Sequential Data - 1D, cont.

Joseph Fourier 1768 - 1830

Three-way conflict – priesthood/math/politics

Jailed in 1794 for speaking out against the terror. Freed 1794.
Ecole Normale – tutors Lagrange and Laplace

1797 – chair Ecole Polytechnique, post-Lagrange

1798 – joined Napoleon's army, invaded Egypt –
Battle of the Nile – oops; stayed in Egypt until 1801,
archaelogy, founded the Cairo Institute

1801- Napoleon patronage: Prefect of Grenoble, overseeing operations to drain the swamps of Bourgogne (Burgundy) and to construct a new highway from Grenoble to Turin.

1804-1807 finally: theory of heat, done with series expansion + controversy:
"The first objection, made by Lagrange and Laplace in 1808, was to Fourier's expansions of functions as trigonometrical series …"

In review of #17 ....

1. We considered why we should wish to **Fourier 1D data** – a scan: baseline assessment, noise property assessment, signal detection, filtering/signal search, period search, cross- or self-correlation for power spectrum.

2. We want to carry out data transformations with **orthogonal basis functions** that must be up to **reverse transformation and discrete sampling**. Basis sets which give a diagonal **C** compress data and isolate noise with max efficiency ⇔the **Karhunen-Loeve equation** from which **Eigenvectors lead to basis functions** – maybe the trig functions of Fourier? Or not.

3. **Fourier is the basis king** – cf the **FFT**, oscillation, rotation, orbits, atomic spectra, period-finding…Many useful properties: **autocorrelation⇔power spectrum, Gaussian ⇔ Gaussian, shift theorem, sine ⇔ delta function, etc.**

4. Discrete Fourier Transform (DFT) – **convolution of fn with comb fn** for uniform sampling; 5 aspects for success – **Nyquist limit ($\pi/\Delta t$), sampling theorem limit, aliasing, scan-length, s/n.**

5. Statistical properties, worked out for Gaussian, no signal case, indicate difficulties with **consistency and bias for power spectra**, **correlation errors** in correlation functions, and the need to sacrifice **detail in estimating response functions**.

Filtering - to Reduce Noise, Compress Data

If noise is shot noise or photon noise, it is **'white' - and its spectrum extends flat to the limit given by the sampling theorem**.

Recall that the FT of a Gaussian is another Gaussian - so that if instrumental response or line-shape is ~Gaussian, there should be little high-frequency information.

=> **tapering off the amplitudes of high frequencies is a winning strategy.**

It is simple **to manipulate the transform of the data to cut out the higher frequencies**.

Whatever we do by chopping out or reducing the amplitudes at high frequencies is bound to decrease the noise - but **it must decrease some signal as well**, particularly on small scales in the spatial domain.

Filtering - to Reduce Noise, Compress Data, cont.

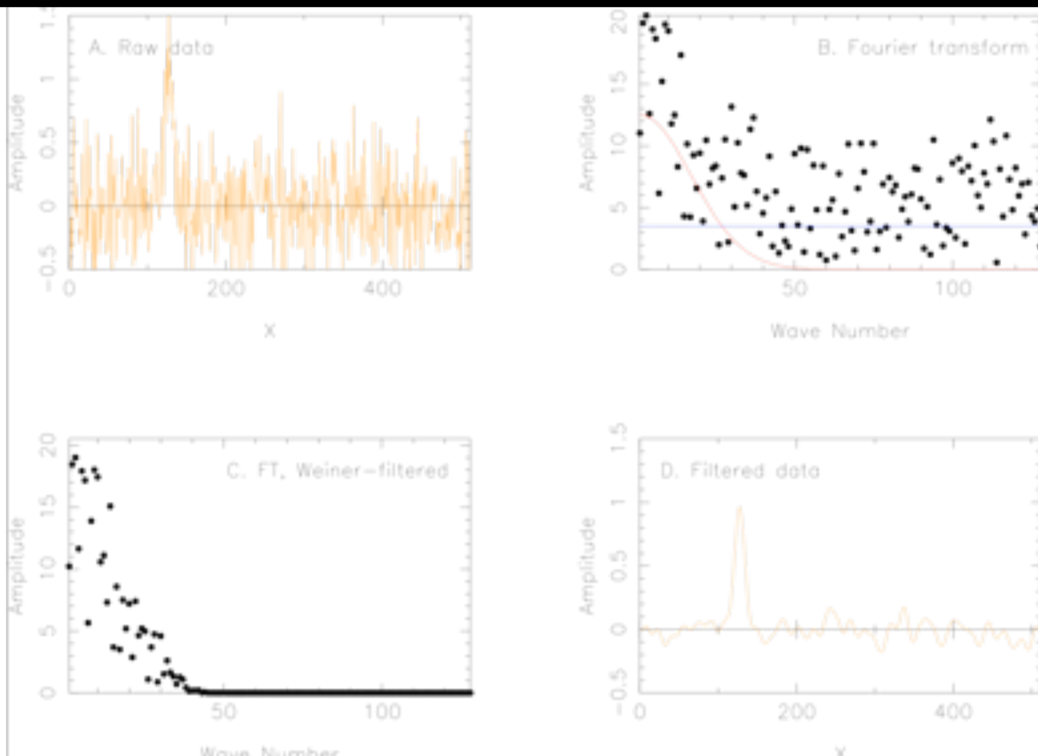
Square filters produce **ringing in the signal**, so that a **tapering** to high frequencies is desirable. There are books full of techniques. A general approach: it is readily shown both by minimizing the variances and by conditional probabilities that an estimate of the optimum filter is given by

$$F(f) = \frac{| S(f) |^2}{| S(f) |^2 + | N(f) |^2}$$

where S is the signal and N the noise.

This is **Wiener filtering**. It requires us **to assess or model the FT of both noise and signal**. This is difficult of course if signal and noise have such similar power spectra - but then, no filter can cope under these circumstances.

Low-Pass Filtering - Weiner Filtering Example



The raw data of A is a Gaussian sitting on a flat baseline, with random Gaussian noise added. The DFT in B shows the signal and noise components, modelled by the Gaussian and horizontal curves respectively. The Wiener filter, applied in the frequency domain, produces the DFT of C, and the reverse transform produces the greatly improved s/n of D.

The procedure is robust; approximate the signal transform with a triangle and the noise with a straight line to get very similar results.

Causal filters (e.g. Kalman) use only 'past' data.
Savitsky-Golay filters: low-order polys fitted to a sliding window; see NumRec.

High-Pass Filtering: Minimum-Component Baselines

High-pass filtering is getting rid of unwanted low frequencies; and this is known in the trade as fitting baselines, or **baseline assessment**.

Heavy smoothing? polynomial fits? spline fits?

The signal is the problem. Note that the transform of a Gaussian is a Gaussian! They are big at low frequencies.

Minimum-component filtering (Wall 1997):

(1) identify regions of clear or possible signal.
(2) patch these to form a segmented baseline array.
(3) end-match with e.g. a linear fit.
(4) form the DFT of the resultant ‘baseline array’.
(5) remove the high frequencies from this with savage multiplicative filter in the FT plane to get rid of any semblance of noise.
(6) reverse-transform to get the real baseline.
(7) restore the gradient removed at (3).

Minimum-Component Baselines, cont.

Error assessment can be carried out with MC analysis. It shows:

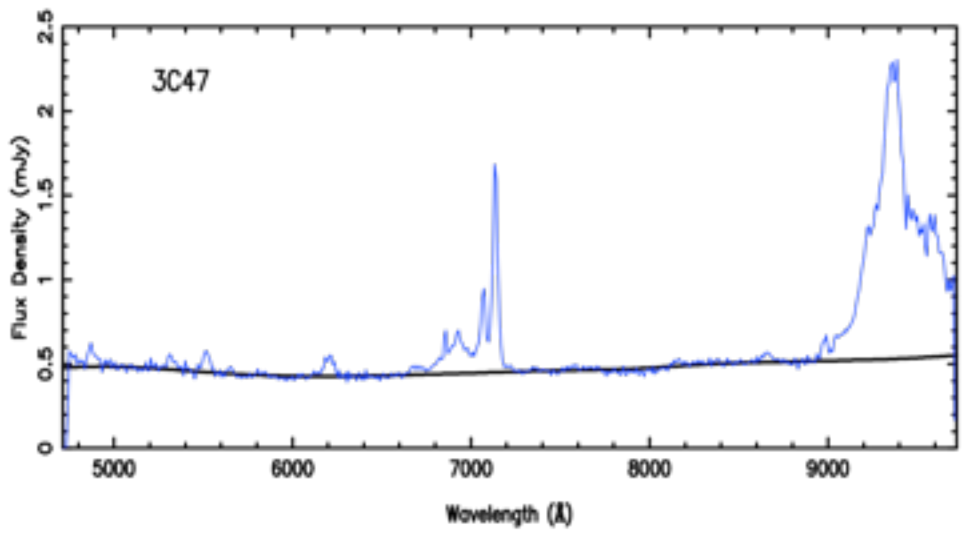
- **patch width is critical**; patches should not extend beyond $\pm 2\sigma_s$ for weak signals, further for strong signals.

- even the **weakest signals must be patched** to get unbiased measurements.

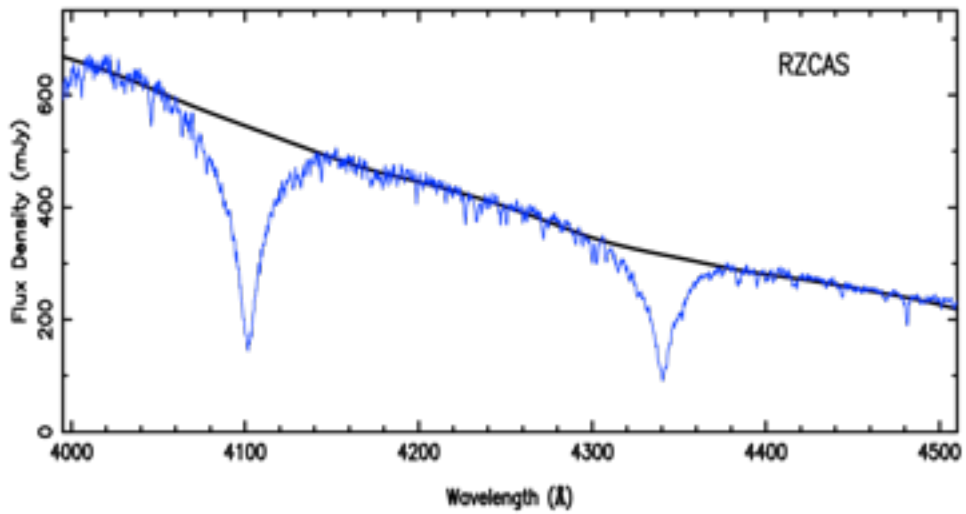
- **heavily curved baselines** rapidly escalate the errors in signal strength measurement.

Minimum-Component Baselines: Example

(Above) A spectrum of 3C47 obtained with the Faint Object Spectrograph of the William Herschel Telescope, La Palma. The redshift is 0.345; broad lines of the hydrogen Balmer series can be seen, together with narrow lines of [OIII].



(Below) A spectrum of the A star RZ Cas (Maxted et al. 1994).
The continuum obtained with the minimum-component technique is shown as the black line superposed on the original data.



The Coherence Function

Functions ***f*** and ***g***: the coherence function is estimated by

$$\hat{\gamma}^2(\omega) = \frac{| \hat{F}(\omega)\hat{G}(\omega)^* |^2}{| \hat{F}(\omega) |^2 \, | \hat{G}(\omega)^* |^2}.$$

with estimation done by either smoothing the power spectrum, or averaging several power spectra derived from separate scans. The coherence function is the correlation coefficient between ***f*** and ***g*** in frequency space.

The coherence is extremely useful in cases where we have an input ***f*** and and output ***g*** and we want to find out out more about the **"black box"** that changes ***f*** into ***g***.

- purely linear? then ***g = f ¤ h*** for some ***h***, and the coherence function ***γ= 1***.
- more likely, noise, so ***g = f ¤ h + ε***
- depending on the frequency content of the noise and the input, we will have structure to ***γ***, which will generally be less than one.
- other interesting reasons for ***γ < 1*** will be that the causal relationship between ***f*** and ***g*** is non-linear, or extra causal factors are in play. The coherence will be lower.

The Coherence Function - Black-Box Example

EXAMPLE: We have a relationship for some synthetic data

$$g(t) = f \otimes h + \epsilon(t) + b(t)$$

in which ***f*** is white Gaussian noise, ***h*** is a Gaussian filter, $\boldsymbol{\varepsilon}$ is noise added at the output side of the box, and ***b*** is an unrelated low-frequency effect (obtained in the following case by vigorous recursive filtering of Gaussian white noise).

Coherence Function - BB Example cont.

a) the input data ***f***,

b) the input convolved, with some noise added ***f ¤ h + ε,***

c) the extraneous effect ***b(t),***

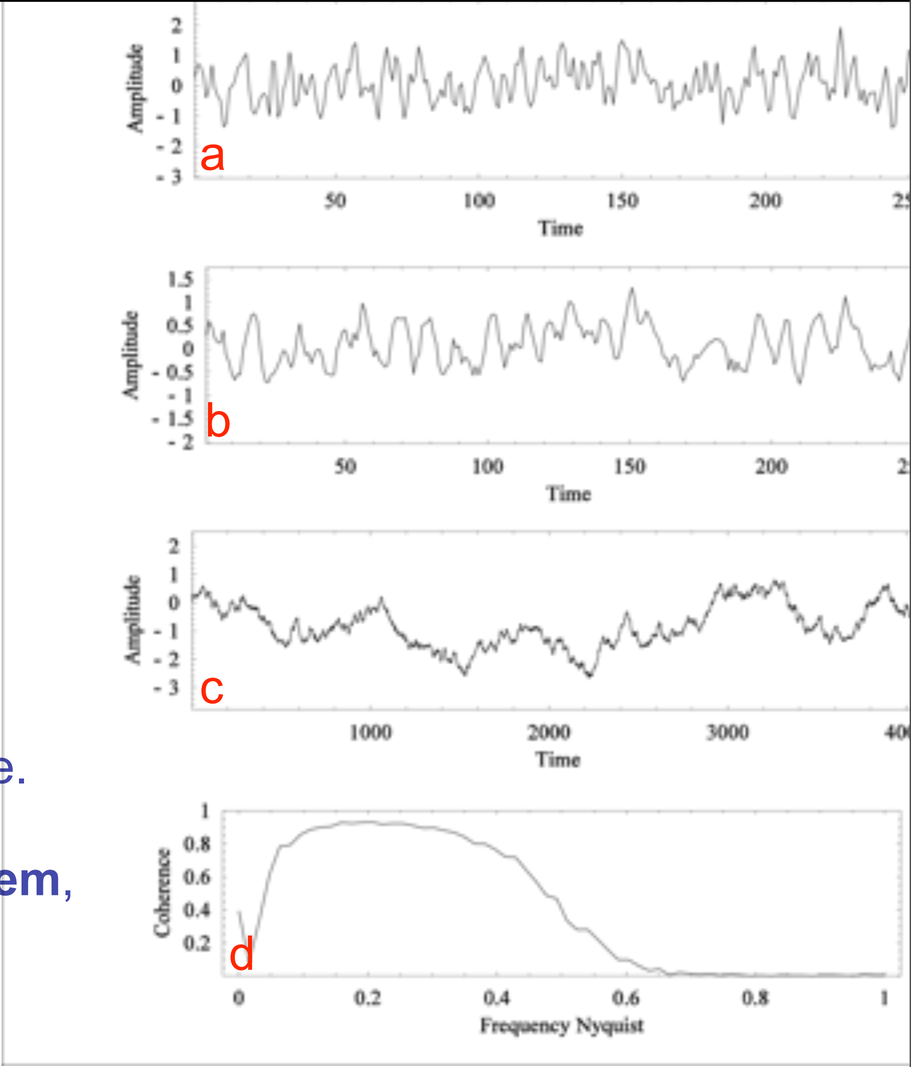
d) the coherence between ***f*** and ***g***.

- loss of coherence at low frequencies (because of the extra effect)

- loss at high frequencies (due to noise + smoothing by instrumental response.

- **intermediate frequencies - a linear system**, where only the input ***f*** affects the output ***g***.

- **can model our box as a simple convolution of input data with an instrumental function**; we also suspect that there must be an **extra causal effect at low frequencies**. This analysis yields the clean part of the spectrum.



END