

Lecture 4:

Time Series Analyses

Shane Elipot

The Rosenstiel School of Marine and Atmospheric Science,
University of Miami

Foreword

This lecture is heavily based on a longer course (The Oslo Lectures) given by Jonathan M. Lilly in Oslo at the invitation of the Norwegian Research School in Climate Dynamics (ResClim), during the week May 23–27, 2016. JML's Oslo Lectures are freely available for online viewing or download at www.jmlilly.net/talks/oslo/index.html



References

- [1] Bendat, J. S., & Piersol, A. G. (2011). *Random data: analysis and measurement procedures* (Vol. 729). John Wiley & Sons.
- [2] Percival, D. B. and Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. Cambridge University Press
- [3] Jenkins, G.M. and Watts, D. G. (1968). *Spectral Analysis and its Applications*. Holden Days



Outline

1. The time domain
2. Stationarity, non-stationarity, and trends
3. Fourier Spectral analysis
4. Bivariate time series
5. Filtering and other topics
6. Multitaper revisited



1. The time domain



The Sample Interval

We have a sequence of N observations

$$x_n, \quad n = 0, 1, 2, \dots, N - 1$$

which coincide with times

$$t_n, \quad n = 0, 1, 2, \dots, N - 1.$$

The sequence x_n is called a *discrete time series*.

It is assumed that the *sample interval*, Δ_t , is constant

$$t_n = n\Delta_t$$

with the time at $n = 0$ defined to be 0. The *duration* is $T = N\Delta_t$.

If the sample interval in your data is not uniform, the first processing step is to interpolate it to be so.



The Underlying Process

A critical assumption is that there exists some “process” $x(t)$ that our data sequence x_n is a *sample of*:

$$x_n = x(n\Delta_t), \quad n = 0, 1, 2, \dots, N - 1.$$

Unlike x_n , $x(t)$ is believed to exist for *all times*.

(i) The process $x(t)$ exists in *continuous time*, while x_n only exists at *discrete times*.

(ii) The process $x(t)$ exists for *all past and future* times, while x_n is only available over a certain time interval.

It is the properties of $x(t)$ that we are trying to estimate, *based on* the available sample x_n .



Measurement Noise

In reality, the measurement device and/or data processing probably introduces some artificial variability, termed *noise*.

It is more realistic to consider that the observations x_n contain samples of the process of interest, $y(t)$, *plus* some noise ϵ_n :

$$x_n = y(n\Delta_t) + \epsilon_n, \quad n = 0, 1, 2, \dots, N - 1.$$

This is an example of the *unobserved components model*. This means that we *believe* that the data is composed of *different components*, but we cannot observe these components individually.

The process $y(t)$ is potentially obscured or degraded by the limitations of data collection in three ways: (i) finite sample interval, (ii) finite duration, (iii) noise.

Because of this, the time series is an *imperfect* representation of the real-world processes we are trying to study.



Time versus Frequency

There are two complementary points of view regarding the time series x_n .

The first regards x_n as being built up as a sequence of discrete values x_0, x_2, \dots, x_{N-1} .

This is the domain of *statistics*: the mean, variance, histogram, etc.

When we look at data statistics, generally, the order in which the values are observed *doesn't matter*.

The second point of view regards x_n as being built up of sinusoids: purely periodic functions spanning the whole duration of the data.

This is the domain of *Fourier spectral analysis*.

In between these two extremes is wavelet analysis which is not covered here, see the Oslo lectures.



Time-Domain Statistics

Time domain statistics consist of the parameters we have considered earlier during the week: sample mean, sample variance, skewness, kurtosis, and higher moments.

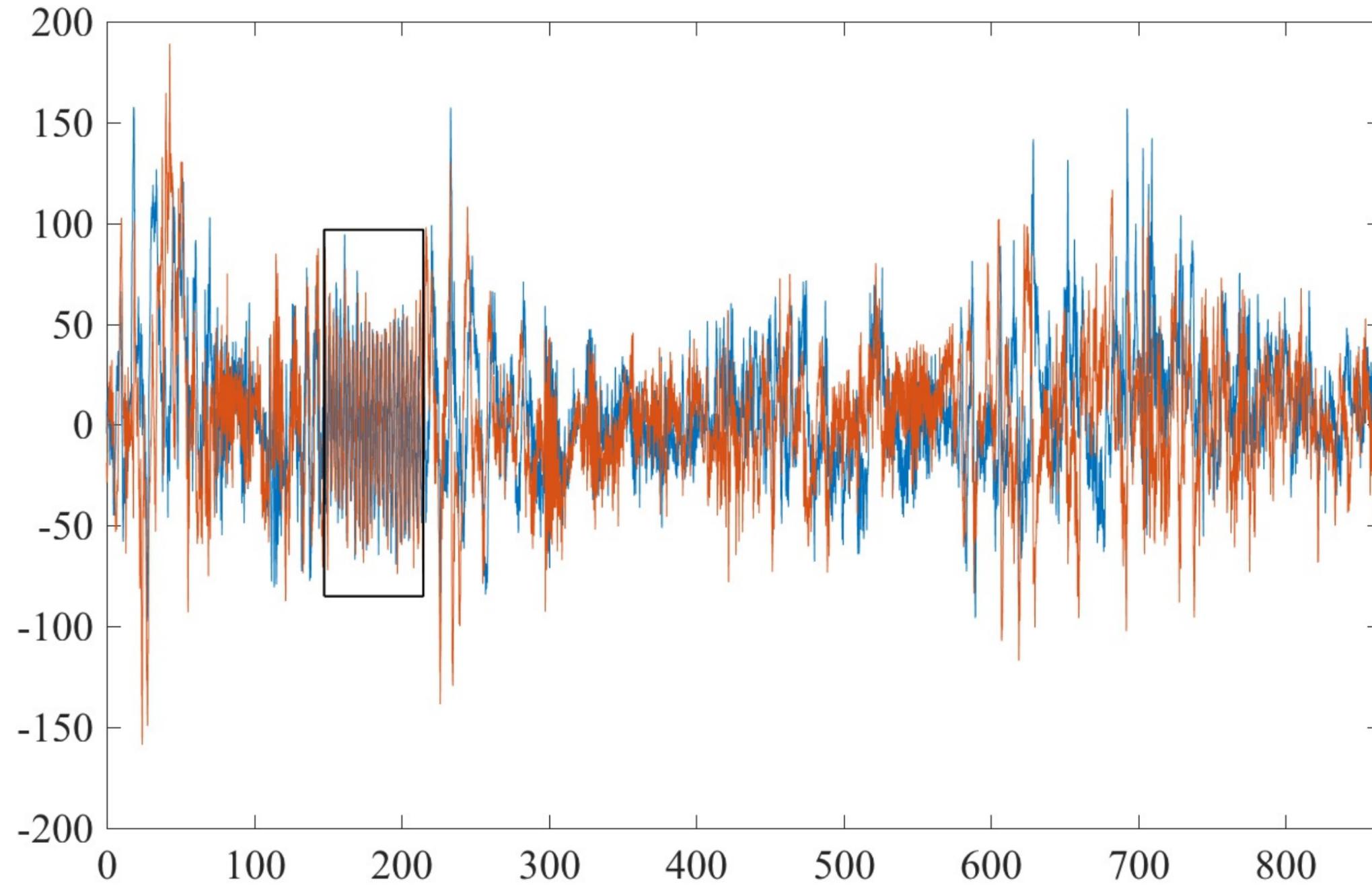
The term *sample* is being used to distinguish these quantities calculated from the data sample from the population, or true, statistics of the assumed underlying process $x(t)$. That's why we use $\widehat{(\cdot)}$



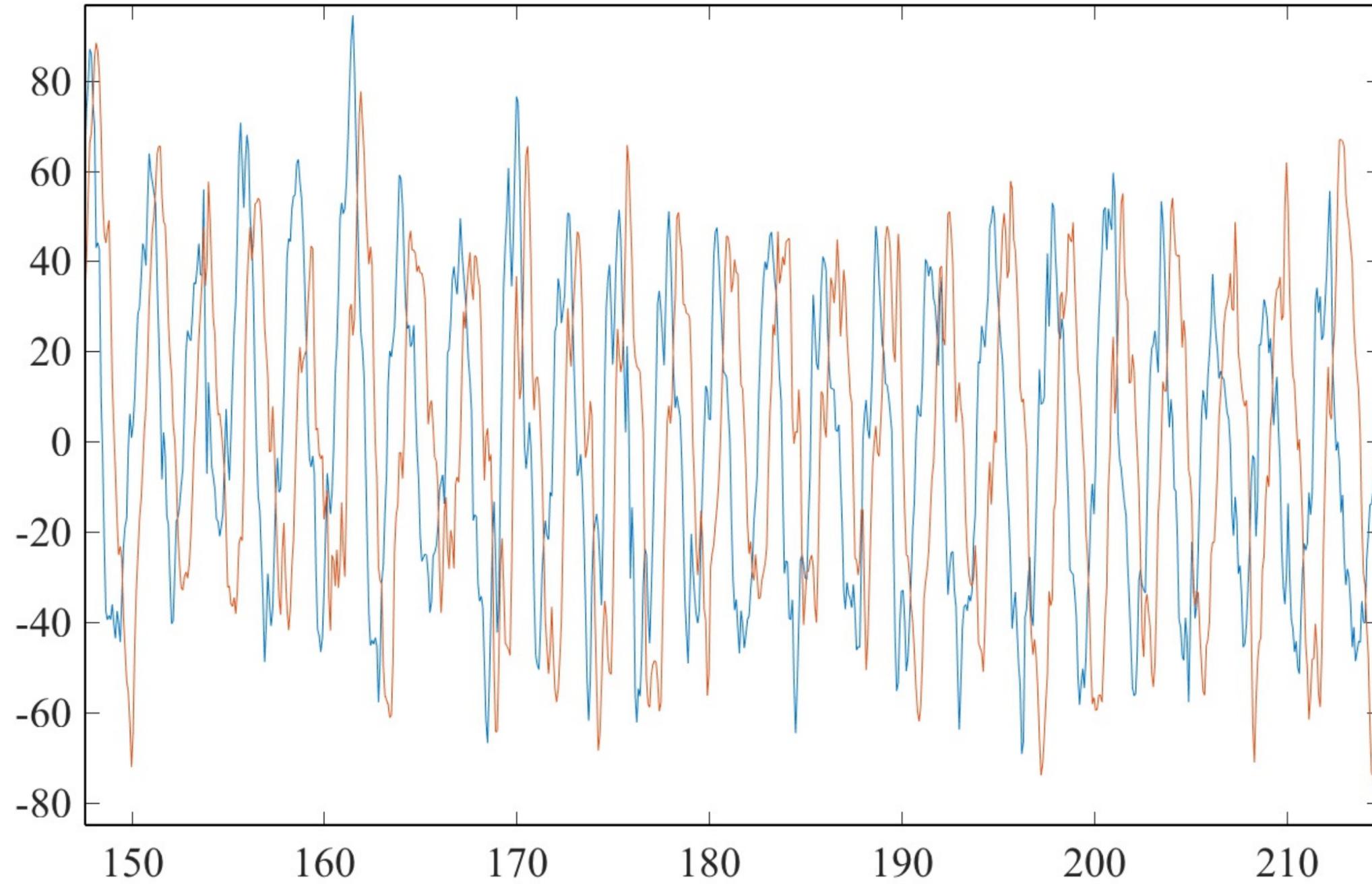
2. Stationarity vs non-stationarity, trends



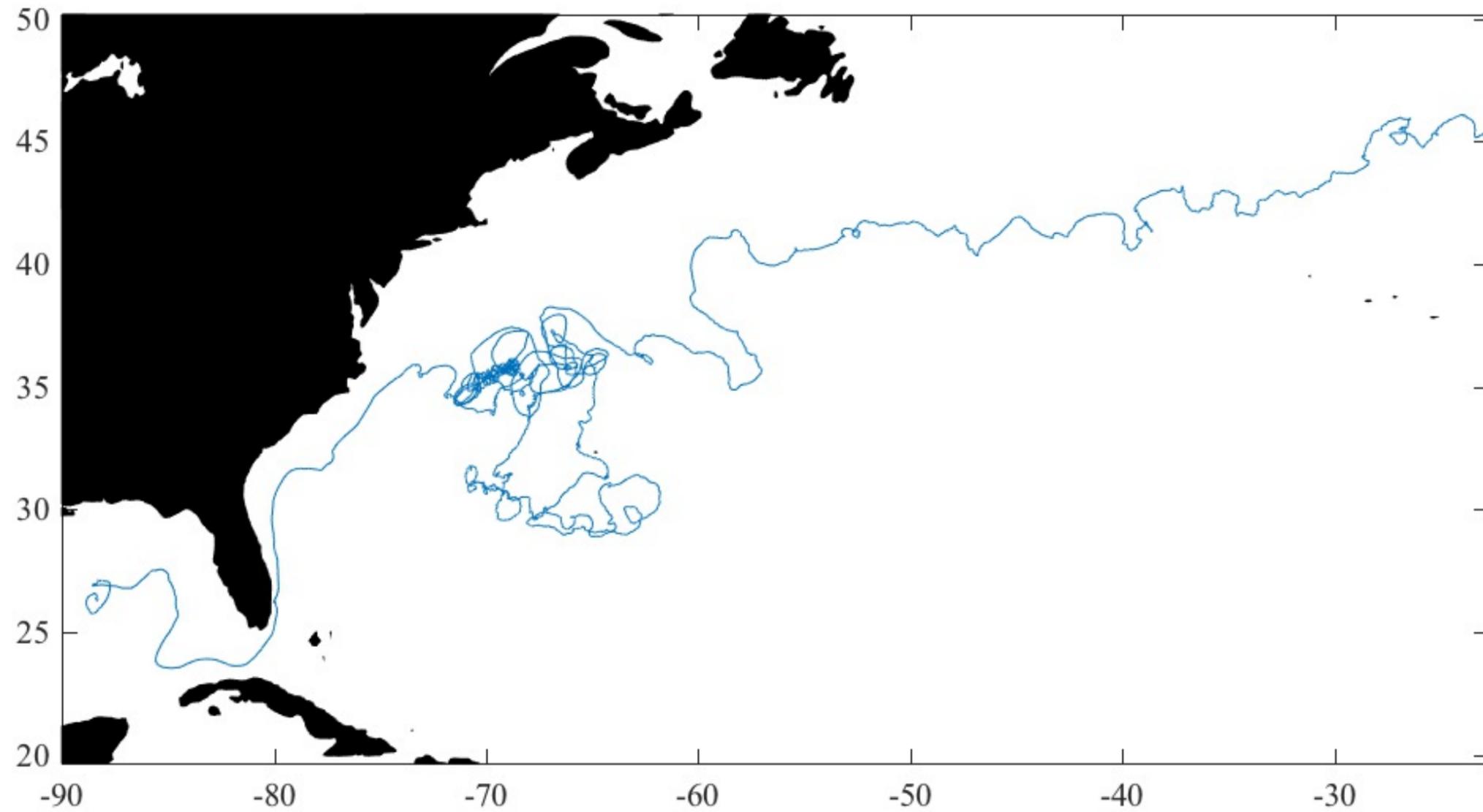
First Example



First Example



First Example



Observable Features

1. The data consists of two time series that are similar in character.
2. Both time series present a superposition of scales and a high degree of roughness.
3. The data seems to consist of different time periods with distinct statistical characteristics—the data is *nonstationary*.
4. Zooming in to one particular period show regular oscillations of roughly uniform amplitude and frequency.
5. The phasing of these show a circular polarization orbited in a counterclockwise direction.
6. The zoomed-in plot shows a fair amount of what appears to be measurement noise superimposed on the oscillatory signal.



Observable Features

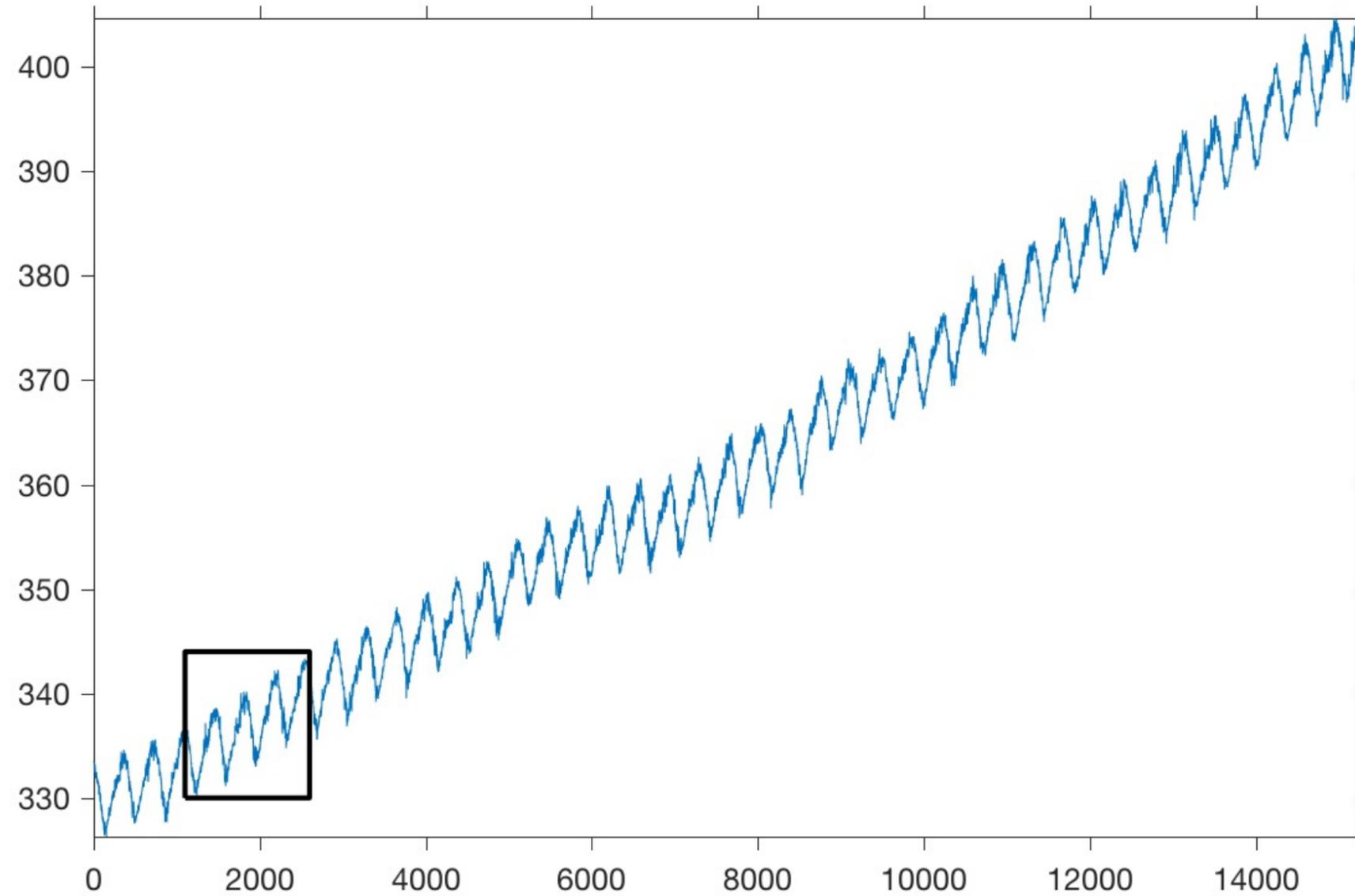
1. The data consists of two time series that are similar in character.
2. Both time series present a superposition of scales and a high degree of roughness.
3. The data seems to consist of different time periods with distinct statistical characteristics—the data is *nonstationary*.
4. Zooming in to one particular period show regular oscillations of roughly uniform amplitude and frequency.
5. The phasing of these show a circular polarization orbited in a counterclockwise direction.
6. The zoomed-in plot shows a fair amount of what appears to be measurement noise superimposed on the oscillatory signal.

This is a record of velocities of a single surface drifter at 6-hour intervals. All Surface drifter data are freely available from the Data Assembly Center of the Global Drifter Program
{www.aoml.noaa.gov/phod/dac/}.

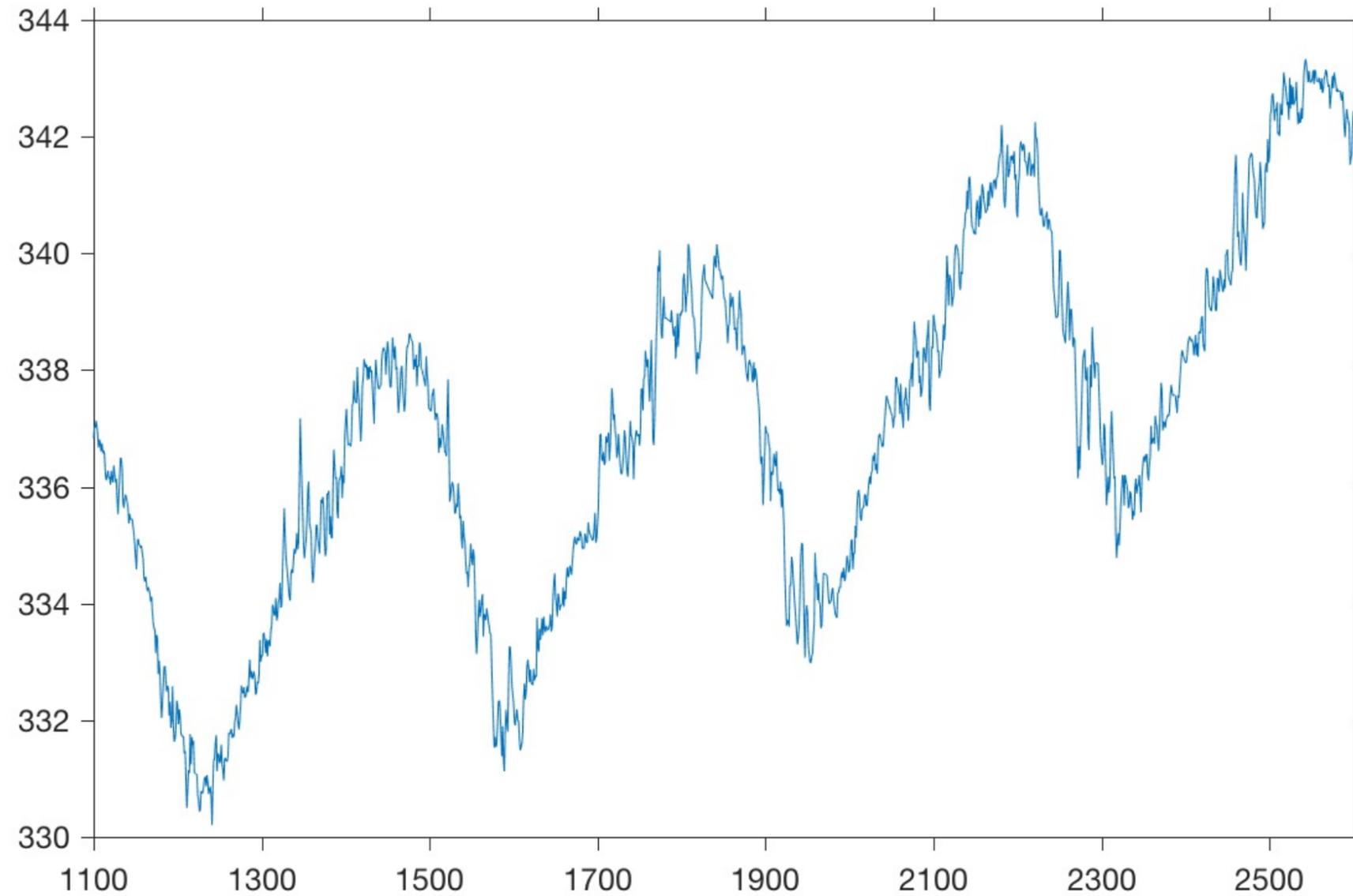


Second Example

We have already encountered this time series ...



Second Example



Observable Features

1. The data exhibit a very strong positive trend, roughly linear with time. Thus, this time series does not present a **mean statistics** that represents a "typical" value.
2. On top of the trend there seems to be a sinusoid-like oscillation that does not appear to change with time.
3. The zoomed-in plot shows noise superimposed on the sinusoidal and trend processes.



Observable Features

1. The data exhibit a very strong positive trend, roughly linear with time. Thus, this time series does not present a **mean statistics** that represents a "typical" value.
2. On top of the trend there seems to be a sinusoid-like oscillation that does not appear to change with time.
3. The zoomed-in plot shows noise superimposed on the sinusoidal and trend processes.

This is a record of daily atmospheric CO₂ measured at Mauna Loa in Hawaii at an altitude of 3400 m. Data from Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (scrippsco2.ucsd.edu).

We will investigate again this time series during the practical session this afternoon, this time using a spectral analysis approach.



Non-stationarity

The sample statistics may be changing with time because the underlying process (that is its statistics) is changing with time. The process is said to be “non-stationary”.

Sometimes we need to re-think our model for the underlying process. As we say in Lecture 3, we can hypothesize that the process $x(t)$ is the sum of an unknown process $y(t)$, plus a linear trend a , plus noise:

$$x(t) = y(t) + at + \epsilon(t),$$



Non-stationarity

The sample statistics may be changing with time because the underlying process (that is its statistics) is changing with time. The process is said to be “non-stationary”.

Sometimes we need to re-think our model for the underlying process. As we say in Lecture 3, we can hypothesize that the process $x(t)$ is the sum of an unknown process $y(t)$, plus a linear trend a , plus noise:

$$x(t) = y(t) + at + \epsilon(t),$$

or maybe the trend is better described as being quadratic with time because of an acceleration:

$$x(t) = y(t) + bt^2 + at + \epsilon(t).$$



Non-stationarity

The sample statistics may be changing with time because the underlying process (that is its statistics) is changing with time. The process is said to be “non-stationary”.

Sometimes we need to re-think our model for the underlying process. As we say in Lecture 3, we can hypothesize that the process $x(t)$ is the sum of an unknown process $y(t)$, plus a linear trend a , plus noise:

$$x(t) = y(t) + at + \epsilon(t),$$

or maybe the trend is better described as being quadratic with time because of an acceleration:

$$x(t) = y(t) + bt^2 + at + \epsilon(t).$$

The goal is then to estimate the unknowns a, b , which consists of methods of analyses generally called “parametric”. It is a bit like analyzing the data in terms of its statistics (with no prior expectations) or assuming a form for the data.



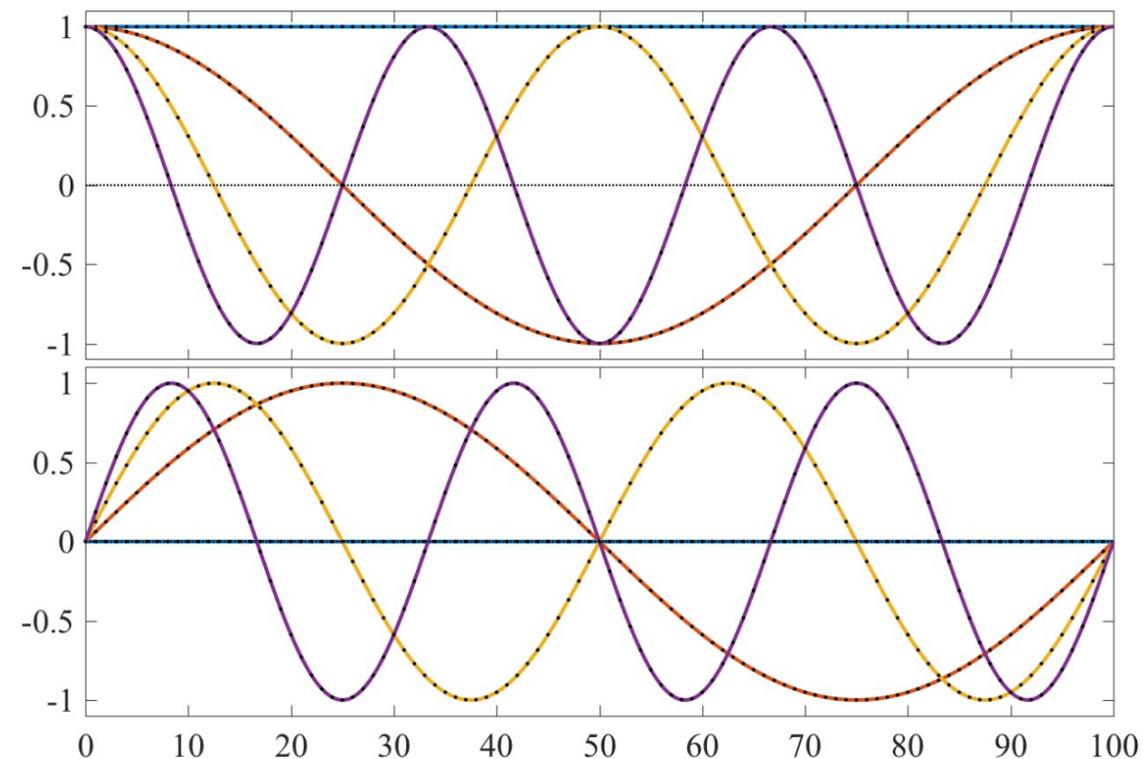
3. Fourier Spectral Analysis



Complex Fourier Series

It is possible to represent a discrete time series x_n as a sum of complex exponentials, a complex Fourier series:

$$x_n = \frac{1}{N\Delta t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N}, \quad n = 0, 1, \dots, N-1$$



We leave out for now how to obtain the complex coefficients X_m ...



About Frequency

You will typically find two frequency notations:

$$\cos(2\pi ft) \quad \text{or} \quad \cos(\omega t)$$

f is called the *cyclic* frequency. Its units are cycles/time. Example: Hz = cycles/sec.

$\omega = 2\pi f$ is called the *radian* or *angular* frequency. Its units are rad/time. The associated *period* of oscillation is $P = 1/f = 2\pi/\omega$.

As t goes from 0 to $1/f = 2\pi/\omega = P$, $2\pi ft$ goes from 0 to 2π and ωt goes from 0 to 2π .

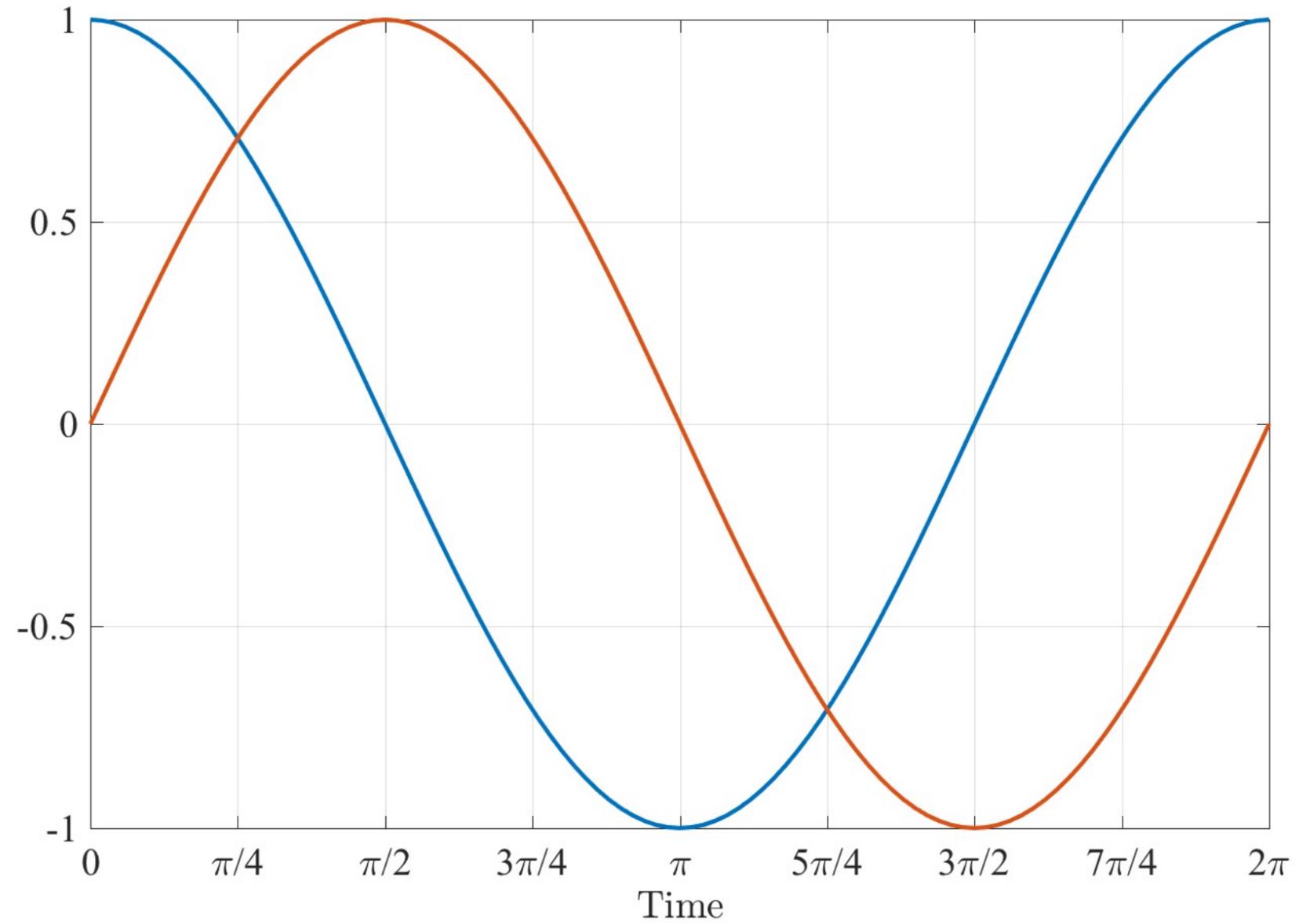
A very common error in Fourier analysis is mixing up cyclic and radian frequencies!

Note: neither “cycles” nor “radians” actually have any units, thus both f and ω have units of 1/time. However, specifying for example 'cycles per day' or 'radians per day' helps to avoid confusion.



Review: Sinusoids

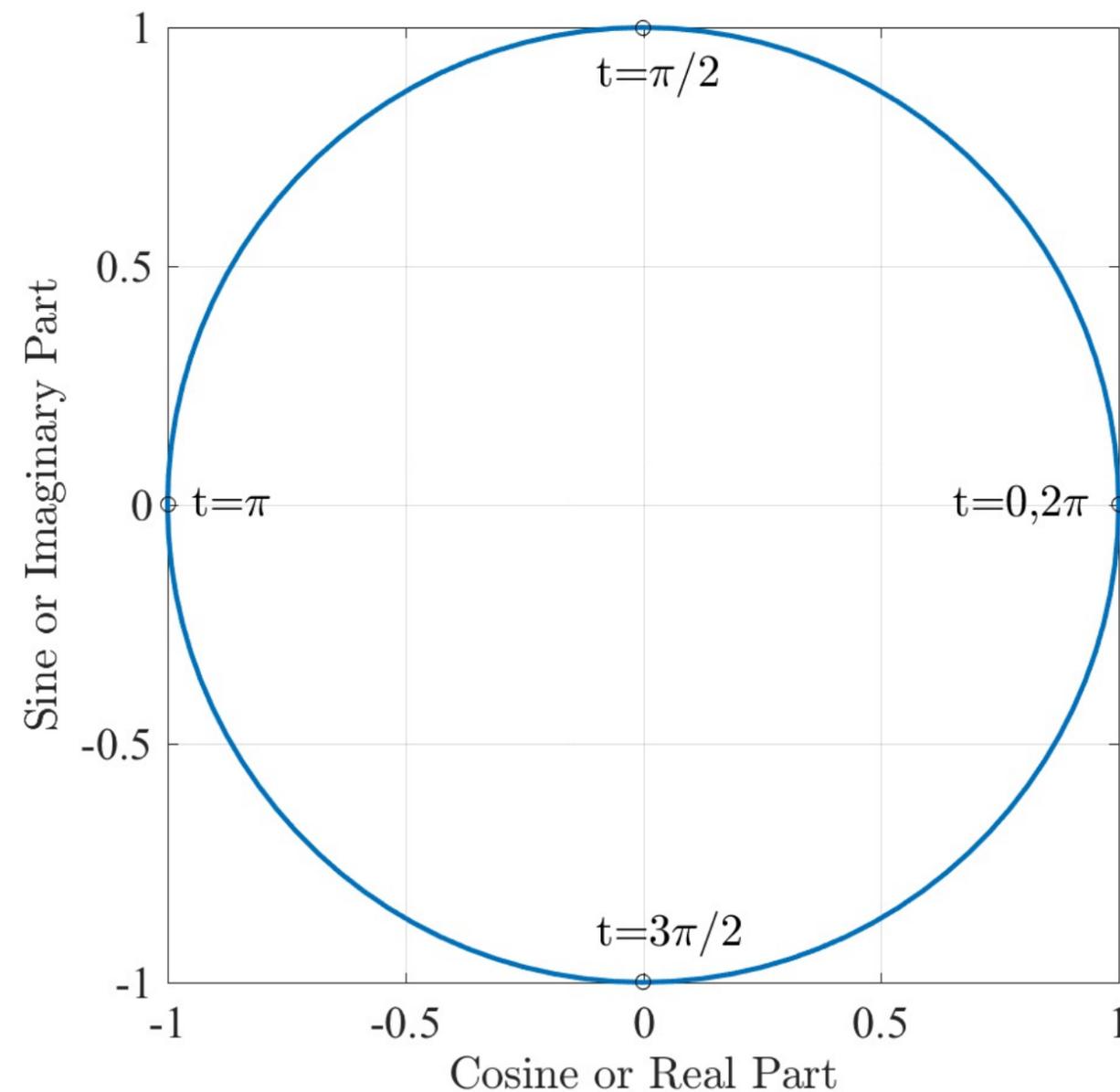
Cosine function (blue) and sine function (orange)



Complex Exponentials, 2D

Now consider a plot $\cos(t)$ vs. $\sin(t)$.

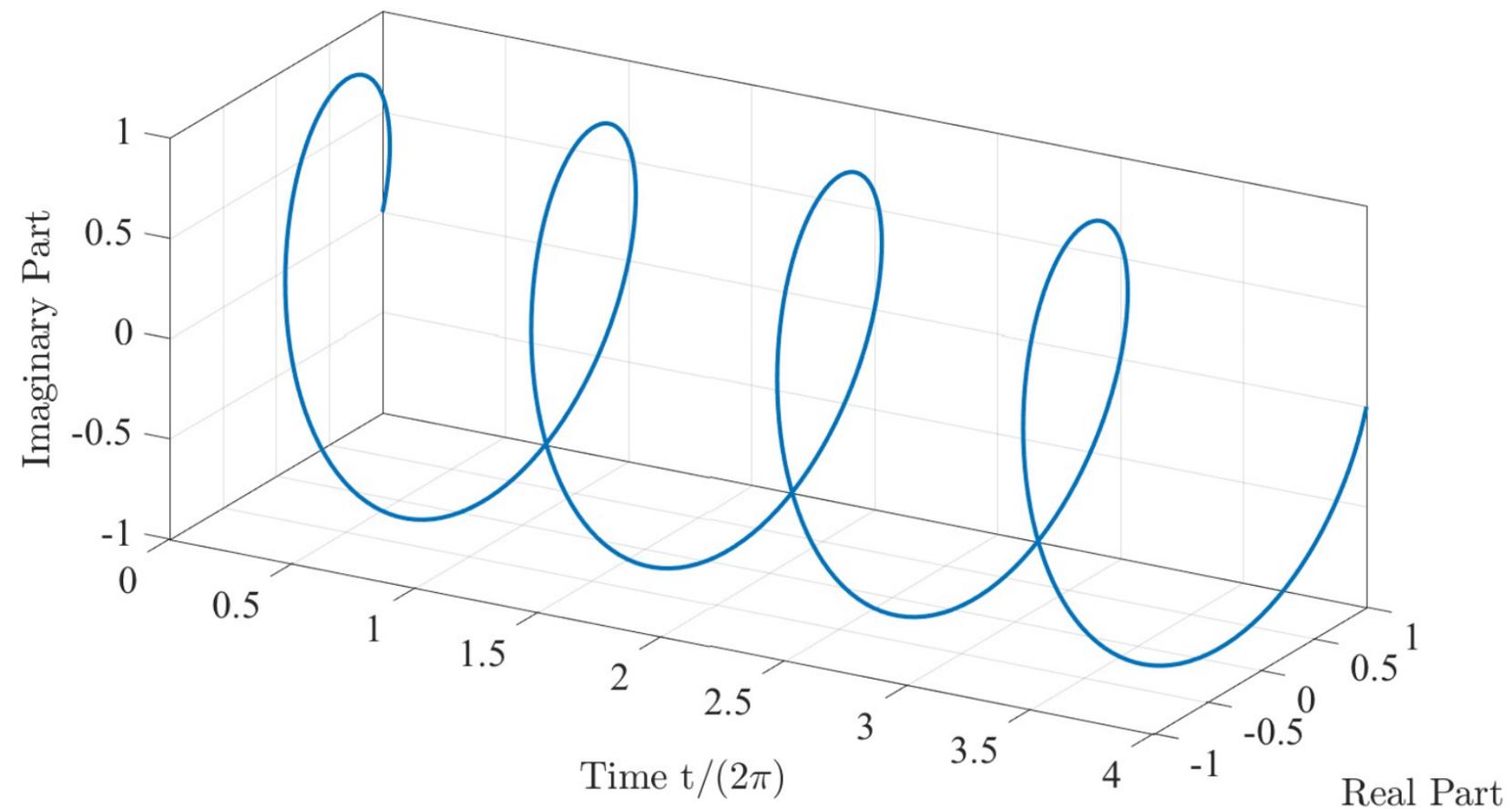
That's the same as $\cos(t) + i \sin(t) = e^{it}$.



Complex Exponentials, 3D

This is better seen in 3D as a *spiral* as time increases.

$$\cos(t) + i \sin(t) = e^{it}$$



The complex Fourier series

The discrete time series x_n can be written as a sum of complex exponentials:

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N} = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi n\Delta_t \cdot (m/N\Delta_t)}, \quad n = 0, 1, \dots$$



The complex Fourier series

The discrete time series x_n can be written as a sum of complex exponentials:

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N} = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi n\Delta_t \cdot (m/N\Delta_t)}, \quad n = 0, 1, \dots$$

The m th term behaves as

$$e^{i2\pi f_m n\Delta_t} = \cos(2\pi f_m n\Delta_t) + i \sin(2\pi f_m n\Delta_t), \text{ where} \\ f_m \equiv m/N\Delta_t.$$

Note that in the literature, Δ_t is often set to one, and thus omitted, leading to a lot of confusion (including for me!).

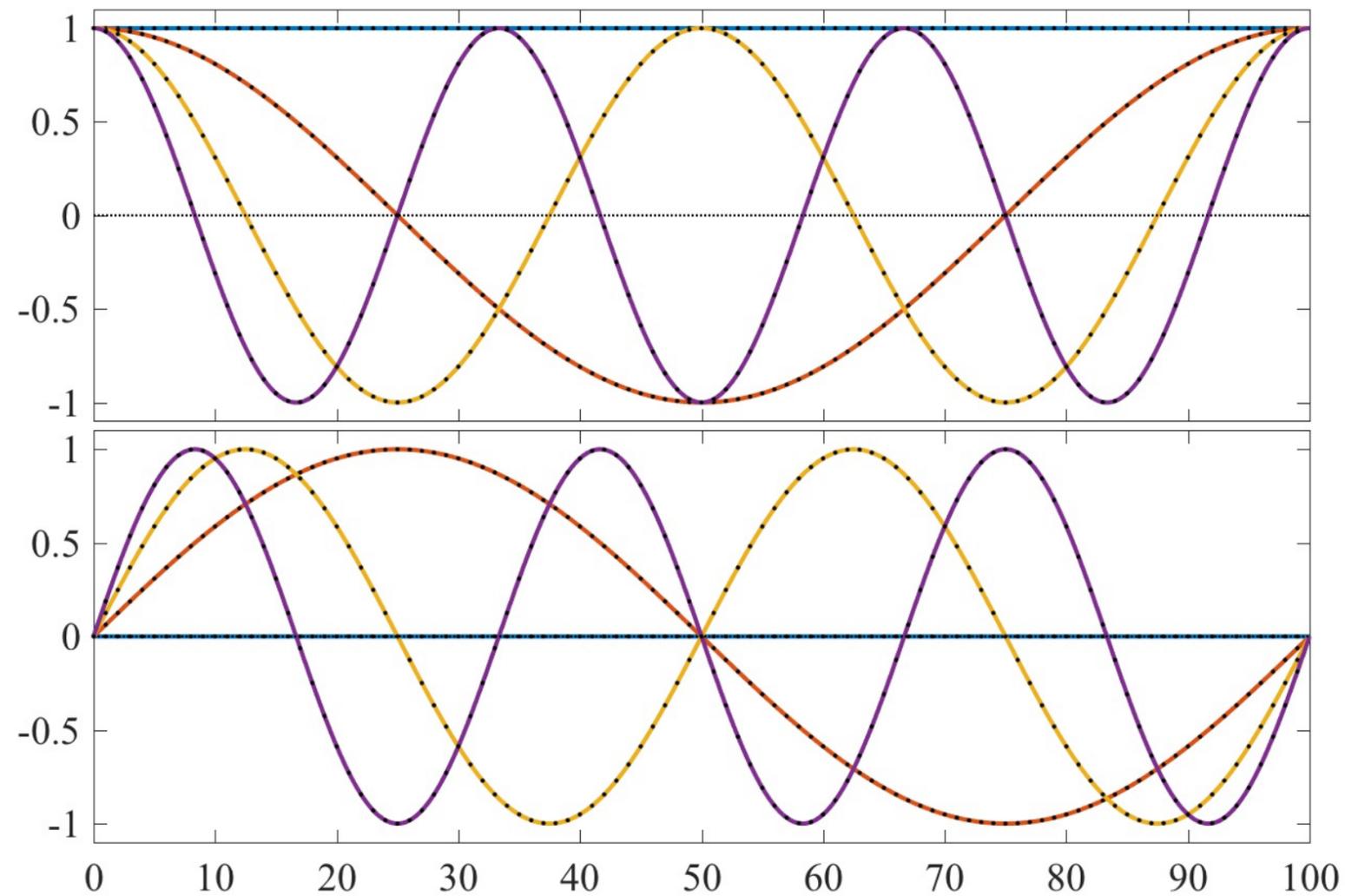
The quantity $f_m \equiv m/N\Delta_t$ is called the m th *Fourier frequency*. The *period* associated with f_m is $1/f_m = \Delta_t N/m$. Thus m tells us the *number of oscillations* contained in the length $N\Delta_t$ time series.



Continuous Time

$$\cos(2\pi f_m t) \text{ and } \sin(2\pi f_m t)$$

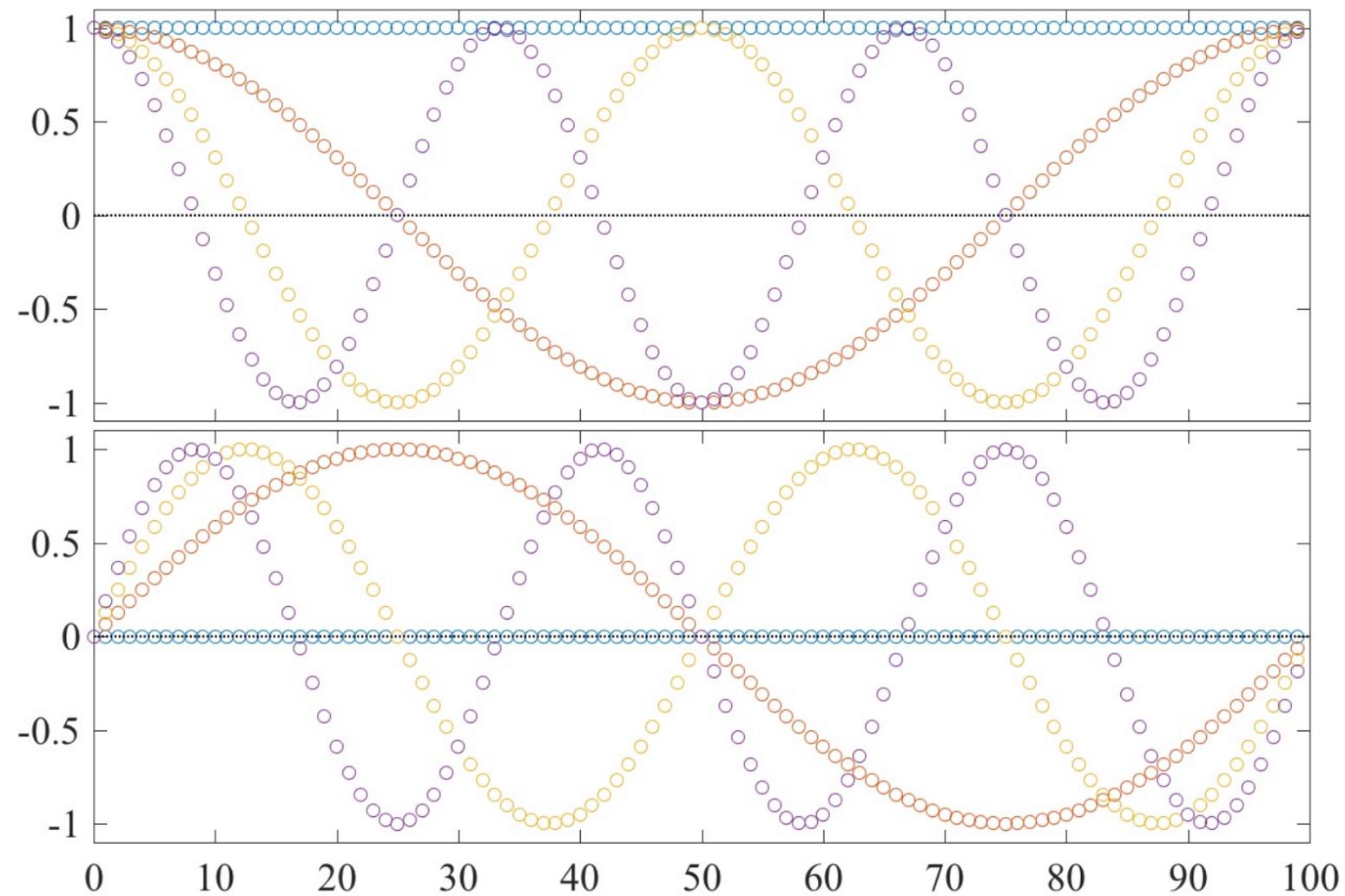
$$f_m = 0, 1/100, 2/100, 3/100 \quad t = [0 \dots 100]$$



Discrete Time

$$\cos(2\pi f_m n \Delta_t) \text{ and } \sin(2\pi f_m n \Delta_t)$$

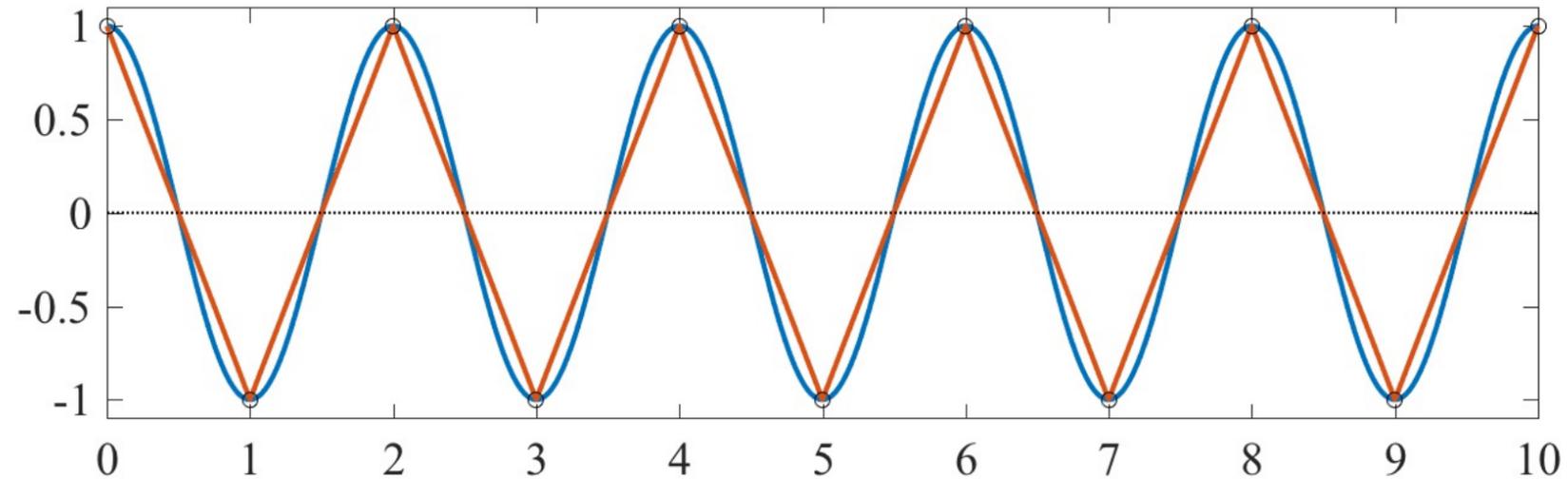
$$f_m = 0, 1/100, 2/100, 3/100 \quad n = 0, 1, 2, \dots, 99 \quad \Delta_t = 1$$



The Nyquist Frequency

The single most important frequency is the *highest resolvable* frequency, the *Nyquist frequency*.

$$f^{\mathcal{N}} \equiv \frac{1}{2\Delta_t} = \frac{1}{2} \cdot \frac{1}{\Delta_t} \quad \omega^{\mathcal{N}} \equiv \frac{1}{2} \cdot \frac{2\pi}{\Delta_t} = \frac{\pi}{\Delta_t}$$



The highest resolvable frequency is *half* the *sampling rate* or *one cycle per two sampling intervals*.

$$e^{i2\pi f^{\mathcal{N}} n\Delta_t} = e^{i2\pi \cdot 1/(2\Delta_t) \cdot n\Delta_t} = e^{i\pi n} = (-1)^n = 1, -1, 1, -1, \dots$$

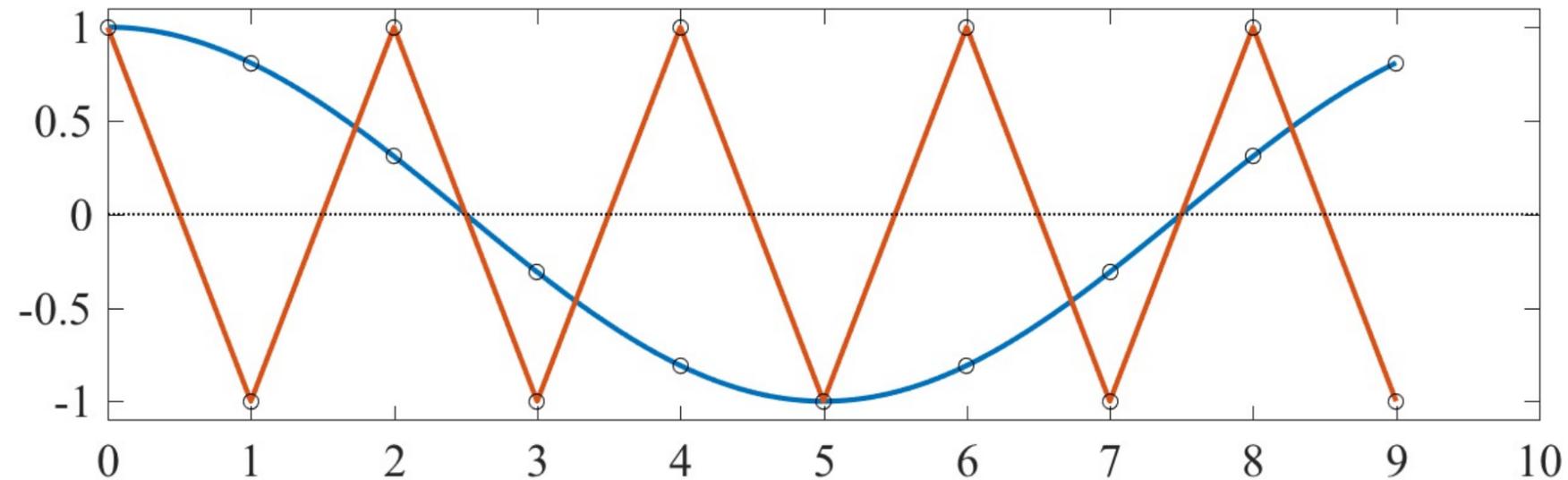
Note that there is no “sine” component at Nyquist in the Fourier series!



The Rayleigh Frequency

The second most important frequency is the *lowest resolvable* frequency, the *Rayleigh frequency*.

$$f^{\mathcal{R}} \equiv \frac{1}{N\Delta_t} \quad \omega^{\mathcal{R}} \equiv \frac{2\pi}{N\Delta_t}$$



The lowest resolvable frequency is one cycle over the *entire record*. Here the sample interval $\Delta_t = 1$ and the number of points is $N = 10$.



Importance of Rayleigh

The Rayleigh frequency $f^{\mathcal{R}}$ is important because it gives the *spacing between* the Fourier frequencies:

$$f_0 = 0, f_1 = \frac{1}{N\Delta_t}, f_2 = \frac{2}{N\Delta_t}, \dots \quad f_n = n f^{\mathcal{R}}, \quad f^{\mathcal{R}} = \frac{1}{N\Delta_t}$$

Thus, it controls the *frequency-domain resolution*. If you want to distinguish between two closely spaced peaks, you need the dataset *duration* to be sufficiently *large* so that the Rayleigh frequency is sufficiently *small*.



Importance of Rayleigh

The Rayleigh frequency $f^{\mathcal{R}}$ is important because it gives the *spacing between* the Fourier frequencies:

$$f_0 = 0, f_1 = \frac{1}{N\Delta_t}, f_2 = \frac{2}{N\Delta_t}, \dots \quad f_n = n f^{\mathcal{R}}, \quad f^{\mathcal{R}} = \frac{1}{N\Delta_t}$$

Thus, it controls the *frequency-domain resolution*. If you want to distinguish between two closely spaced peaks, you need the dataset *duration* to be sufficiently *large* so that the Rayleigh frequency is sufficiently *small*.

As an example, the two principal semi-diurnal tidal “species” have period of 12 h (M_2) and 12.4206012 h (S_2). The minimum record length to distinguish the two frequencies is thus

$$N\Delta_t = \frac{1}{f^{\mathcal{R}}} = \frac{1}{f^{S_2} - f^{M_2}} = \frac{1}{\frac{1}{12} - \frac{1}{12.4206012}} = 354.36 \text{ hours.}$$



Rayleigh and Nyquist frequencies

The ratio of the Rayleigh to Nyquist frequencies tells you how many *different frequencies* you can resolve.

$$\frac{f^{\mathcal{N}}}{f^{\mathcal{R}}} = \frac{N\Delta_t}{2\Delta_t} = \frac{N}{2}$$

So why do we have N frequencies in the sum for the complex Fourier series?

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N}$$



The Fourier Frequencies

The first few Fourier frequencies $f_m = m/(N\Delta_t)$ are:

$$f_0 = \frac{0}{N\Delta_t}, \quad f_1 = \frac{1}{N\Delta_t}, \quad f_2 = \frac{2}{N\Delta_t}, \dots$$

while the last few are

$$\dots, f_{N-2} = \frac{N-2}{N\Delta_t} = \frac{1}{\Delta_t} - \frac{2}{N\Delta_t}, \quad f_{N-1} = \frac{N-1}{N\Delta_t} = \frac{1}{\Delta_t} - \frac{1}{N\Delta_t}.$$

But notice that the last Fourier exponential term is

$$e^{i2\pi f_{N-1} n \Delta_t} = e^{i2\pi(N-1)n/N} = e^{i2\pi n} e^{-i2\pi n/N} = e^{-i2\pi n/N} = e^{-i2\pi f_1 n \Delta_t}$$

because $e^{i2\pi n} = 1$ for all integers n ! Frequencies higher than the Nyquist cannot appear due to our sample rate. Therefore, these terms instead specify terms that have a frequency *less than* the Nyquist but that rotate in the *negative* direction.



The Fourier Frequencies

In the vicinity of $m = N/2$, for even N , we have

$$f_{N/2-1} = \frac{1}{2\Delta_t} - \frac{1}{N\Delta_t}, \quad f_{N/2} = \frac{1}{2\Delta_t}, \quad f_{N/2+1} = \frac{1}{2\Delta_t} + \frac{1}{N\Delta_t}, \dots$$

but actually the first frequency higher than the Nyquist is the *highest negative frequency*:

$$f_{N/2-1} = \frac{1}{2\Delta_t} - \frac{1}{N\Delta_t}, \quad f_{N/2} = \frac{1}{2\Delta_t}, \dots$$

$$f_{N/2+1} = -f_{N/2-1} = -\left(\frac{1}{2\Delta_t} - \frac{1}{N\Delta_t}\right), \dots$$

Thus the positive frequencies and negative frequencies are both increasing *toward the middle* of the Fourier transform array.

For this reason Matlab provides `fftshift`, to shift the zero frequency, *not* the Nyquist, to be in the middle of the array.



One-Sided vs. Two-Sided

There exists two strictly equivalent representations, *two-sided* and *one-sided*, of the discrete Fourier transform:

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N},$$

$$x_n = \frac{1}{N\Delta_t} X_0 + \frac{2}{N\Delta_t} \sum_{m=1}^{N/2-1} A_m \cos(2\pi mn/N + \Phi_m) + X_{N/2} (-1)^n,$$

where A_m and Φ_m are an *amplitude* and *phase*, with
 $X_m = A_m e^{i\Phi_m}$.

The two-sided representation is more compact mathematically.

For real-valued x_n , the one-sided representation is more intuitive as it expresses x_n as a sum of *phase-shifted* cosinusoids.



One-Sided vs. Two-Sided

There exists two strictly equivalent representations, *two-sided* and *one-sided*, of the discrete Fourier transform:

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N},$$

$$x_n = \frac{1}{N\Delta_t} X_0 + \frac{2}{N\Delta_t} \sum_{m=1}^{N/2-1} A_m \cos(2\pi mn/N + \Phi_m) + X_{N/2} (-1)^n,$$

where A_m and Φ_m are an *amplitude* and *phase*, with $X_m = A_m e^{i\Phi_m}$.

The two-sided representation is more compact mathematically.

For real-valued x_n , the one-sided representation is more intuitive as it expresses x_n as a sum of *phase-shifted* cosinusoids. A price of the one-sided form is that even and odd N are somewhat different! The expression above is for even-valued N .



The Forward DFT

So? How do we know the values of the Fourier coefficients X_m ? It can be shown that:

$$X_m = \Delta_t \sum_{n=0}^{N-1} x_n e^{-i2\pi mn/N}$$

This is called the *discrete Fourier transform* of x_n .

The DFT *transforms* x_n from the time domain to the frequency domain. The DFT *defines* a sequence of N complex-valued numbers, X_m , for $m = 0, 1, 2, \dots, N - 1$, which are termed the *Fourier coefficients*.

In Matlab, the discrete Fourier transform defined above is computed by $\text{fft}(\mathbf{x}) \times \Delta_t$.



The Inverse DFT

In fact,

$$x_n \equiv \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N}$$

is called the *inverse discrete Fourier transform*. It expresses how x_n may be *constructed* using the Fourier coefficients multiplying complex exponentials—or, as we saw earlier, phase-shifted cosinusoids.



The Spectrum

One of several definitions of the *spectrum*, or *spectral density function*, at frequency f_m , is:

$$S(f_m) \equiv \lim_{N \rightarrow \infty} E \left\{ \frac{|X_m|^2}{N} \right\}.$$

$E\{\cdot\}$ is called the *expectation*, it is a conceptual “average” over a statistical ensemble, and it cannot be obtained in practice.



The Spectrum

One of several definitions of the *spectrum*, or *spectral density function*, at frequency f_m , is:

$$S(f_m) \equiv \lim_{N \rightarrow \infty} E \left\{ \frac{|X_m|^2}{N} \right\}.$$

$E\{\cdot\}$ is called the *expectation*, it is a conceptual “average” over a statistical ensemble, and it cannot be obtained in practice.

Formally, the function S is defined for all frequencies f , not only f_m , but as $N \rightarrow \infty$, the Rayleigh frequency $1/N \Delta_t$ becomes infinitesimally small, and all frequencies are obtained.



The Spectrum

One of several definitions of the *spectrum*, or *spectral density function*, at frequency f_m , is:

$$S(f_m) \equiv \lim_{N \rightarrow \infty} E \left\{ \frac{|X_m|^2}{N} \right\}.$$

$E\{\cdot\}$ is called the *expectation*, it is a conceptual “average” over a statistical ensemble, and it cannot be obtained in practice.

Formally, the function S is defined for all frequencies f , not only f_m , but as $N \rightarrow \infty$, the Rayleigh frequency $1/N \Delta_t$ becomes infinitesimally small, and all frequencies are obtained.

However, $N \rightarrow \infty$ is not achievable ... Therefore, one aspect of spectral analysis is to find an acceptable **estimate** of the true, unknown, spectrum $S(f)$ of the process $x(t)$.



The Parseval Theorem

A very important theorem is *Parseval's theorem* which takes the following form for the discrete case:

$$\Delta_t \sum_{n=0}^{N-1} |x_n|^2 = \frac{1}{N \Delta_t} \sum_{m=0}^{N-1} |X_m|^2.$$

When $\hat{\mu}_x = 0$, this theorem shows that the total **variance** of x_n is recoverable from the sum of absolute Fourier coefficients squared.

Which can be interpreted as saying that the spectrum gives you the distribution of the variance as a function of frequency.



Spectral Estimates

The simplest way to estimate the *spectrum* $S(f)$ function of frequency f is to simply take the modulus squared of the Fourier transform,

$$\hat{S}(f_m) = \hat{S}_m \equiv \frac{1}{N} |X_m|^2, \quad m = 0, 1, 2, \dots, (N - 1).$$

This quantity is known as the *periodogram*.

Note that the Matlab `fft(x)` command assumes $\Delta_t = 1$ so you need to plot $\text{abs}(\text{fft}(x))^2 \times \Delta_t / N$.

As we shall see, the periodogram is *not* the spectrum! It is an *estimate* of the spectrum—and generally speaking, a very poor one.

It is also said to be the *naive* spectral estimate, meaning it is the spectral estimate that you get if you don't know that there is something better. Please do not use the periodogram in your publications.



The Multitaper Method

An alternate spectral estimate called the *multitaper* method. Here is a quick sketch of this method.

We form a set of K different sequences the same length as the data, that is, having N points in time. These sequences are chosen from a special family of functions that is closely related to familiar orthogonal functions, e.g. the Hermite functions.

These K different sequences are denoted as $\psi_n^{\{k\}}$ for $k = 1, 2, \dots, K$. For each of these sequence, we form a spectral estimate as

$$\hat{S}_m^{\{k\}} \equiv \left| \Delta_t \sum_{n=0}^{N-1} \psi_n x_n e^{-i2\pi mn/N} \right|^2, \quad n = 0, 1, 2, \dots, (N - 1).$$

which involves *multiplying* the data by the sequence $\psi_n^{\{k\}}$ *before* taking the Fourier transform.



The Multitaper Method

The action of multiplying the data by some sequence before Fourier transforming it, as in

$$\hat{S}_m^{\{k\}} \equiv \left| \Delta_t \sum_{n=0}^{N-1} \psi_n x_n e^{-i2\pi mn/N} \right|^2, \quad n = 0, 1, 2, \dots, (N - 1)$$

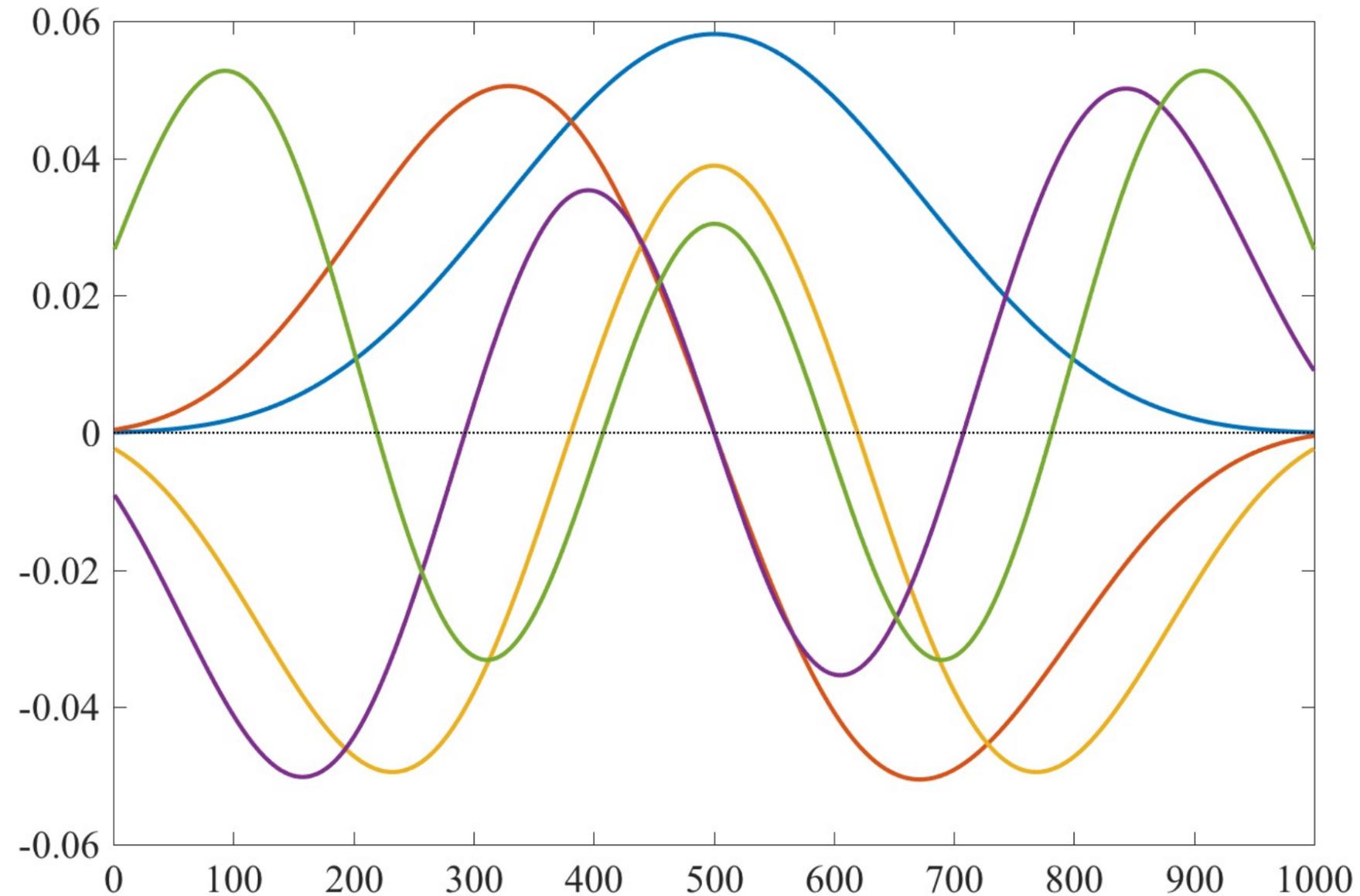
is called *tapering*. The goal is to reduce the **bias** (systematic error) of the spectral estimate. These K different individual estimates (aka *eigenspectra*), are combined into one *average* spectral estimate, in order to reduce the **variance** (random error) of the estimate

$$\hat{S}_m^{\psi} \equiv \frac{1}{K} \sum_{k=1}^K \hat{S}_m^{\{k\}}.$$

The multitaper method therefore involves two steps: (i) *tapering* the data, and (ii) *averaging* over multiple individual spectral estimates.



The Taper Functions



Here $K = 5$ Slepian tapers are shown. These are orthogonal functions that become more oscillatory for increasing K .



The Multitaper Method

The multitaper method controls the degrees of spectral *smoothing* and *averaging* through changing the properties of the tapers.

The multitaper method is generally the favorite spectral analysis method among those researchers who have thought the most about spectral analysis.

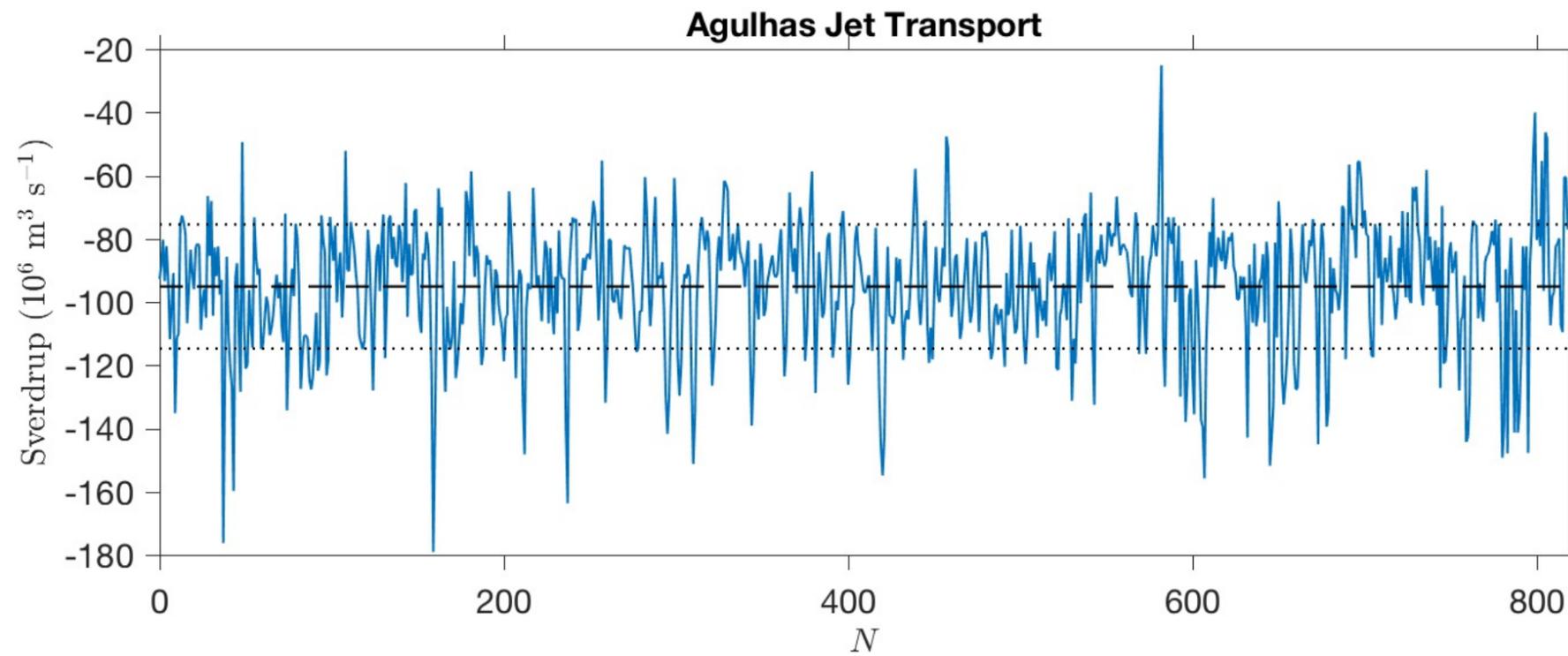
It is recommended because **(i)** it avoids the deficiencies of the periodogram, **(ii)** it has, in a certain sense, provable optimal properties, **(iii)** it is very easy to implement and adjust, **(iv)** it allows an estimate of the spectrum for the period equal to the length of your time series (no need to divide up your time series as for the Welch's method!).

See *Thomson (1982)*, *Park et al. (1987)*, and *Percival and Walden, Spectral Analysis for Physical Applications*.



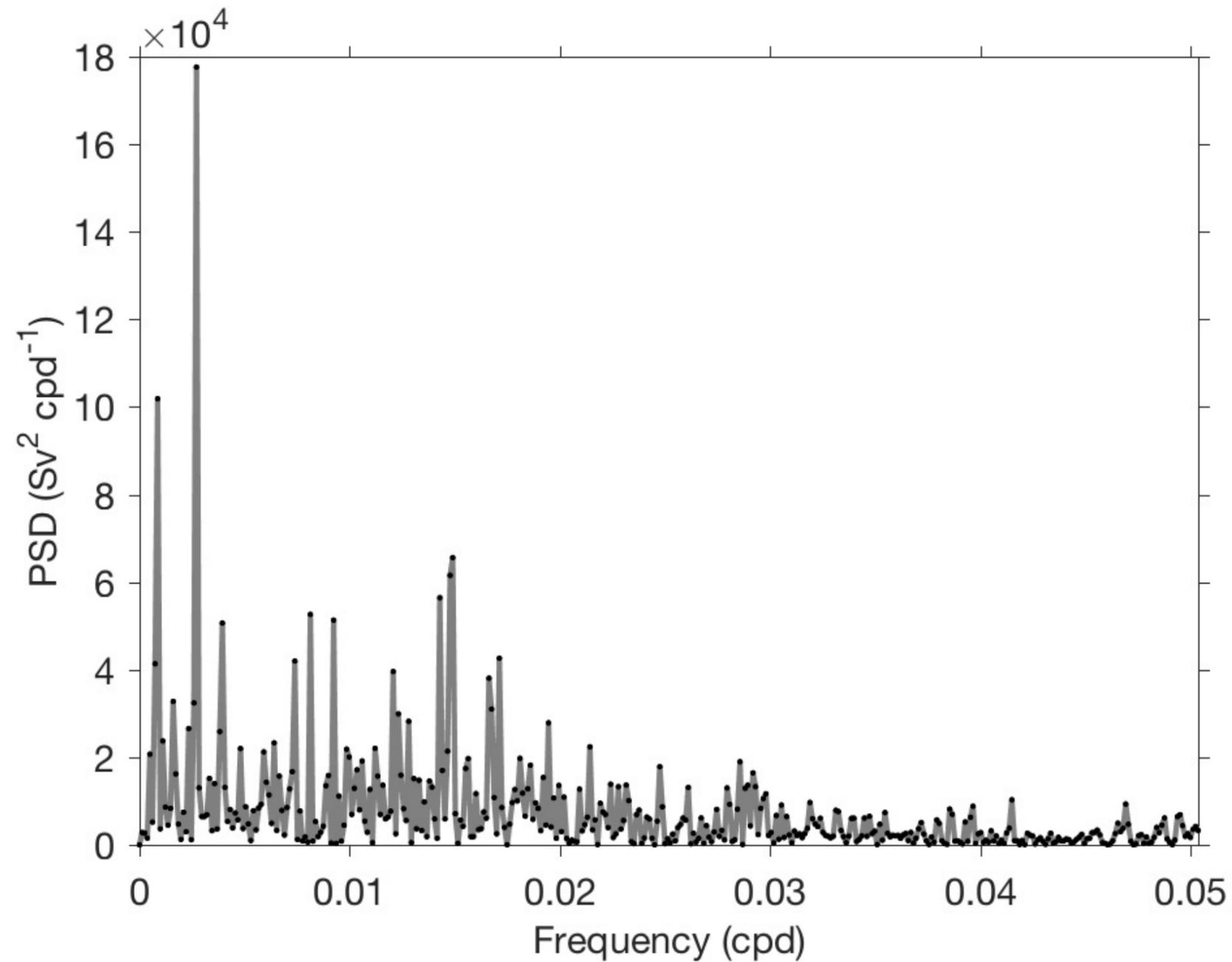
Example

Agulhas current boundary transport from *Beal, L. M. and S. Elipot, Broadening not strengthening of the Agulhas Current since the early 1990s, Nature, 540, 570573, doi:10.1038/nature19853*



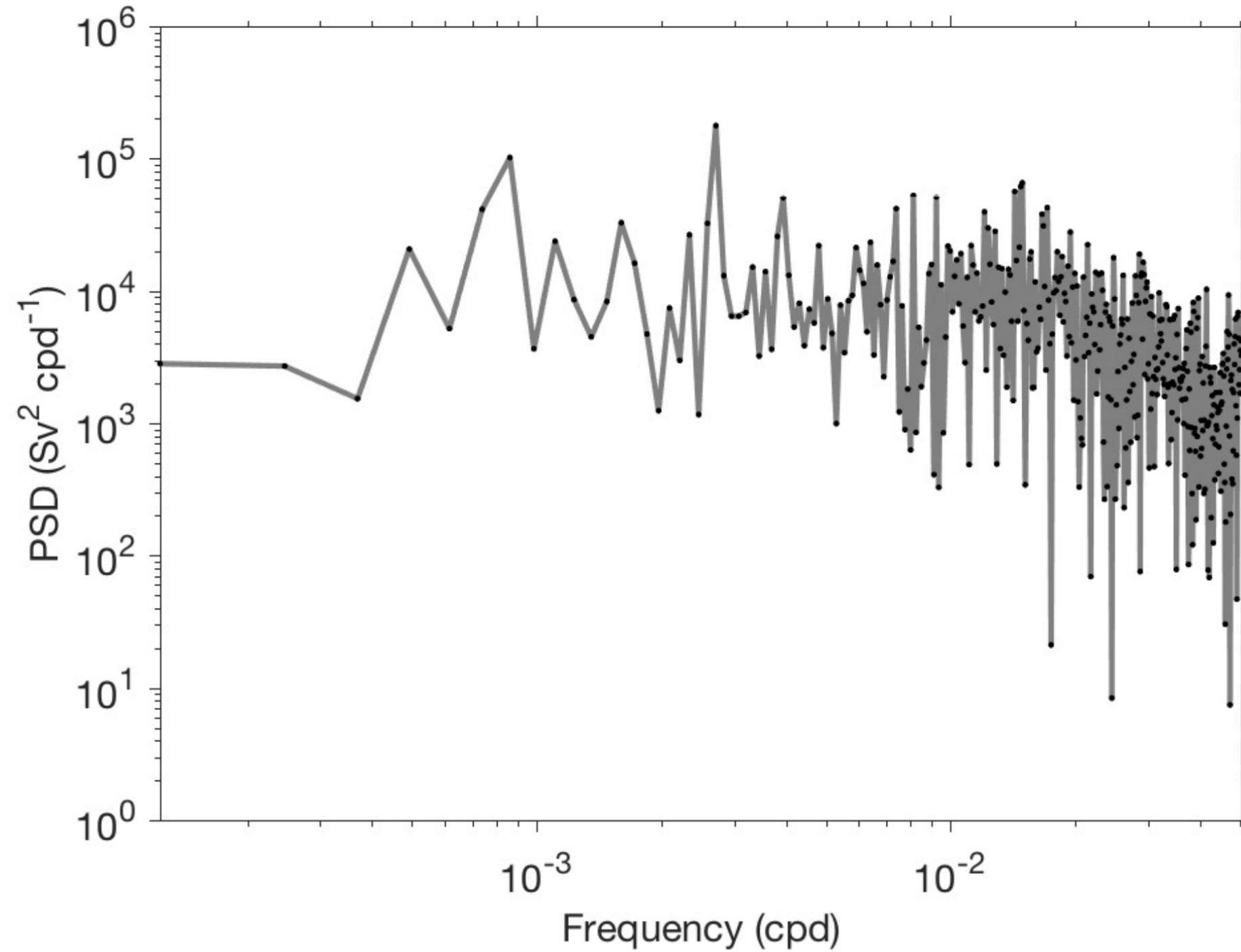
Example: periodogram

Linear plot



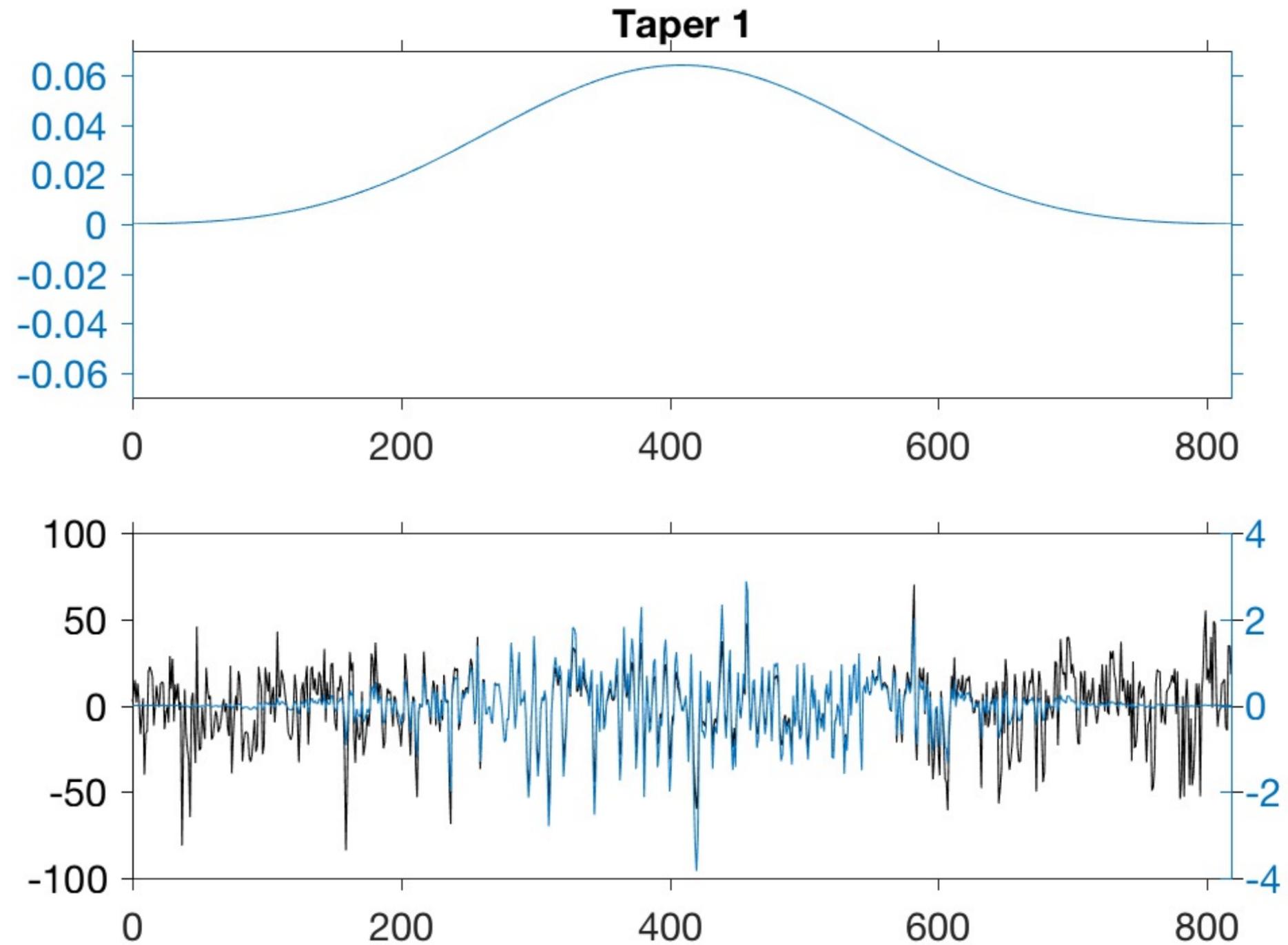
Example: periodogram

Log-log plot



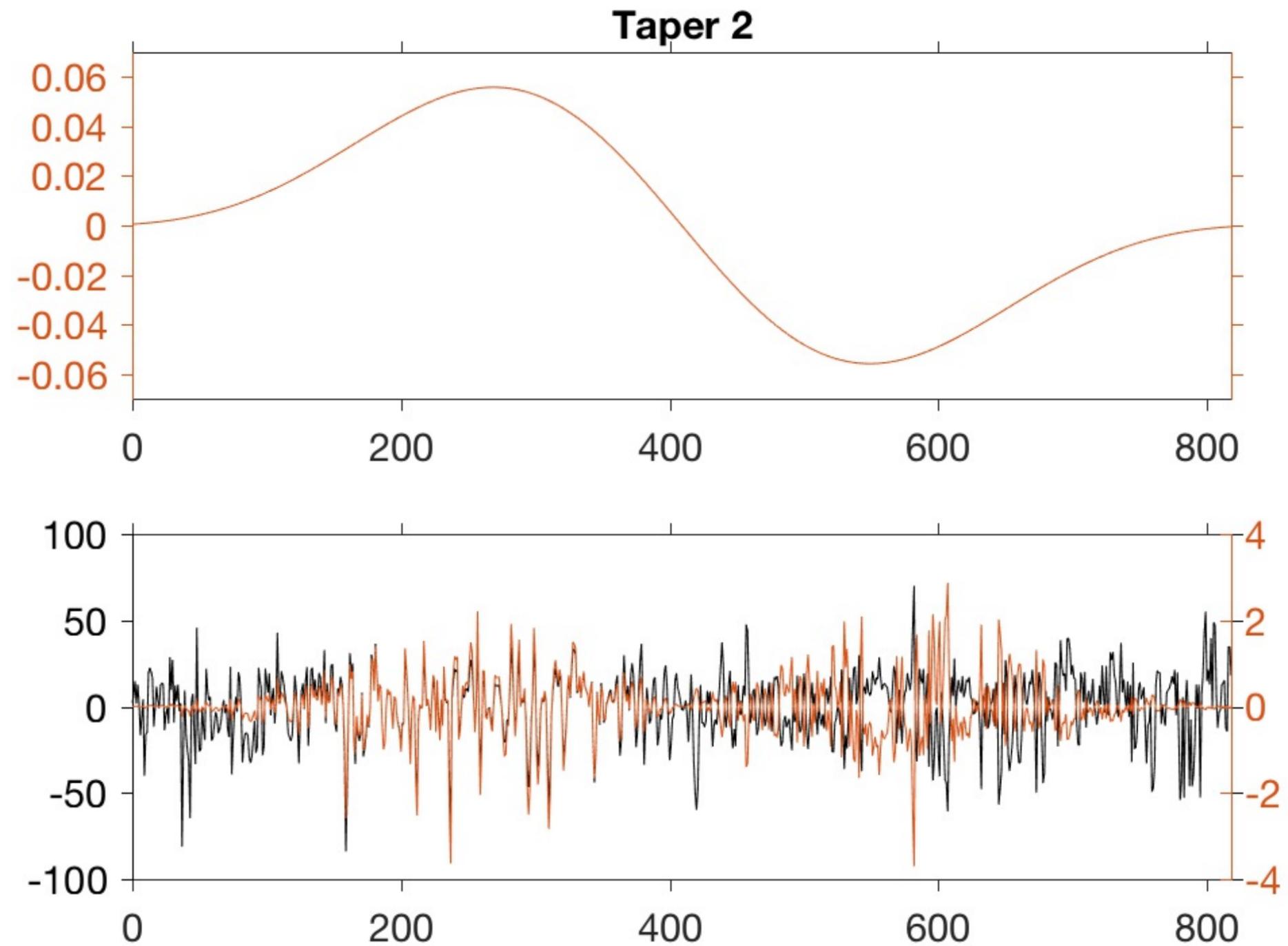
Example: multitaper

Effect of first taper



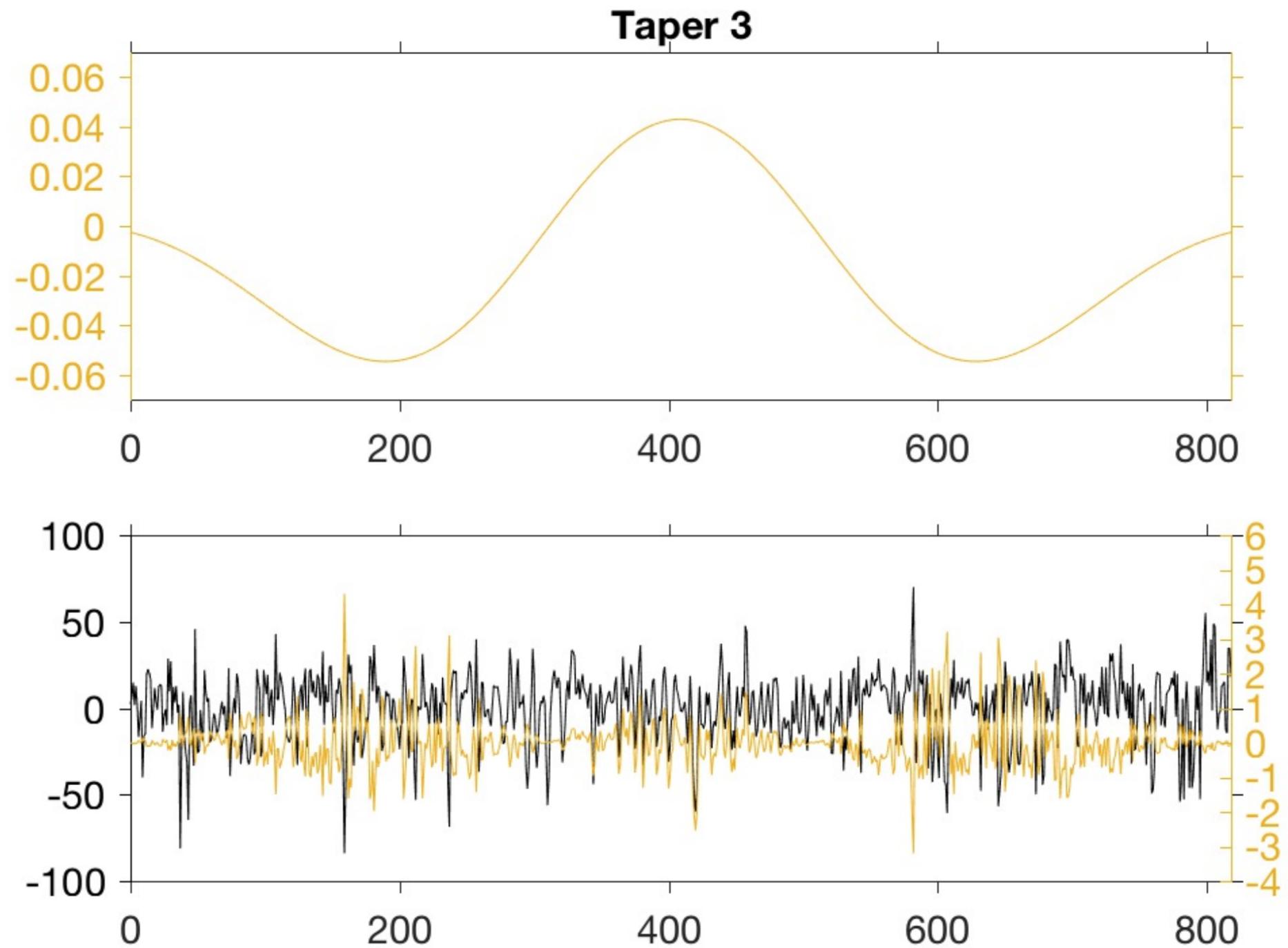
Example: multitaper

Second taper



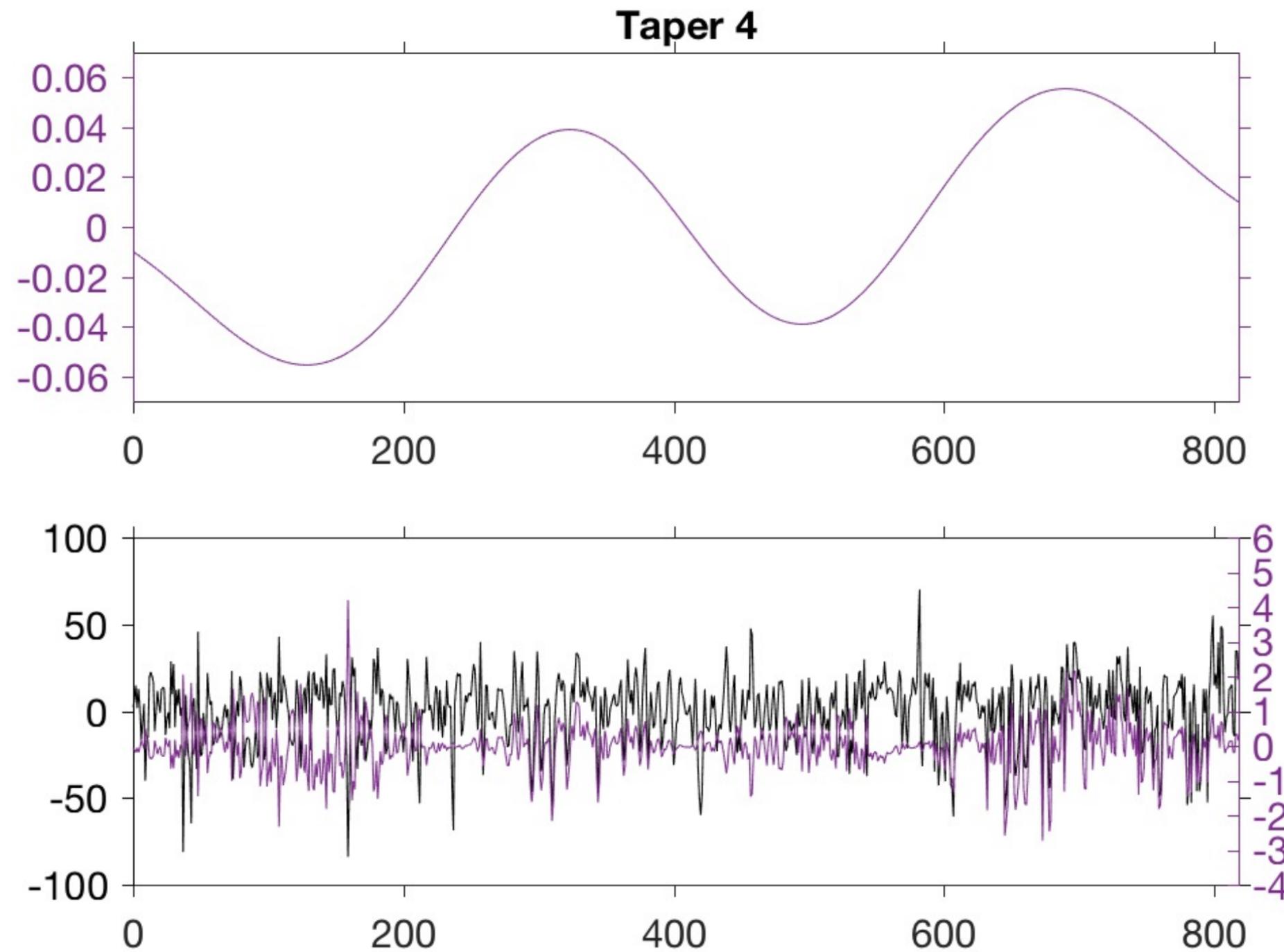
Example: multitaper

Third taper



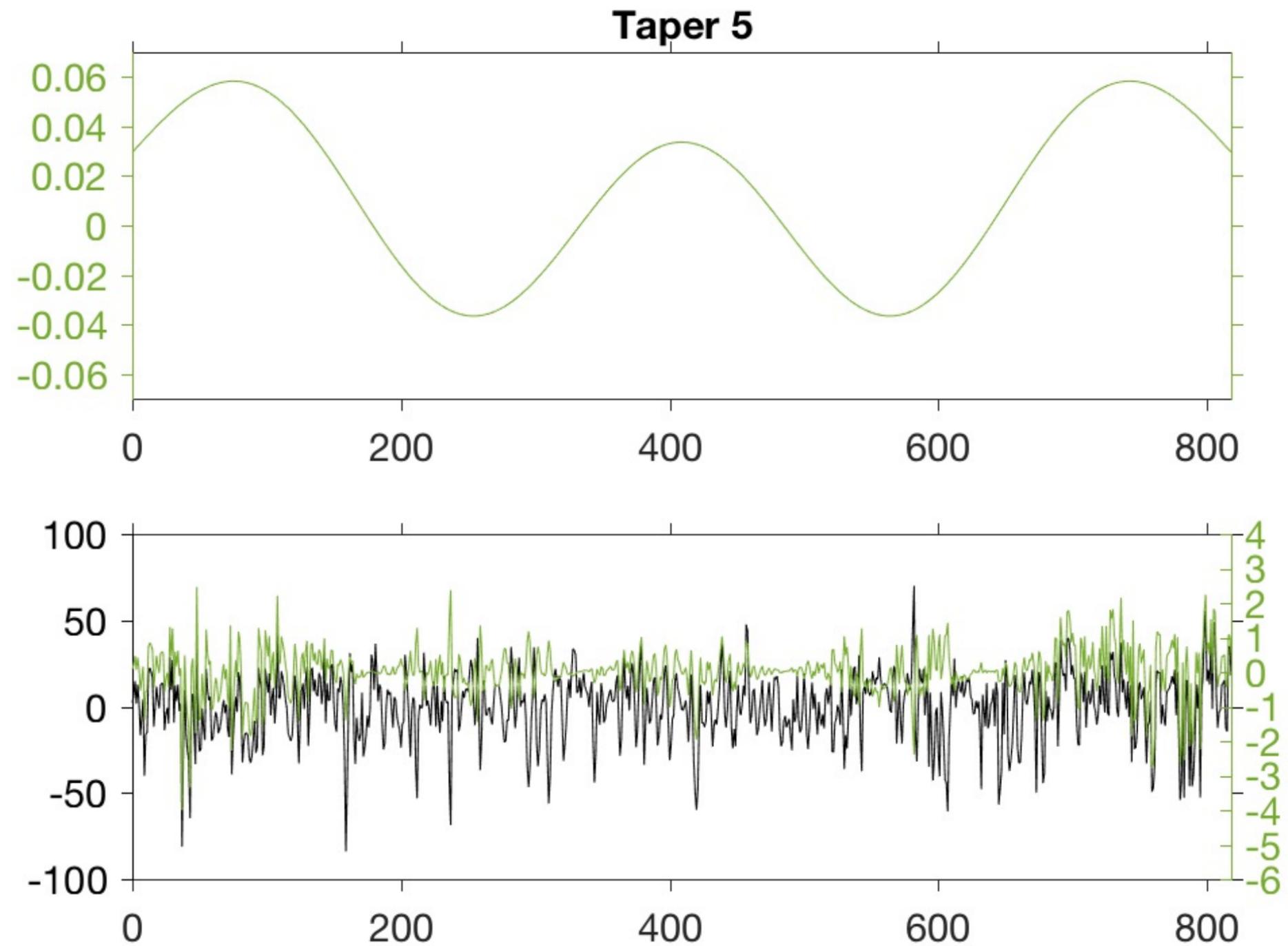
Example: multitaper

Fourth taper



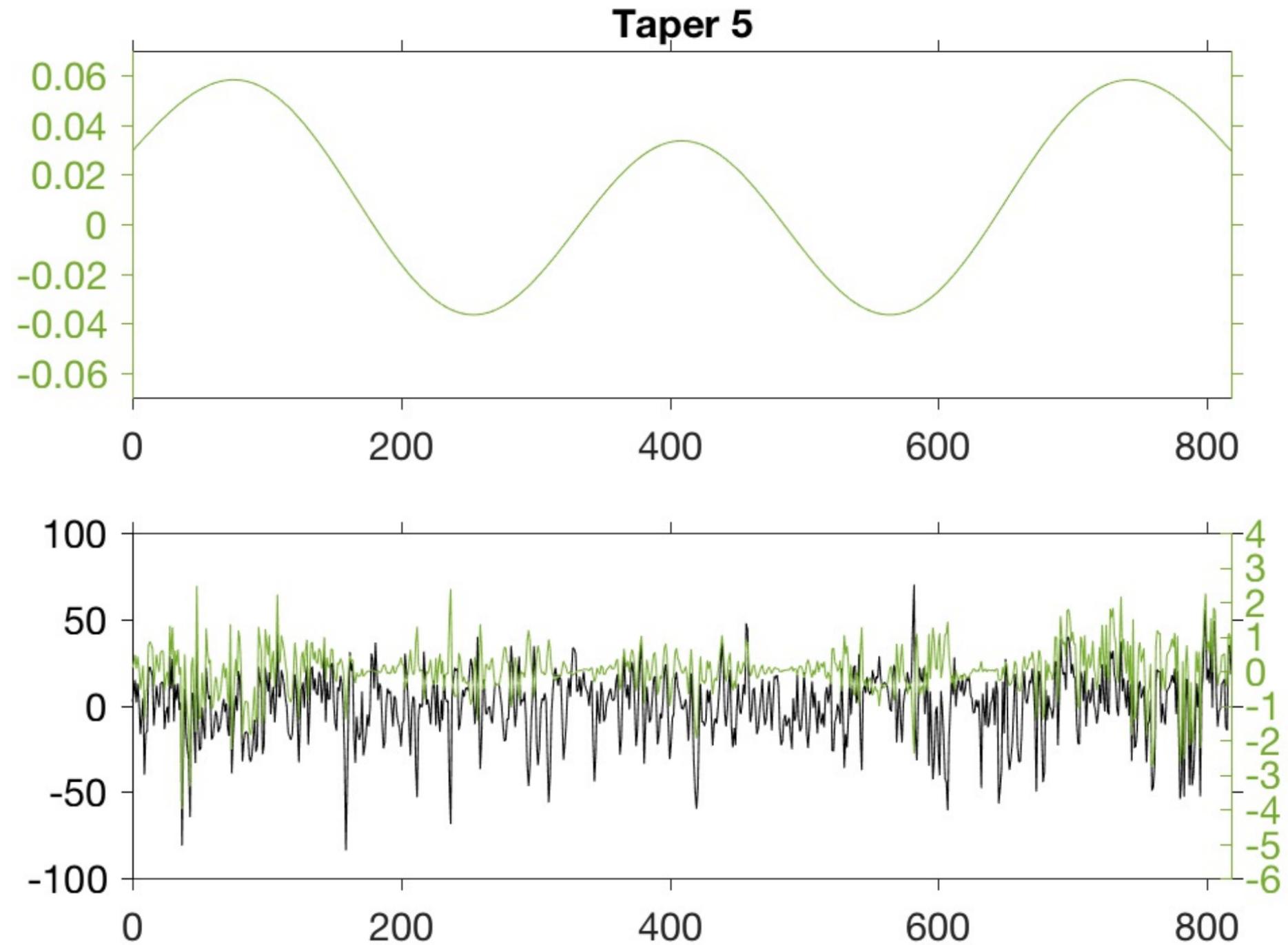
Example: multitaper

Fifth taper



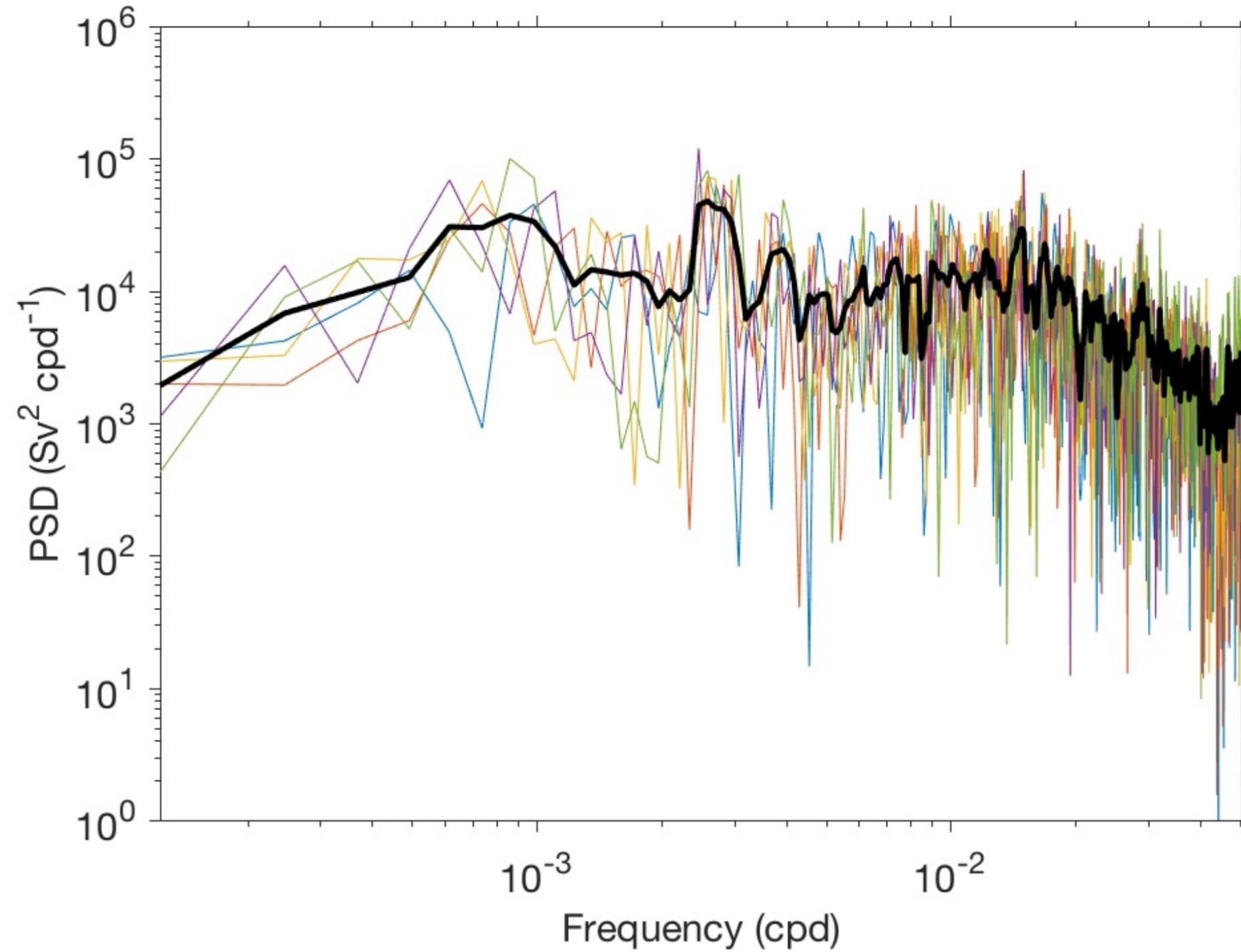
Example: multitaper

Fifth taper



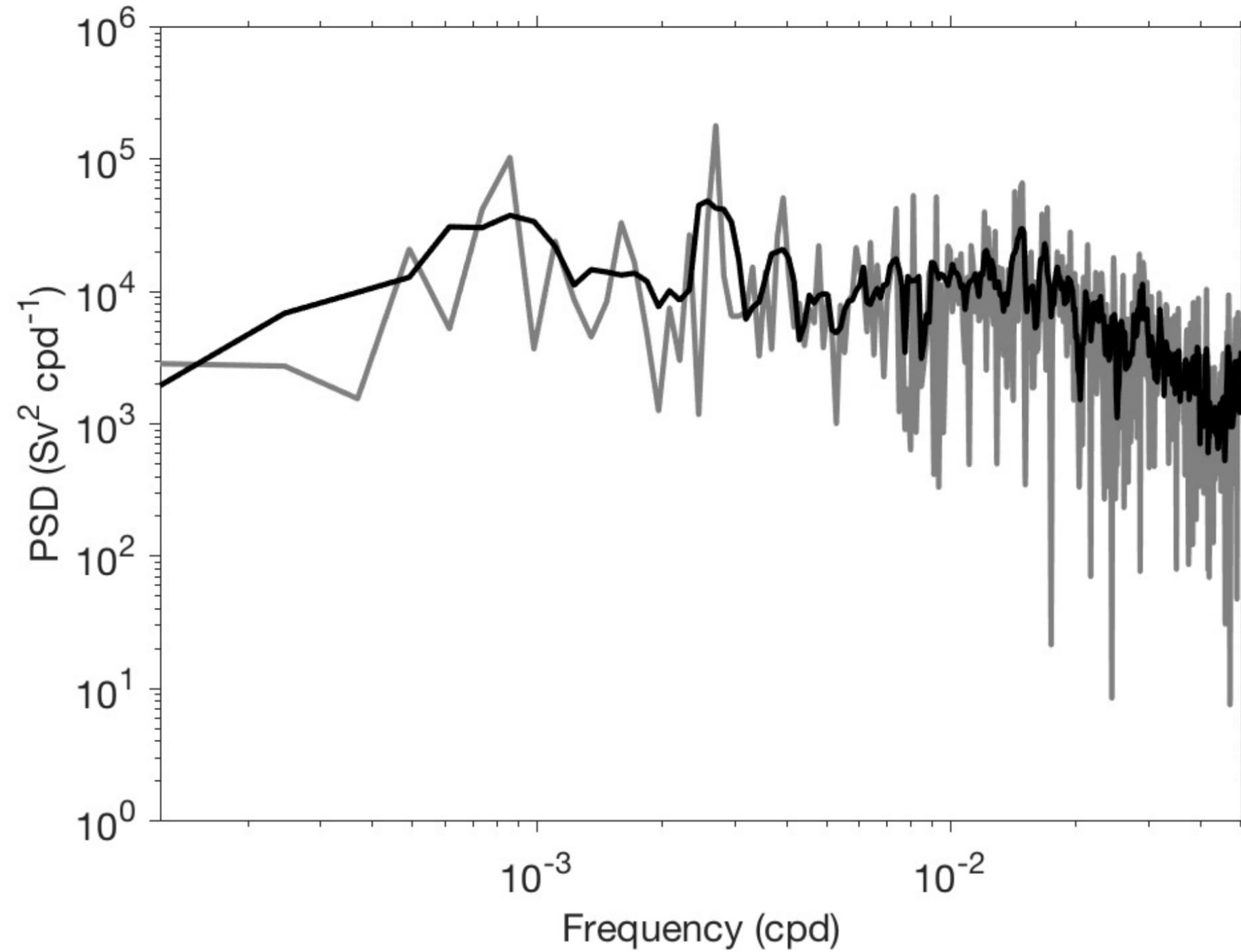
Example: multitaper

Eigen spectra (colors) and multitaper estimate (black)



Example: period. vs mt

Periodogram (gray) vs multitaper (black)



Uncertainty of the spectrum

It can be shown (not here) that the estimate of the spectrum with K tapers

$$\hat{S}(\omega) \sim S(\omega) \frac{\chi_{2K}^2}{2K}$$



Uncertainty of the spectrum

It can be shown (not here) that the estimate of the spectrum with K tapers

$$\hat{S}(\omega) \sim S(\omega) \frac{\chi_{2K}^2}{2K}$$

As such, a $(1 - \alpha)100\%$ CI is

$$\left[\frac{2K\hat{S}(\omega)}{\chi_{2K;\alpha/2}^2} < S(\omega) < \frac{2K\hat{S}}{\chi_{2K;1-\alpha/2}^2} \right]$$



Uncertainty of the spectrum

It can be shown (not here) that the estimate of the spectrum with K tapers

$$\hat{S}(\omega) \sim S(\omega) \frac{\chi_{2K}^2}{2K}$$

As such, a $(1 - \alpha)100\%$ CI is

$$\left[\frac{2K\hat{S}(\omega)}{\chi_{2K;\alpha/2}^2} < S(\omega) < \frac{2K\hat{S}}{\chi_{2K;1-\alpha/2}^2} \right]$$

This means that you **multiply** $\hat{S}(\omega)$ by $2K / \chi_{2K;\alpha/2}^2$ to get the lower bound and similarly for the upper bound.



Uncertainty of the spectrum

It can be shown (not here) that the estimate of the spectrum with K tapers

$$\hat{S}(\omega) \sim S(\omega) \frac{\chi_{2K}^2}{2K}$$

As such, a $(1 - \alpha)100\%$ CI is

$$\left[\frac{2K \hat{S}(\omega)}{\chi_{2K; \alpha/2}^2} < S(\omega) < \frac{2K \hat{S}}{\chi_{2K; 1-\alpha/2}^2} \right]$$

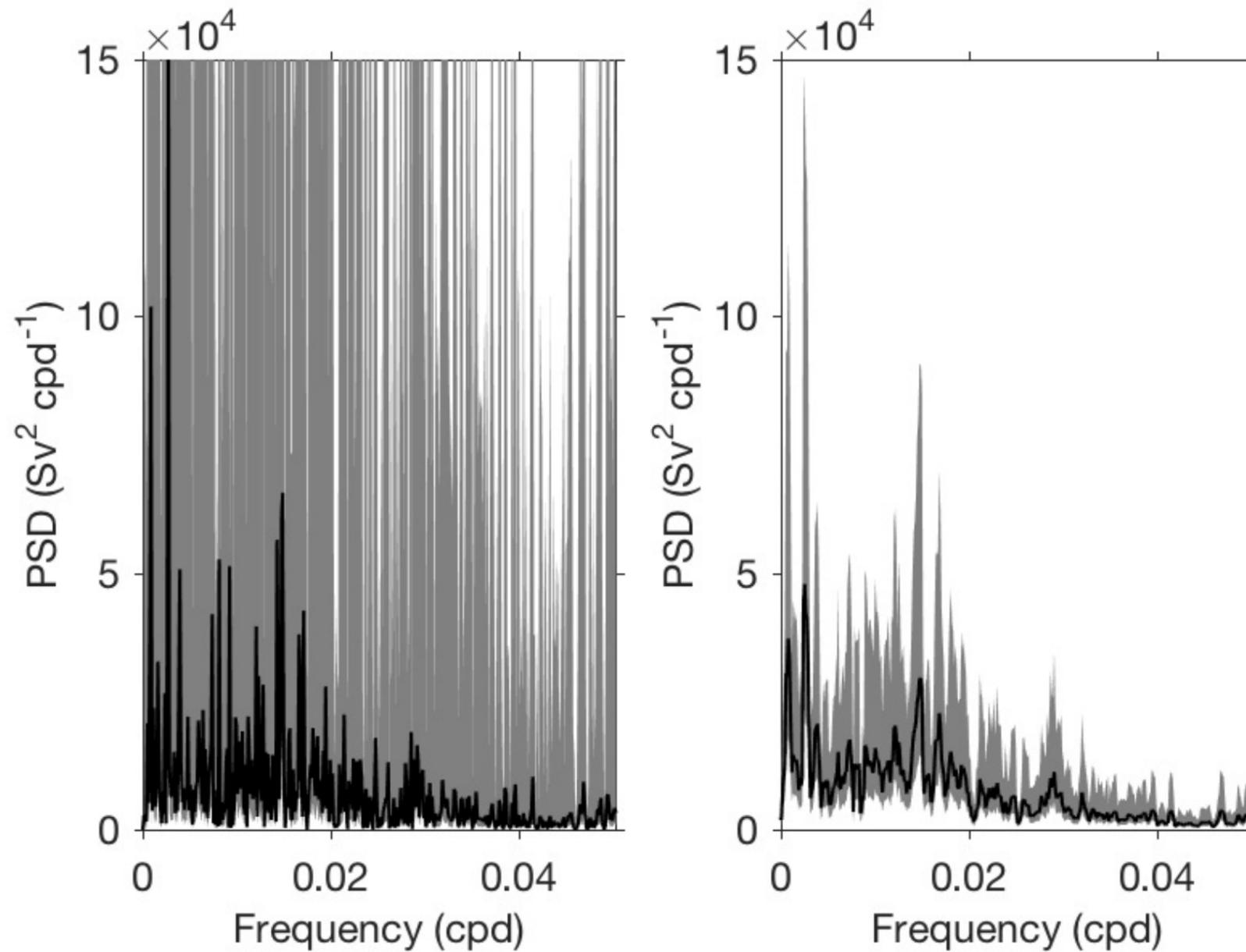
This means that you **multiply** $\widehat{S}(\omega)$ by $2K / \chi_{2K; \alpha/2}^2$ to get the lower bound and similarly for the upper bound. If you plot your estimates on a logarithmic scale, you obtain

$$\left[\log \left(\frac{2K}{\chi_{2K; \alpha/2}^2} \right) + \log \hat{S} < \log S < \log \left(\frac{2K}{\chi_{2K; 1-\alpha/2}^2} \right) + \log \hat{S} \right]$$



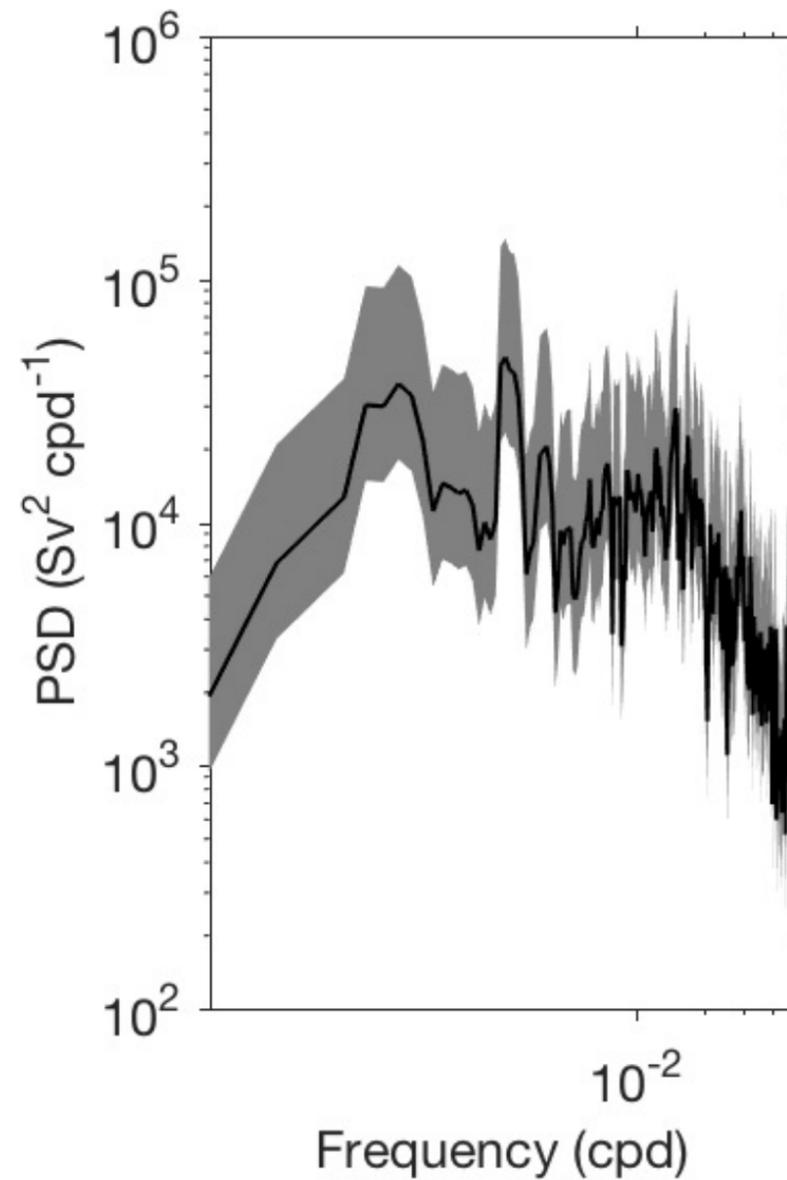
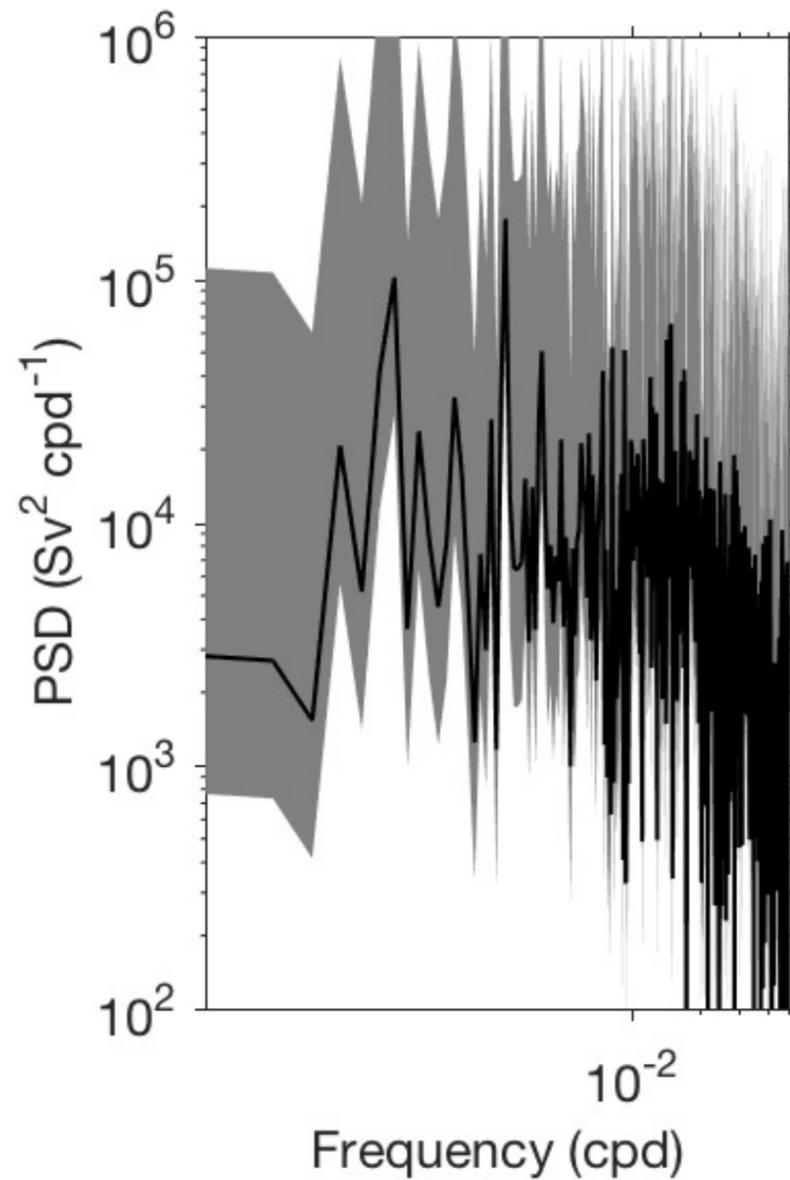
Uncertainty of the spectrum

Periodogram (left) and multitaper (right) estimate with CIs on linear-linear scales



Uncertainty of the spectrum

Multitaper estimate with CIs on log-log scales Periodogram (left) and multitaper (right) estimate with CIs on log-log scales



4. Bivariate time series



Vector and complex notations

What if your process of interest is composed of two time series, let's say $x(t)$ and $y(t)$? As in the vector components of ocean currents or atmospheric winds:

$$\mathbf{z}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

Often, a *bivariate* time series is conveniently written as a complex-valued time series:

$$z(t) = x(t) + iy(t) = |z(t)|e^{i \arg(z)},$$

where $i = \sqrt{-1}$ and $\arg(z)$ is the *complex argument* (or polar angle) of z in the interval $[-\pi, +\pi]$.

The Mean of Bivariate Data

The sample mean of the vector time series \mathbf{z}_n is also a vector,

$$\mu_{\mathbf{z}} \equiv \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{z}_n = \begin{bmatrix} \hat{\mu}_x \\ \hat{\mu}_y \end{bmatrix}$$

that consists of the *sample means* of the x_n and y_n components of \mathbf{z}_n .



Variance of Bivariate Data

The *variance* of the vector-valued times series \mathbf{z}_n is not a scalar or a vector, it is a 2×2 *matrix*

$$\mathbf{\Sigma} \equiv \frac{1}{N} \sum_{n=0}^{N-1} (\mathbf{z}_n - \mu_{\mathbf{z}}) (\mathbf{z}_n - \mu_{\mathbf{z}})^T$$

where “ T ” is the *matrix transpose*, $\mathbf{z}_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$, $\mathbf{z}_n^T = [x_n \quad y_n]$.

Carrying out the matrix multiplication leads to

$$\mathbf{\Sigma} = \frac{1}{N} \sum_{n=0}^{N-1} \begin{bmatrix} (x_n - \hat{\mu}_x)^2 & (x_n - \hat{\mu}_x)(y_n - \hat{\mu}_y) \\ (x_n - \hat{\mu}_x)(y_n - \hat{\mu}_y) & (y_n - \hat{\mu}_y)^2 \end{bmatrix}$$

The diagonal elements of $\mathbf{\Sigma}$ are the sample variances σ_x^2 and σ_y^2 , while the off-diagonal gives the *covariance* between x_n and y_n . Note that the two off-diagonal elements are identical.



Fourier transform

The Fourier theory presented earlier for *scalar* time series is completely applicable to complex-valued time series, in discrete form (N even):

$$\begin{aligned} z_n &= \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} Z_m e^{i2\pi mn/N} + \frac{1}{N\Delta_t} \sum_{m=N/2+1}^{N-1} Z_m e^{i2\pi mn/N} \\ &= \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} Z_m^+ e^{i2\pi mn/N} + \frac{1}{N\Delta_t} \sum_{m=1}^{N/2-1} Z_m^- e^{-i2\pi mn/N}. \end{aligned}$$

The first sum corresponds to **positive** frequencies, and the second sum to the associated **negative** frequencies (except the zero and Nyquist frequencies for $m = 0, N/2$).



Rotary Spectra

$$z_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} Z_m^+ e^{i2\pi mn/N} + \frac{1}{N\Delta_t} \sum_{m=1}^{N/2-1} Z_m^- e^{-i2\pi mn/N}.$$

This introduces the concept of *rotary spectrum*:

$$S(f_m > 0) \equiv S^+(f_m) \equiv \lim_{N \rightarrow \infty} E \left\{ \frac{|Z_m^+|^2}{N} \right\} \quad \text{counterclockwise spectrum}$$

$$S(f_m < 0) \equiv S^-(-f_m) \equiv \lim_{N \rightarrow \infty} E \left\{ \frac{|Z_m^-|^2}{N} \right\} \quad \text{clockwise spectrum.}$$

$$f_m = \frac{m}{N\Delta_t}, \quad m = 0, \dots, N/2$$

This is very useful in geophysical fluid mechanics because counterclockwise motions are *cyclonic* in the northern hemisphere and clockwise motions are *anticyclonic*, and vice-versa in the southern hemisphere.



Rotary variance

Imagine you have only two opposite components present in your time series at frequency f_k :

$$\begin{aligned}z_n &= Z_k^+ e^{i2\pi kn/N} + Z_k^- e^{-i2\pi kn/N} \\ &= \left\{ A^+ e^{i\phi^+} \right\} e^{i2\pi kn/N} + \left\{ A^- e^{i\phi^-} \right\} e^{-i2\pi kn/N} \\ &= e^{i\theta} \left\{ A \cos(2\pi kn/N + \phi) + iB \sin(2\pi kn/N + \phi) \right\}\end{aligned}$$

where

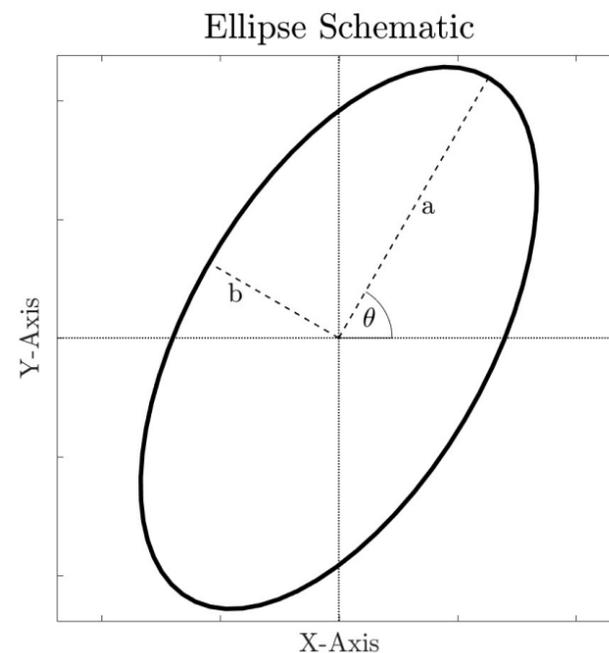
$$\begin{aligned}\theta &= (\phi^+ - \phi^-)/2 \\ \phi &= (\phi^+ + \phi^-)/2 \\ A &= A^+ + A^- \\ B &= A^+ - A^-.\end{aligned}$$



Elliptic variance

$$z_n = e^{i\theta} \{A \cos(2\pi kn/N + \phi) + iB \sin(2\pi kn/N + \phi)\}$$

This is the equation for an ellipse oriented at an angle θ from the x axis, with semi-major and semi-minor axes A and B , respectively, rotating at frequency $f_k = k/(N\Delta_t)$, in the direction given by the sign of B .



See more details about elliptic variance in JML's Oslo lectures.



Cartesian Spectra

Rotary and Cartesian spectra are two alternate representation of the variance of the complex time series:

$$z_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} Z_m^+ e^{i2\pi mn/N} + \frac{1}{N\Delta_t} \sum_{m=1}^{N/2-1} Z_m^- e^{-i2\pi mn/N}.$$

$$\hat{S}_m^+ \equiv \frac{1}{N} |Z_m^+|^2, \quad \hat{S}_m^- \equiv \frac{1}{N} |Z_m^-|^2 \quad \text{Rotary spectra estimates}$$

$$z_n = x_n + iy_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{i2\pi mn/N} + i \left\{ \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} Y_m e^{i2\pi mn/N} \right\}.$$

$$\hat{S}_m^x \equiv \frac{1}{N} |X_m|^2, \quad \hat{S}_m^y \equiv \frac{1}{N} |Y_m|^2 \quad \text{Cartesian spectra estimates}$$



Parseval theorem

For bivariate data, the discrete form of the Parseval theorem takes the form:

$$\begin{aligned}\Delta_t \sum_{n=0}^{N-1} |z_n|^2 &= \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |Z_m|^2 = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |X_m|^2 + \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |Y_m|^2 \\ &= \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} |Z_m^+|^2 + \frac{1}{N\Delta_t} \sum_{m=1}^{N/2-1} |Z_m^-|^2\end{aligned}$$



Parseval theorem

For bivariate data, the discrete form of the Parseval theorem takes the form:

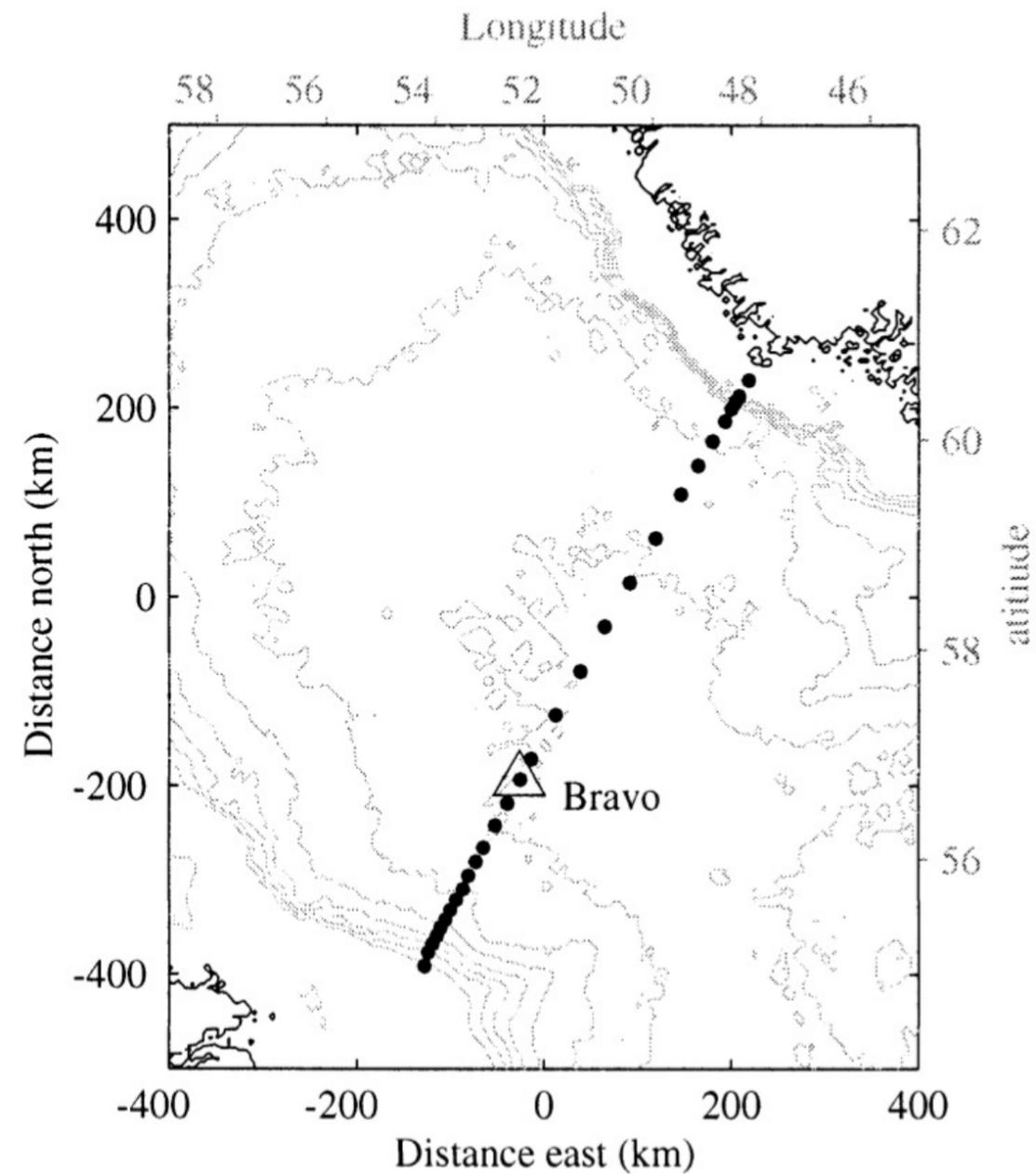
$$\begin{aligned}\Delta_t \sum_{n=0}^{N-1} |z_n|^2 &= \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |Z_m|^2 = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |X_m|^2 + \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} |Y_m|^2 \\ &= \frac{1}{N\Delta_t} \sum_{m=0}^{N/2} |Z_m^+|^2 + \frac{1}{N\Delta_t} \sum_{m=1}^{N/2-1} |Z_m^-|^2\end{aligned}$$

This shows that the total variance of the bivariate process is recovered completely by the Cartesian, or rotary Fourier representation.



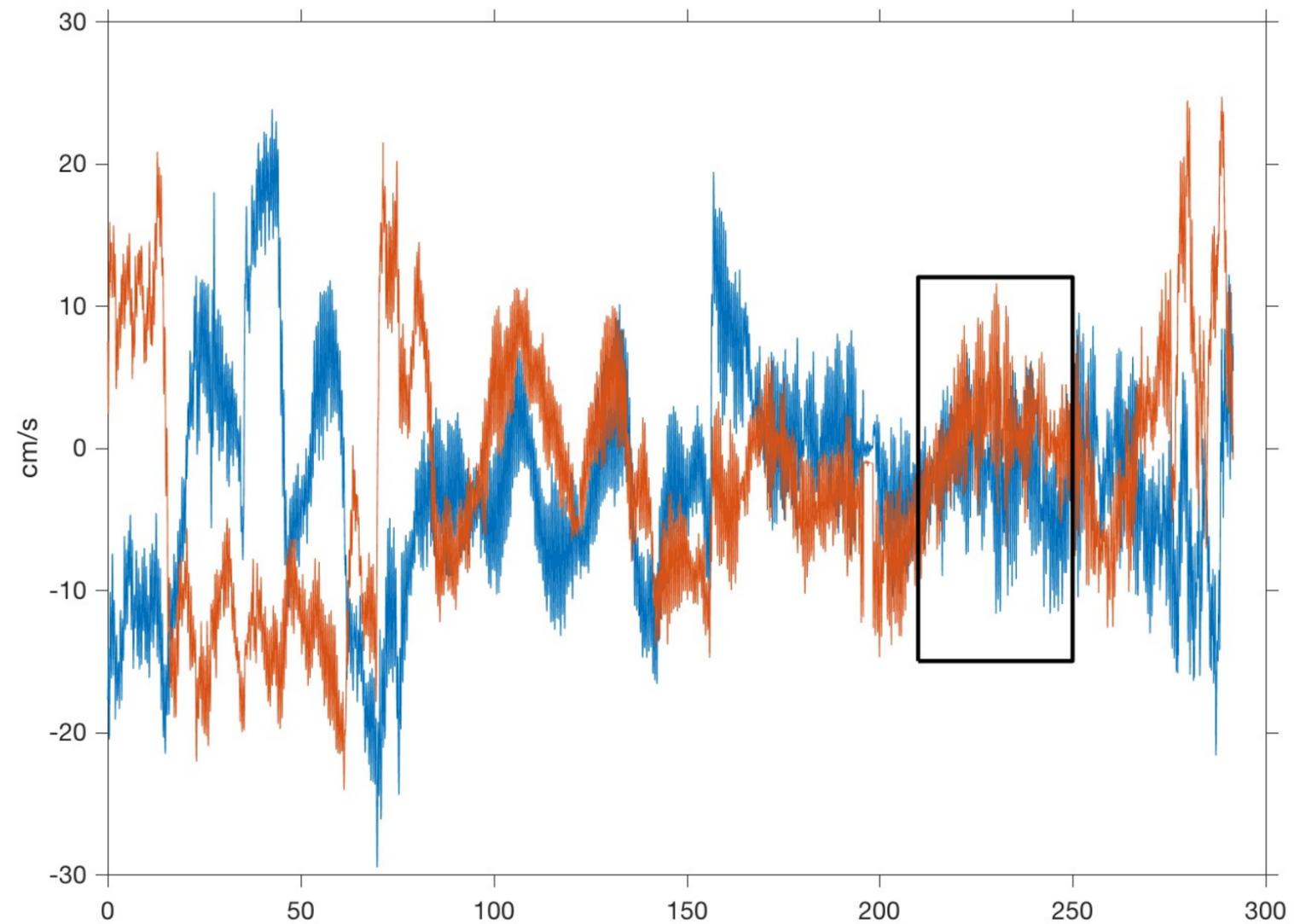
Example

Hourly current meter record at 110 m depth from the Bravo mooring in the Labrador Sea, *Lilly and Rhines (2002)*



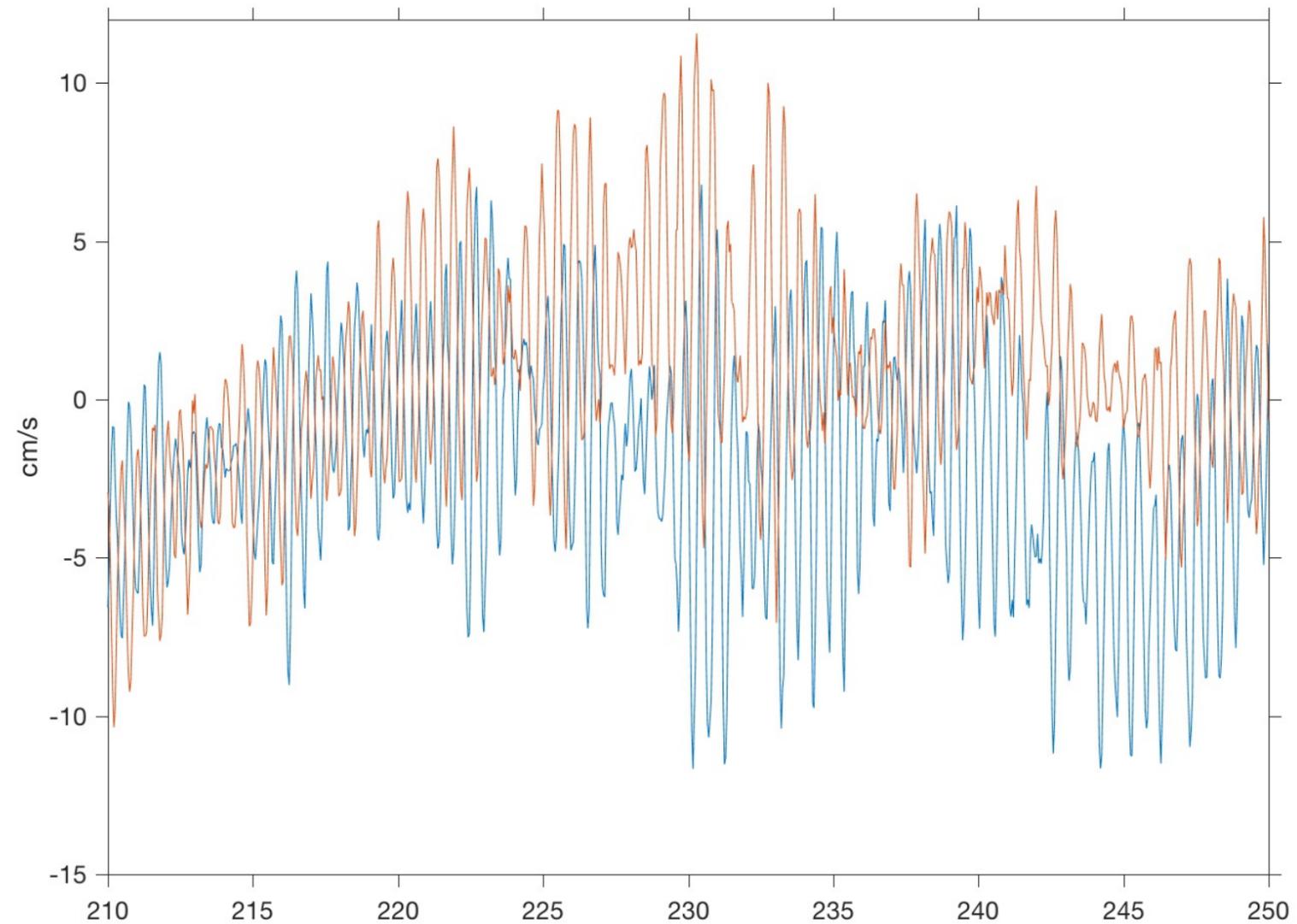
Example 1

Hourly current meter record at 110 m depth from the Bravo mooring in the Labrador Sea, *Lilly and Rhines (2002)*



Example 1

Hourly current meter record at 110 m depth from the Bravo mooring in the Labrador Sea, *Lilly and Rhines (2002)*



Observable Features

1. The data consists of two time series that are similar in character.
2. Both time series present a superposition of scales.
3. At the smallest scale, there is an apparently oscillatory roughness which changes its amplitude in time.
4. A larger scale presents itself either as localized features, or as wavelike in nature.
5. Several sudden transitions are associated with isolated events.
6. Zooming in, we see the small-scale oscillatory behavior is sometimes 90° degrees out of phase, and sometimes 180° .
7. The amplitude of this oscillatory variability changes with time.

The fact that the oscillatory behavior is not consistently 90° out of phase removes the possibility of these features being purely inertial oscillations. The amplitude modulation suggests tidal beating.



Observable Features

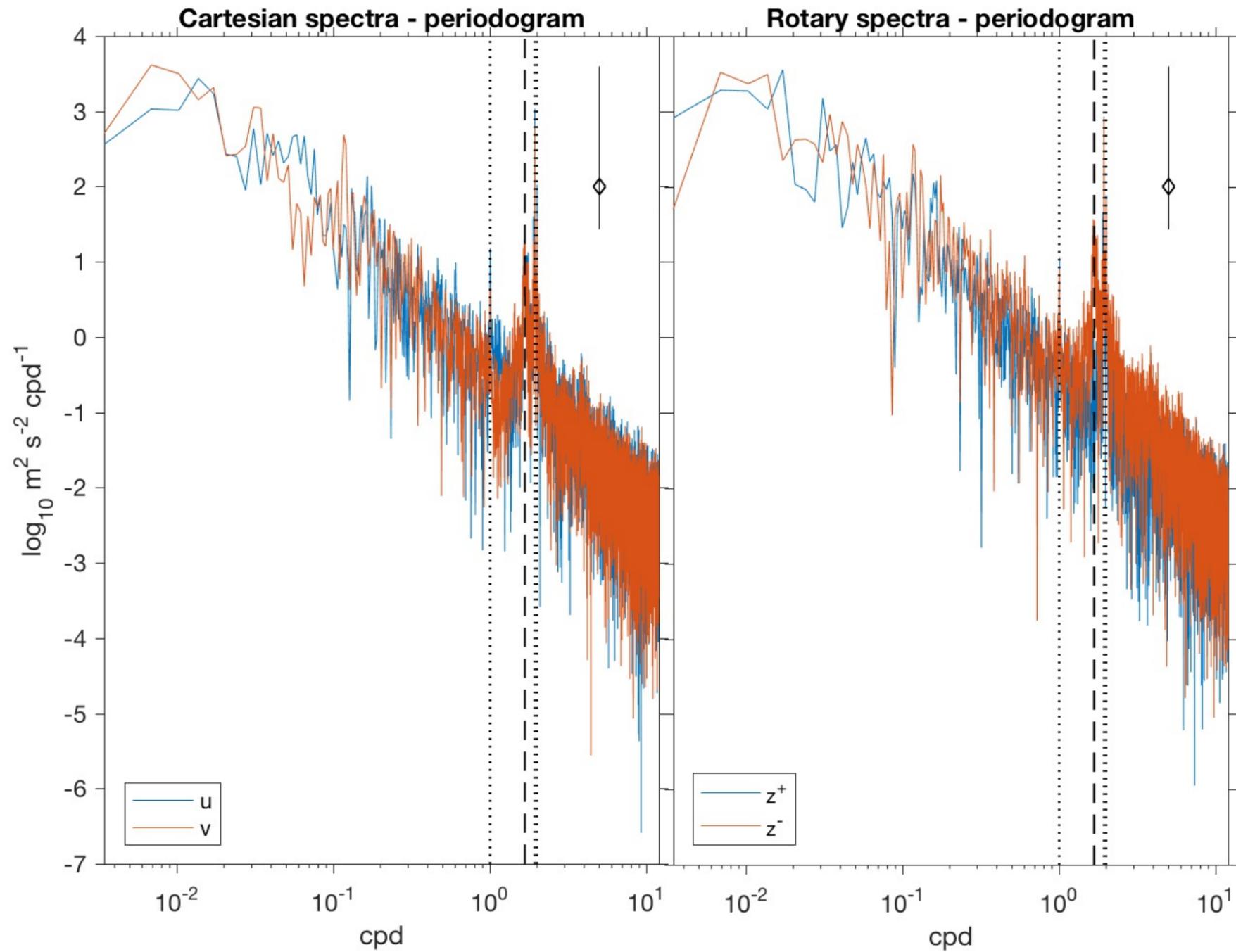
1. The data consists of two time series that are similar in character.
2. Both time series present a superposition of scales.
3. At the smallest scale, there is an apparently oscillatory roughness which changes its amplitude in time.
4. A larger scale presents itself either as localized features, or as wavelike in nature.
5. Several sudden transitions are associated with isolated events.
6. Zooming in, we see the small-scale oscillatory behavior is sometimes 90° degrees out of phase, and sometimes 180° .
7. The amplitude of this oscillatory variability changes with time.

The fact that the oscillatory behavior is not consistently 90° out of phase removes the possibility of these features being purely inertial oscillations. The amplitude modulation suggests tidal beating.

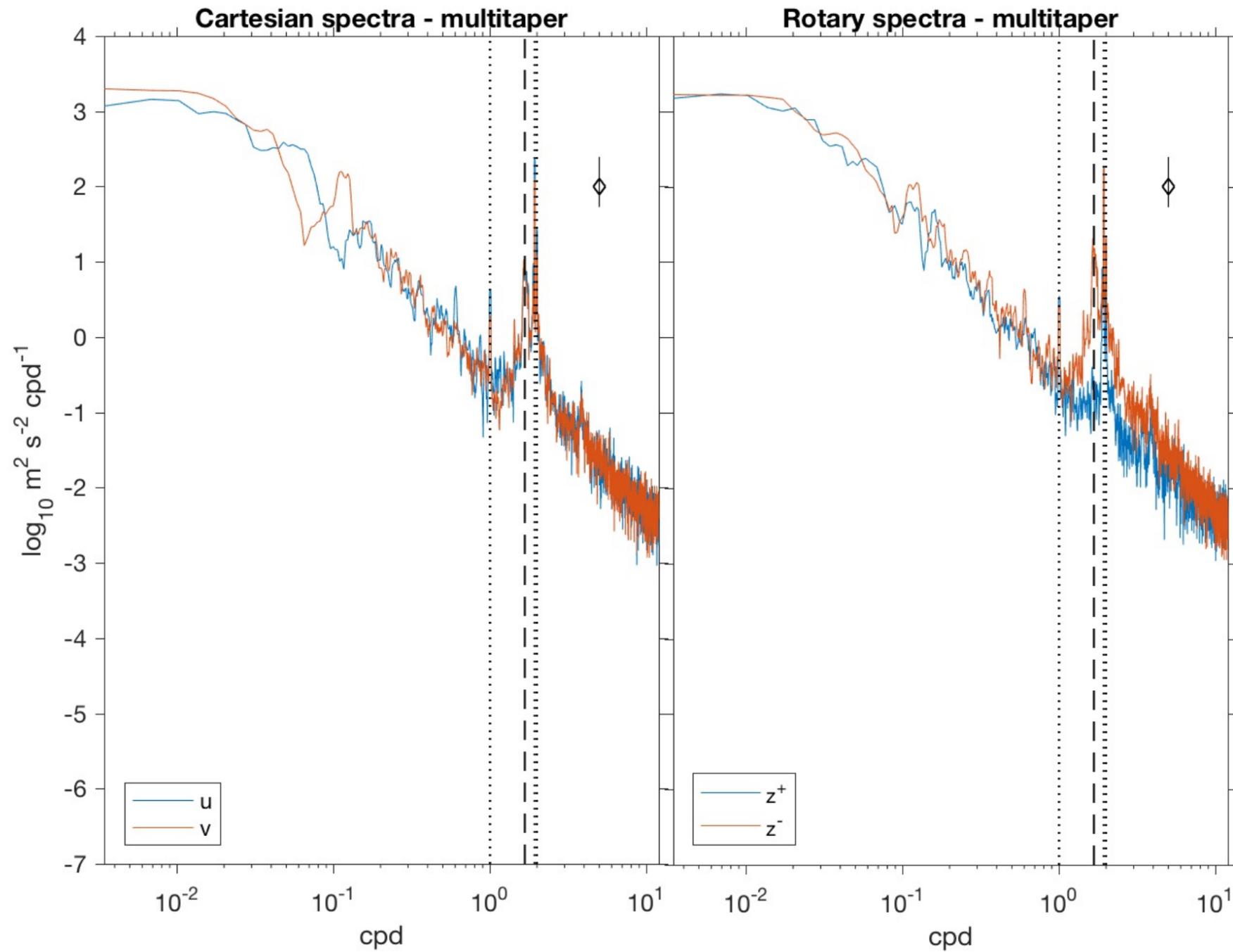
The isolated events are eddies, which cause the currents to suddenly rotate as they pass by. The oscillations are due to tides and internal waves.



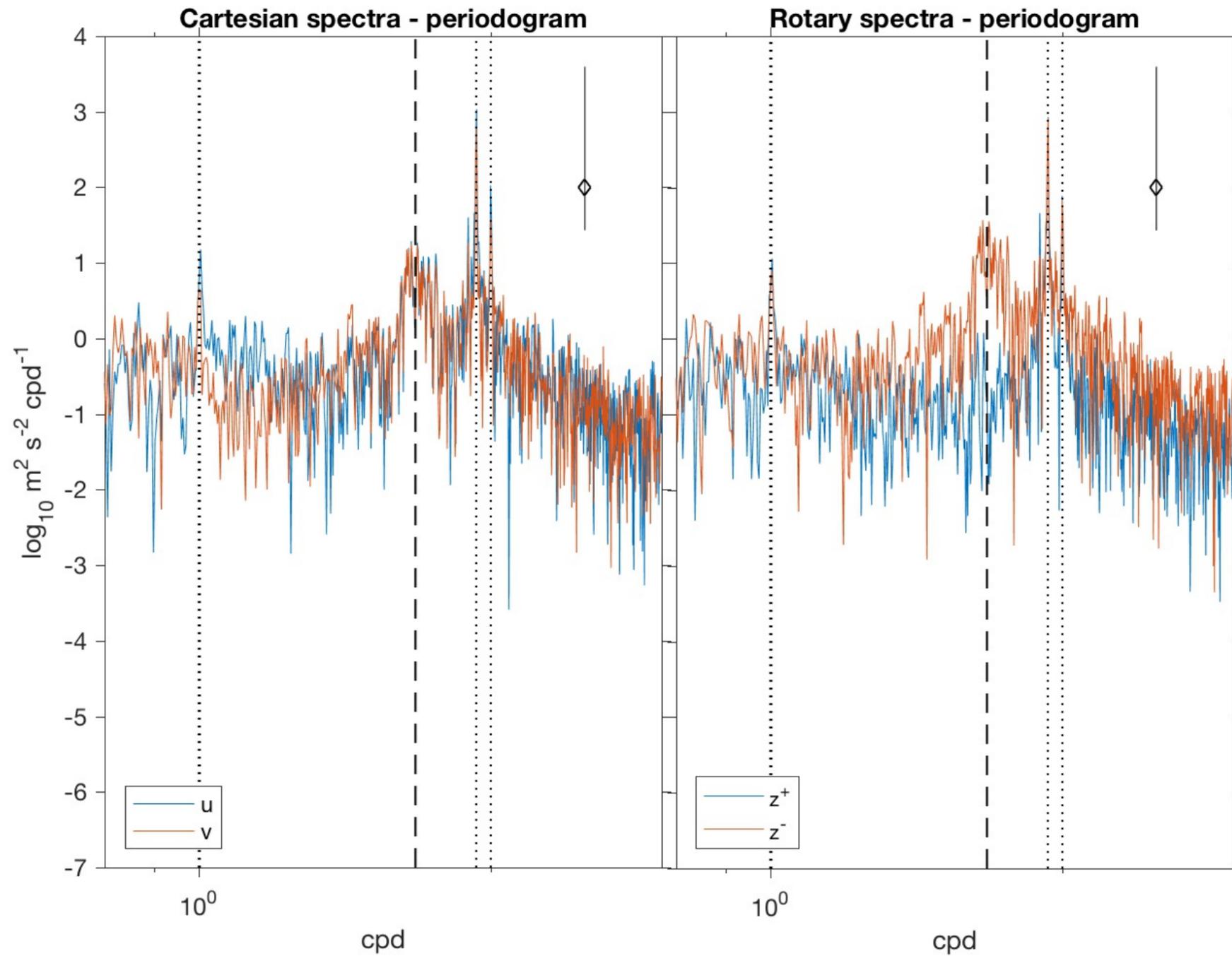
Cartesian vs Rotary Spectra



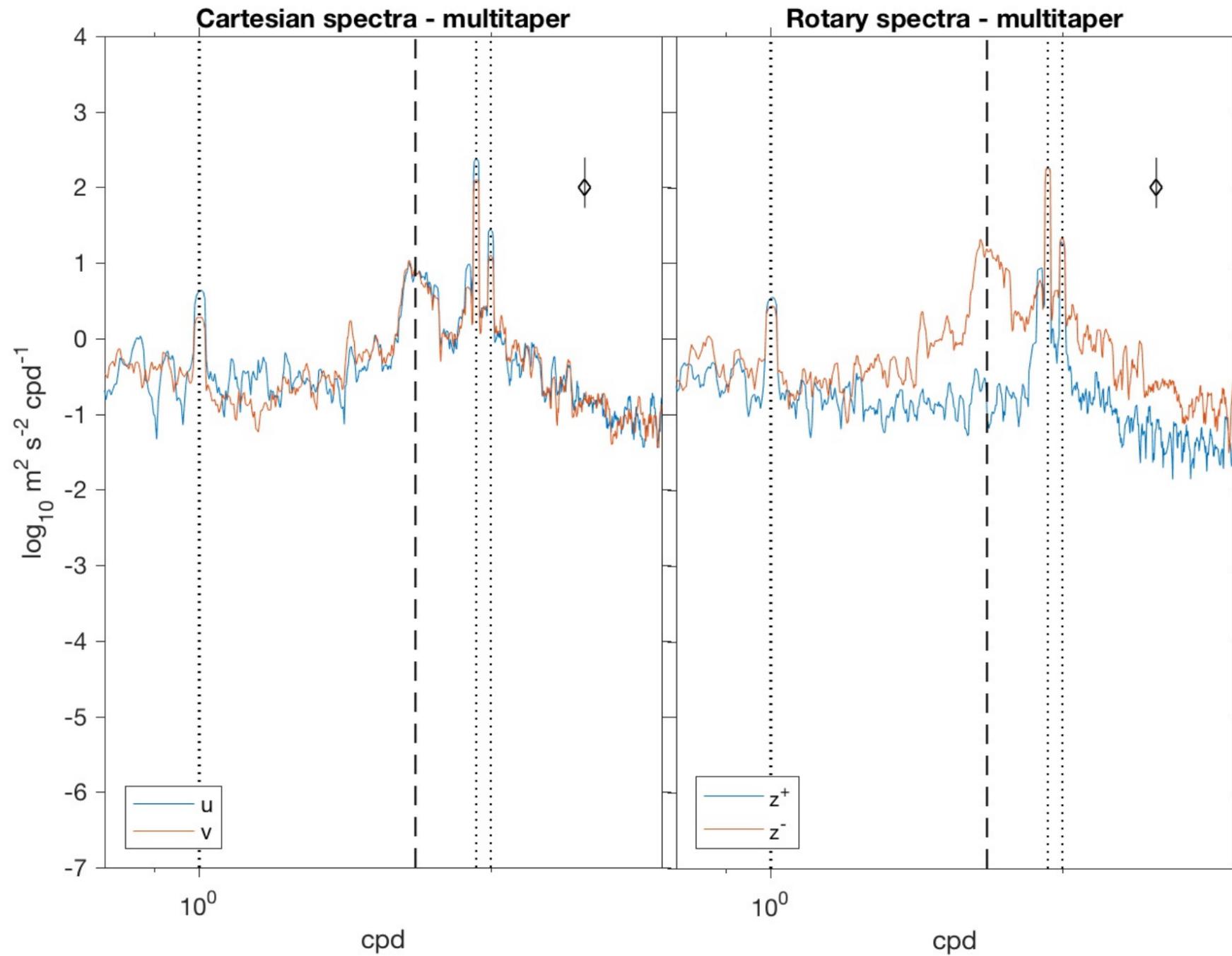
Cartesian vs Rotary Spectra



Cartesian vs Rotary Spectra

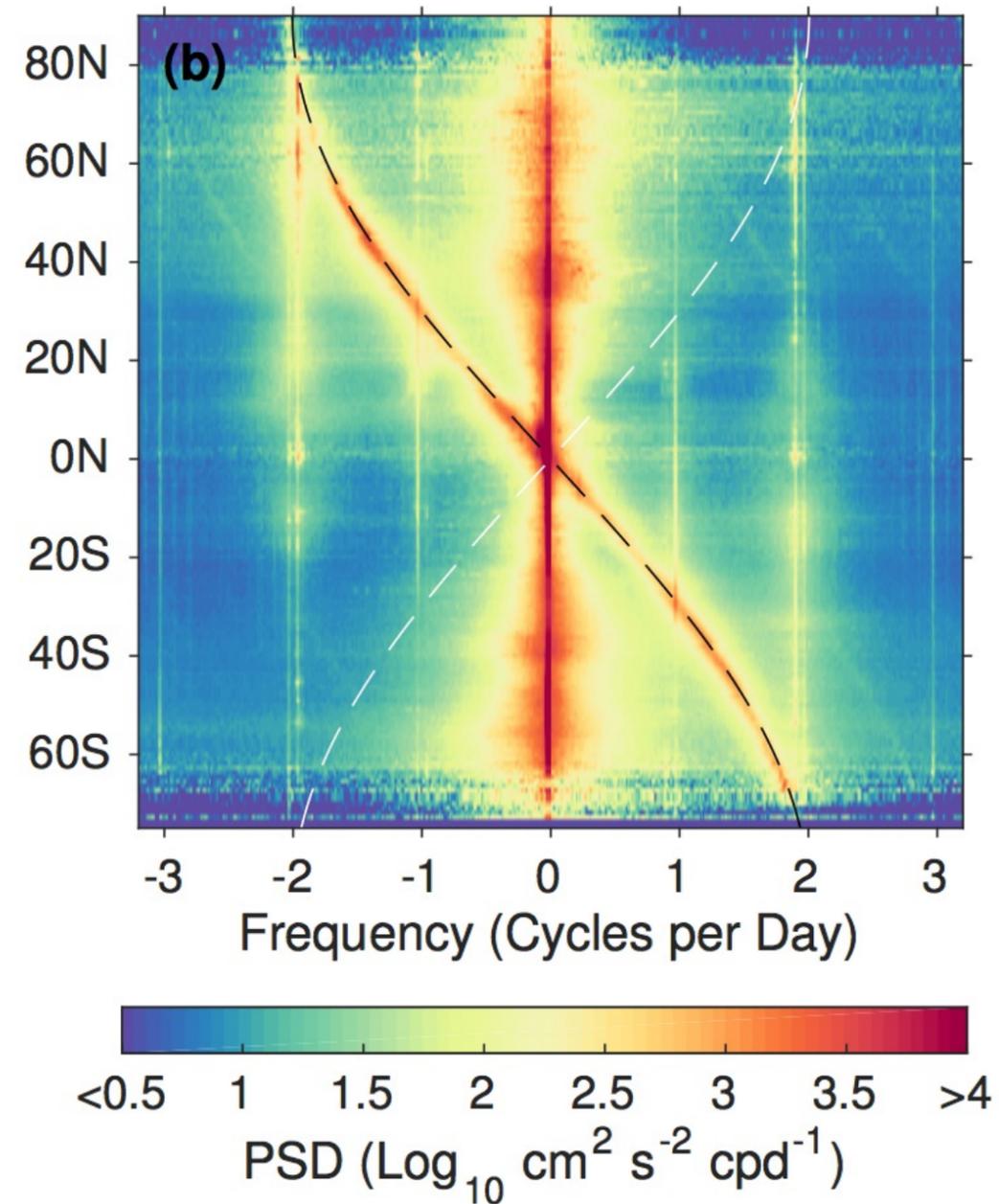


Cartesian vs Rotary Spectra



Example

Global zonally-averaged rotary spectra from hourly drifter velocities, see Elipot et al. 2016.



5. Filtering and other topics



Continuous Fourier

We have considered the FT for a discrete time series x_n :

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{2\pi mn/N}, \quad X_m = \Delta_t \sum_{n=0}^{N-1} x_n e^{-i2\pi mn/N}$$



Continuous Fourier

We have considered the FT for a discrete time series x_n :

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{2\pi mn/N}, \quad X_m = \Delta_t \sum_{n=0}^{N-1} x_n e^{-i2\pi mn/N}$$

but it extends to continuous time series $x(t)$:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega, \quad X(\omega) \equiv \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt.$$

Note that here we are using radian frequency ω .



Continuous Fourier

We have considered the FT for a discrete time series x_n :

$$x_n = \frac{1}{N\Delta_t} \sum_{m=0}^{N-1} X_m e^{2\pi mn/N}, \quad X_m = \Delta_t \sum_{n=0}^{N-1} x_n e^{-i2\pi mn/N}$$

but it extends to continuous time series $x(t)$:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega, \quad X(\omega) \equiv \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt.$$

Note that here we are using radian frequency ω .

The continuous notation is easier to understand the mechanics of *filtering* a time series, as well as spectral *blurring*.



The Spectrum (revisited)

An alternative definition of the spectrum $S(\omega)$ is that it is the Fourier transform of the autocorrelation function $R(\omega)$:

$$S(\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau, \quad R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega$$



The Spectrum (revisited)

An alternative definition of the spectrum $S(\omega)$ is that it is the Fourier transform of the autocorrelation function $R(\omega)$:

$$S(\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau, \quad R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega$$

This is called the *Wiener–Khinchine theorem*. The spectrum and the autocorrelation function are *Fourier transform pairs*. While both are essentially equivalent in that they capture the *same* second-order statistical information in different forms, the spectrum turns out to generally be far more illuminating, as well as easier to work with in practice.



The Spectrum (revisited)

An alternative definition of the spectrum $S(\omega)$ is that it is the Fourier transform of the autocorrelation function $R(\omega)$:

$$S(\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau, \quad R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega$$

This is called the *Wiener–Khinchine theorem*. The spectrum and the autocorrelation function are *Fourier transform pairs*. While both are essentially equivalent in that they capture the *same* second-order statistical information in different forms, the spectrum turns out to generally be far more illuminating, as well as easier to work with in practice.

But the true autocorrelation function is not observable unless we have (i) infinite time and (ii) access to an abstract set of other universes where things might have happened differently!



The Convolution Integral

The *convolution* $h(t)$ of a function $f(t)$ and $g(t)$ is defined as:

$$h(t) \equiv \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau.$$

Note that in convolution, the order does not matter and we can show that

$$h(t) \equiv \int_{-\infty}^{\infty} g(\tau)f(t - \tau)d\tau$$

This mathematical operation is actually what is being done when “smoothing” data. (It is like sliding the iron on the tablecloth, or pulling the tablecloth under a static iron).



Convolution Theorem

The *convolution theorem* states convolving $f(t)$ and $g(t)$ in the time domain is the same as a *multiplication* in the frequency domain.

Let $F(\omega)$ and $G(\omega)$ be the Fourier transforms of $f(t)$ and $g(t)$, respectively. It can be shown that if

$$h(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

then the fourier transform of $h(t)$ is

$$H(\omega) = F(\omega)G(\omega).$$



Convolution Theorem

The *convolution theorem* states convolving $f(t)$ and $g(t)$ in the time domain is the same as a *multiplication* in the frequency domain.

Let $F(\omega)$ and $G(\omega)$ be the Fourier transforms of $f(t)$ and $g(t)$, respectively. It can be shown that if

$$h(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

then the fourier transform of $h(t)$ is

$$H(\omega) = F(\omega)G(\omega).$$

This result is key to understand what happens in the Fourier domain when you perform a time-domain smoothing.



Smoothing

Now we consider what happens when we *smooth* the time series $x(t)$ by the filter $g(t)$ to obtain a smoothed version $\tilde{x}(t)$ of your time series:

$$\begin{aligned}\tilde{x}(t) &= \int_{-\infty}^{\infty} x(t - \tau)g(\tau)d\tau \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{X}(\omega) e^{i\omega t} d\omega. \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)G(\omega) e^{i\omega t} d\omega,\end{aligned}$$

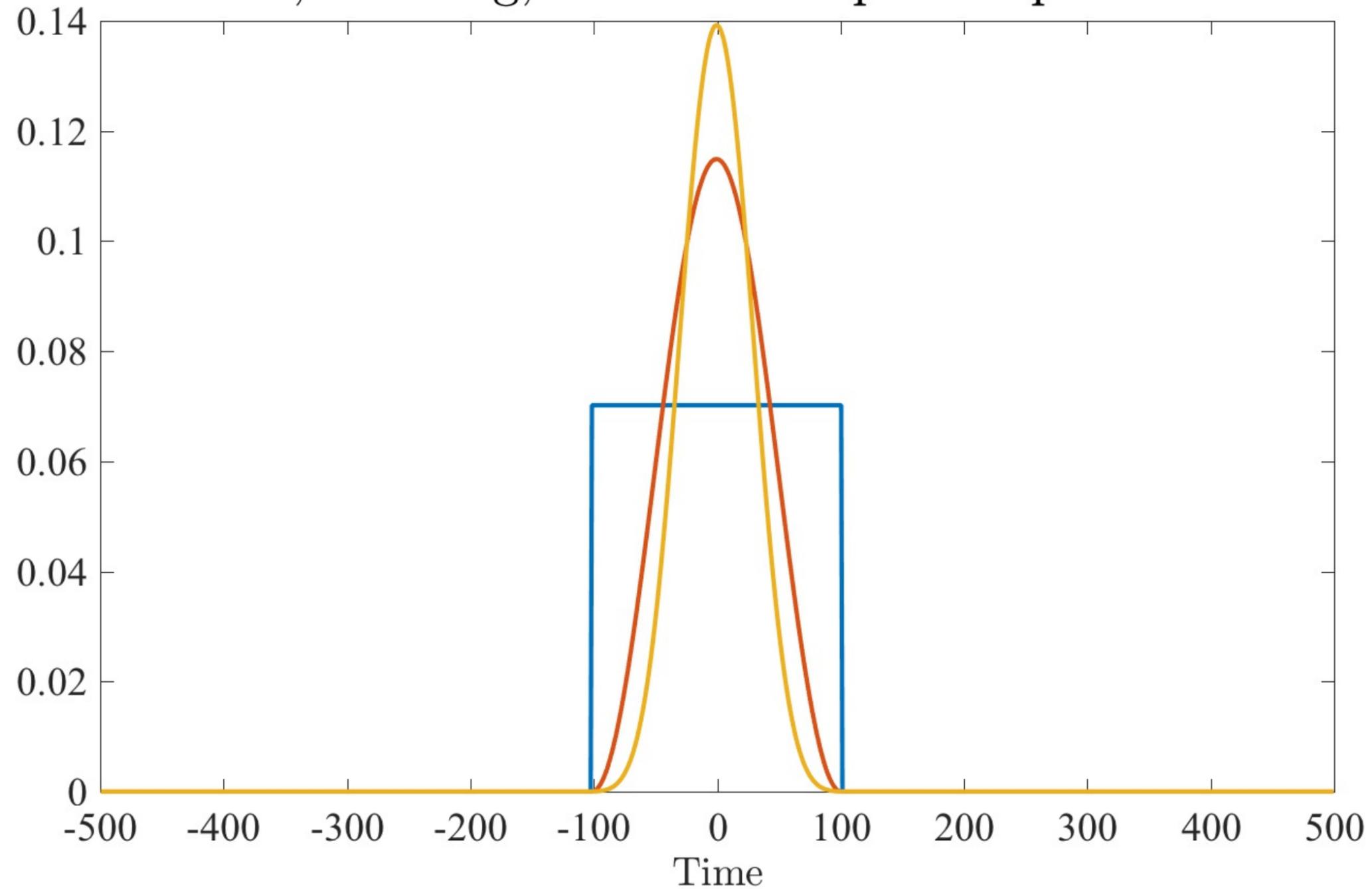
by the convolution theorem.

When we perform simple smoothing, we are also reshaping the Fourier transform of the signal by *multiplying* its Fourier transform by that of the smoothing window.



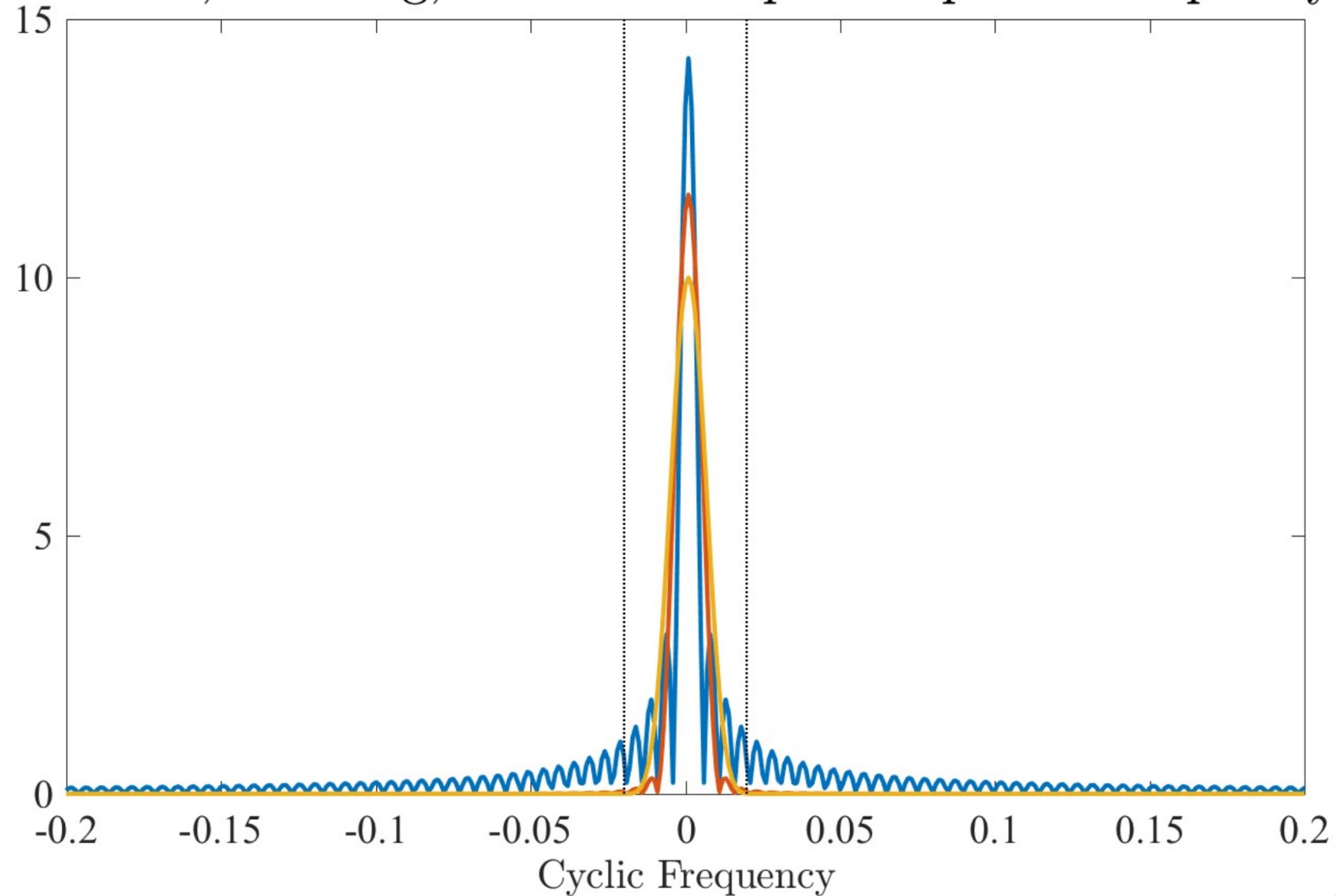
Three Window examples

Boxcar, Hanning, and First Slepian Taper in Time



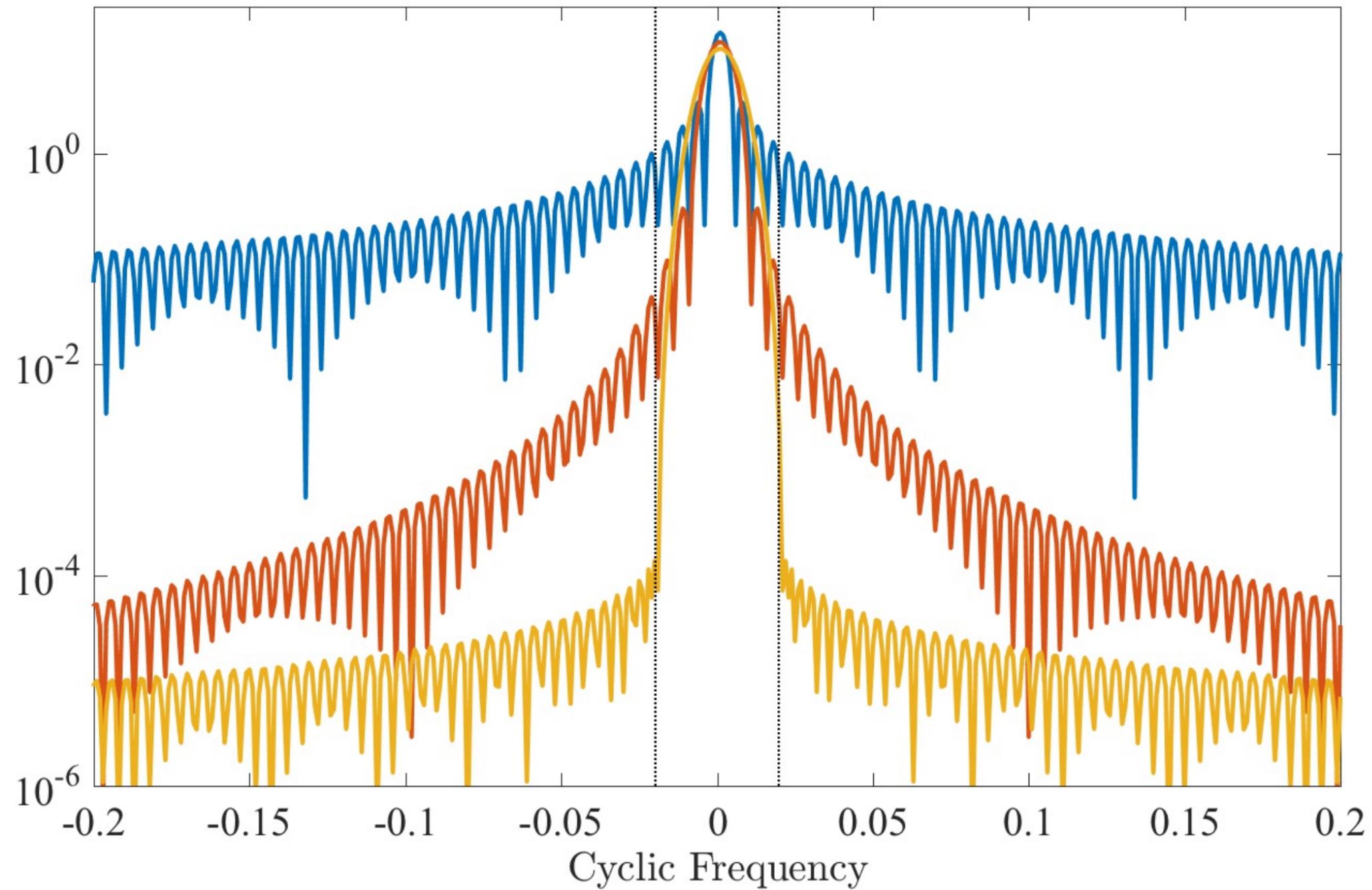
Three Tapering Windows

Boxcar, Hanning, and First Slepian Taper in Frequency



Three Tapering Windows

Boxcar, Hanning, and First Slepian Taper in Frequency



Lowpass & Highpass Filters

From the convolution theorem, we understand that filtering will keep the frequencies near zero but reject higher frequencies. For this reason they are called *low-pass filters*.

The reverse type of filtration, rejecting the low frequencies but keeping the high frequencies, is called *high-pass filtering*.

The residual $\check{x}(t) \equiv x(t) - \tilde{x}(t)$ is an example of a high-pass filtered time series.

In practice, to find the frequency form of your filter, you *pad it with zeros* so that it becomes the same length as your time series, and then you take its discrete Fourier transform.



Convolution Theorem II

This theorem is reciprocal: if you *multiply* in the time domain, you *convolve* in the Fourier domain.

It can be shown that if

$$h(t) = f(t)g(t)$$

then the fourier transform of $h(t)$ is

$$H(\omega) = \int_{-\infty}^{\infty} F(\nu)G(\omega - \nu)d\nu.$$

This result is key to understand what happens in the Fourier domain when you try to estimate spectra, i.e. *spectral blurring*, or to design band-pass filters.



Bandpass filtering

We can use the convolution theorem to build a band-pass filter.

We want to modify the lowpass filter $g(t)$ so that its Fourier transform is localized not about zero, but about some non-zero frequency ω_o . To do this, we multiply $g(t)$ by a complex exponential $g(t)e^{i\omega_o t}$.

It can be shown (see Oslo lectures) that the Fourier transform of $g(t)e^{i\omega_o t}$ is $G(\omega - \omega_o)$, which is localized around ω_o .

Thus, a convolution with $g(t)e^{i\omega_o t}$ will *bandpass* the data in the vicinity of ω_o .

In fact, a lowpass filter is a particular type of bandpass in which the center of the *pass band* has been chosen as zero frequency.



Effect of Truncation

Now imagine instead that we have a *continuously* sampled time series of length T , that is, we have $z(t)$ but only between times $-T/2$ and $T/2$. This is like multiplying $z(t)$ by a function $g(t)$ which is equal to one between $-T/2$ and $T/2$ and 0 otherwise:

$$z_T(t) = g(t) \times z(t)$$

We will denote this truncated version of $z(t)$ by $z_T(t)$.

How does the spectrum compare of $z_T(t)$ compare with that of $z(t)$?



Effect of Truncation

Now imagine instead that we have a *continuously* sampled time series of length T , that is, we have $z(t)$ but only between times $-T/2$ and $T/2$. This is like multiplying $z(t)$ by a function $g(t)$ which is equal to one between $-T/2$ and $T/2$ and 0 otherwise:

$$z_T(t) = g(t) \times z(t)$$

We will denote this truncated version of $z(t)$ by $z_T(t)$.

How does the spectrum compare of $z_T(t)$ compare with that of $z(t)$?

Q. Using one of the windows encountered today, how can we express the relationship between $z(t)$ and $z_T(t)$?



Effect of Truncation

Now imagine instead that we have a *continuously* sampled time series of length T , that is, we have $z(t)$ but only between times $-T/2$ and $T/2$. This is like multiplying $z(t)$ by a function $g(t)$ which is equal to one between $-T/2$ and $T/2$ and 0 otherwise:

$$z_T(t) = g(t) \times z(t)$$

We will denote this truncated version of $z(t)$ by $z_T(t)$.

How does the spectrum compare of $z_T(t)$ compare with that of $z(t)$?

Q. Using one of the windows encountered today, how can we express the relationship between $z(t)$ and $z_T(t)$?

Q. Therefore, using another theorem learned today, what is the difference between their spectra?



Spectral Blurring

The spectrum of the truncated time series is *blurred* through smoothing with a function $F_T(\omega)$ that is the *square* of the Fourier transform of a boxcar:

$$\tilde{S}(\omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\nu) F_T(\nu - \omega) d\nu.$$

The smoothing function, which is known as the *Fejér kernel*

$$F_T(\omega) \equiv \int_{-T}^T \left(1 - \frac{|\tau|}{T}\right) e^{i\omega\tau} d\tau = \frac{1}{T} \frac{\sin^2(\omega T/2)}{(\omega/2)^2}$$

is essentially a squared version of the “sinc” or $\sin(x)/x$ function.

However, $\sin(x)/x$ is not a very smooth function at all!



Multitapering Revisited

We can now understand the purpose of multitapering.

Doing nothing in your spectral estimate is equivalent to *truncating* your data, thus implicitly *smoothing* the true spectrum by an extremely undesirable function!

The Fejér kernel has a major problem in that it is not *well concentrated*. Its “side lobes” are large, leading to a kind of error called *broadband bias*.

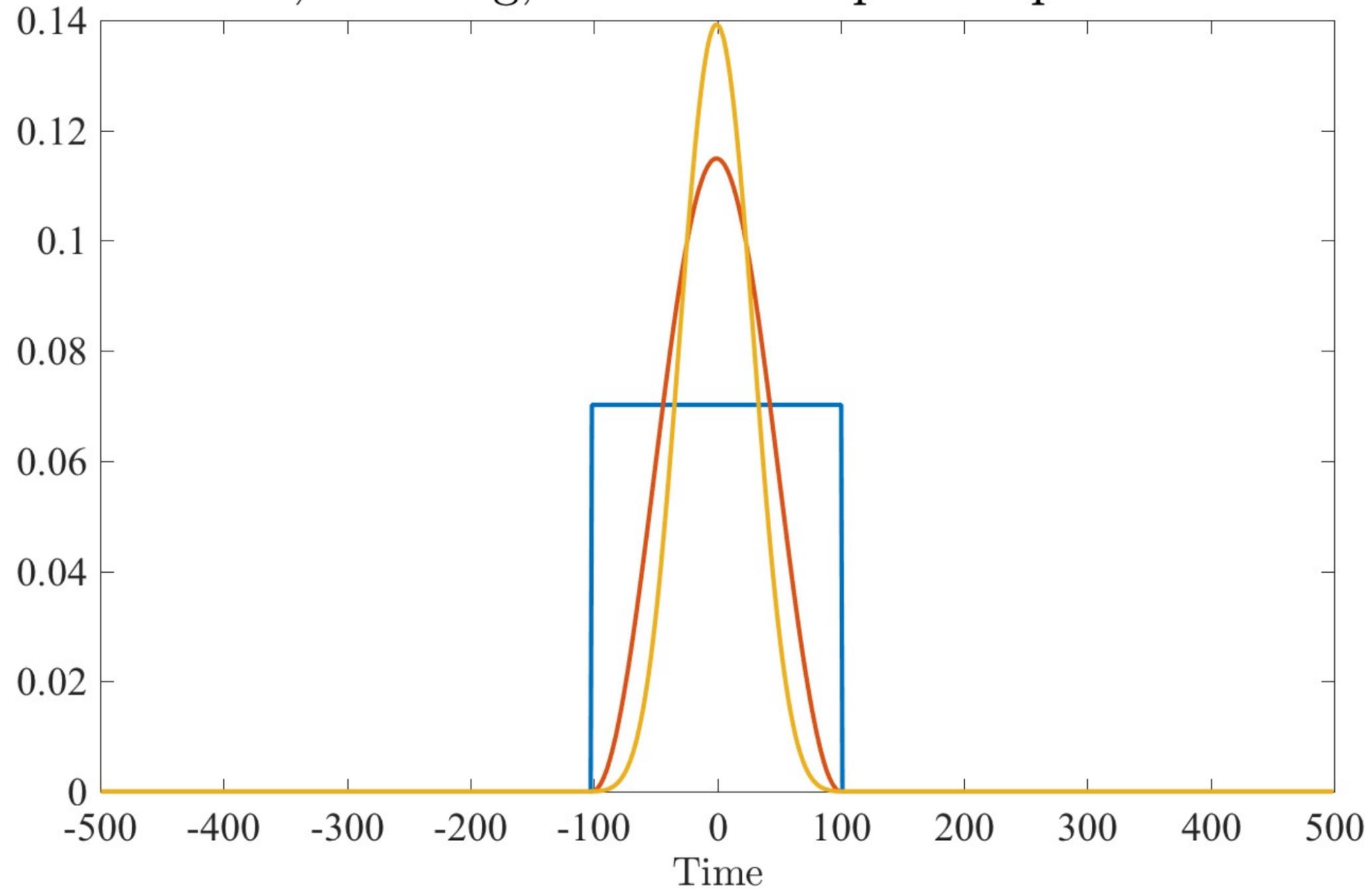
This is the source of the error shown in the motivating example.

Next we take a look at three different tapering functions and their squared Fourier transforms. The broadband bias is most clear if we use logarithmic scaling for the y -axis.



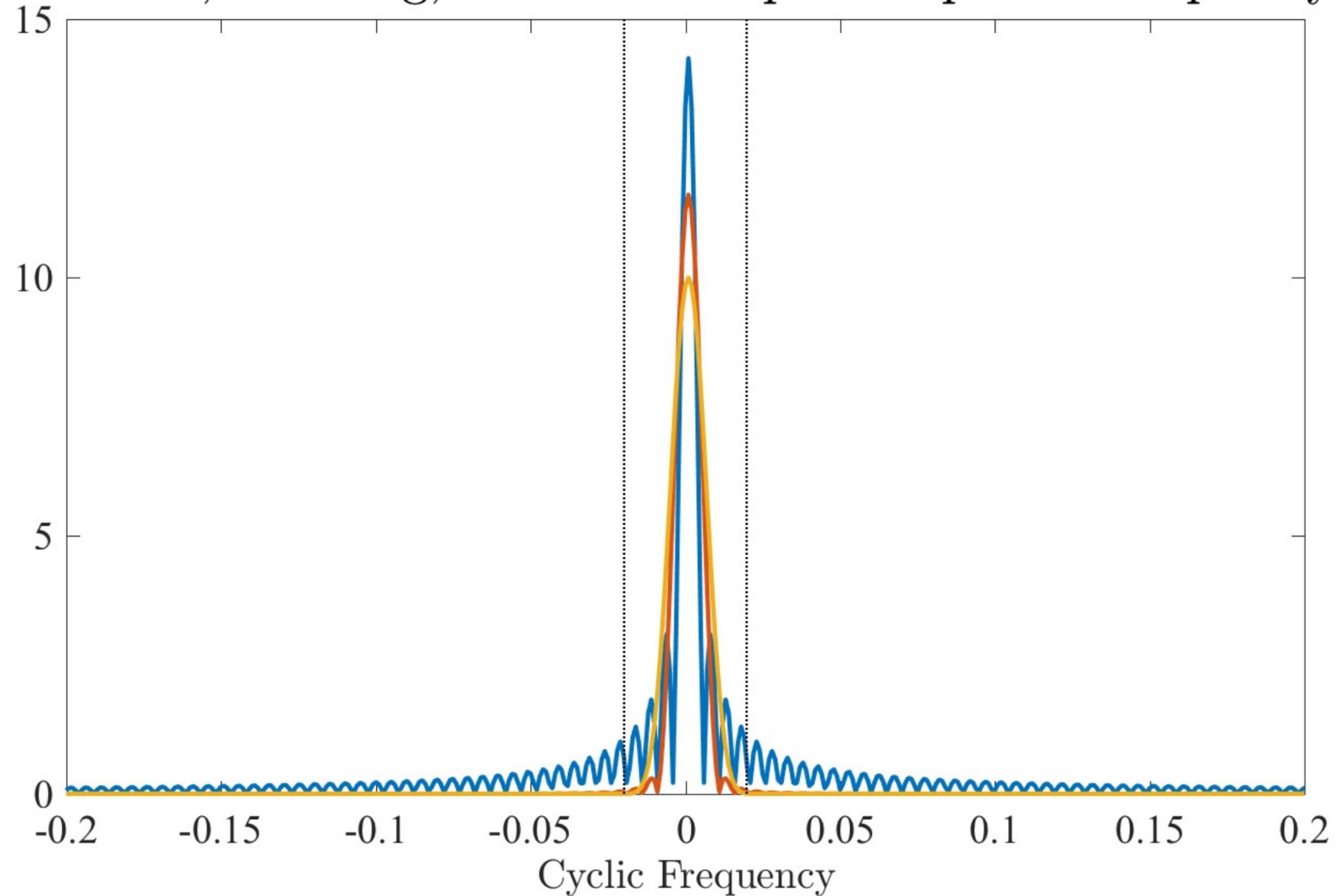
Three Tapering Windows

Boxcar, Hanning, and First Slepian Taper in Time



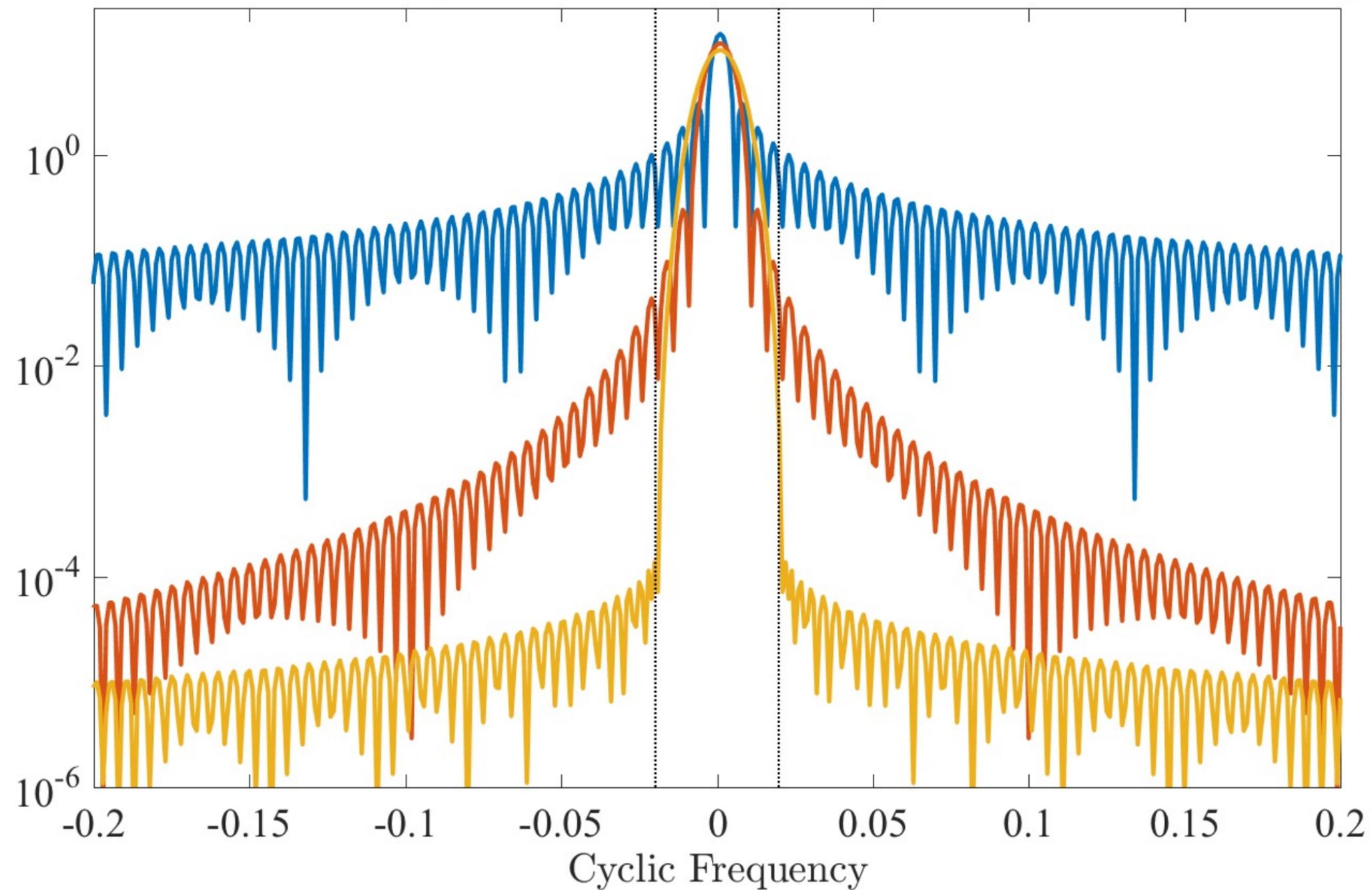
Three Tapering Windows

Boxcar, Hanning, and First Slepian Taper in Frequency



Three Tapering Windows

Boxcar, Hanning, and First Slepian Taper in Frequency



Epilogue

During the practical session this afternoon we will cover the material presented this morning, as well cover some of the topic of filtering.

Thank you!

Shane Elipot

email: selipot@rsmas.miami.edu

