Today's course

Chapter 1.3, 1.4 & 2.2 OAF-2
Numerical Recipes Chapter 13.3

Topics:

- Fourier transformations
- Aliasing & Nyquist frequency
- (Optimal) Filtering
- measuring moments of a stochastic process

Recap lecture 1

- detection of astronomical signals (s.p.) plus noise is convoluted by instrument transfer function and data sampling
- statistical moments characterise the signal (plus noise)
- Noise can be due to the detector, background, and/or intrinsic to the signal
- Assume WSS s.p. (mean does not depend on time, or much slower than measuring process, auto-correlation depends on offset only)
- Convolutions and cross-correlations

More on the auto-correlation

$$R(x) = E\{\underbrace{f(x)}_{f(x_1)}\underbrace{f(x+t)}_{f(x_2)}\}$$

if $x_1 = x_2$

$$R(x) = R(x,x) = \mathbf{E}\{f^2(x)\} = \mathbf{E}\{|f(x)|^2\}$$

Average generally not zero

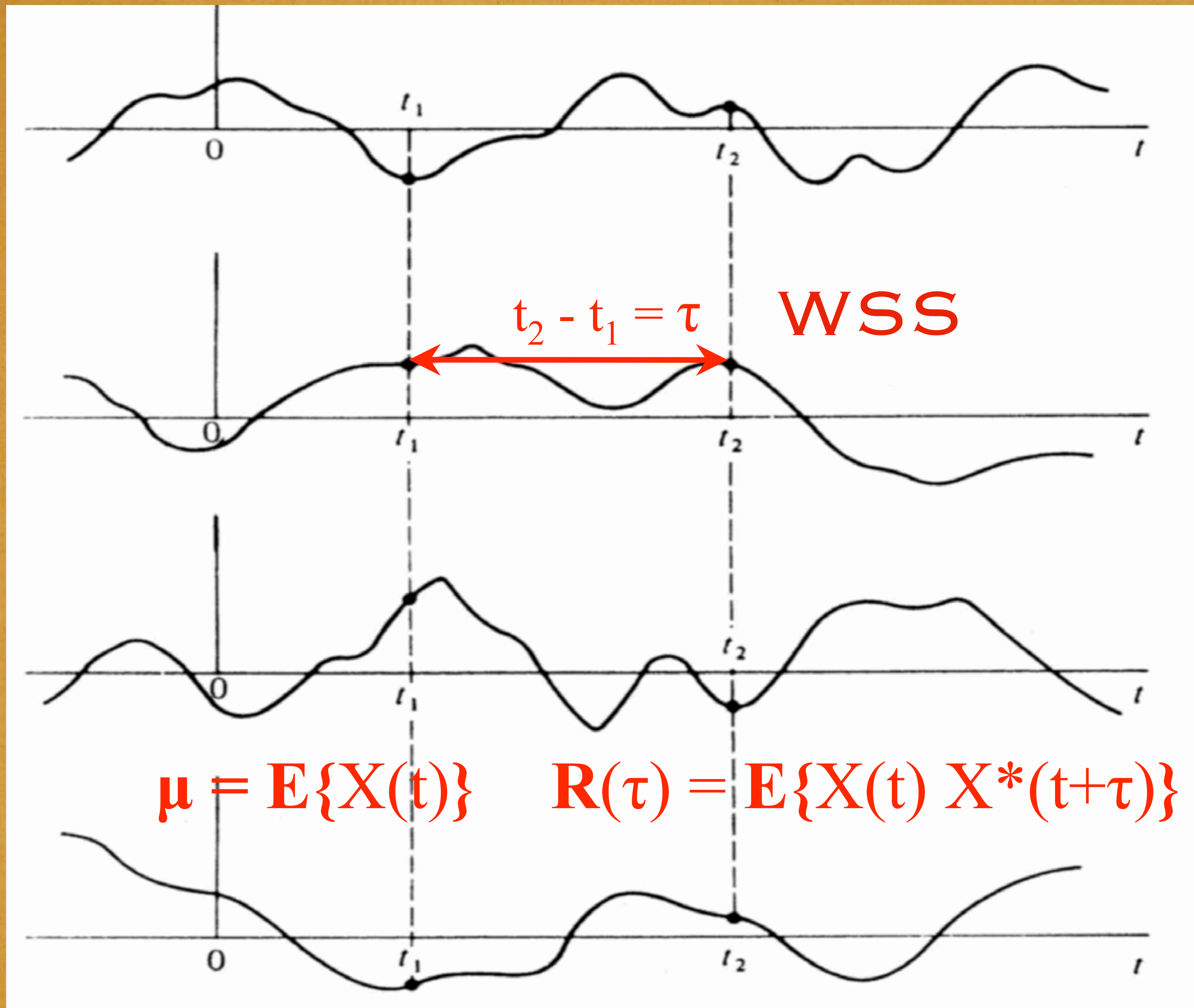
→ autocovariance

$$C(x_1, x_2) = \mathbf{E}\{(f(x_1) - \eta(x_1))(f(x_2) - \eta(x_2))^*\}$$

$$C(x) = R(x) - |\eta(t)|^2 = \sigma^2(x)$$

C(x) → average power in the fluctuations around the mean

Wide-sense stationary S.P.



WSS: mean time independent

& autocorrelation depends on

time difference

Not all signals are WSS:

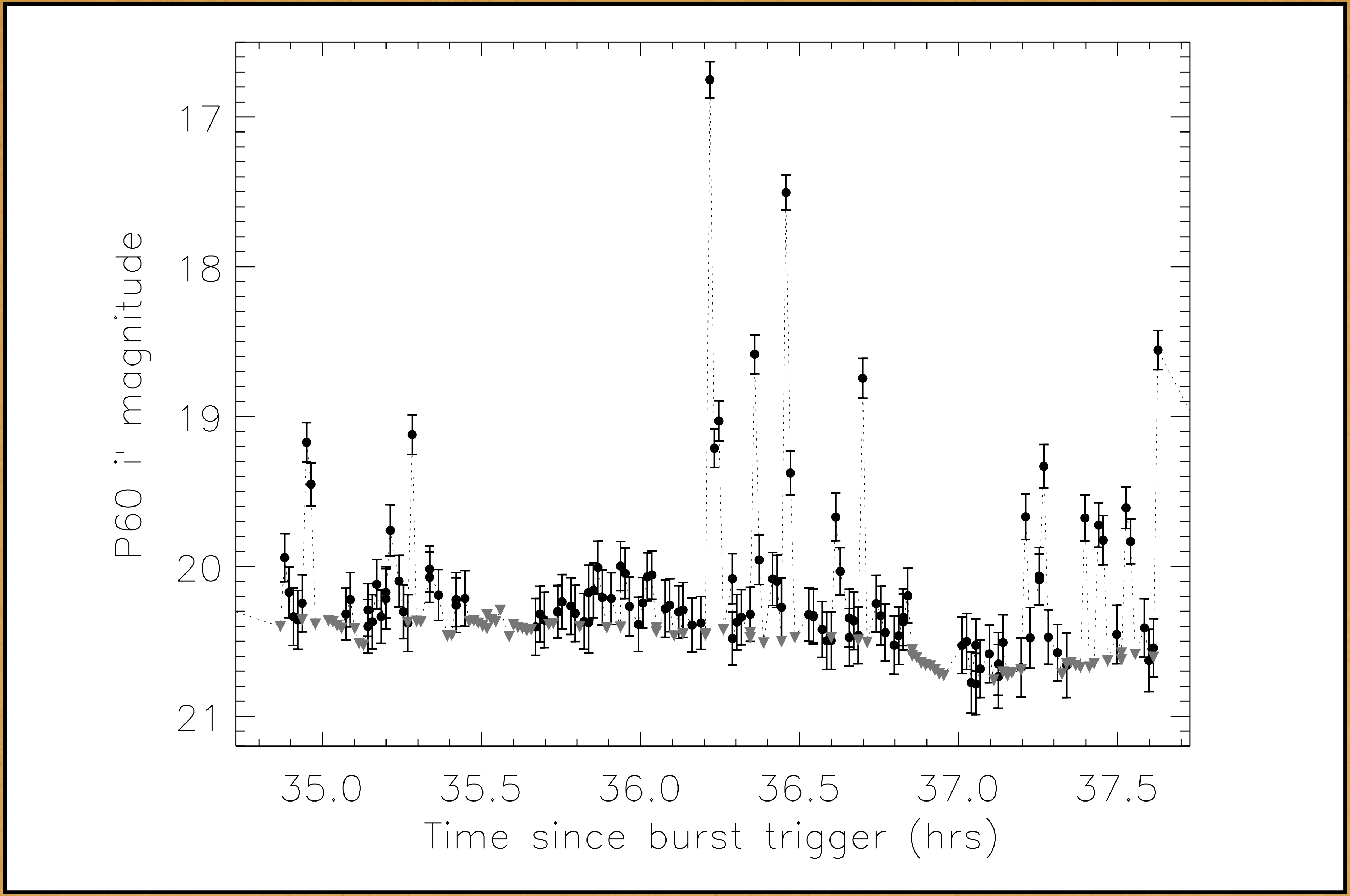
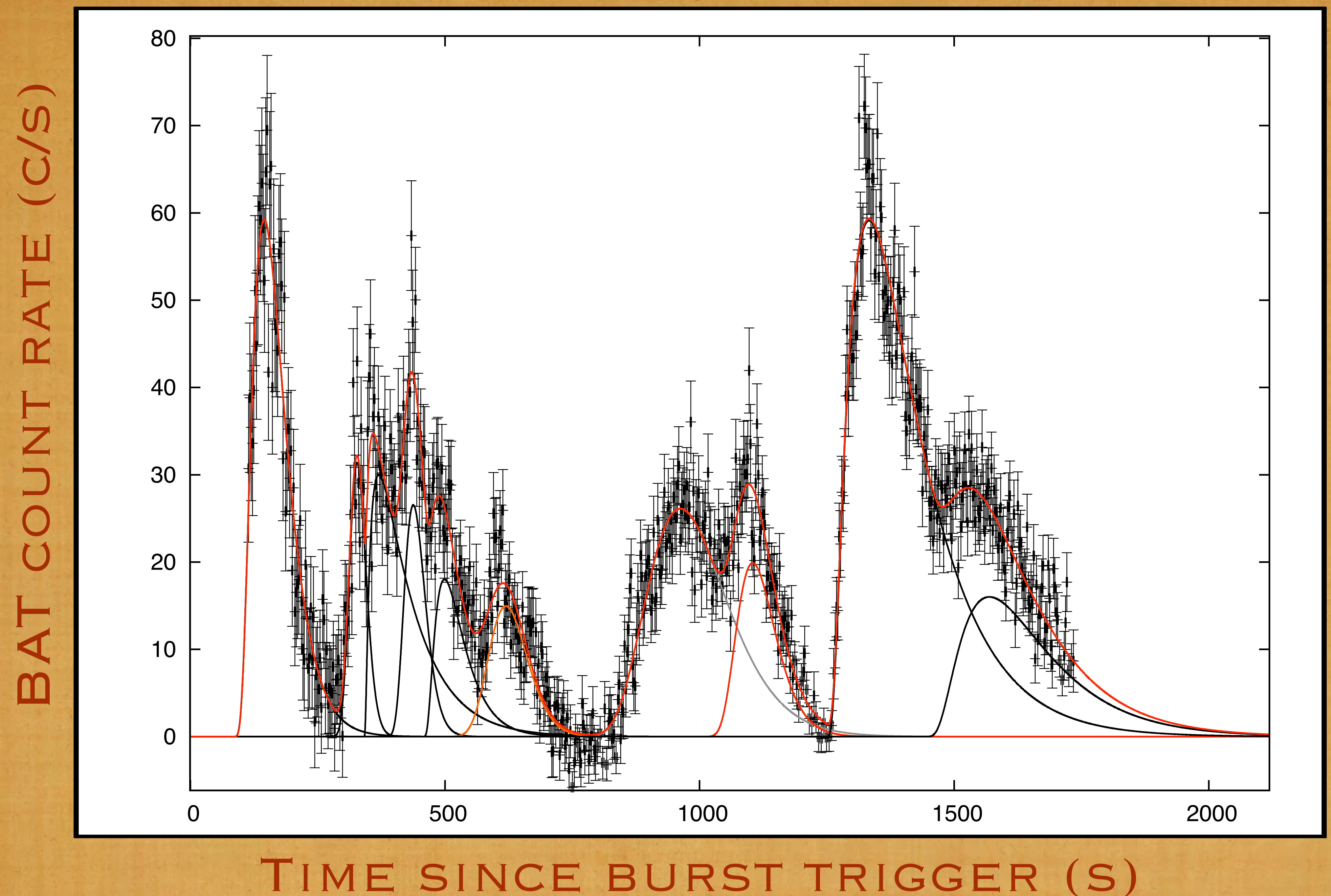


Figure from Kasliwal et al. 2007

Nature of the source uncertain:
GRB, SGR, BH-X-ray binary?

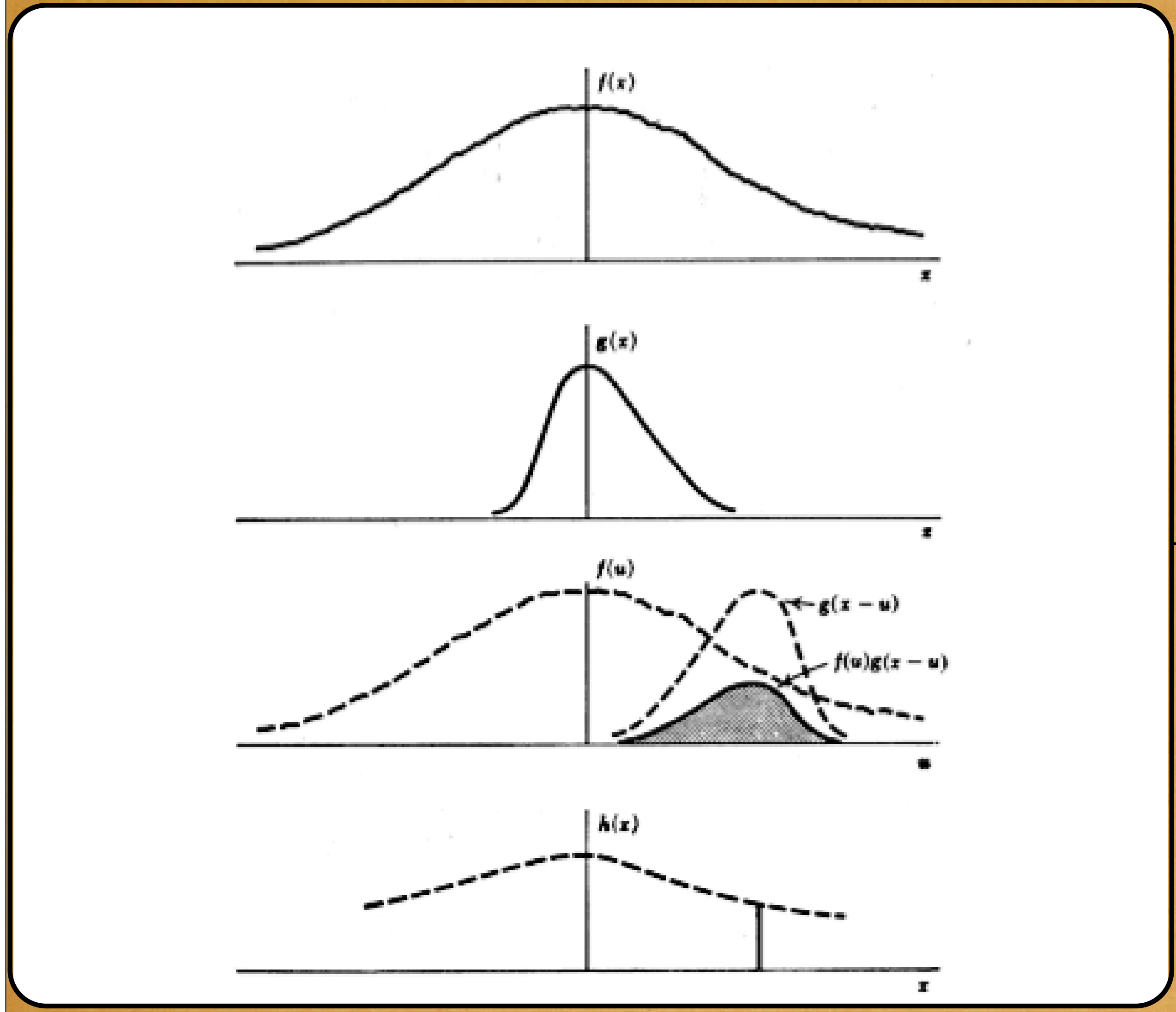
Not all signals are WSS:



Swift BAT detector light curve
of a Gamma Ray Burst
Figure from Chincarini et al. 2008

CONVOLUTION:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x) g(x_1 - x) dx$$



Convolution Theorem

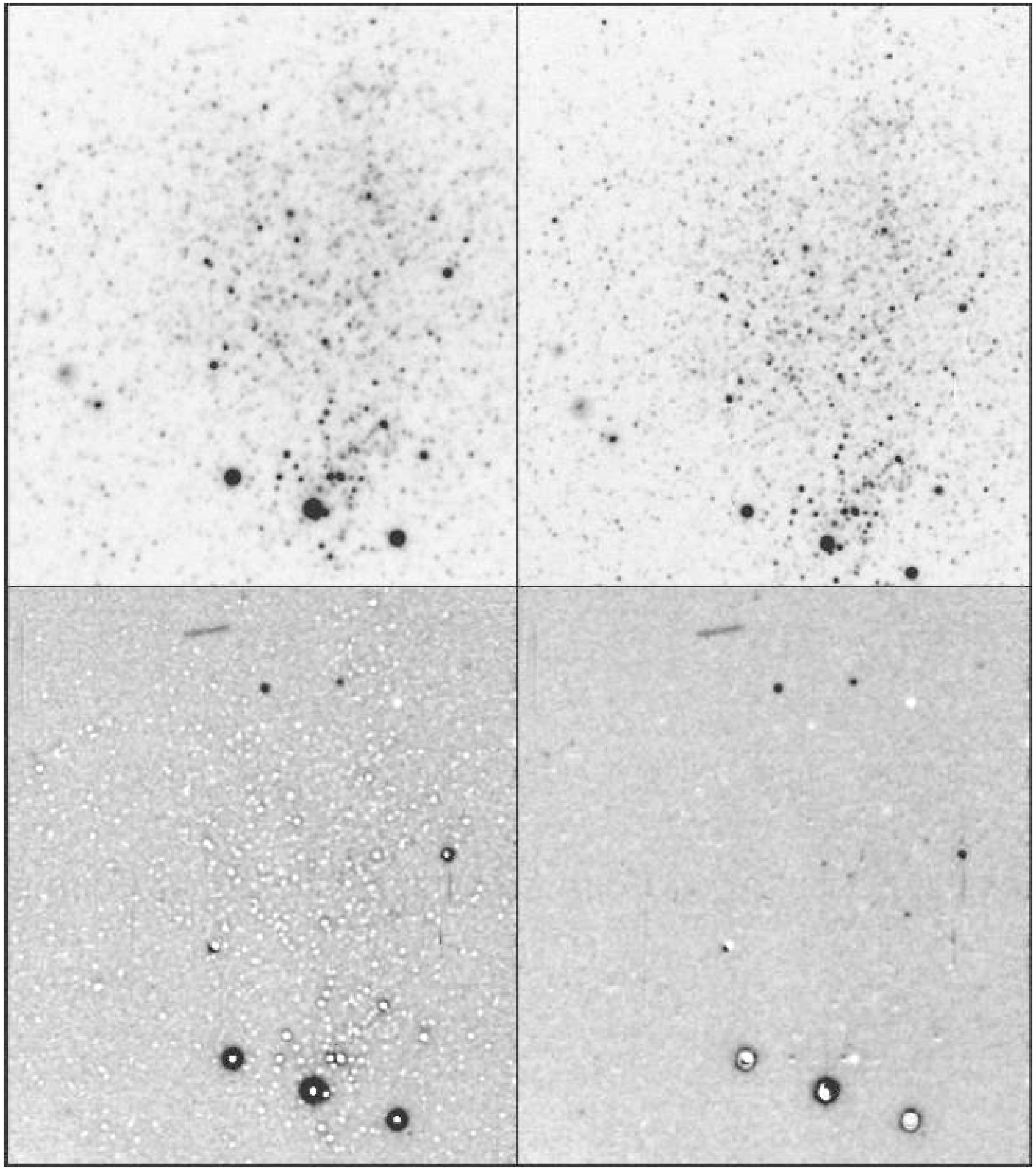
$$F(f(x) * g(x) = F(f(x))F(g(x))$$

$$f(x) * g(x) \Leftrightarrow F(s)G(s)$$

SIMILARLY FOR CROSS CORRELATIONS

$$F(f \otimes g) = F(f)F(g)$$

Convolute images



2 B-band images of the Phoenix dwarf galaxy

Difference image after convoluting the better-seeing image with a smoothing kernel and scaling the fluxes

From Phillips & Davis 1995

Continuous Fourier Transformations

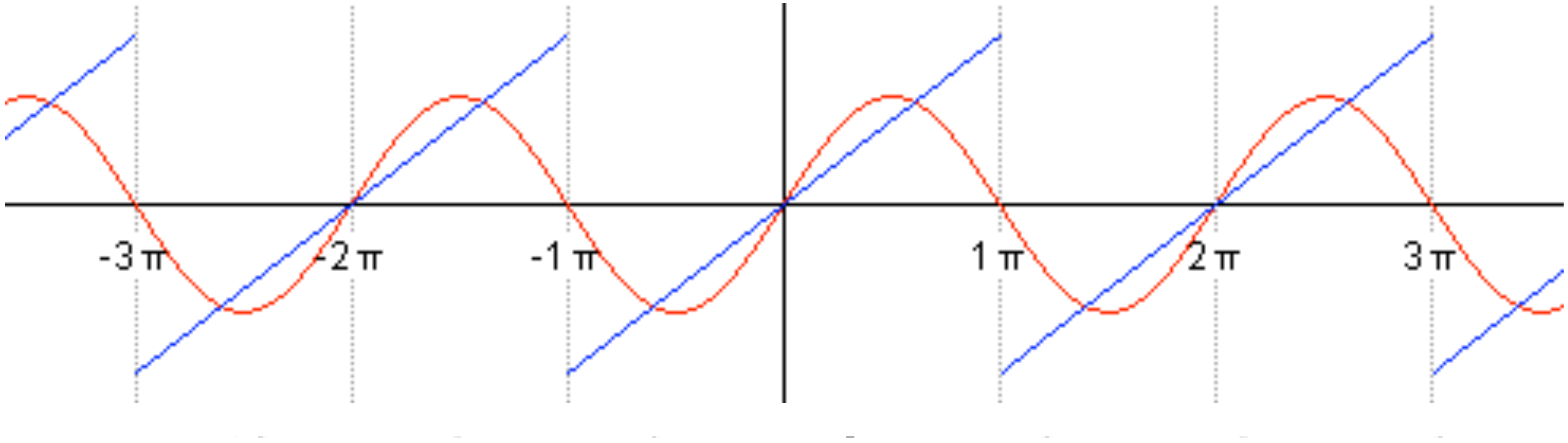


Figure from Wikipedia

$$F(t) \Leftrightarrow f(x)$$

$$F(t) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x t} dx$$

$$f(x) = \int_{-\infty}^{\infty} F(t) e^{2\pi i x t} dt$$

$$Euler's\ relation: \quad e^{ix} = cos x + i sin x$$

Used in restoration and/or spectral analysis of the signal

Convolution using Fourier transformations

Convolution theorem

$$M(\lambda) = S(\lambda) * R(\lambda)$$

$$F(M(\lambda)) = F(S(\lambda)) \cdot F(R(\lambda))$$

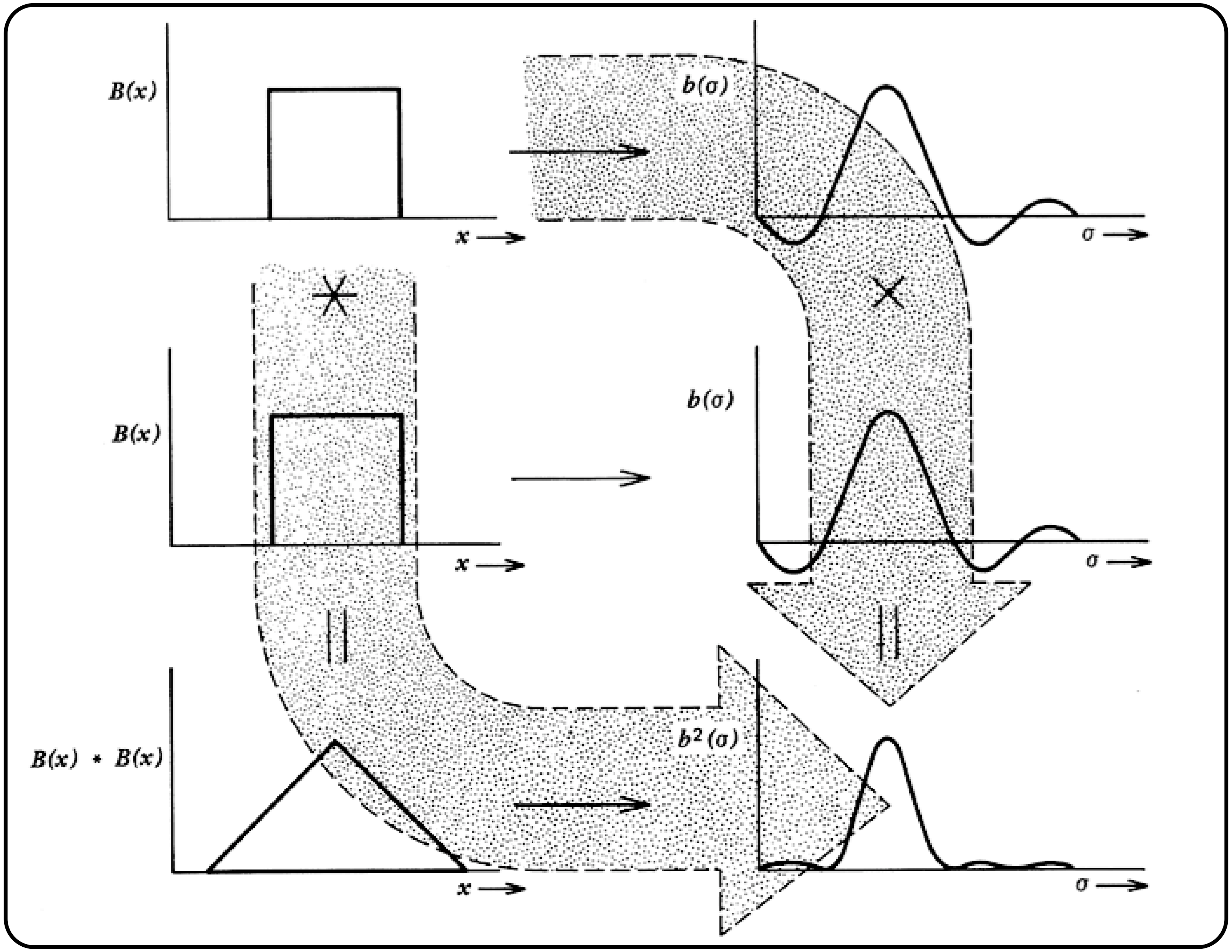


figure from Gray

Convolution using FTs in practice

Convolution

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x) g(x_1 - x) dx$$

Convolution always broadens the input function

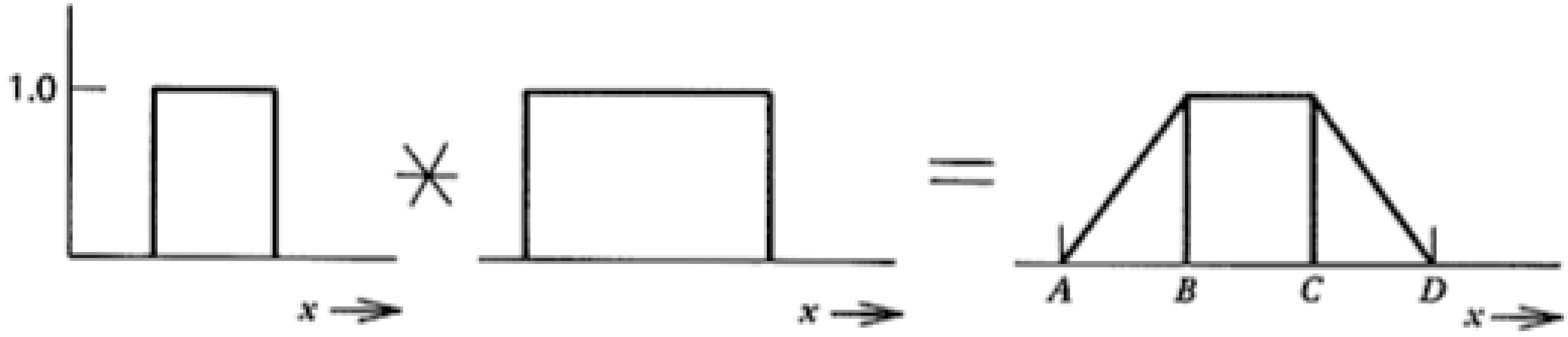


figure from Gray

Central limit theorem

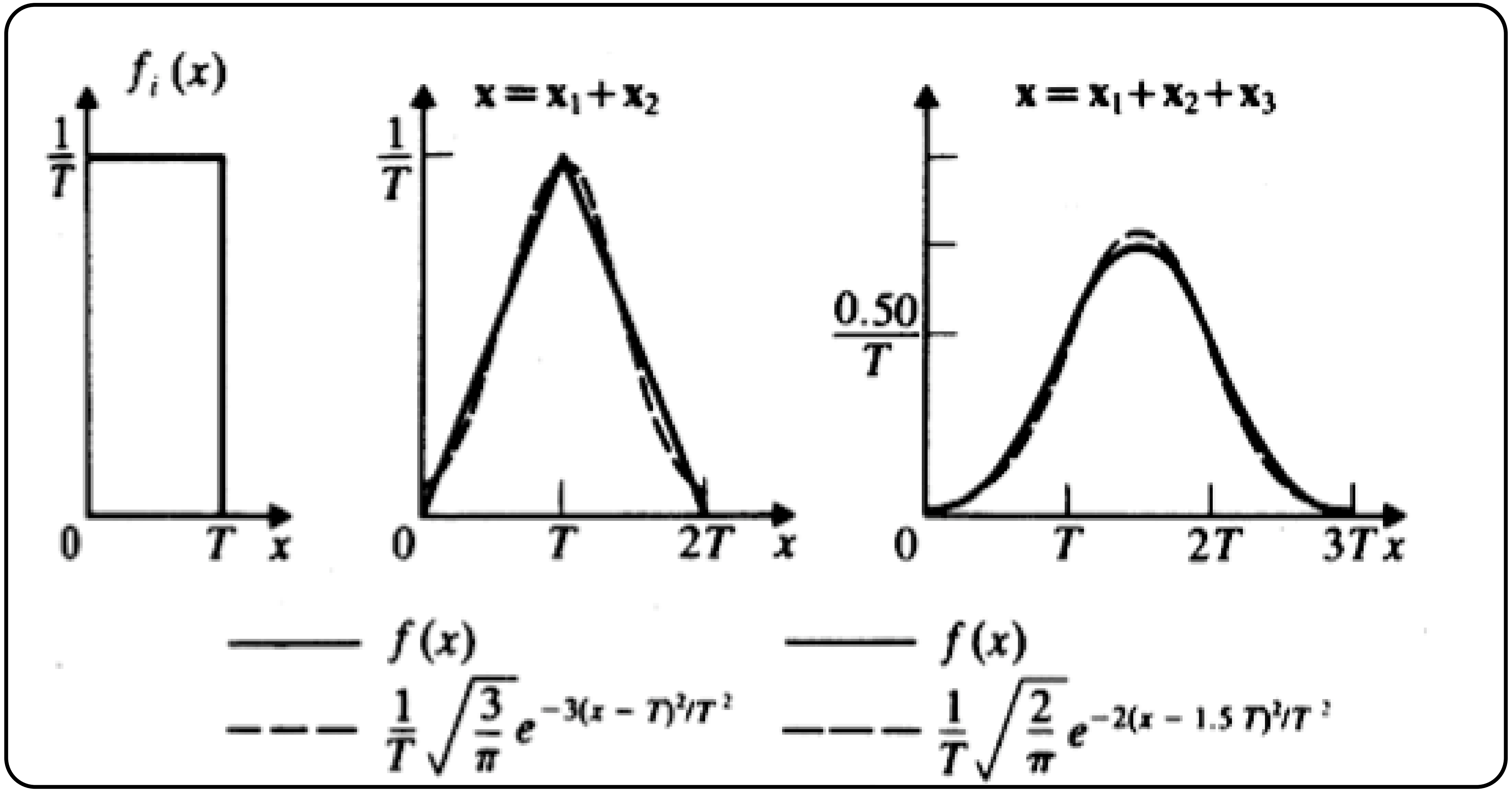
many convolutions → smoothing until Gaussian PDF

$$p_X(x) = p_{X_1}(x) * p_{X_2}(x) * p_{X_3}(x) * \cdots p_{X_n}(x)$$

$$\lim_{n\to\infty} p_X(x) = \frac{1}{\sqrt{(2\pi)}\sigma} exp - \frac{(x-\eta)^2}{2\sigma^2}$$

$$\eta = \sum_{i=1}^{n} \eta_i$$

$$\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$$



many physical processes/measurements yield a Gaussian probability density function

Reconstruction of the input=source spectrum

$$M(\lambda) = S(\lambda) * R(\lambda)$$

Convolution theorem

$$F(M(\lambda)) = F(S(\lambda)) \cdot F(R(\lambda))$$

$$F(M(\lambda)) \equiv M(s)\,(etc)$$

$$M(s) = S(s) \cdot R(s)$$

$$S(s) = \frac{M(s)}{R(s)}$$

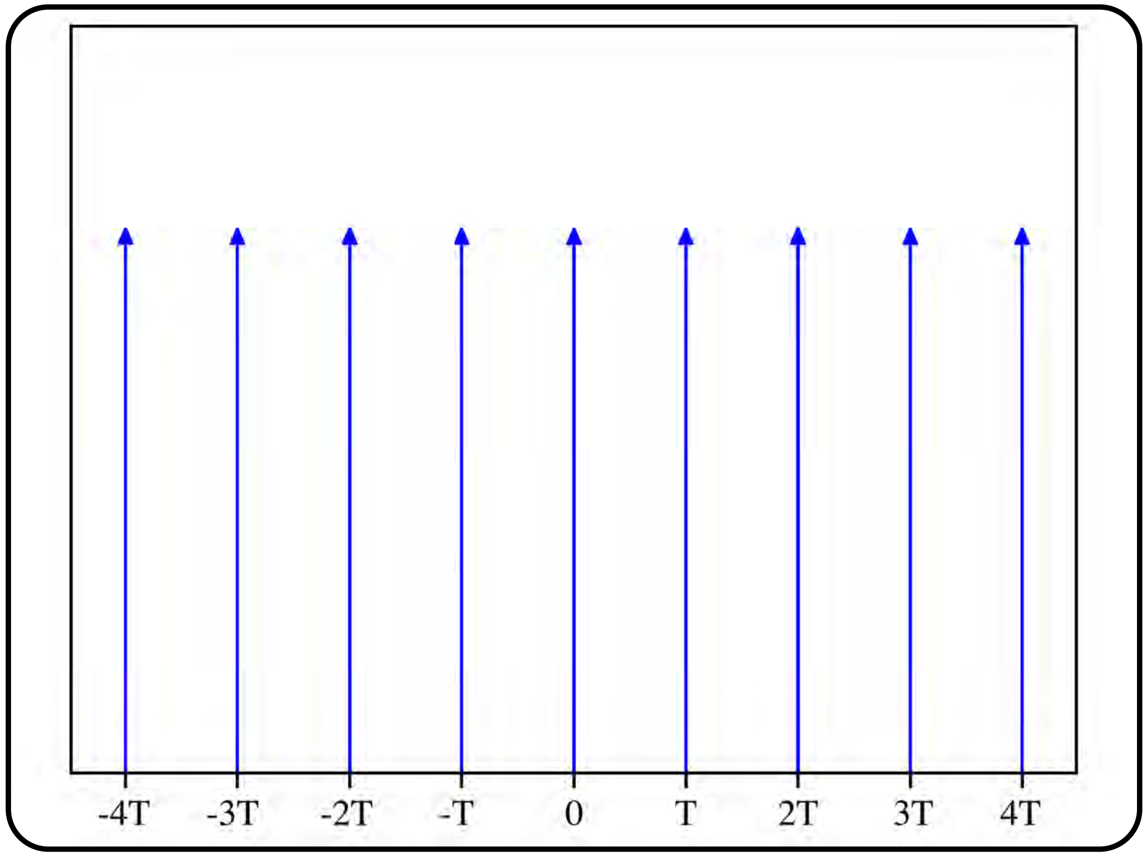
$$S(\lambda) = F^{-1}\left(\frac{M(s)}{R(s)}\right)$$

SOME SPECIAL FUNCTIONS:

Sampling

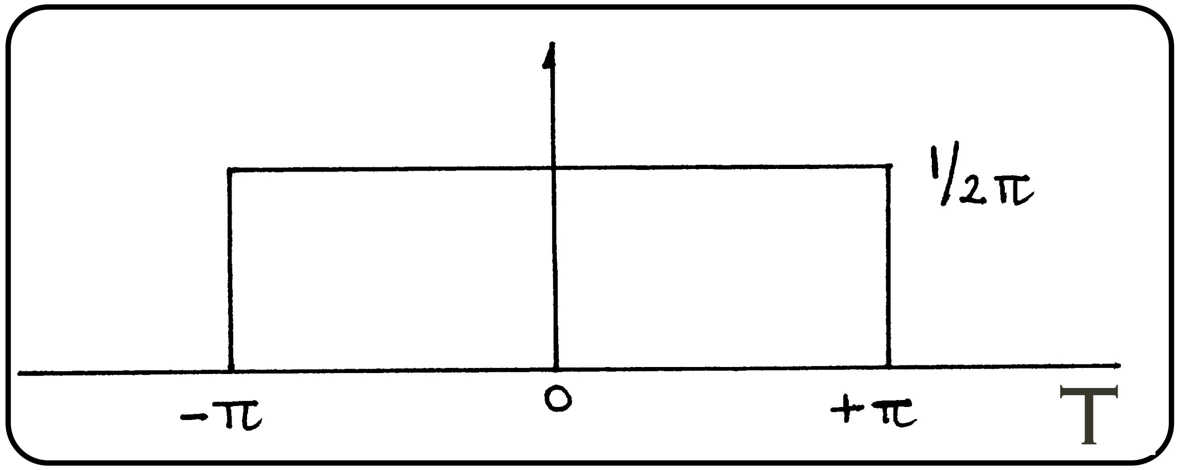
SHAH'S FUNCTION/DIRAC COMB

$$III(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$$



Finite stretch of data

BOX/WINDOW FUNCTION

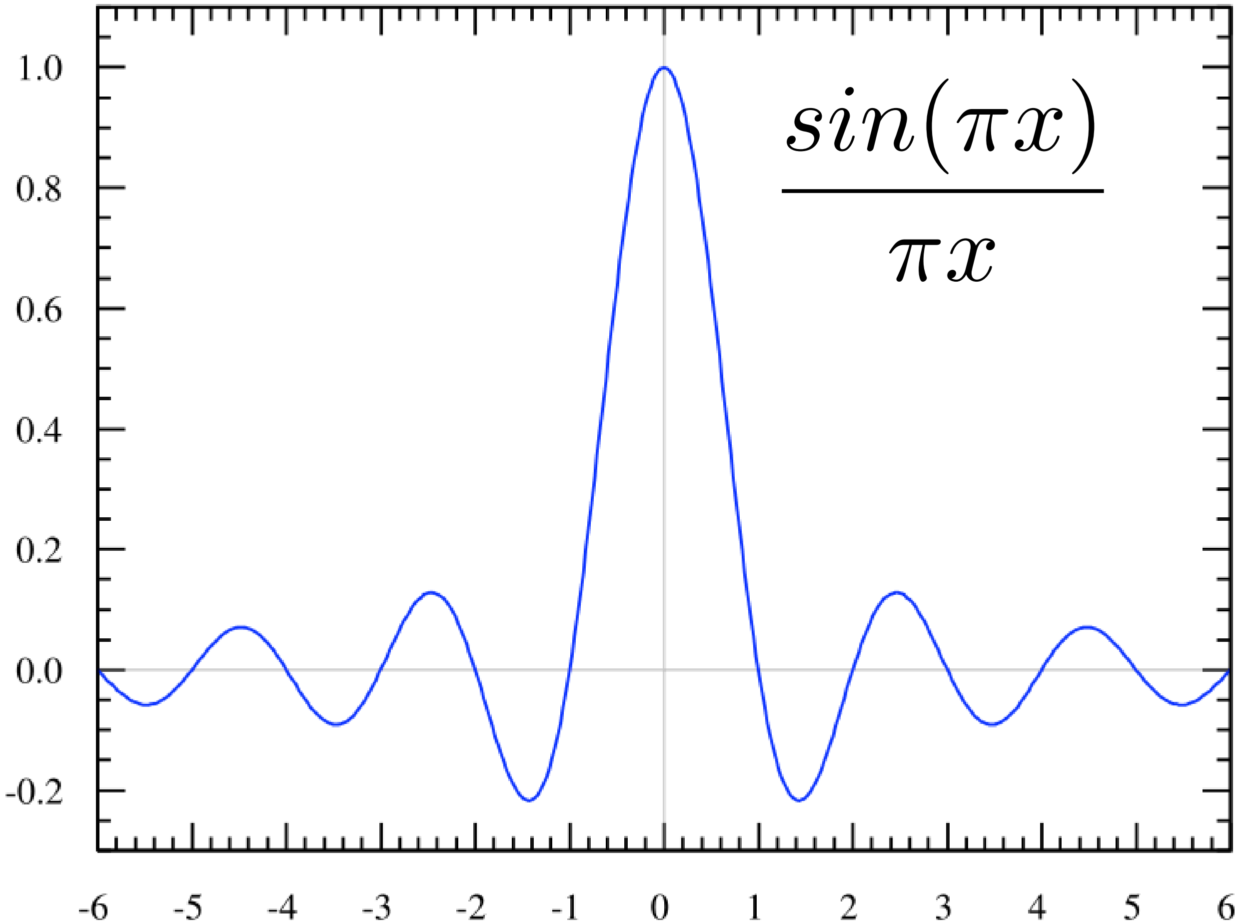


$$B(t) = 0 \ for \ -\frac{W}{2} > t > \frac{W}{2}$$

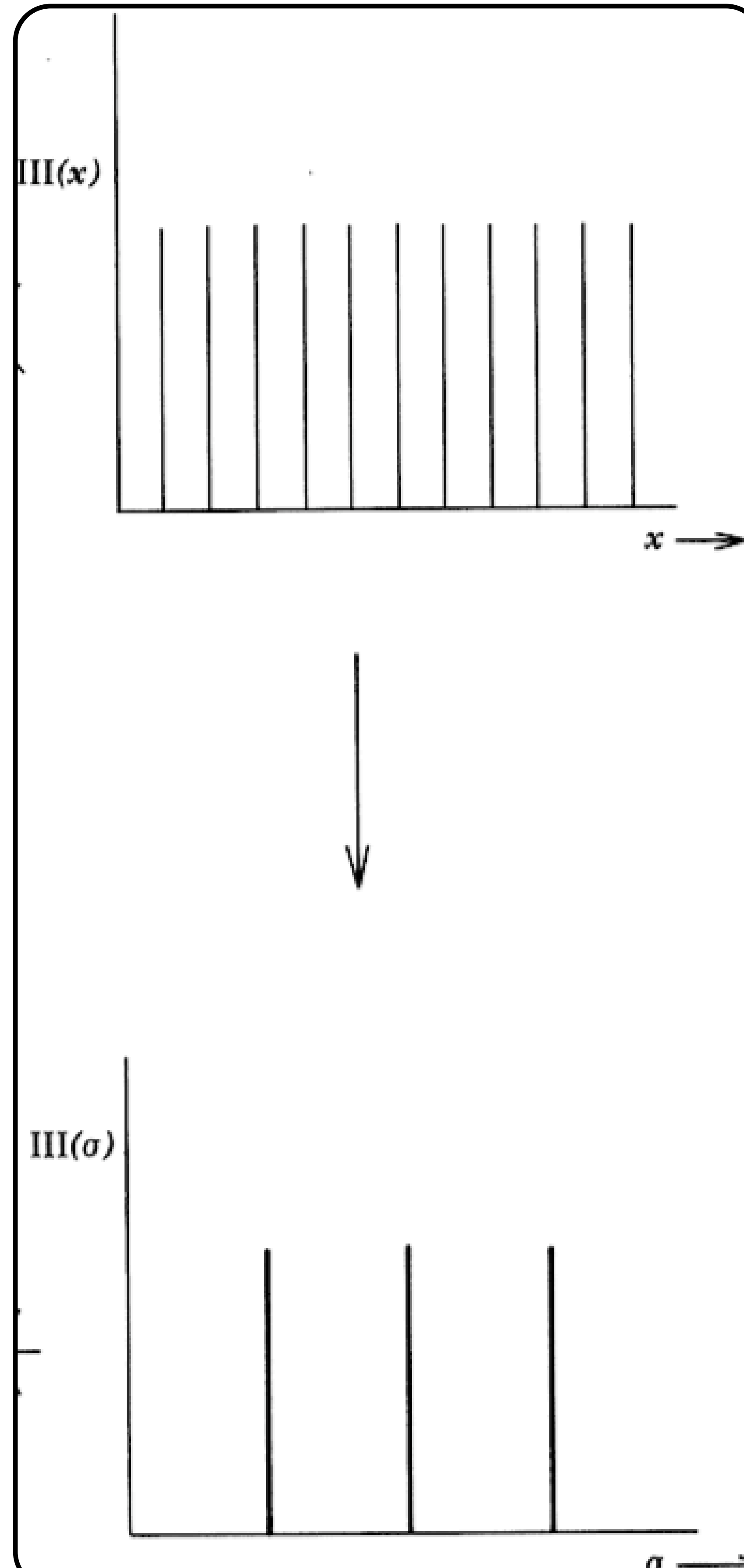
$$B(t) = 1 \ for \ -\frac{W}{2} < t < \frac{W}{2}$$

Fourier transformations of these special functions

Sinc function



$$\frac{sin(\pi x)}{\pi x}$$



$$\sum_{n=-\infty}^{\infty} \delta(t - nT) \Leftrightarrow \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{T})$$

From Fourier animation of harmonic decomposition remember that:

A sharp narrow signal needs more/higher frequencies to be described in the Fourier Transform than broad shallow signal

cf. the number of sin+cos necessary to describe the signal

Optical spectra: bandwidth set by the width of the spectral lines

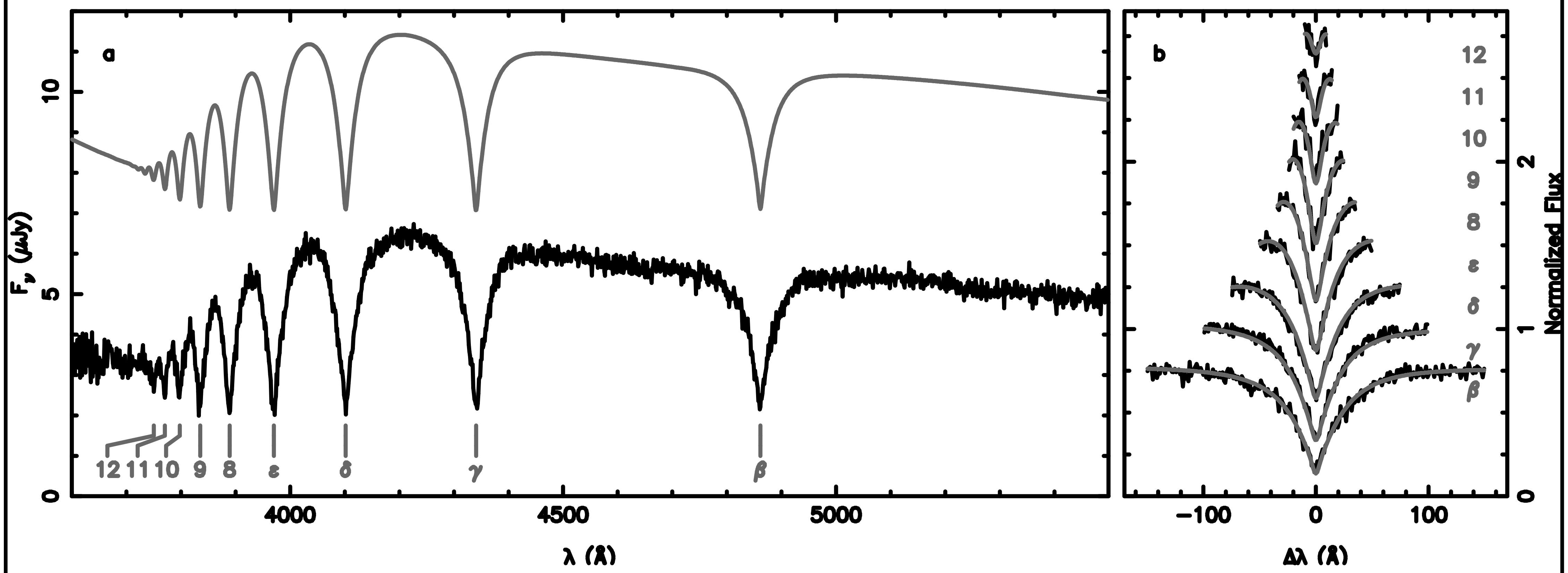
Sampling theorem

Sampling: no loss of information if the input process has no frequencies > $\frac{1}{2\Delta t_{crit}} \equiv \mathrm{f}_{Nyquist}$

Continuous signal h(t) fully described by the samples

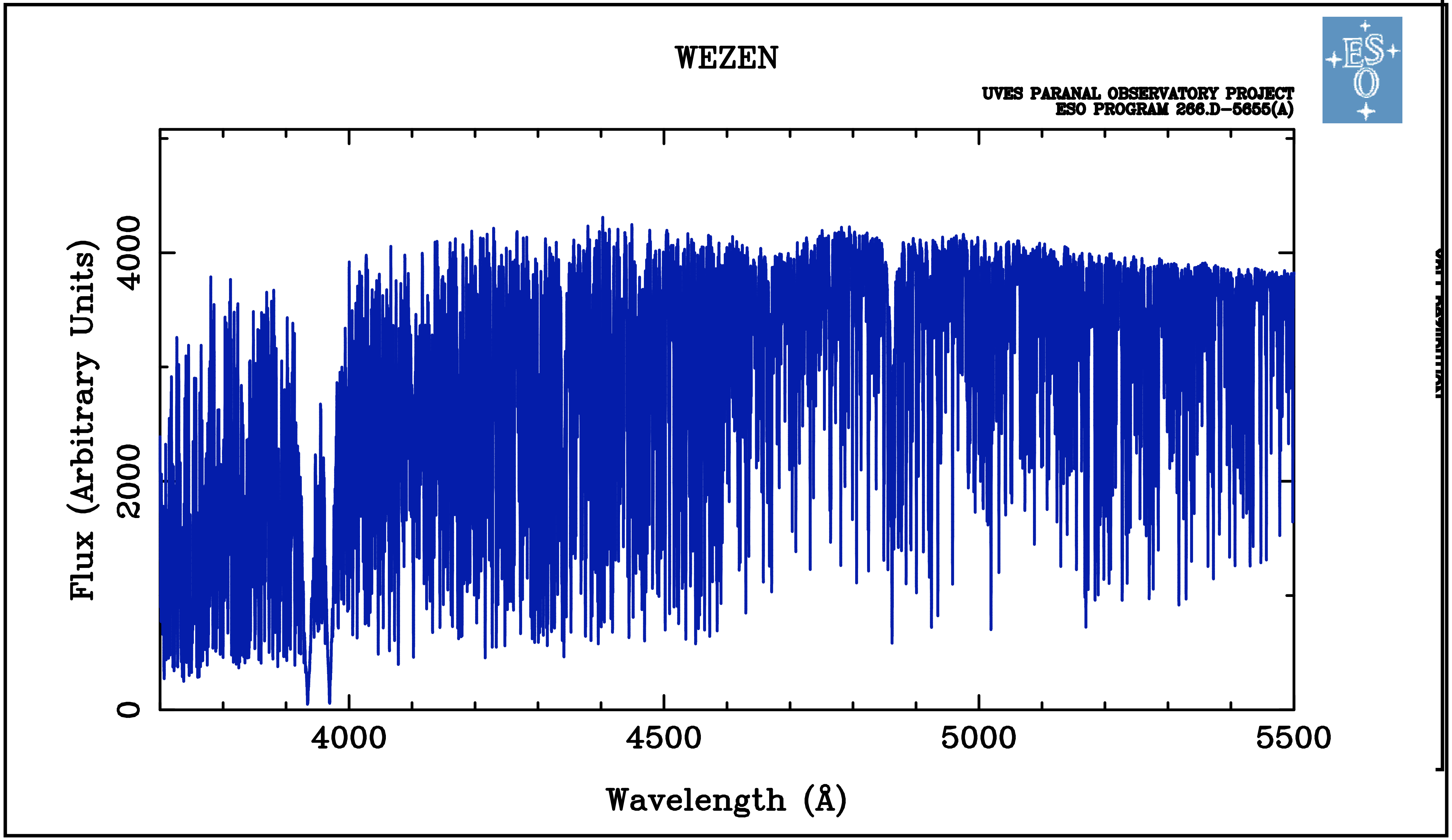
Optical spectra: bandwidth set by the width of the spectral lines

Low-mass White Dwarf spectrum

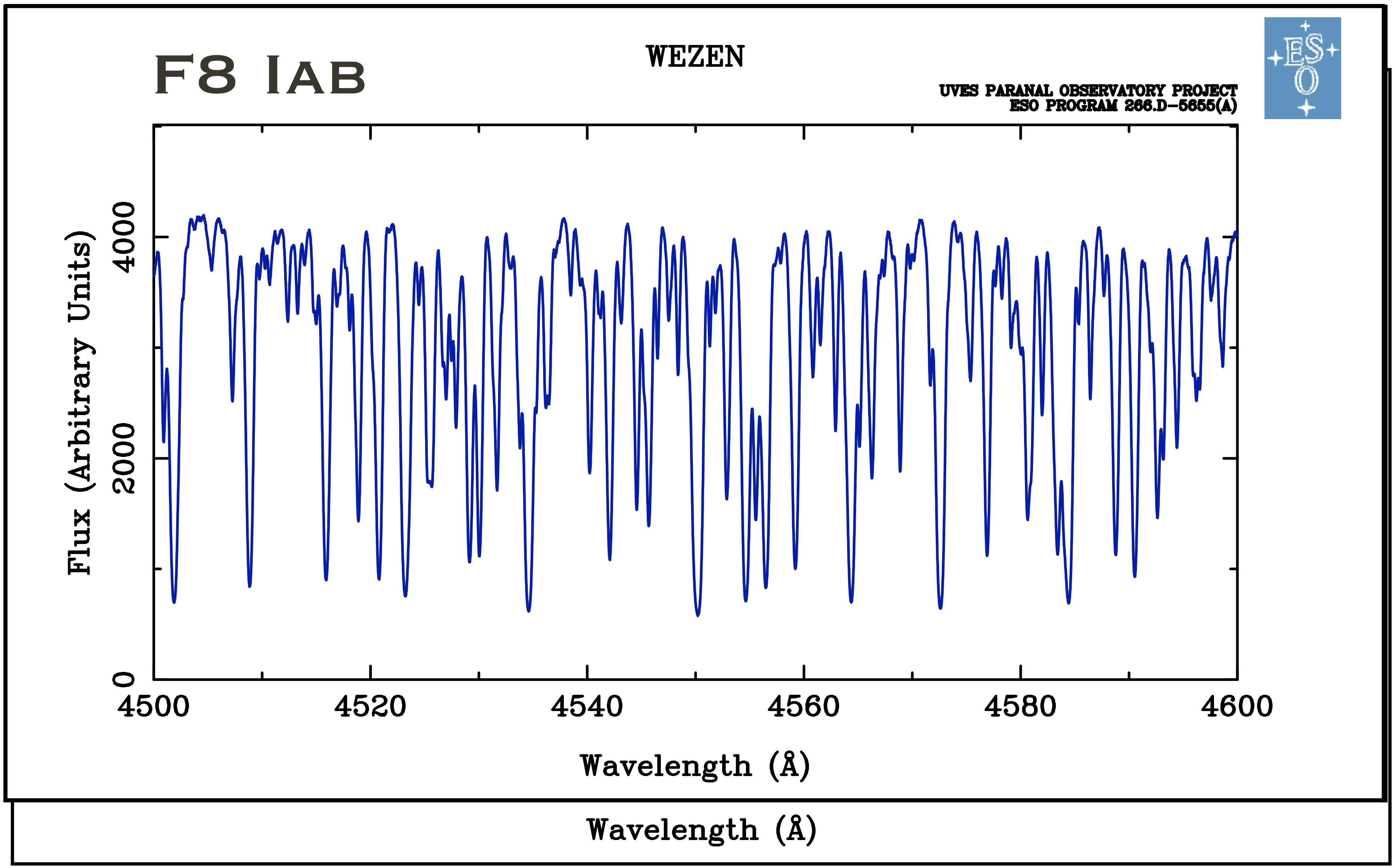


Bassa et al. 2006

Optical spectra: bandwidth set by the width of the spectral lines



Optical spectra: bandwidth set by the width of the spectral lines



Mazur: Peer Instruction

Besides spectral frequencies there are also temporal and angular frequencies

Discuss with your neighbor a possible example of these

Another math tool
Power Spectral Density

(∝ amplitude of individual sinusoids)

(will return in more depth in Chapter 6)

Continuous FT: $$F(f) = \int_{-\infty}^{\infty} f(t)\ e^{-2\pi i f t} dt$$

Continuous PSD: $$P(f) = F\tilde{(f)} F\tilde{(f)}^*$$

for WSS signals: $$P(f) = \int_{-\infty}^{\infty} R(\tau) e^{-2\pi i f \tau} d\tau$$

hence:

$$F\widetilde{(f)} F\widetilde{(f)} = |F(f)|^2 = \int_{-\infty}^{\infty} R(\tau) e^{-2\pi i f \tau} d\tau$$

Discrete Fourier Transform

- Now estimate Fourier transform of discretely sampled function with N consecutive sample values with interval Δ

$$h_k \equiv h(t_k), \qquad t_k \equiv k\Delta, \qquad k = 0, 1, 2, \ldots, N-1$$

- N input values → no more than N output values, seek estimates of Fourier transform only at discrete frequencies values in the range $[-f_c, f_c]$:

$$f_n \equiv \frac{n}{N\Delta}, \qquad n = -\frac{N}{2}, \ldots, \frac{N}{2}$$

(N+1 values of n, but values of n at boundaries are not independent) in fact they are the same

- Approximate integral by discrete sum: $$H(f_n) = \int_{-\infty}^{\infty} h(t) e^{2\pi i f_n t} dt \approx \sum_{k=0}^{N-1} h_k \, e^{2\pi i f_n t_k} \Delta = \Delta \sum_{k=0}^{N-1} h_k \, e^{2\pi i k n/N}$$

$$H_n \equiv \sum_{k=0}^{N-1} h_k \, e^{2\pi i k n/N}$$

Discrete Fourier Transform of h(t) : mapping N (complex) numbers (h_k's) onto N complex H_n's

- We also have the discrete inverse Fourier Transform, which recovers the N h_k's from the N H_n's (using periodicity of n with period N)

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} H_n \, e^{-2\pi i k n/N}$$

- Discrete form of Parseval's theorem: $$\sum_{k=0}^{N-1} |h_k|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |H_n|^2$$

Measuring process in the time / in the frequency domain

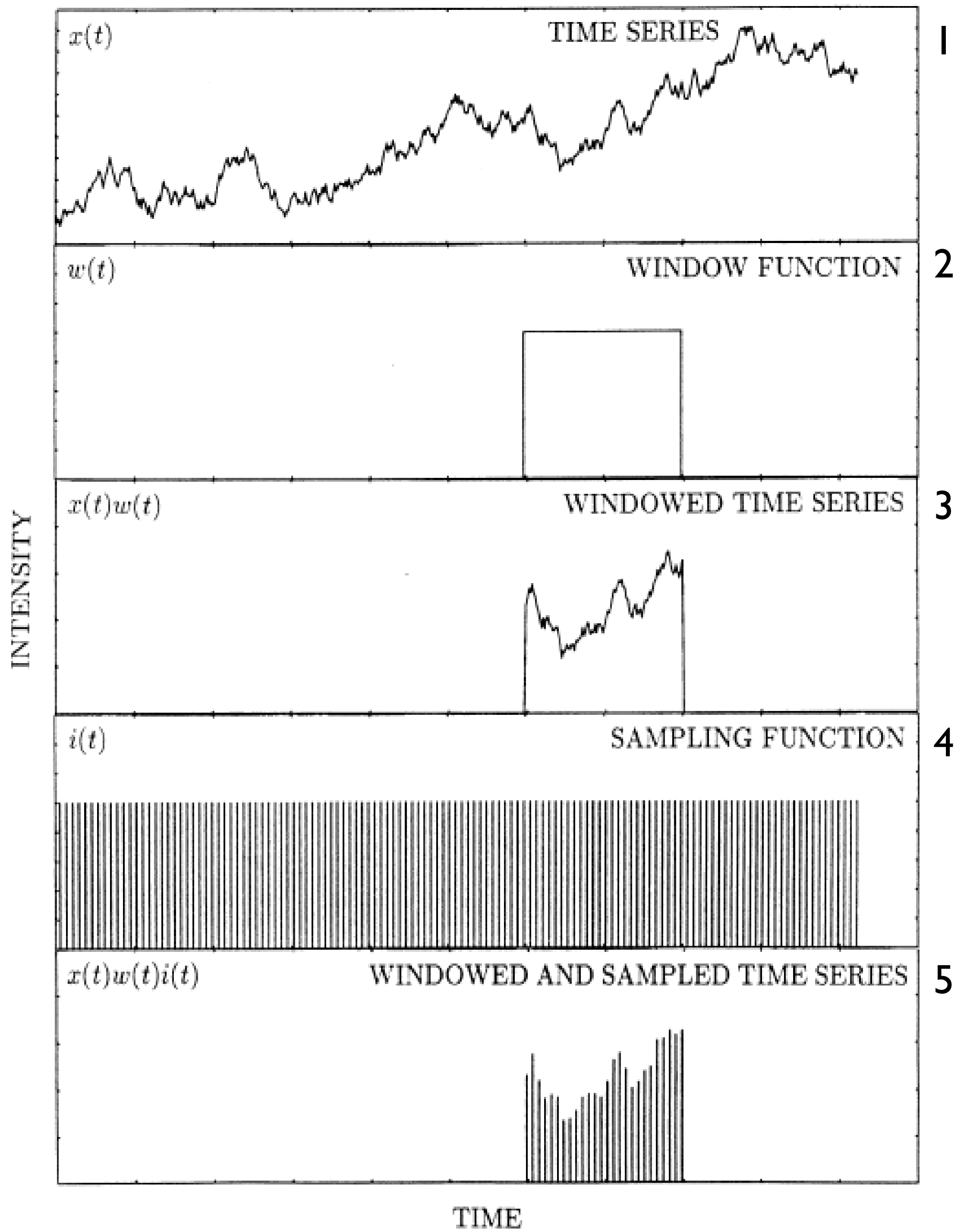


Fig. 2.2. *a) Obtaining the discrete time series x_k as a discretely sampled section of $x(t)$ involves a double multiplication.*

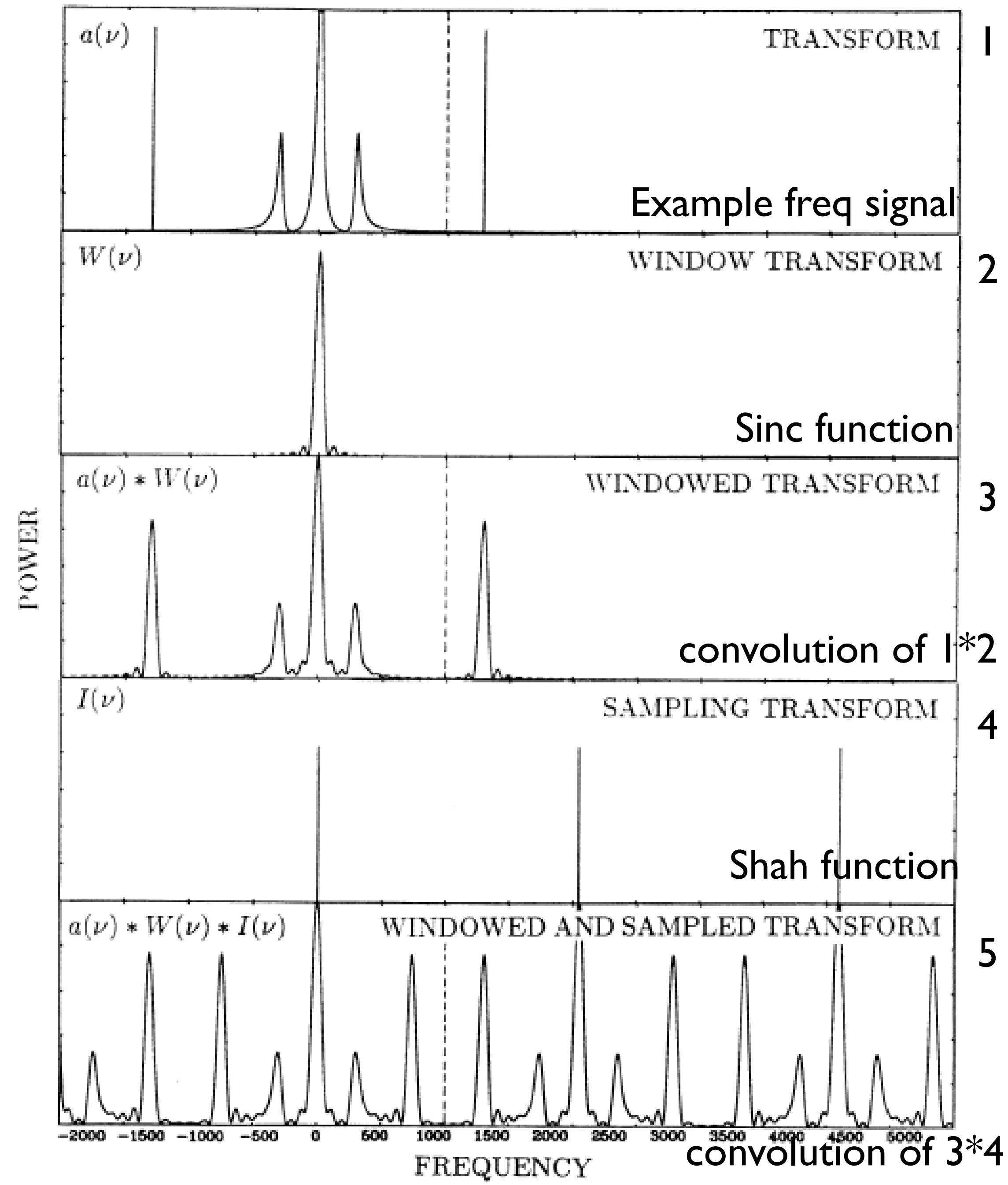
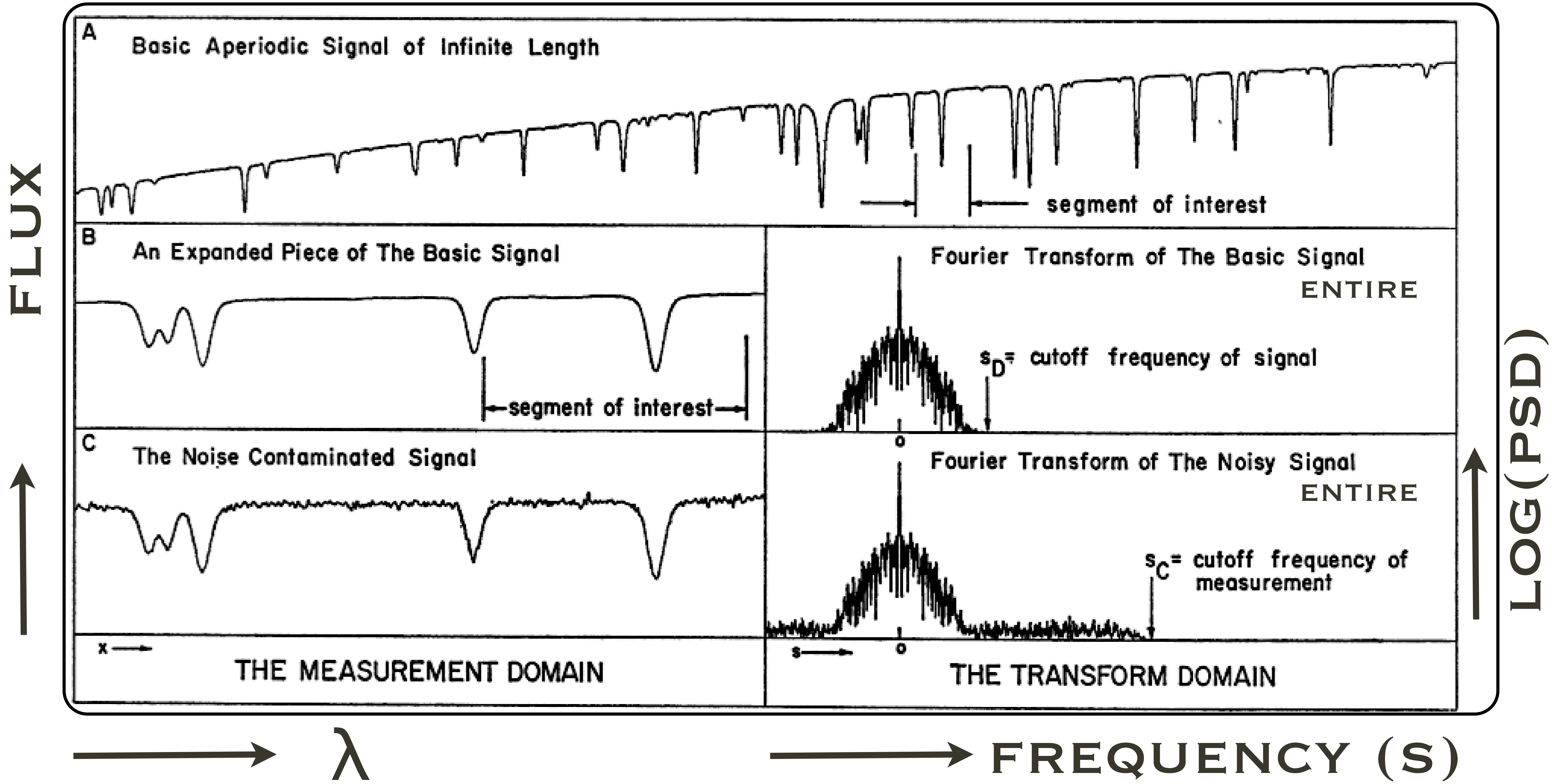


Fig. 2.2. *b) The discrete Fourier transform a_j of x_k is obtained out of the continuous Fourier transform $a(\nu)$ by a double convolution. The figure shows the power spectra corresponding to the various Fourier transforms. Vertical dashed lines indicate the Nyquist frequency.*

Timing domain (t) ⟺ Frequency (ν) domain

SRON

Maximum signal frequency vs maximum passed frequency of measuring apparatus



WINDOWING & NOISE, BRAULT & WHITE 1971, A&A

Data sampling

Data is discrete not continuous

TIME DOMAIN MULTIPLY S.P. WITH SHAH FUNCTION

$$m_{samp,n} = m_s(x) = m(x)\frac{1}{\tau}\text{III}(\frac{x}{\tau}) = \sum_n m(n\tau)\delta(x - n\tau)$$

DISCRETE FT:

$$M_{samp,k} = \sum_{n=0}^{N-1} m_{samp,n}\ e^{2\pi ink/N}$$

DISCRETE PSD:

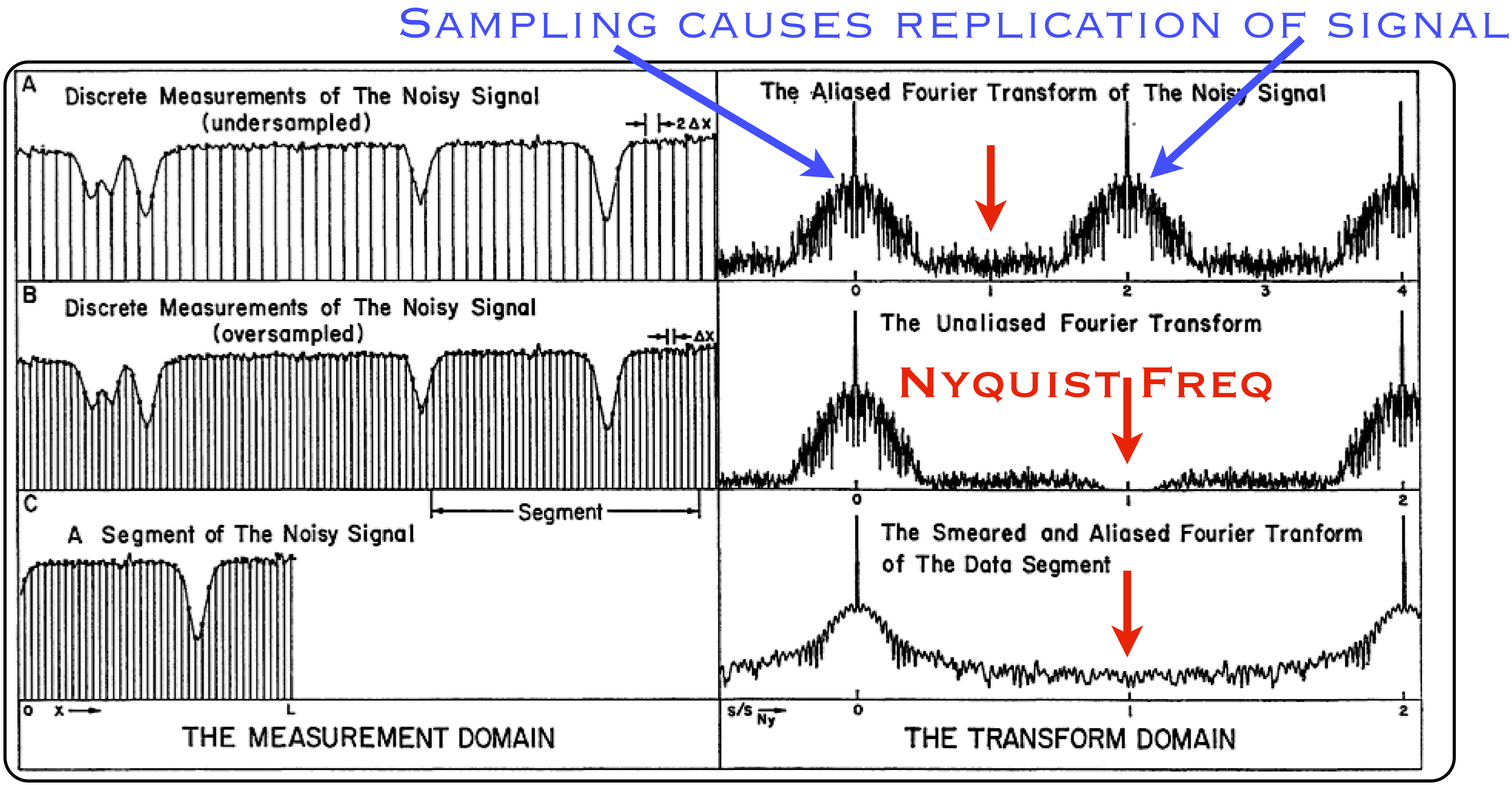
$$P_j = \frac{2}{a_0}|a_j|^2$$

POWER ∝ AMPLITUDE SQUARED:

$$a_0 = M_{samp,k=0} = \sum_{n=0}^{N-1} m_{samp,n} \equiv N_0$$

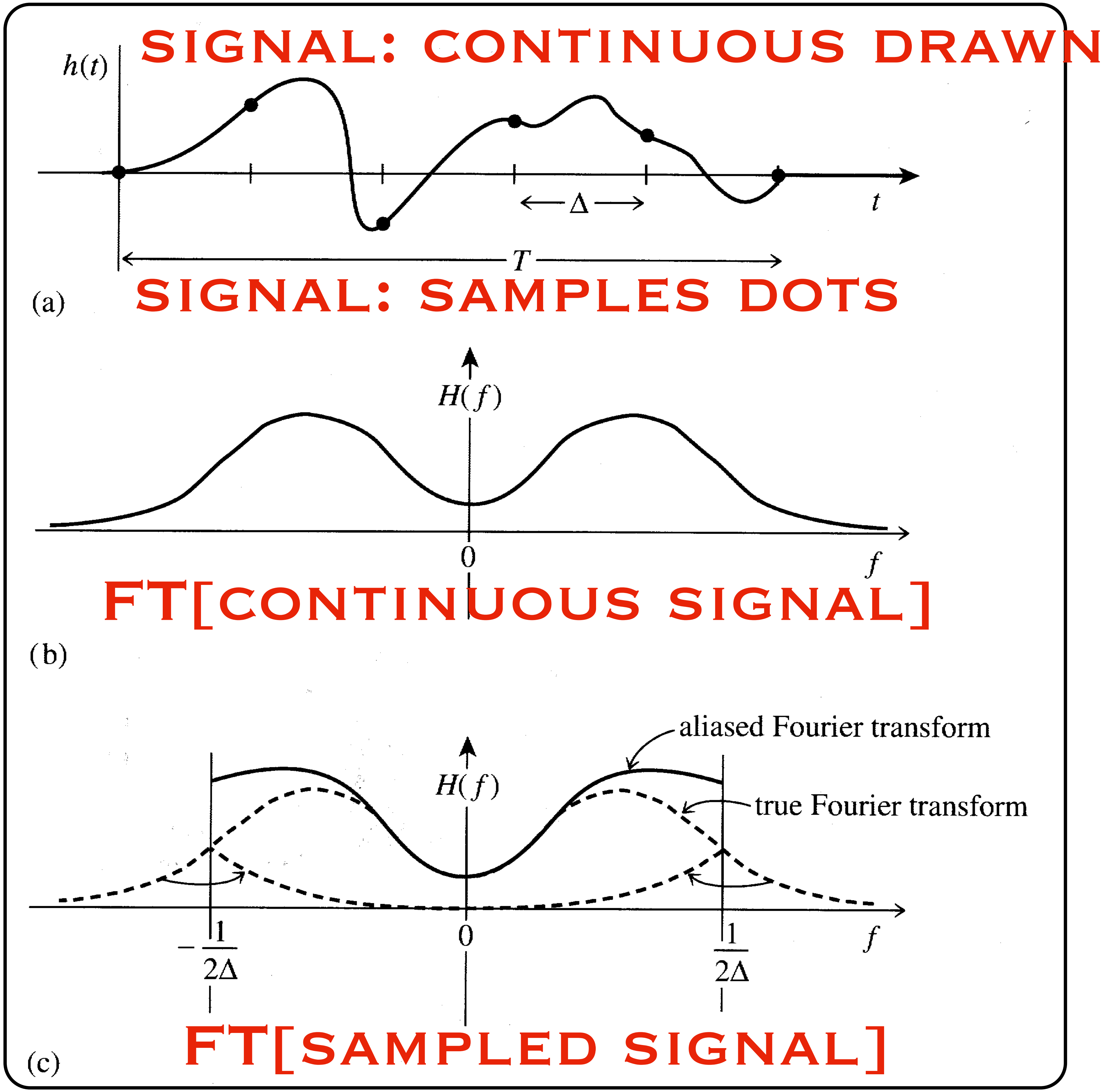
$$a_k = M_{samp,k} = \sum_{n=0}^{N-1} m_{samp,n}\ e^{2\pi ink/N}$$

Nyquist theorem: cont'd



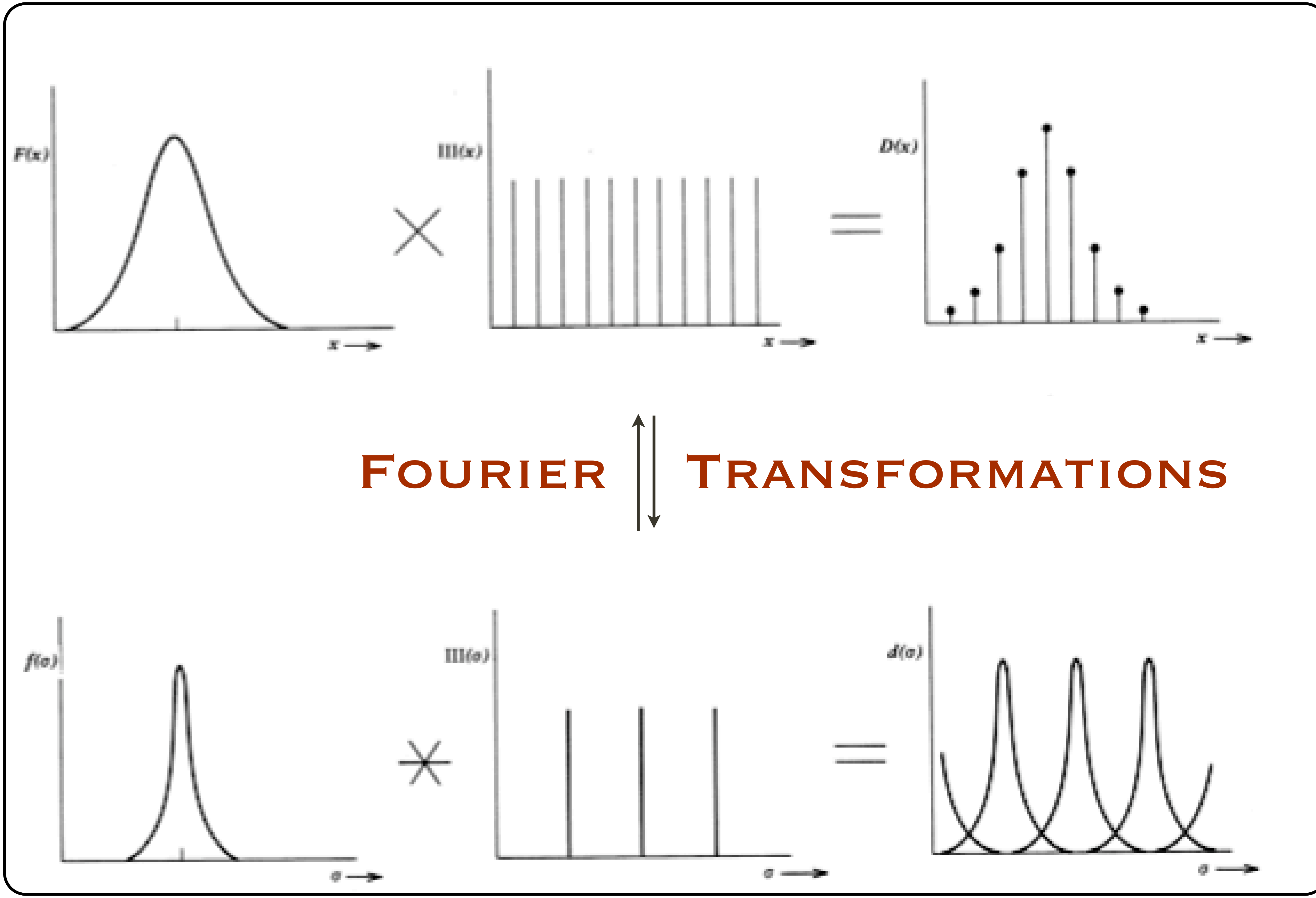
Sampling; Brault & White 1971, A&A, 13, 169 (in list of presentation papers!)

Aliasing



SIGNAL: CONTINUOUS DRAWN
h(t)
Δ
t
T
(a)
SIGNAL: SAMPLES DOTS
H(f)
0
f
FT[CONTINUOUS SIGNAL]
(b)
aliased Fourier transform
H(f)
true Fourier transform
$-\frac{1}{2\Delta}$
0
$\frac{1}{2\Delta}$
f
(c)
FT[SAMPLED SIGNAL]

Aliasing: cont'd



Convolution with shah function in freq space: replication

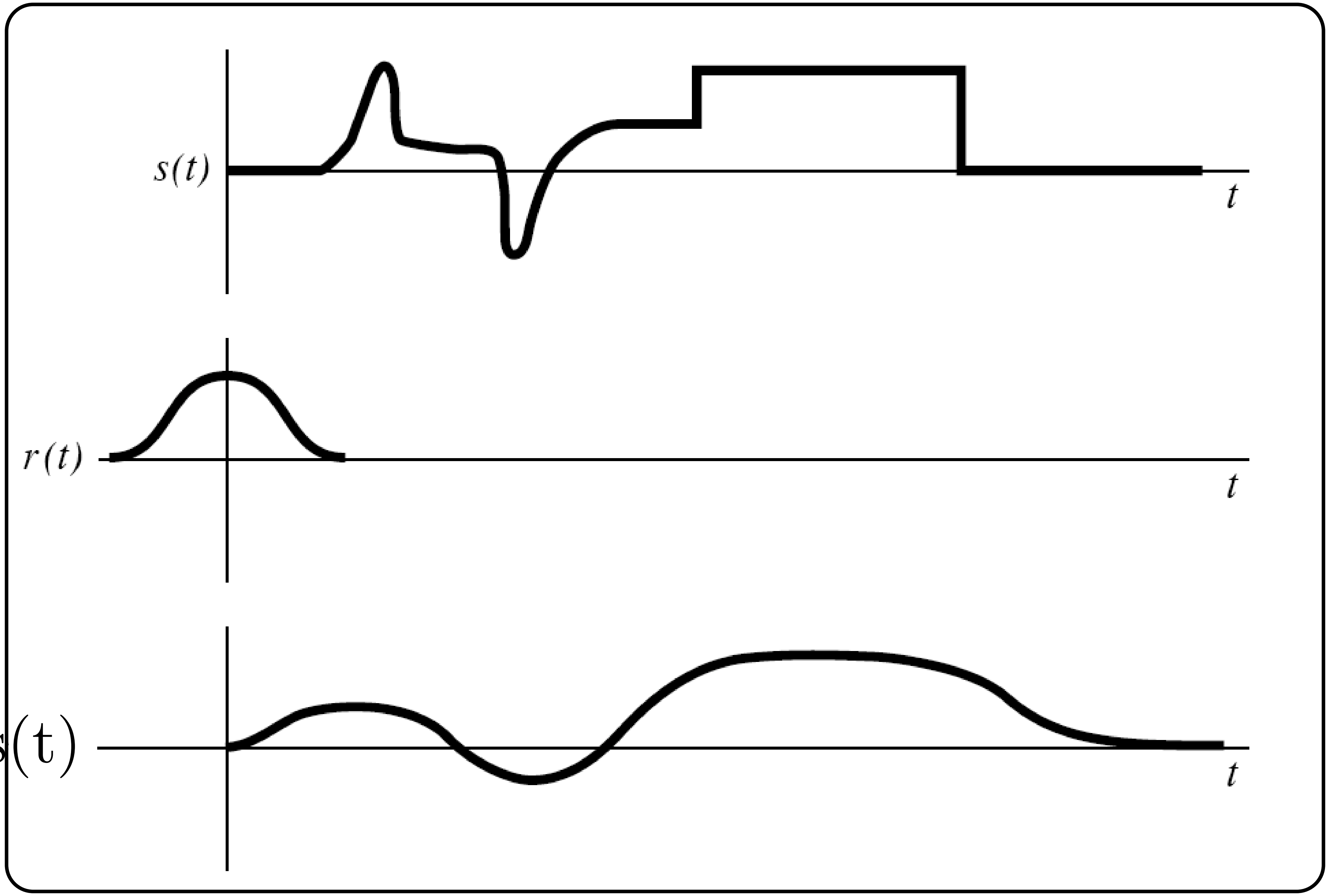
Mazur: Peer Instruction

Explain to your neighbor the process of aliasing

(de)convolution of sampled date

Num Res Chapter 13.0-13.3

DECONVOLVE MEASURED SIGNAL AND RESPONSE FUNCTION OF SAMPLED DATA



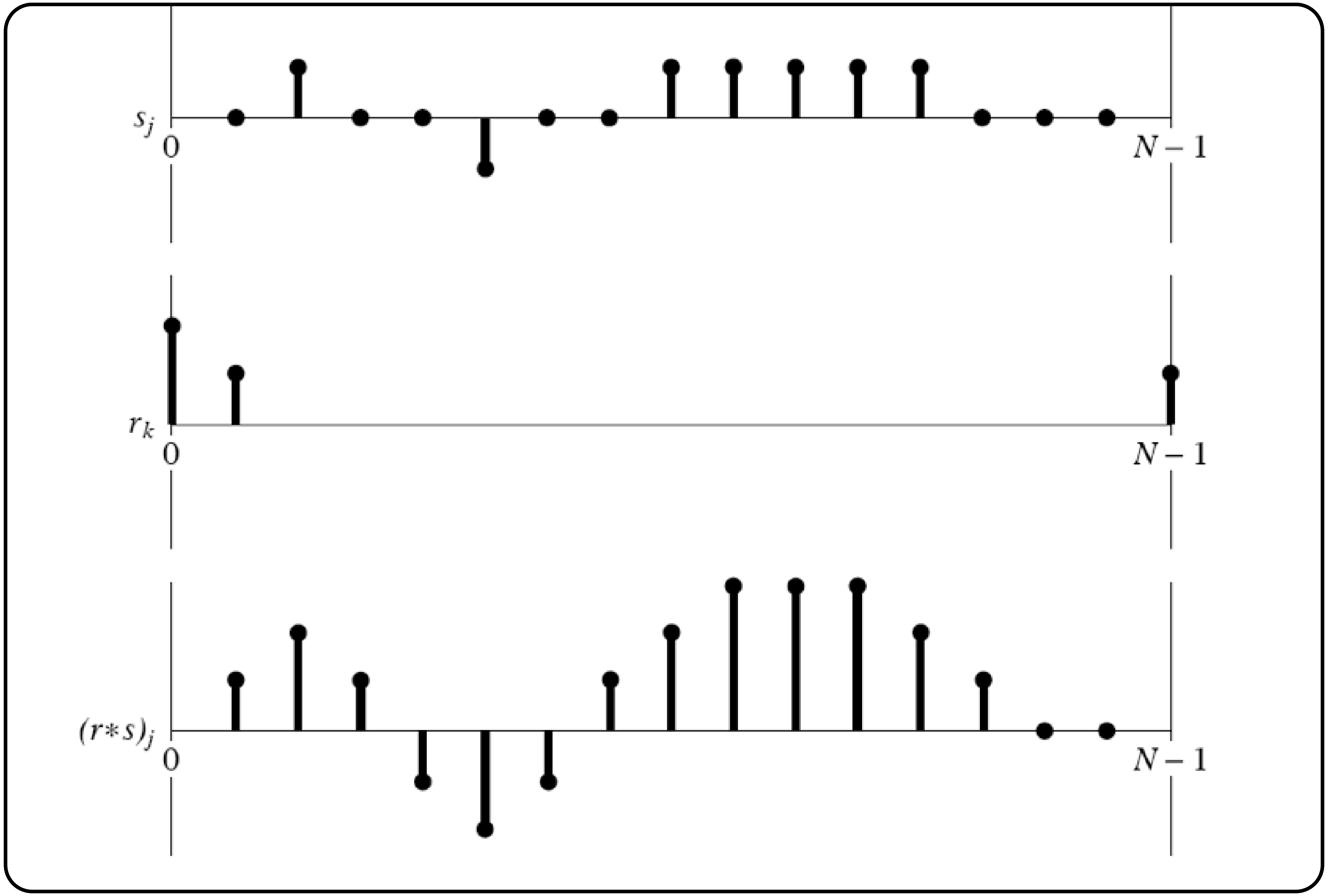
$\mathrm{r(t) * s(t)}$

Num Res Chapter 13.1

(Optimal) Filtering

Num Res Chapter 13.0-13.3

deconvolve measured signal and response function of sampled data

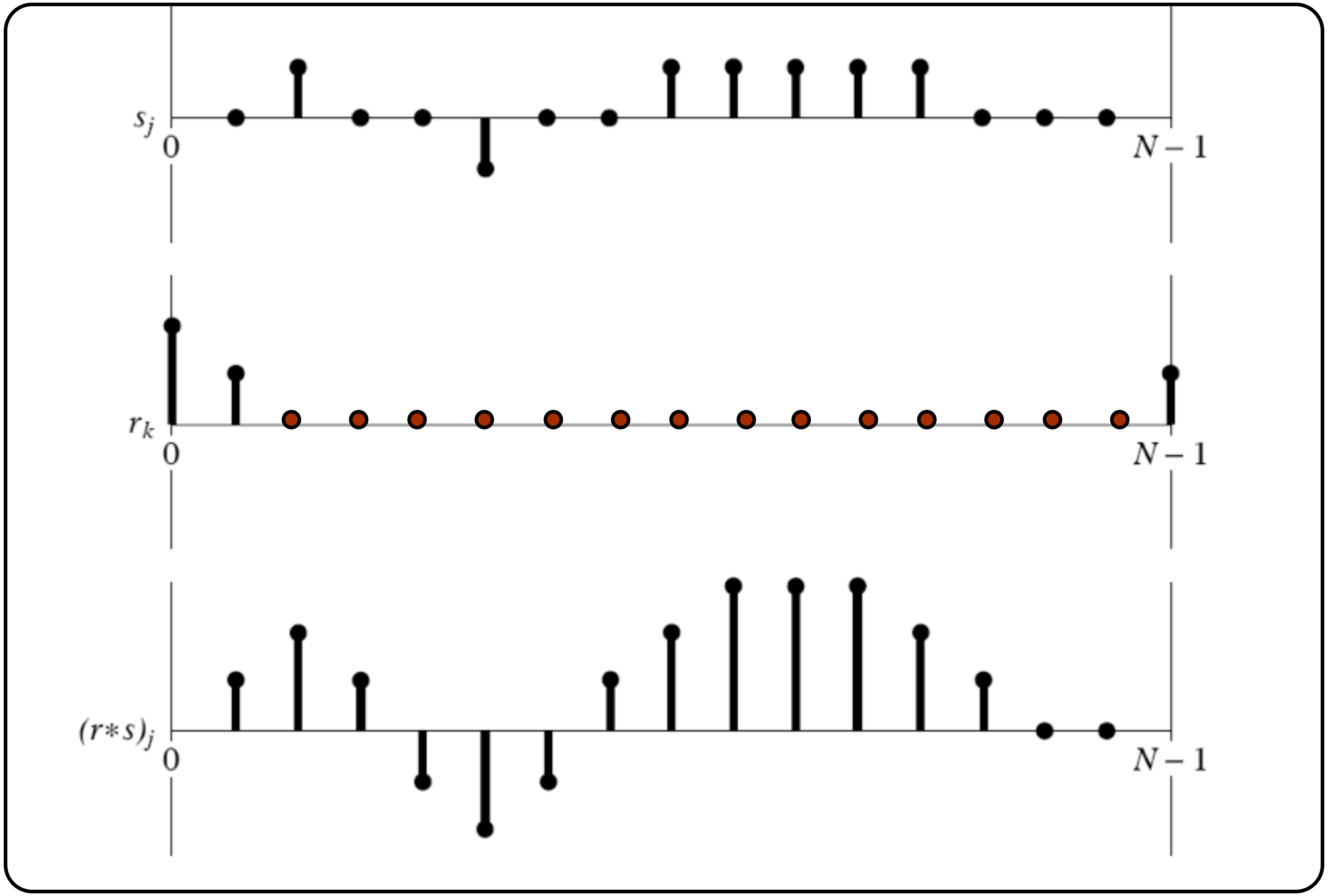


Num Res Chapter 13.1

(Optimal) Filtering

Num Res Chapter 13.0-13.3

deconvolve measured signal and response function of sampled data

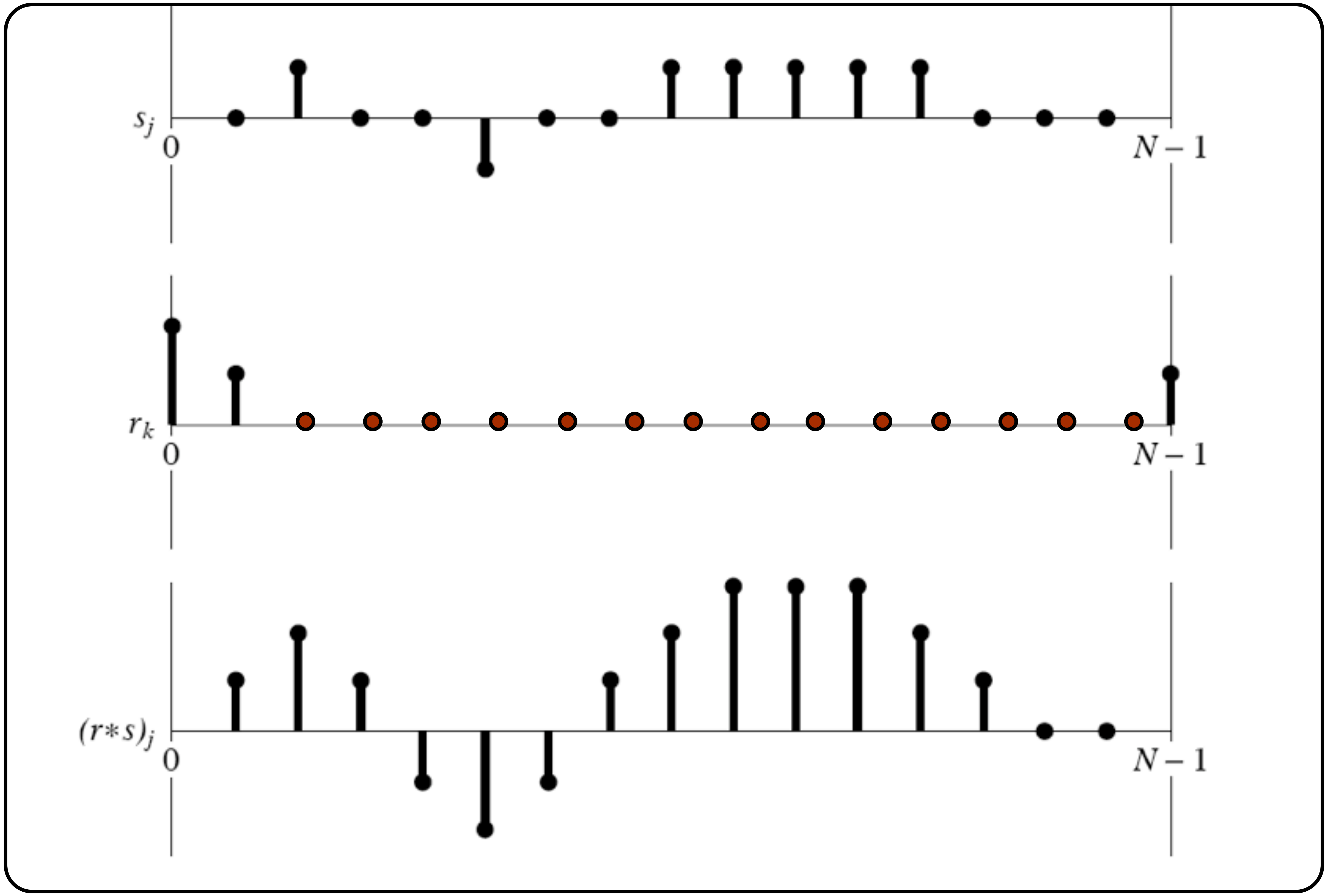


Num Res Chapter 13.1

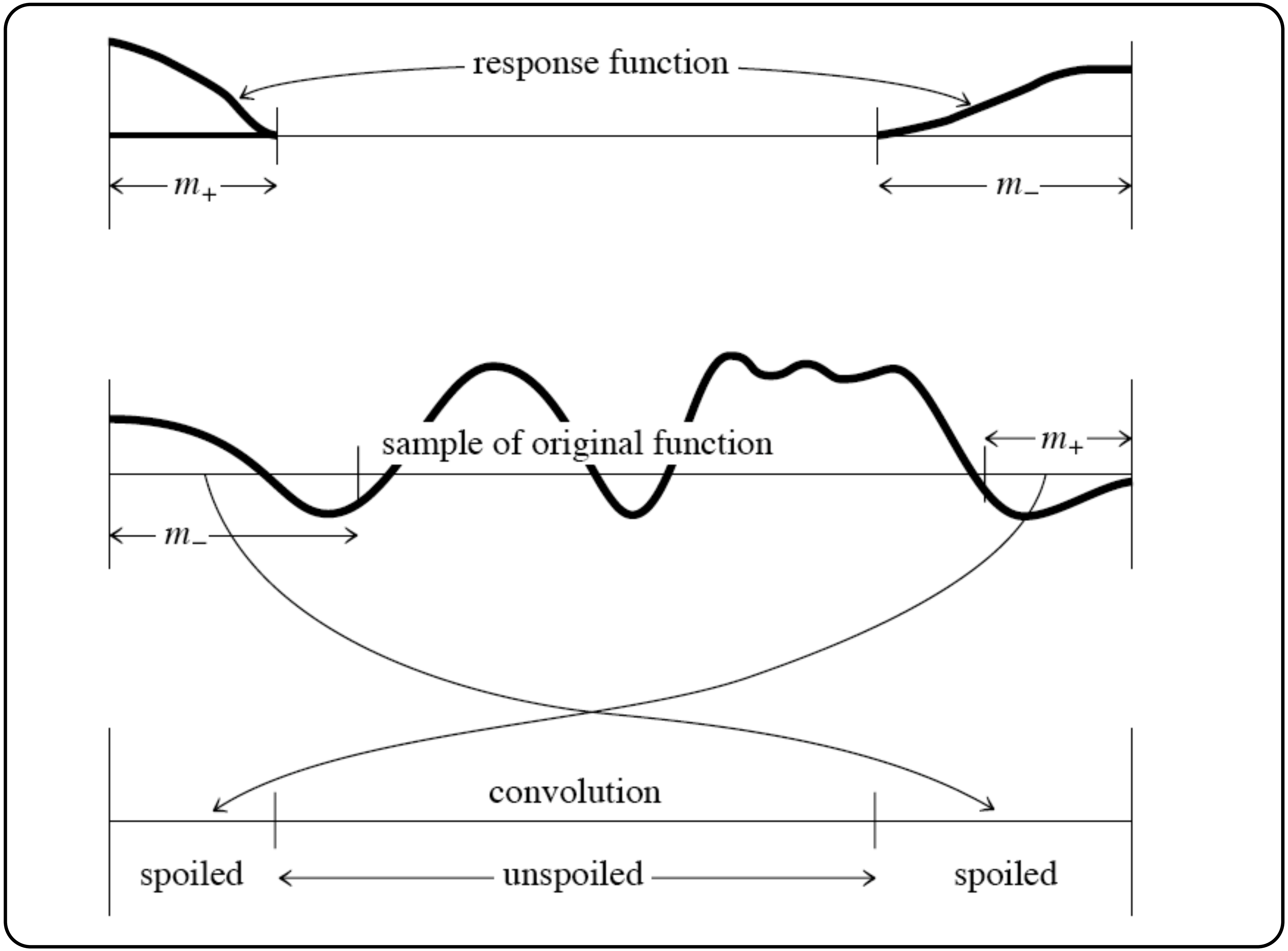
Discrete convolution theorem

$$(r * s)_j \equiv \sum_{k=-M/2+1}^{M/2} s_{j-k} r_k \Leftrightarrow S_n R_n$$

M only non-zero values of r_k



Discrete convolution theorem



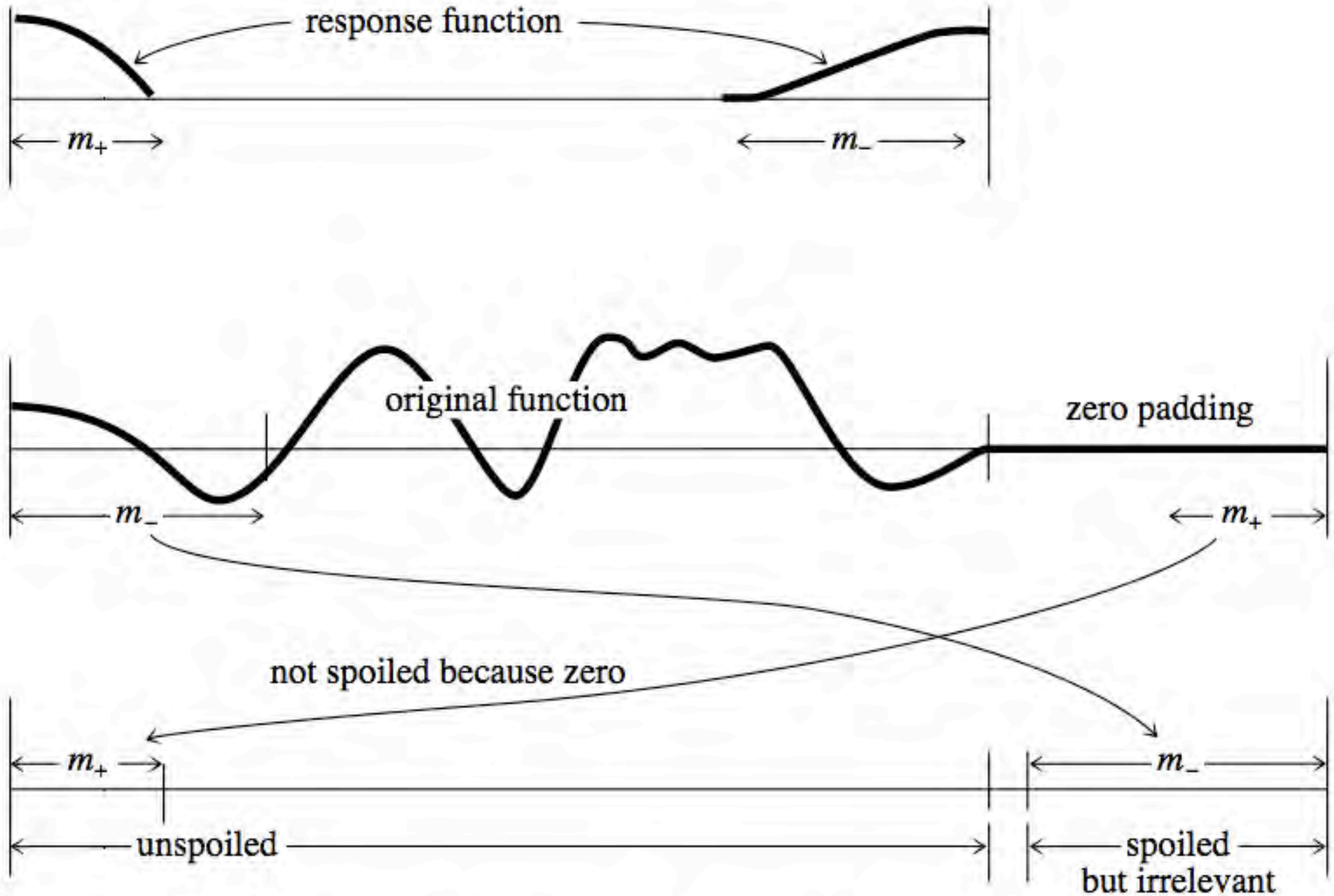


Figure 13.1.4. Zero padding as solution to the wrap-around problem. The original function is extended by zeros, serving a dual purpose: When the zeros wrap around, they do not disturb the true convolution; and while the original function wraps around onto the zero region, that region can be discarded.

Discrete convolution theorem

$$(r * s)_j \equiv \sum_{k=-M/2+1}^{M/2} s_{j-k} r_k \Leftrightarrow S_n R_n$$

Discrete deconvolution

$$\frac{\mathcal{F}(\mathrm{r} * \mathrm{s})_{\mathrm{j}}}{\mathrm{R_n}} = \mathrm{S_n}$$

However noise and uncertainties in response can make this process unreliable

Recap

Sampling: high frequencies are filtered out

Window: low frequencies are filtered out

→ leads to band limited data

Aka Filtering

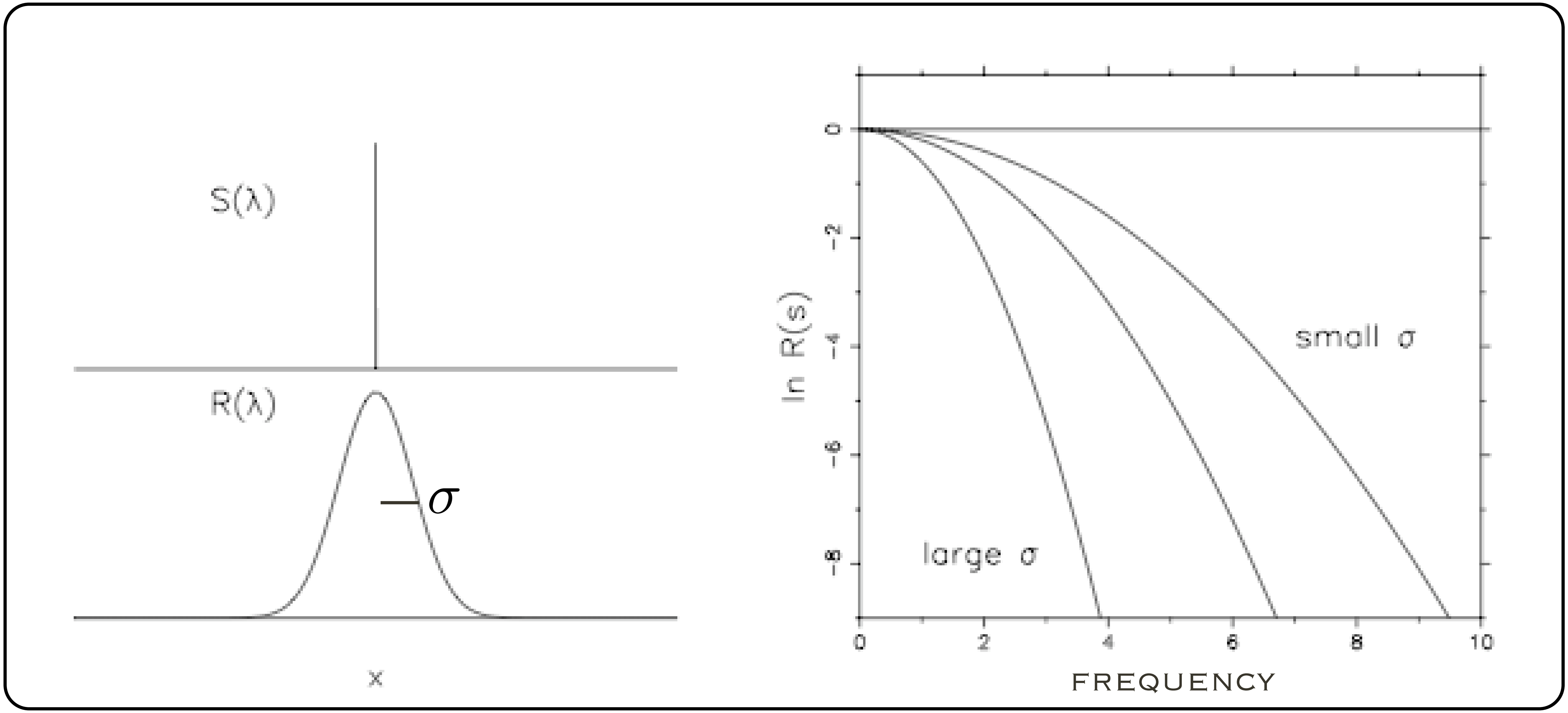
→ Frequency Filtering $Y(f) = X(f)H(f)$

$$y(t) = \int_{-\infty}^{\infty} x(t-\theta)h(\theta)d\theta$$

$$y(t) = x(t) * h(t)$$

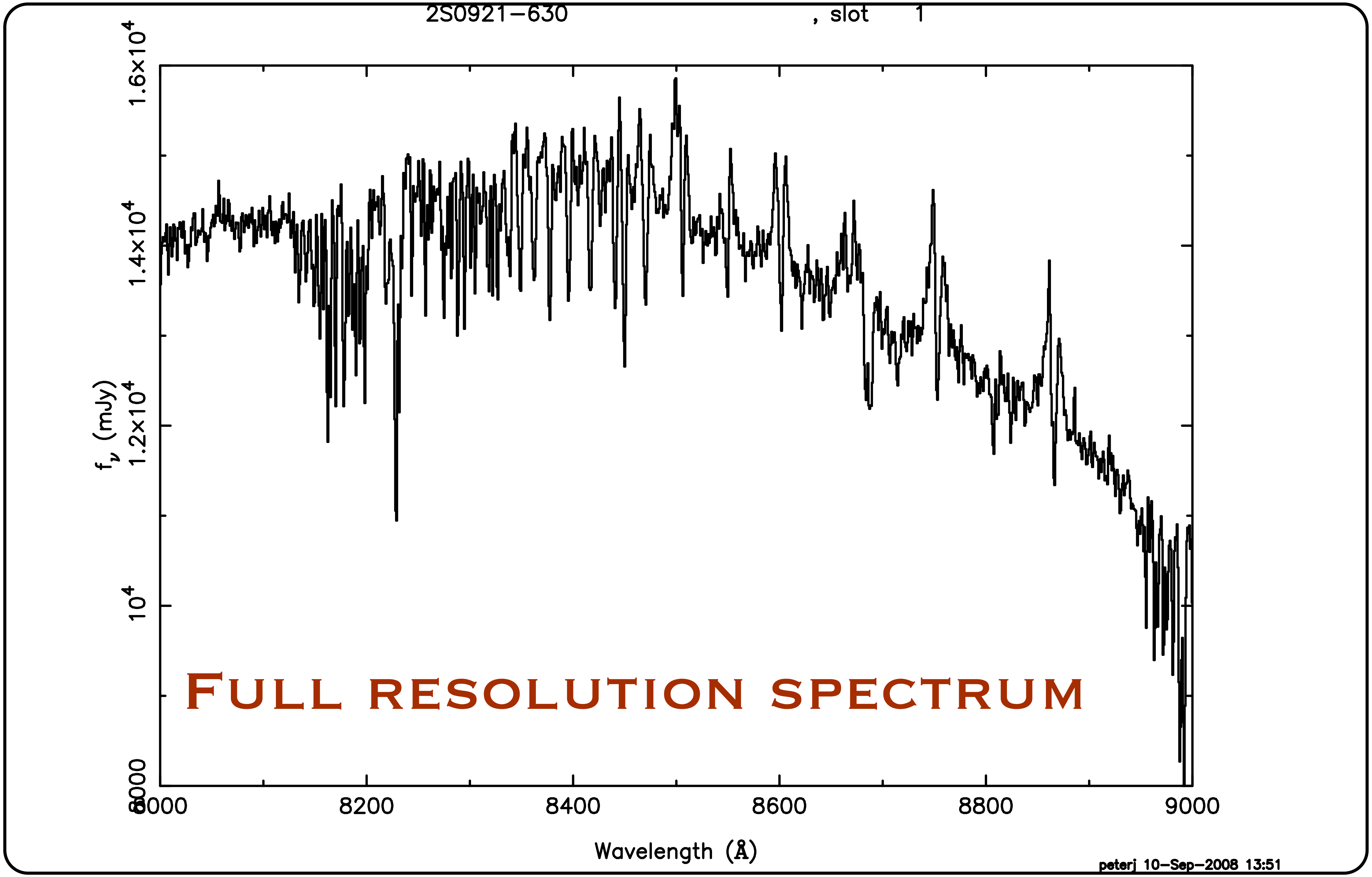
Filtering of process x with filter h

Gaussian response function

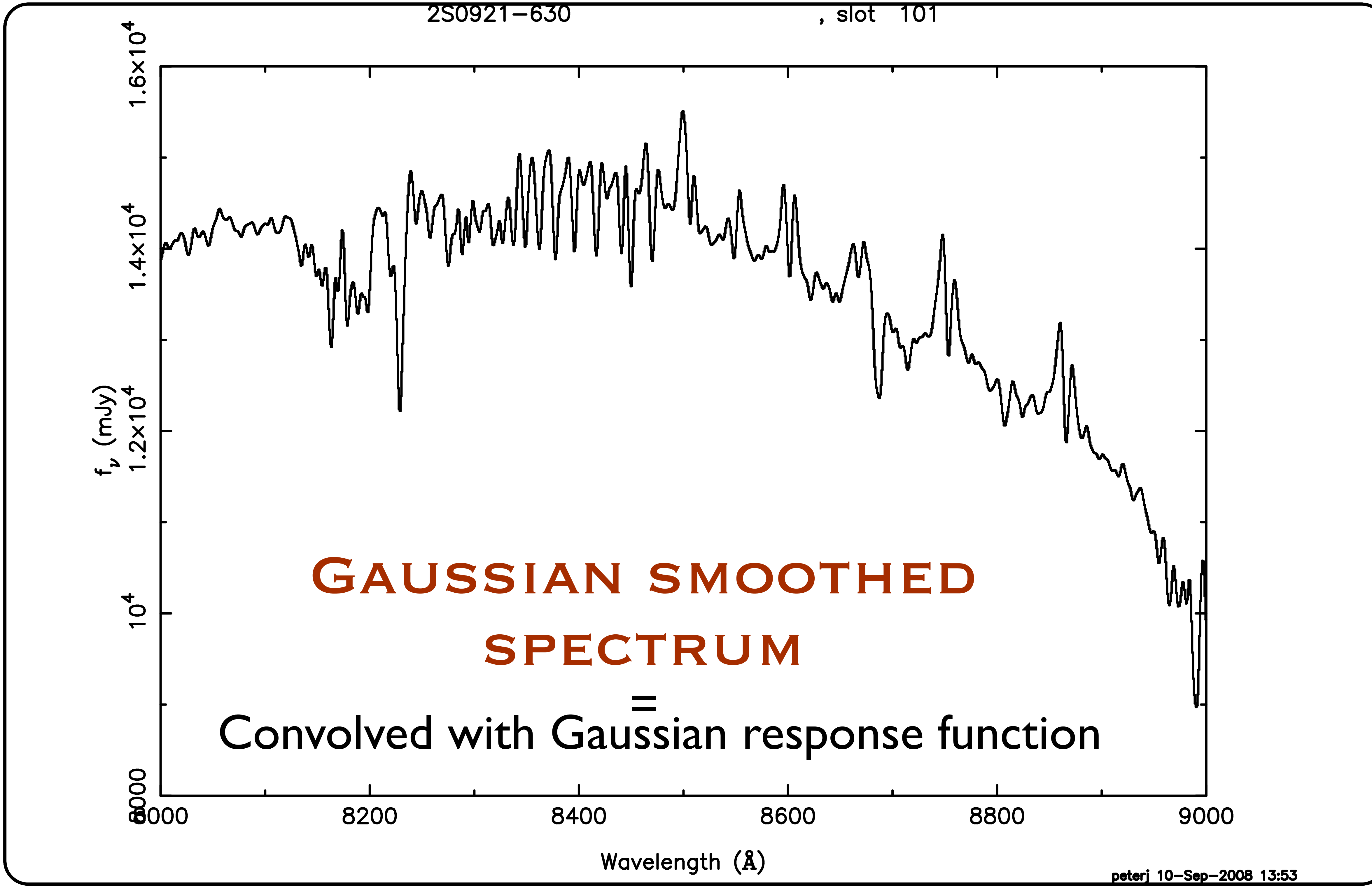


$$R(\lambda)\frac{1}{\sqrt{(2\pi)}\sigma}exp-(\frac{\lambda^2}{2\sigma^2})$$

Gaussian response function



Gaussian response function



Gaussian smoothed spectrum

=

Convolved with Gaussian response function

Recap

SAMPLING: HIGH FREQUENCIES ARE FILTERED OUT

WINDOW: LOW FREQUENCIES ARE FILTERED OUT

→

LEADS TO BAND LIMITED DATA

FILTERING

→ TIME FILTERING

MEASURE A PROCESS X(T) OVER INTERVAL T ASSUMED ZERO OUTSIDE T

$$\equiv y(t) = \Pi(\frac{t}{T})x(t)$$

$$Y(f) = X(f) * Tsinc(Tf)$$

ALL INFORMATION ABOUT FREQUENCIES <1/T IS LOST!

Recap

Data sampling: ideal case Nyquist criterium is fulfilled → sampling does not lead to loss of information RV completely described by the samples

Conditions **either:**

→ Band-limited response of the detector removes highest noise powers and the sampling is fast enough to cover the band limit of the detector **or:**

→ Signal is band-limited **and**

$$\nu_{\text{sampling}} > \nu_{\text{max,detector}} > \nu_{\text{max,signal}}$$