FOURIER TRANSFORMS

$$F(w) = \int_0^T f(t)\ e^{-iwt} dt \qquad f(t) = \frac{1}{2\pi} \int_{-w_N}^{w_N} F(w)\ e^{iwt} dw$$

$$F_k = \sum_{n=0}^{N-1} f_n\ e^{-2\pi ikn/N} \qquad f_n = \frac{1}{N} \sum_{k=0}^{N-1} F_k\ e^{2\pi ikn/N}$$

Properties

1. Shift origin by Δt $\quad F \rightarrow F\ e^{iw\Delta t}$

   $\mid F \mid$ invariant $\qquad \mid F \mid^2 =$ power spectrum

2. If $f(t)$, f_n real then $F(w) = F^*(-w)$ and

   $F_k = F^*_{-k} = F^*_{N-k} = F_{N+k}$ ie. cyclic

3. Parseval's theorem

$$\frac{1}{2\pi} \int_{-w_N}^{w_N} \mid F(w) \mid^2 dw = \int_0^T f^2(t)\ dt$$

$$\frac{1}{N} \sum_{k=0}^{N-1} \mid F_k \mid^2 = \sum_{n=0}^{N-1} f_n^2$$

Compute DFT using an FFT algorithm $\rightarrow N \times log_2(N)$

See Brault & White 1971, A&A, 13 169, for a tutorial introduction.

CONVOLUTION

$$g(x) = \int_{-\infty}^{\infty} f(u)\ h(x-u)\ du \quad \rightarrow\ f \otimes h$$

FT $\updownarrow$

$$G(w) = F(w)\ H(w)$$

CORRELATION

$$g(x) = \int_{-\infty}^{\infty} f(u)\ h(x+u)\ du$$

FT $\updownarrow$

$$G(w) = F(w)\ H^*(w)$$

If h symmetric – correlation and convolution are the same.

POWER SPECTRUM

$$Power\ spectrum = \mid F(w) \mid^2$$

FT $\updownarrow$

$$Autocorrelationfunction = \int_{-\infty}^{\infty} f(u)\ f(x+u)\ du$$

WINDOW FUNCTIONS

$$f_{obs}(t) = f(t)\ w(t)$$

where $w(t) = 1 \quad 0 \leq t < T; \quad = 0 \quad$ elsewhere

$$F_{obs}(w) = F(w) \otimes W(w)$$

<u>SHANNON'S SAMPLING THEOREM</u>

Shannon 1949

If $f(t)$ is a continuous signal bandlimited such that $-\nu_N \leq \nu \leq \nu_N$ then $f(t)$ can be **completely** specified by sampling at an interval

$$\Delta t = 1/2\nu_N$$

$$f(t) = \sum_n x_n \, \frac{sin(2\pi\nu_N t - n\pi)}{2\pi\nu_N t - n\pi} \;\; \equiv \;\; x_n \otimes \frac{sin(2\pi\nu_N t)}{2\pi\nu_N t}$$

where $x_m = f(m\Delta t)$

Expanding $f(t)$ as a series of orthogonal functions. If $0 \leq t \leq T$ then $f(t) \equiv$ point in an $2T\nu_N$ dimensional space.

Also known as the perfect interpolation formula.

Note that real signals are bandlimited in both the signal & Fourier domain.

Commonly used interpolation methods include:

nearest neighbour; binlinear; bicubic spline; and assorted Lanczos variants.

COMBINING SIGNALS AND NOISE

Consider a series of zero-mean random variables (ie. noise in some signal) $\{x_1, x_2, x_3, ...x_n\}$, form a linear combination

$$y = \sum_{k=1}^{n} a_k \; x_k$$

what is the *noise* in y ?

From CLT $y \rightarrow$ Gaussian distribution, in this case with zero-mean, and variance σ^2.

$$var\{y\} = < y^2 > = < \sum_k a_k \; x_k \; \sum_j a_j \; x_j >= \; \sum_{k \; j} a_k \; a_j \;\; < x_k x_j >$$

$$\sigma^2 = var\{y\} = \mathbf{a^T C \; a}$$

where **a** is the vector of coefficients, **C** is the covariance matrix with elements $c_{kj} =< x_k \; x_j >= \sigma^2_{kj}$ and $^{\mathbf{T}}$ denotes transpose.

Special cases:-

For independent variables $\sigma^2 = \sum_k \; a_k^2 \; \sigma^2_{kk}$

For $\sigma_{kk} = \sigma_{noise}$ and $\sum_k \; a_k = 1$, (a filter), noise is reduced by $\sum_k \; a_k^2$.

If a_k are coefficients derived from a normalised Gaussian filter, the "noise" variance is reduced by $1/\sqrt{4\pi\sigma_g}$ in 1D and $1/4\pi\sigma_g^2$ in 2D.

MAXIMISING SIGNAL:TO:NOISE

.... or aim to keep noise in output to minimum.

Assume all signal levels are suitably normalised to the same average level (by pre-scaling). Form a new signal by

$$r_{new} \;=\; \sum_k w_k \; r_k \;=\; \sum_k \; w_k(s_k + n_k)$$

Constrain weights $\sum_k \; w_k = 1$ $\quad\Rightarrow$ signal level unchanged

Output noise variance $= \sum_k w_k^2 \sigma_k^2$ $\quad$ minimise subject to $\sum_k w_k = 1$

Implies the optimum weights are given by

$$w_{j_{opt}} = \frac{1/\sigma_j^2}{\sum_j 1/\sigma_j^2}$$

and the output noise in this case is

$$\sigma_{out}^2 = \frac{1}{\sum_j 1/\sigma_j^2}$$

PERIODICITY ESTIMATION I

Estimate period; form of periodic component; degree of periodicity

Classical methods: autocorrelation function / power spectrum

$$\phi(\tau) = \int s(t)\ s(t+\tau)\ dt$$

$$\text{FT} \updownarrow$$

$$\Phi(\omega) = \mid S(\omega) \mid^2$$

Least-squares method (Friedman, 1978 IEEE.....): – model problem

$$s(t) = s_o(t) + a(t); \qquad s_o(t) = s_o(t + k\tau)$$

where $s_o(t)$ is periodic component, $0 \leq t < \tau$, and $a(t)$ is aperiodic component.

$$minimise \qquad I(\tau) = \int_0^T w(t)\ [s(t) - s_o(t)]^2\ dt$$

for all periods τ of interest, where $w(t)$ is the sample window.

$$\approx \quad maximum \quad \sum_{k=1}^{K-1} \frac{\phi(k\tau)}{(K-1)\ \phi(0)}$$

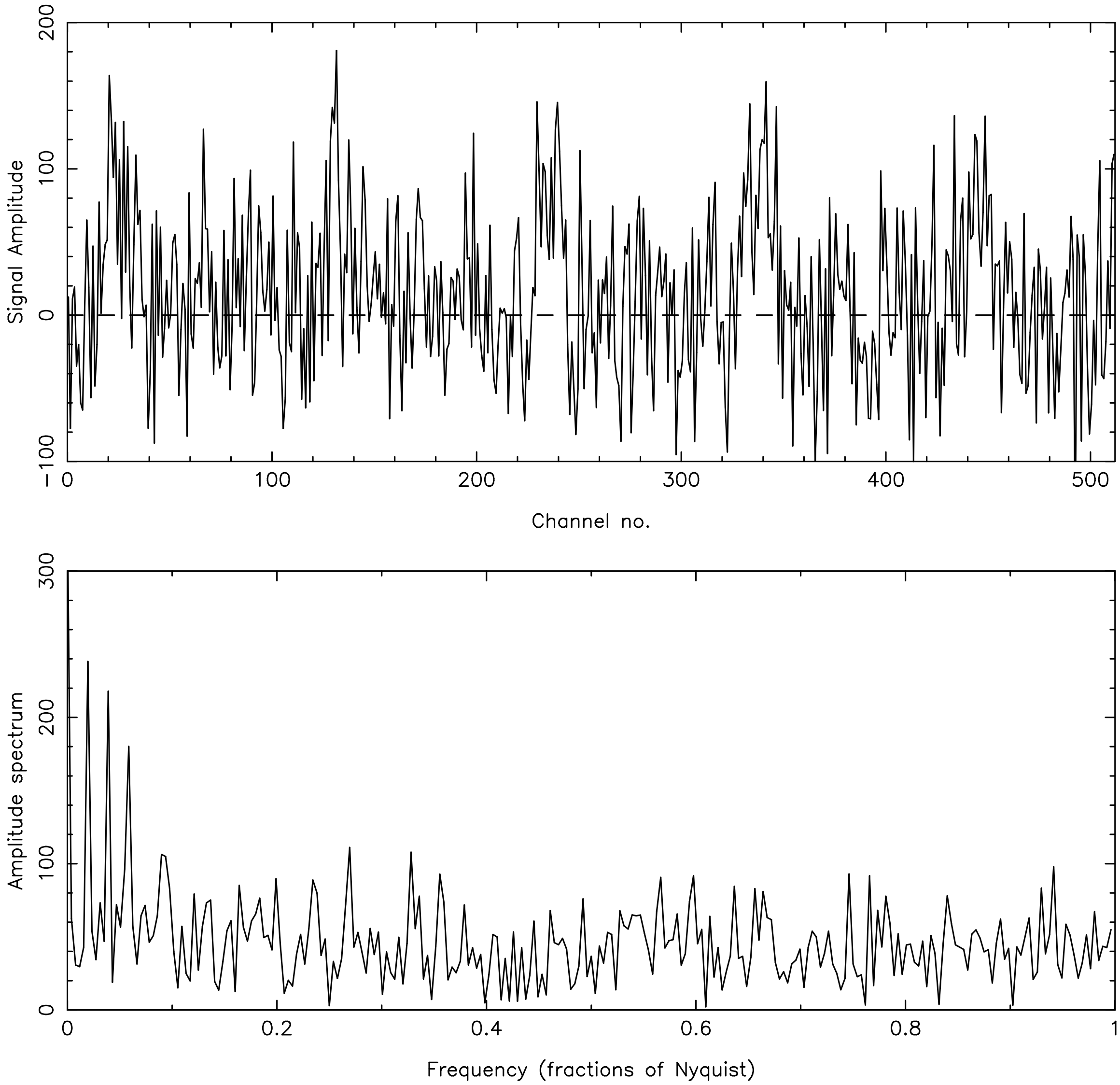

Figure 1: Example of periodic signal with peak signal:noise = 2:1 and the Fourier amplitude spectrum. The fundamental period and the first two harmonics clearly stand out.

PERIODICITY ESTIMATION II

Sparsely sampled data methods

Phase minimisation – Lafler & Kinman 1965 ApJS p216;

Stellingworth 1978 AJ

Minimise "smoothness" of phase-folded light curve

$$\theta = \sum_i (m_i - m_{i+1})^2 / \sum_i (m_i - \bar{m})^2$$

where $m_i \& m_{i+1}$ are adjacent phase magnitudes

Sine-curve model fitting

$$f_t = A\ sin(\frac{2\pi t}{\tau} - \phi) + B$$

Note that the DFT and sine-curve modelling are **exactly** the same method **iff** either $T >> \tau$ or $T/\tau =$ integer and the sampling is complete.

Sine curve fitting is also known as the Lomb-Scargle method, see Press & Rybicki (1989) for fast implementation and references.

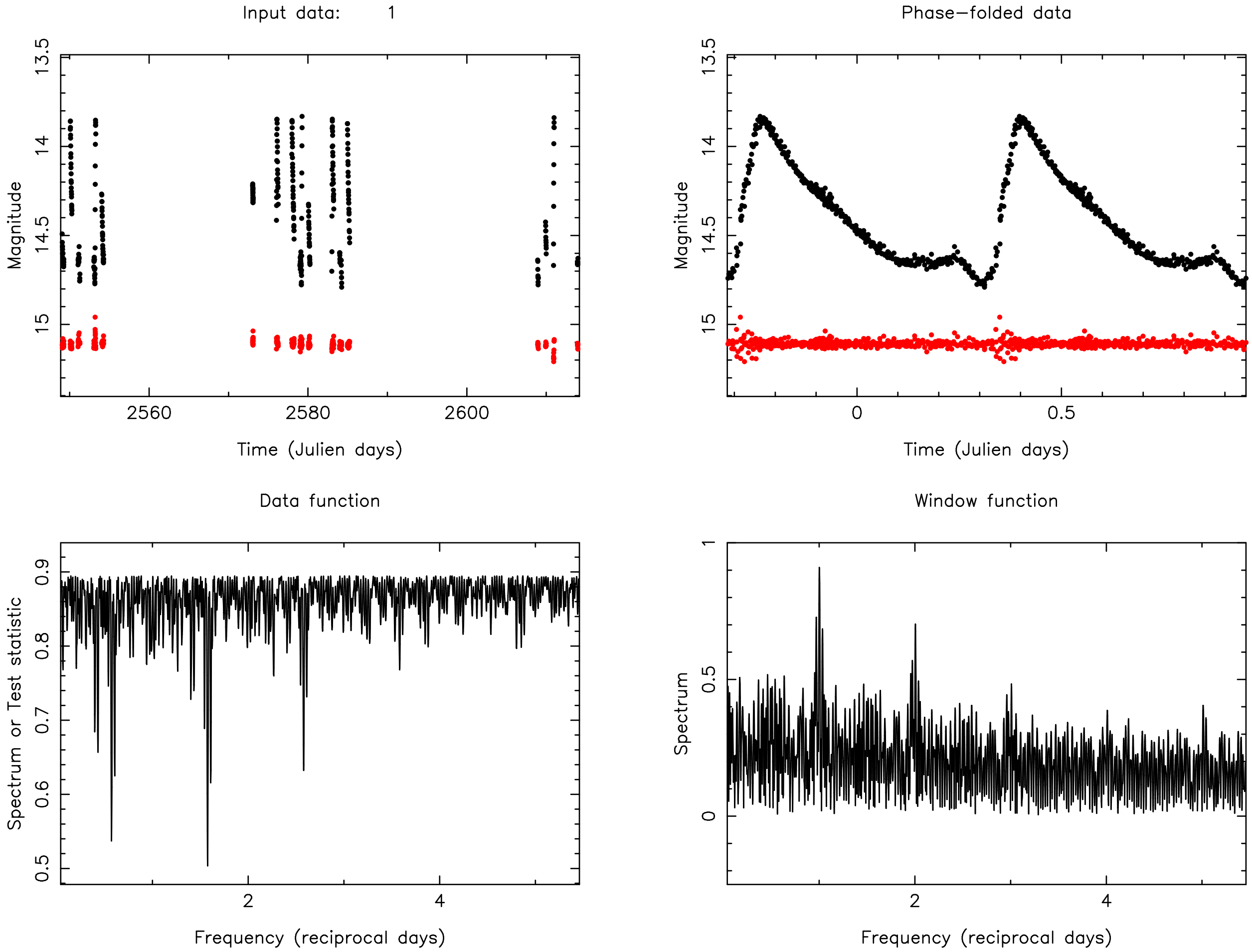

Figure 2: Example of sine curve fitting to determine the period (0.6343 days) and lightcurve of an RRLyrae star: top left input data (black) and residuals after period fitting (red); top right the phase-folded lightcurve (black) and residuals (red); bottom left the period finding statistic; bottom right the Fourier amplitude spectrum of the window function.

CROSS-CORRELATION & MAXIMUM LIKELIHOOD

$$\phi_\tau = \sum_t \; y_t \quad x_{t+\tau}$$

cross-correlation function data reference signal

Aims: <u>detect</u> signal in data; accurately estimate position τ.

Rephrase the problem as model fitting (e.g. radar pulse echo location Woodward & Davies MNRAS 1958)

$$y_t \quad = \quad x_{t+\tau} \quad + \quad \epsilon_t$$

data model residual or noise

Then the Likelihood function is

$$L = \prod_t \; P(y_t \mid \theta_{\mathbf{t}}) = \prod_t \; P(\epsilon_t)$$

Assume *initially* independent Gaussian noise with variance σ_t^2, then the likelihood of the data is given by

$$L(y \mid x, \tau, \sigma_t^2) = \prod_{t=1}^{N} \; P(\epsilon_t)$$

$$L(\tau) = (2\pi)^{-N/2} \; (\prod_{t=1}^{N} \sigma_t^2)^{-1/2} \; exp\left[-\sum_{t=1}^{N} (y_t - x_{t+\tau})^2 / 2\sigma_t^2 \right]$$

$$ln(L) = const - \sum_{t=1}^{N} (y_t - x_{t+\tau})^2 / 2\sigma_t^2$$

Maximum likelihood $\equiv$ least-squares $\equiv$ cross–correlation

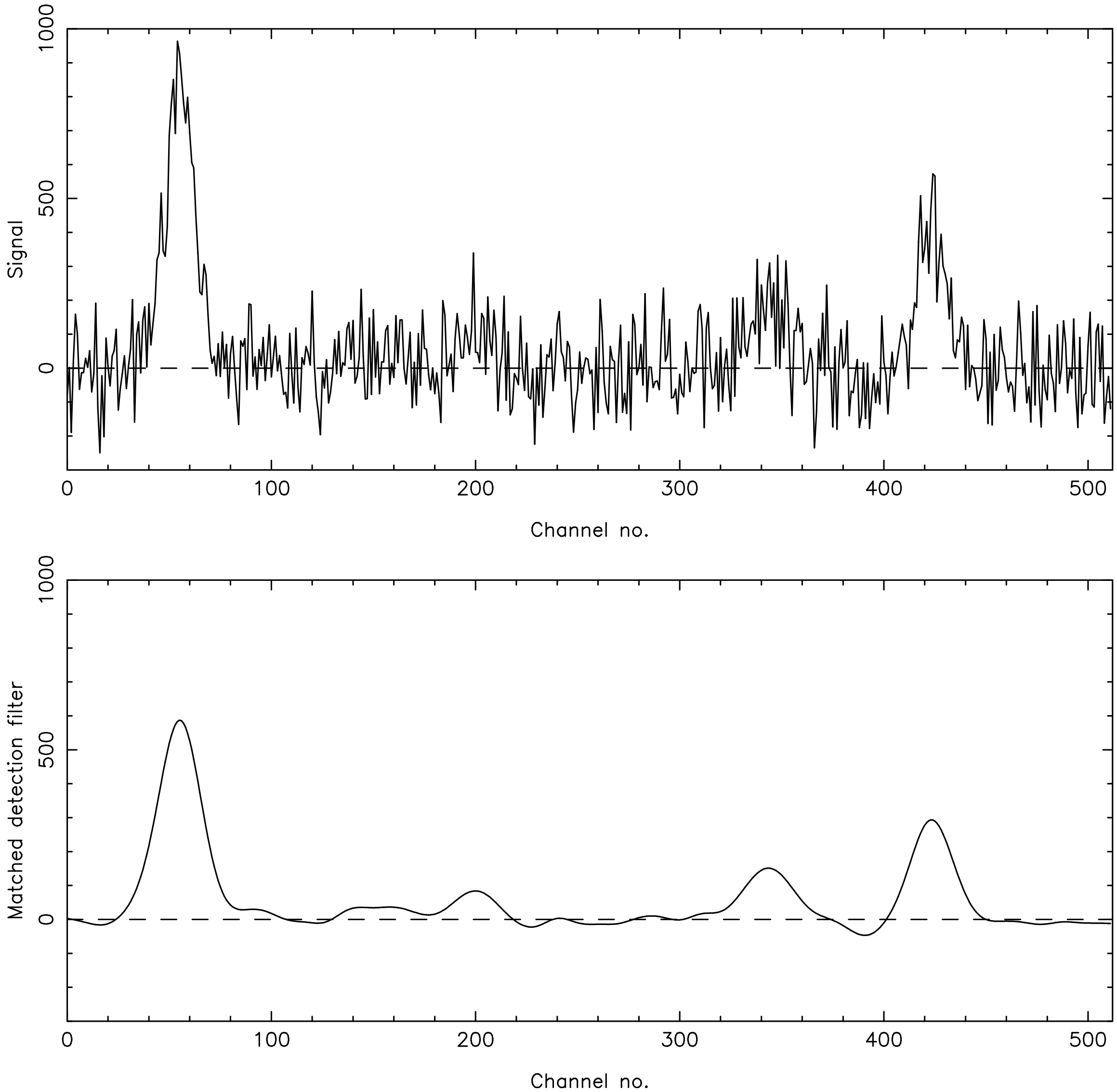

Figure 3: Simulation of radar pulse echo location using pulses with peak signal:noise of 1,2,4,8. The top panel shows the input data and the bottom panel the results of applying a matched detection filter (cross-correlation).

CROSS-CORRELATION

Aims: detect signal in noise and accurately estimate relative shift

$$\phi_\tau = \sum_t y_t \; x_{t+\tau}$$

Normalise to lie in range ± 1 since

$$\phi_\tau \leq \sqrt{\sum y_t^2 \;\; \sum x_t^2}$$

Used most often in astronomy for estimating redshifts and intrinsic velocity dispersions eg. galaxies, quasars, stellar clusters

Fourier quotient method (Sargent et al. 1977 ApJ 212 326)

$$g(\lambda') \propto s(\lambda') \otimes b(\lambda') \otimes \delta(\Delta\lambda')$$

$$\text{FT} \updownarrow$$

$$G(k) = \gamma \; S(k) \; exp\left[-\frac{1}{2}(\frac{2\pi k\sigma}{N})^2 + \frac{2\pi ik\delta}{N}\right]$$

Direct method (Tonry & Davies 1979 AJ 84 1511) – maximise

$$g(\lambda') \otimes t(\lambda')$$

* Note – λ' denotes $log(\lambda)$ binning

Practicalities of Cross-Correlation

- cross-correlation $\equiv$ convolution $\equiv$ optimal matched detection // faint spectral features, images in 2D data
- choosing radial velocity standards $\leftrightarrow$ template matching, classification
- rebinning to $log(\lambda)$ necessary in general since $\lambda_{obs} = (1 + z)\ \lambda_{ref}$
- continuum removal (ie. slowly varying spatial components such as DC level, slopes etc....) usually necessary, also called rectifying
- Apodizing/ windowing to deal with "edges" and
- Fourier computation -v- spatial computation $O(n^2)$ -v- $O(nlogn)$
- correcting to Helio-centric, LSR, Galactocentric velocity systems
- position of objects in slit can cause velocity shifts
- sky absorption lines or residuals from sky lines during spectral extraction in low signal:noise data $\Rightarrow$ lock on to wrong velocity
- error estimation *e.g.* Tonry & Davis (1979) messy, but treat as ML or LS problem $\rightarrow$ alternative