.
- To understand the physical processes that give rise to each of these oscillations.

To achieve the above, we need to:

- Identify the regular cycle (autocovariance; harmonic analysis)
- Estimate the importance of these cycles (power spectral analysis)
- Isolate or remove these cycles (filtering)

Autocovariance Function

- Originally, autocorrelation/autocovariance function is used to estimate the dominant periods in the time series.
- The autocovariance is the covariance of a variable with itself at some other time, measured by a time lag (or lead) τ.
- The autocovariance as a function of the time lag (τ and L):

$$\phi(\tau) = \frac{1}{t_2 - t_1 - \tau} \int_{t_1}^{t_2-\tau} x'(t)x'(t+\tau)dt$$

Autocorrelation Function

- The Autocorrelation function is the normalized autocovariance function:

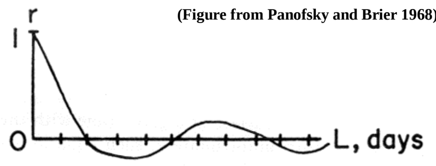
$$r(\tau) = \frac{\phi(t)}{\phi(0)}$$

- Symmetric property of the autocovarinace/autocorrelation function:

$\phi(-\tau)=\phi(\tau)$ and **$r(-\tau)=r(\tau)$.**

Typical Autocorrelation Function

(Figure from Panofsky and Brier 1968)



- If the lag is small, the autocorrelation is still positive for many geophysical variables.
- This means there is some "persistence" in the variables.
- Therefore, if there are N observations in sequence, they can not be considered independent from each other.
- This means the degree of freedom is less than N.

Degree of Freedom

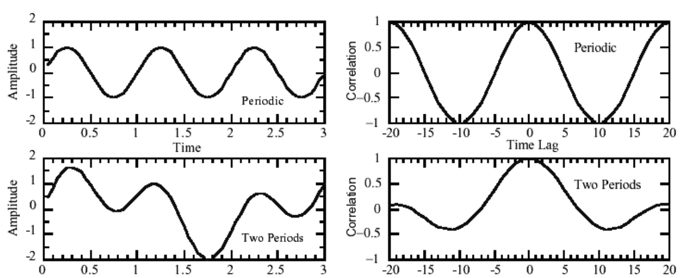
- The typical autocorrelation function tells us that data points in a time series are not independent from each other.
  ➔ The degree of freedom is less than the number of data points (N).
- Can we estimate the degree of freedom from the autocorrelation function?
- For a time series of red noise, it has been suggested that the degree of freedom can be determined as following:

  $N^* = N\ \Delta t\ /\ (2T_e)$.

  Here T_e is the e-folding decay time of autocorrelation (where autocorrelation drops to 1/e). Δt is the time interval between data.
- The number of degrees is only half of the number of *e*-folding times of the data.

Example for Periodic Time Series

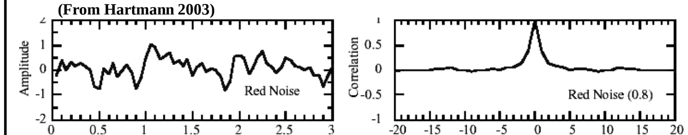
Time Series | **Autocorrelation Function**



(From Hartmann 2003)

Example – Red Noise

(From Hartmann 2003)



- The mathematic form of red noise is as following:

  $$x(t) = a\,x(t-\Delta t) + (1-a^2)^{1/2}\,\varepsilon(t)$$

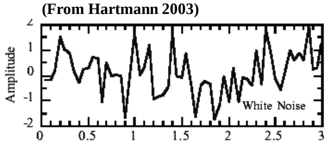
  a: the degree of memory from previous states (0 < *a* < 1)
  ε: random number
  Δt: time interval between data points
  x: standardized variable (mean =0; stand deviation = 1)
- It can be shown that the autocorrelation function of the red noise is:

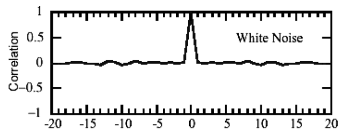
  $$r(\tau) = \exp\left(-\frac{\tau}{T_e}\right) \text{ where } T_e = \frac{-\Delta t}{\ln a}$$

  T_e is the e-folding decay time.

Example – White Noise

(From Hartmann 2003)





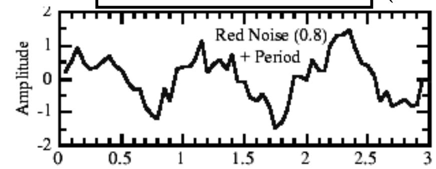
- ❑ If a = o in the red noise, then we have a white noise:
  $x(t) = \varepsilon(t)$ ➔ a series of random numbers
- ❑ The autocorrelation function of white noise is:
  $r(\tau)=\delta(0)$ ➔ non-zero only at $\tau=0$
- ❑ White noise has no prediction value.
  Red noise is useful for persistence forecasts.

ESS210B
Prof. Jin-Yi Yu

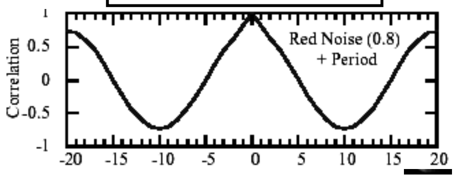
Example – Noise + Periodic

Time Series

(From Hartmann 2003)



Autocorrelation Function



ESS210B
Prof. Jin-Yi Yu

Harmonic Analysis

- ❑ Harmonic analysis is used to identify the periodic (regular) variations in geophysical time series.
- ❑ If we have N observations of (x_i, y_i), the time series $y(t)$ can be approximated by cosine and sine functions :

$$y(t) = A_o + \sum_{k=1}^{\frac{N}{2}} \left(A_k \cos 2\pi k \frac{t}{T} + B_k \sin 2\pi k \frac{t}{T} \right)$$

t: Time
T: Period of observation = $N\Delta t$
A_k, B_k: Coefficients of kth harmonic

- ❑ ***How many harmonics (cosine/sine functions) do we need?***

  In general, if the number of observations is ***N***, the number of harmonic equal to ***N/2*** (pairs of cosine and sine functions).

ESS210B
Prof. Jin-Yi Yu

What Does Each Harmonic Mean?

As an example, if the time series is the monthly-mean temperature from January to December:

- ❑ N=12, Δt =1 month, and T=12× Δt = 12 month = one year
- ❑ 1s harmonic (k=1) ➔ annual harmonic (one cycle through the period)

$$\sin\left(2\pi k \frac{t}{T}\right) = \sin\left(\frac{2\pi i \Delta t}{N \Delta t}\right) = \sin\left(\frac{2\pi i}{N}\right)$$

Period = NΔt = 12 months

- ❑ 2nd harmonic (k=2) ➔ semi-annual harmonic (2 cycles through the period)

$$\sin\left(2\pi k \frac{t}{T}\right) = \sin\left(\frac{2\pi i \Delta t}{\frac{N}{2} \Delta t}\right) = \sin\left(\frac{2\pi i}{\frac{N}{2}}\right)$$

Period = 0.5NΔt = 6 months

- ❑ Last harmonic (k=N/2) ➔ the smallest period to be included.

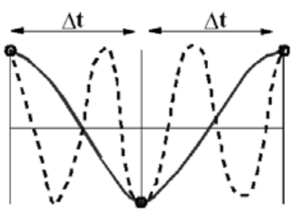
$$\sin\left(2\pi k \frac{t}{T}\right) = \sin\left(\frac{2\pi i \Delta t}{N \times \frac{2}{N} \Delta t}\right) = \sin\left(\frac{2\pi i}{2}\right)$$

Period = 2Δt = 2 months

Nyquist frequency

ESS210B
Prof. Jin-Yi Yu

Aliasing



True Period = 2/3 Δt
Aliased Period = 2 Δt

- The variances at frequency higher than the Nyquist frequency (***k>k****) will be "aliased" into lower frequency (***k<k****) in the power spectrum. This is the so-called "aliasing problem".
- This is a problem if there are large variances in the data that have frequencies smaller than k*.

Orthogonality

- In Vector Form:

$$(\mathbf{f}, \mathbf{g}) = \sum_{n=1}^{N} f_n \bullet g_n = 0$$

- In Continuous Function Form

$$(f, g) = \int_0^L f(x) g(x) dx = 0$$

- A Set of Orthogonal Functions $f_n(x)$

$$(f_n, f_m) = \int_0^L f_n(x) f_m(x) dx = \begin{cases} 0 & if\ m \neq n \\ 1 & if\ m = n \end{cases}$$

(1) The inner product of orthogonal vectors or functions is zero.

(2) The inner product of an orthogonal vector or function with itself is one.

Multiple Regression (shown before)

- If we want to regress y with more than one variables ($x_1, x_2, x_3, \ldots x_n$):

$$y = a_o + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$$

- After perform the least-square fit and remove means from all variables:

$$a_1 \overline{x_1^2} + a_2 \overline{x_1 x_2} + a_3 \overline{x_1 x_3} + \ldots + a_n \overline{x_1 x_n} = \overline{x_1 y}$$
$$a_1 \overline{x_1 x_2} + a_2 \overline{x_2^2} + a_3 \overline{x_2 x_3} + \ldots + a_n \overline{x_2 x_n} = \overline{x_2 y}$$
$$a_1 \overline{x_1 x_3} + a_2 \overline{x_2 x_3} + a_3 \overline{x_3^2} + \ldots + a_n \overline{x_3 x_n} = \overline{x_3 y}$$

- Solve the following matrix to obtain the regression coefficients: $a_1, a_2, a_3, a_4, \ldots, a_n$:

$$\begin{bmatrix} \overline{x_1^2} & \overline{x_1 x_2} & \overline{x_1 x_3} & .. \\ \overline{x_2 x_1} & \overline{x_2^2} & \overline{x_2 x_3} & .. \\ \overline{x_3 x_1} & \overline{x_3 x_2} & \overline{x_3^2} & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \ldots \end{bmatrix} = \begin{bmatrix} \overline{x_1 y} \\ \overline{x_2 y} \\ \overline{x_3 y} \\ \ldots \end{bmatrix}$$

Fourier Coefficients

- Because of the orthogonal property of cosine and sine function, all the coefficients A and B can be computed independently (you don't need to know other $A_{i=2,3,\ldots N/2}$ or $B_{i=1,2,3\ldots N/2}$ in order to get A_1, for example).
- This is a multiple regression case. Using least-square fit, we can show that:

$$A_0 = \frac{1}{N} \Sigma y_i$$
$$B_0 = 0$$
$$A_{N/2} = \frac{1}{N} \sum_{i=1}^{N} y_i \cos \pi N i \Delta t / T$$

(no $B_{N/2}$ component)

For k=1,N/2

$$\begin{cases} A_k = \frac{2}{N} \sum_{i=1}^{N} y_i \cos 2\pi k i \Delta t / T \\ B_k = \frac{2}{N} \sum_{i-1}^{N} y_i \sin 2\pi k i \Delta t / T \end{cases}$$

Amplitude of Harmonics

- Using the following relation, we can combine the sine and cosine components of the harmonic to determine the amplitude of each harmonic.

$$A\cos\theta + B\sin\theta = C\cos(\theta - \theta_0)$$

$C^2 = A^2 + B^2$ ➔ (amplitude)^2 of the harmonic

θ_0 ➔ the time (phase) when this harmonic has its largest amplitude

- With this combined form, the harmonic analysis of $y(t)$ can be rewritten as:

$$y(t) = \bar{y} + \sum_{k=1}^{\frac{N}{2}-1} C_k \cos\left\{\frac{2\pi k}{T}(t - t_k)\right\} + A_{N/2}\cos\left(\frac{\pi N t}{T}\right)$$

$$C_k^2 = A_k^2 + B_k^2 \text{ and } t_k = \frac{T}{2\pi k}\tan^{-1}\left(\frac{B_k}{A_k}\right)$$

ESS210B
Prof. Jin-Yi Yu

Fraction of Variance Explained by Harmonics

- What is the fraction of the variance (of y) explained by a single harmonic?
- Remember we have shown in the regression analysis that the fraction is equal to the square of the correlation coefficient between this harmonic and y:

$$r^2(y, x_k) = \frac{\overline{x'_k y'}^2}{\overline{x'^2_k}\ \overline{y'^2}}$$

- It can be shown that this fraction is

$$r^2 = 0.5 \times \frac{(A_k^2 + B_k^2)}{\sigma_y^2} = \frac{0.5 \times C_k^2}{\sigma_y^2} \quad \text{for k=1, 2, 3, ..., N/2-1}$$

$$= \frac{\frac{A_N^2}{2}}{\sigma_y^2} = \frac{\frac{C_N^2}{2}}{\sigma_y^2} \quad \text{for k=N/2}$$

ESS210B
Prof. Jin-Yi Yu

How Many Harmonics Do We Need?

- Since the harmonics are all uncorrelated, no two harmonics can explain the same part of the variance of y.
- In other words, the variances explained by the different harmonics can be added.
- We can add up the fractions to determine how many harmonics we need to explain most of the variations in the time series of y.

ESS210B
Prof. Jin-Yi Yu

Time Domain .vs. Frequency Domain

time

frequency

f(t) Fourier Transform → ***F(w)***

Inverse Fourier Transform

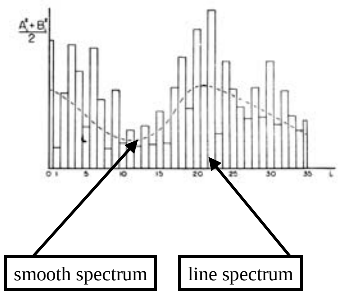
Time Series

Power Spectrum

ESS210B
Prof. Jin-Yi Yu

Power Spectrum

(Figure from Panofsky and Brier 1968)



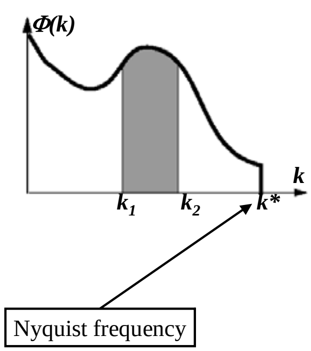
- By plotting the amplitude of the harmonics as a function of k, we produce the "power spectrum" of the time series y.
- The meaning of the spectrum is that it shows the contribution of each harmonic to the total variance.
- If t is time, then we get the frequency spectrum.
- If t is distance, then we get the wavenumber spectrum.

Problems with Line Spectrum

The C_k^2 is a "line spectrum" at a specific frequency and wavenumber (*k*). We are not interested in these line spectra. Here are the reasons:

- Integer values of ***k*** have no specific meaning. They are determined based on the length of the observation period T (=NΔt): ***k*** = (0, 1, 2, 3,..N/2) cycles during the period T.
- Since we use N observations to determine a mean and N/2 line spectra, each line spectrum has only about 2 degrees of freedom. With such small dof, the line spectrum is not likely to be reproduced from one sampling interval to the other.
- Also, most geophysical "signals" that we are interested in and wish to study are not truly "periodic". A lot of them are just "quasi-periodic", for example ENSO. So we are more interested in the spectrum over a "band" of frequencies, not at a specific frequency.

Continuous Spectrum

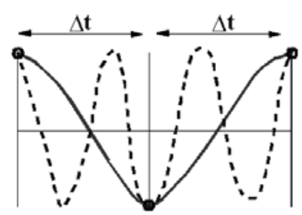


- So we need a "continuous spectrum" that tells us the variance of ***y(t)*** per unit frequency (wavenumber) interval:

$$\overline{y'^2} = \int_0^{k^*} \Phi(k)\,dk$$

- k* is called the "Nyquist frequency", which has a frequency of one cycle per 2Δt. (This is the k=N/2 harmonics).
- The Nyquist frequency is the highest frequency can be resolved by the given spacing of the data point.

Aliasing



True Period = 2/3 Δt
Aliased Period = 2 Δt

- The variances at frequency higher than the Nyquist frequency (***k>k****) will be "aliased" into lower frequency (***k<k****) in the power spectrum. This is the so-called "aliasing problem".
- This is a problem if there are large variances in the data that have frequencies smaller than k*.

How to Calculate Continuous Spectrum

❑There are two ways to calculate the continuous spectrum:

(1)(1) Direct Method (use Fourier transform)
(2)(2) Time-Lag Correlation Method (use autocorrelation function)

(1) Direct Method (a more popular method)

Step 1: Perform Fourier transform of y(t) to get $C^2(k)$
Step 2: smooth $C^2(k)$ by averaging a few adjacent frequencies together.
or by averaging the spectra of a few time series together.

➔ both ways smooth a line spectrum to a continuous spectrum and increase the degrees of freedom of the spectrum.

ESS210B
Prof. Jin-Yi Yu

Examples

❑ Example 1 – smooth over frequency bands

A time series has 900 days of record. If we do a Fourier analysis then the bandwidth will be 1/900 day^{-1}, and each of the 450 spectral estimates will have 2 degrees of freedom. If we averaged each 10 adjacent estimates together, then the bandwidth will be 1/90 day^{-1} and each estimate will have 20 d.o.f.

❑ Example 2 – smooth over spectra of several time series

Suppose we have 10 time series of 900 days. If we compute spectra for each of these and then average the individual spectral estimates for each frequency over the sample of 10 spectra, then we can derive a spectrum with a bandwidth of 1/900 $days^{-1}$ where each spectral estimate has 20 degrees of freedom.

ESS210B
Prof. Jin-Yi Yu

Time-Lag Correlation Method

(2) Time-Lag Correlation Method

It can be shown that the autocorrelation function and power spectrum are Fourier transform of each other. So we can obtain the continuous spectrum by by performing harmonic analysis on the lag correlation function on the interval $-T_L \le \tau \le T_L$.

$$\Phi(k) = \int_{-T_L}^{T_L} r(\tau) e^{-ik\tau} d\tau$$

$$r(\tau) = \frac{1}{2\pi} \int_{-k_*}^{k_*} \Phi(k) e^{ik\tau} dk$$

Φ(k): Power Spectrum in frequency (k)
r(τ): Autocorrelation in time lag (τ)

ESS210B
Prof. Jin-Yi Yu

Resolution of Spectrum - Bandwidth

❑ Bandwidth (Δf)= width of the frequency band ➔ ***resolution of spectrum***
Δf = 1/N (cycle per time interval)

❑ For example, if a time series has 900 monthly-mean data:
bandwidth = 1/900 (cycle per month).
Nyquist frequency = ½ (cycle per month)
Total number of frequency bands = (0-Nyquist frequency)/bandwidth
= (0.5)/(1/900) = 450 = N/2
Each frequency band has about 2 degree of freedom.

❑ If we average several bands together, we increase the degrees of freedom but reduce the resolution (larger bandwidth).

ESS210B
Prof. Jin-Yi Yu

Bandwidth in Time-Lag Correlation Method

- With the time-lag correlation method, the bandwidth of the power spectrum is determined by the maximum time lag (L) used in the calculation:

  $\Delta f = 1$ cycle/$(2L\Delta t)$.

- Number of frequency band = (Nyquist frequency – 0) / Δf

  $= (2\,\Delta t)^{-1} / (2L\,\Delta t)^{-1} = L$

- Number of degrees of freedom = N/(number of bands)

  = N/L

ESS210B
Prof. Jin-Yi Yu

An Example

- For red noise, we know:

  $r(\tau)=\exp(-\tau/T_e) \Rightarrow T_e = -\tau / \ln(r(\tau))$

- If we know the autocorrelation at $\tau=\Delta t$, then we can find out that

  $$\frac{N^*}{N} = -\frac{1}{2}\ln[r(\Delta t)] \;\; ; \frac{N^*}{N} \le 1$$

- For example:

$r(\Delta t)$	< 0.16	0.3	0.5	0.7	0.9
N^*/N	1	0.6	0.35	0.18	0.053

(From Hartmann 2003)

ESS210B
Prof. Jin-Yi Yu

Example – Spectrum of Red Noise

- Let's use the Parseval's theory to calculate the power spectrum of red noise.
- We already showed that the autocorrelation function of the red noise is:

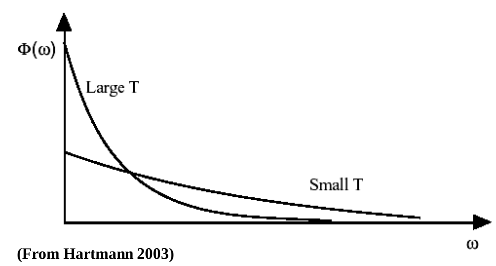
$$r(\tau) = \exp\left(-\frac{\tau}{T_e}\right) \text{ where } T_e = \frac{-\Delta t}{\ln a}$$

- By performing the Fourier transform of the autocorrelation function, we obtain the power spectrum of the red noise:

$$\Phi(\omega) = \int_{-\infty}^{\infty} \exp\left(\frac{-\tau}{T_e}\right) e^{-i\omega\tau} d\tau = \frac{2T_e}{1+\omega^2 T_e^2}$$

ESS210B
Prof. Jin-Yi Yu

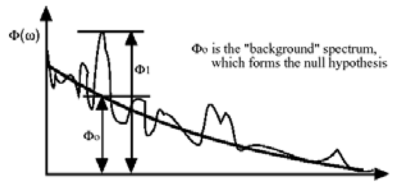
Power Spectrum of Red Noise



(From Hartmann 2003)

ESS210B
Prof. Jin-Yi Yu

Significance Test of Spectral Peak



(From Hartmann 2003)

- Null Hypothesis : the time series is not periodic in the region of interest, but simply noise.
- We thus compare amplitude of a spectral peak to a background value determined by a red noise fit to the spectrum.
- Use ***F***-Test: $F = \frac{\Phi_1}{\Phi_0}$

Calculate the Red Noise Spectrum for Test

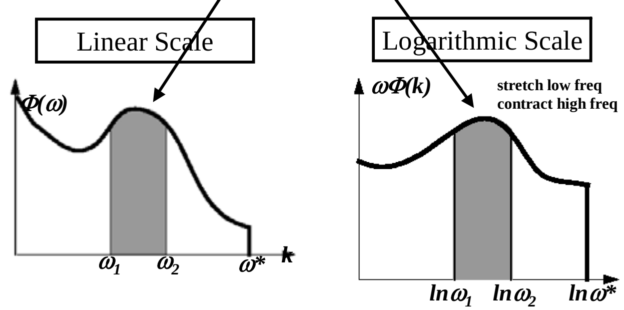
- The red noise power spectrum can be calculated using the following formula:

$$P(h,\rho,M) = \frac{1-\rho^2}{1-2\rho\cos\left(\frac{h\pi}{M}\right)+\rho^2} \times \frac{\textbf{Power of the Tested Spectrum}}{\textbf{Power of the Red Noise}}$$

- P(h, ρ, M) is the power spectrum at frequency ***h***
  h = 0, 1, 2, 3, …., M
  ρ = autocorrelation coefficient at one time lag
- We would normally obtain the parameter ρ from the original time series as the average of the one-lag autocorrelation and the square root of the two-lag autocorrelation.
- We then make the total power (variance) of this red noise spectrum equal to the total power (variance) of the power spectrum we want to test.

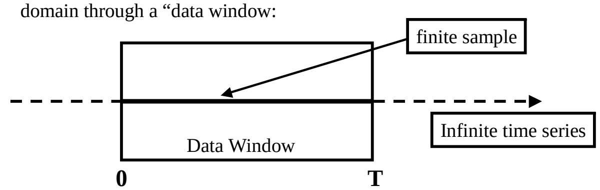
How To Plot Power Spectrum?

$$\int_{\omega_1}^{\omega_2} \Phi(\omega)\, d\omega = \int_{\ln\omega_1}^{\ln\omega_2} \omega\Phi(\omega)\, d\ln\omega$$



Data Window

- The Fourier transform obtains the "true" power spectrum from a time series with a infinite time domain. In real cases, the time series has a finite length.
- It is like that we obtain the finite time series from the infinite time domain through a "data window:



- How does the data window affect the power spectrum?

Parseval's Theorem

- This theory is important for power spectrum analysis and for time filtering to be discussed later.
- The theory states that the square of the time series integrated over time is equal to the square (inner product) of the Fourier transform integrated over frequency:

$$\int_{-\infty}^{\infty} f_1(t) f_2(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_2(\omega) F_1^*(\omega)\, d(\omega)$$

- Here $F_1(\omega)/F_2(\omega)$ is the Fourier transform of f1(t)/f2(t).

ESS210B
Prof. Jin-Yi Yu

Power Spectrum of Finite Sample

- If the infinite time series is f(t) and the sample time series is g(t), then the power spectrum calculated from the sample is related to the true spectrum in the following way:

$$G(\omega) = \int_{-\infty}^{\infty} g(t) e^{iwt} dt = \int_{-\infty}^{\infty} f(t) w(t) e^{iwt} dt$$

- Based on the "Convolution Theory"

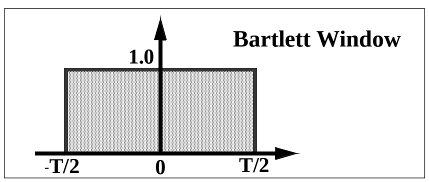
$$\int_{-\infty}^{\infty} f_1(t) f_2(t) e^{-i\omega t} dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_1(\overline{\omega}) F_2(\omega - \overline{\omega})\, d\overline{\omega}$$

- The sample spectrum is not equal to the true spectrum but weighted by the spectrum of the data window used:

$$G(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega) W(\omega - \overline{\omega})\, d\overline{\omega}$$

ESS210B
Prof. Jin-Yi Yu

Square Data Windows



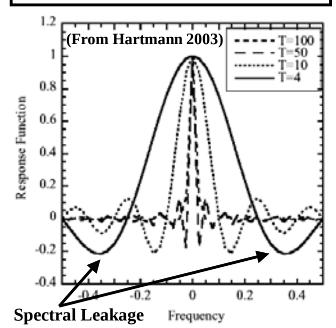
- Square data window is:
  w(t) = 1 within the window domain
  = 0 everywhere else.
- The data window has the following weighting effects on the true spectrum:

$$W(\omega) = \frac{2}{\pi}\int_{-0}^{0} \delta(\omega - \overline{\omega}) \sin(\omega - \overline{\omega}) \frac{T}{2}$$

ESS210B
Prof. Jin-Yi Yu

The Weighting Effect of Square Window

Response Function of Square Window



- The square window smooth the true spectrum.
- The degree of the smoothing is determined by the window length (T).
- The shorter the window length, the stronger the smoothing will be.
- In addition to the smoothing effect, data window also cause "spectral leakage".
- This leakage will introduce spurious oscillations at higher and lower frequencies and are out of phase with the true oscillation.

ESS210B
Prof. Jin-Yi Yu

We Wish the Data Window Can…

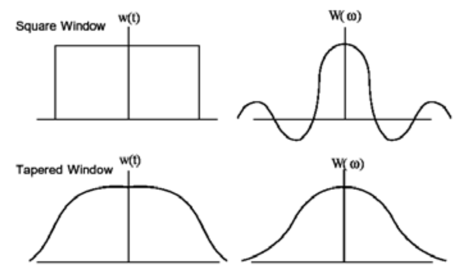
- ❑ Produce a narrow central lobe
  - ➔ require a larger T (the length of data window)

  Produce a insignificant side lobes
  - ➔ require a smooth data window without sharp corners
- ❑ A rectangular or Bartlett window leaves the time series undistorted, but can seriously distort the frequency spectrum.

  A tapered window distorts the time series but may yield a more representative frequency spectrum.

Tapered Data Window

- ❑ How do we reduce the side lobes associated with the data window?
  - ➔ A tapered data window.



(from Hartmann 2003)

Bartlett Window

- ❑ Bartlett (square or rectangular) window

$$W(t)=\begin{cases}\dfrac{1}{T} & 0\le |t|\le T\\ 0 & |t|>T\end{cases}$$

$$W(\omega)=\frac{2\sin\left(\dfrac{\omega T}{2}\right)}{\dfrac{\omega T}{2}}=2\,\mathrm{sinc}\left(\frac{\omega T}{2\pi}\right)$$

- ❑ This is the most commonly used window, but we use it without knowing we are using it.
- ❑ The Bartlett window has a serious side lobe problem. Frequencies that are outside the range of frequencies actually resolved can have too strong an influence on the power spectra at the frequencies resolved.

Hanning Window (Cosine Bell)

- ❑ The cosine bell window is perhaps the most frequently used window in meteorological applications.

$$w(t)=\begin{cases}\dfrac{1}{T}\cdot\left(1+\cos\dfrac{2\pi t}{T}\right) & ;0\le |t|\le T/2\\ 0 & ;|t|>T/2\end{cases}$$

$$W(\omega)=\mathrm{sinc}\left(\frac{\omega T}{\pi}\right)+\frac{1}{2}\left[\mathrm{sinc}\left(\frac{\omega T}{\pi}+1\right)+\mathrm{sinc}\left(\frac{\omega T}{\pi}-1\right)\right]$$

The same as Bartlett window

Partially cancel out Side lobs, but also Broaden the central lobe

Filtering of Time Series

- Time filtering technique is used to remove or to retain variations at particular bands of frequencies from the time series.
- There are three types of filtering:

(1) High-Pass Filtering

keep high-frequency parts of the variations and remove low-frequency parts of the variations.

(2) Low-Pass Filtering

keep low-frequency and remove high-frequency parts of the variations.

(3) Band-Pass Filtering

remove both higher and lower frequencies and keep only certain frequency bands of the variations.

Response Function

- Time filters are the same as the data window we have discussed earlier.

$$G(\omega) = \int_{-\infty}^{\infty} g(t)e^{iwt}dt = \int_{-\infty}^{\infty} f(t)w(t)e^{iwt}dt$$

filter or data window

- By performing Fourier transform, we know that:

$$G(\omega) = F(\omega)\cdot W(\omega)$$

- The ration between the filtered and original power spectrum is called the "response function":

$$R(\omega) = \frac{G(\omega)}{F(\omega)}$$

power spectrum after filtering

original power spectrum

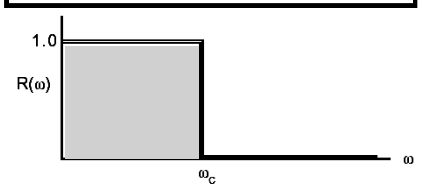
- If R(ω)=1 ➔ the original amplitude at frequency ω is kept.

  R(ω)=0 ➔ the original amplitude at frequency ω is filtered out.

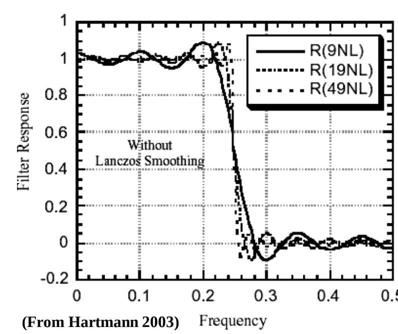
An *Perfect* Filter

- The ideal filter should have a response of 1 over the frequency bands we want to keep and a response of zero over the frequency bands we want to remove:

A Perfect Square Response Function



A Sharp-Cutoff Filter



(From Hartmann 2003)