Chapter 8: Time series analysis
- Power spectrum and periodogram

- Time series
- Spectral density
- The periodogram for evenly spaced data
- The (Lomb-Scargle) periodogram for unevenly spaced data

Time series and periodogram analysis

A ***time series*** is an ordered sequence of data points, typically some variable quantity measured at successive times.

Examples:

- temperature measurements at a certain location every 3rd hour
- a sampled and digitally converted microphone signal
- the daily stock market index
- the magnitude or radial velocity of a star measured at irregular points in time

Time series can be regular (evenly sampled) or irregular (unevenly sampled).

A ***periodogram*** is a quantitative description of the amount of variation in a time series, separated into frequency components. Other related terms:

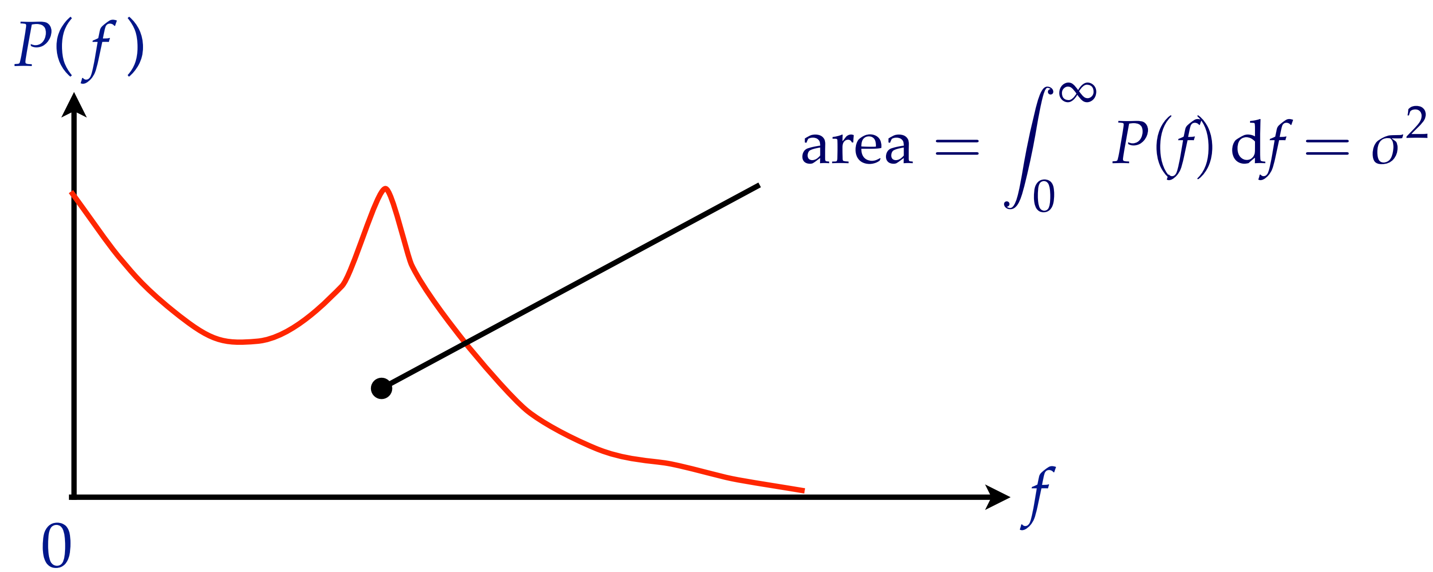
- spectrum
- power spectrum
- power spectral density
- amplitude spectrum

Power spectral density (PSD)

Let $h(t)$ be a continuous stationary process with $\mathrm{E}[h] = 0$ and $\mathrm{Var}[h] = \sigma^2$ at any t. (These are **ensemble averages**, taken over an infinite set of different possible realizations of the same random process.) In practice we only have the single realization $h(t)$, but for an **ergodic** process the time averages equal the ensemble averages:

$$\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} h(t)\,\mathrm{d}t = 0\,, \quad \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} h(t)^2\,\mathrm{d}t = \sigma^2$$

The power spectral density specifies how the variance is divided among different frequencies:



Note: In general both positive and negative frequencies must be considered, and $P(-f) \neq P(f)$. However, for **real-valued** processes $P(-f) = P(f)$, and $P(f)$ is then often doubled such that only $f \geq 0$ should be included in the integral — this is known as the "one-sided power spectrum" and is always used here.

Periodogram

In practice we can only get an **estimate** of the power spectral density of the process. The **periodogram** is such an estimate based on a finite set of data, covering the time interval T sampled in N discrete points.

The periodogram has important limitations compared with the theoretical PSD:

- Finite T implies finite frequency resolution $\Delta f = 1/T$ (and a minimum frequency).
- Finite number of data points N implies that at most N (independent) frequency components can be estimated (and a maximum frequency) - aliasing
- Truncation at the endpoints cause distortion of the PSD (frequency leakage).

From a statistical viewpoint the periodogram is a statistic (it is computed from the data), and like other sample statistics is has uncertainties.

To decide if a signal is periodic or not can be treated as a hypothesis test.

We will consider separately:

- Periodograms for evenly sampled data $\Rightarrow$ discrete Fourier transform
- Periodograms for unevenly sampled data $\Rightarrow$ Lomb-Scargle periodogram

The periodogram for evenly sampled data

Let $h(t)$ be a **real-valued** continuous process sampled at discrete times t_j ($j = 0, 1, \ldots, N-1$): $h_j = h(t_j)$.

- The set of paired values $\{(t_j, h_j), j = 0, 1, \ldots, N-1\}$ is called a **time series**.
- For an evenly sampled time series, $t_j = t_0 + j\Delta t$, where Δt is the **sampling interval**.
- The sampling frequency is $f_s = 1/\Delta t$, and the Nyquist frequency is $f_{\rm Ny} = f_s/2 = 1/(2\Delta t)$
- The total **length** of the time series is defined as $T = N\Delta t$. (Note that $t_{N-1} - t_0 = (N-1)\Delta t < T$.)
- The **periodogram** is the continuous function

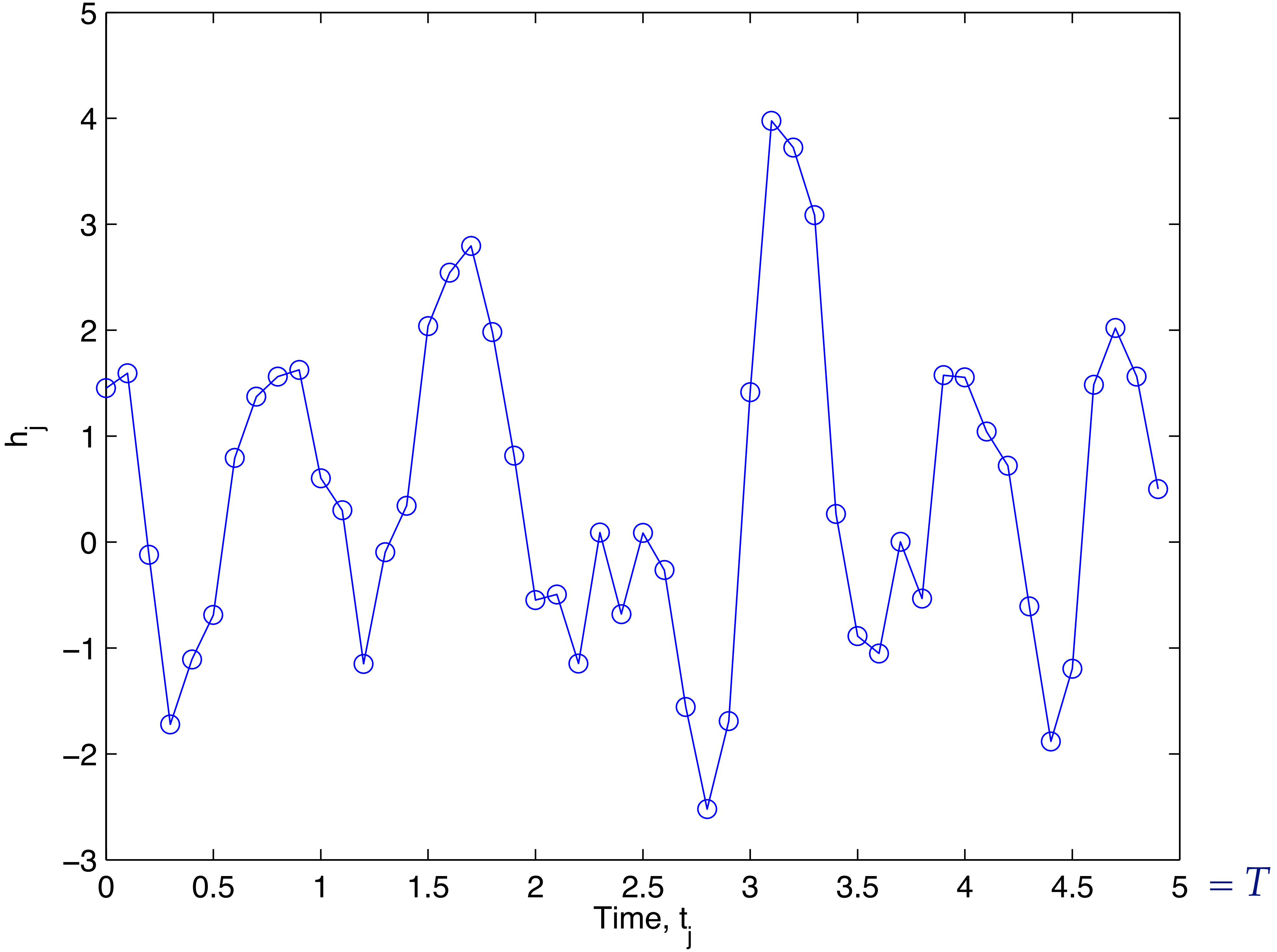
$$\hat{P}(f) = \frac{2\Delta t}{N}\left|\sum_{j=0}^{N-1} h_j \exp(-i\,2\pi j\Delta t\, f)\right|^2$$

where i is the imaginary unit. Some important properties of the periodogram:

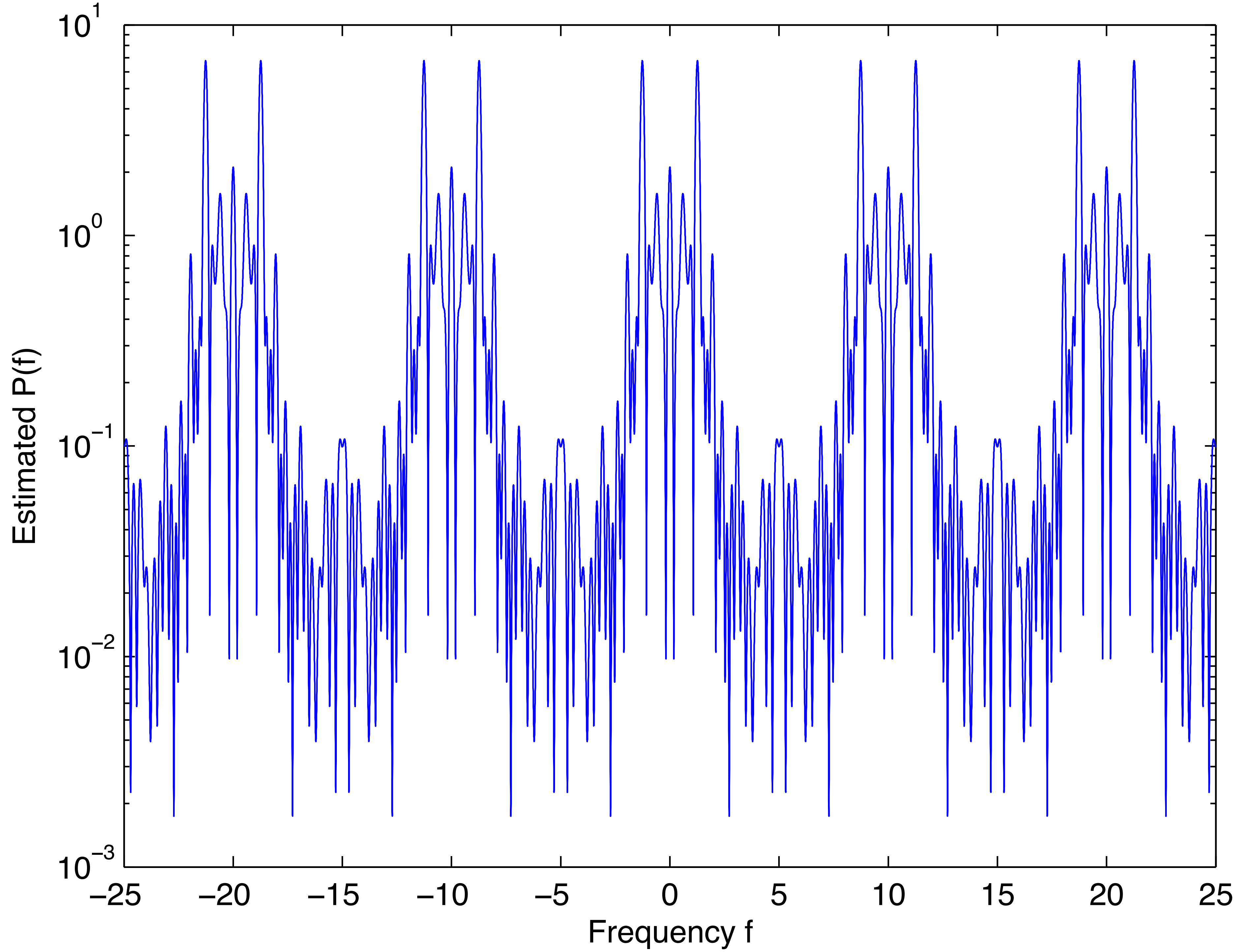
$\hat{P}(\pm f + k f_{\rm s}) = \hat{P}(f)$ (illustrated on next pages)

$$\mathrm{E}\left[\int_0^{f_{\rm Ny}} \hat{P}(f)\,\mathrm{d}f\right] = \sigma^2 \qquad \lim_{N\to\infty}\int_{f_1}^{f_2}\hat{P}(f)\,\mathrm{d}f = \sum_{k=-\infty}^{\infty}\int_{f_1}^{f_2} P(f+kf_{\rm s})\,\mathrm{d}f$$

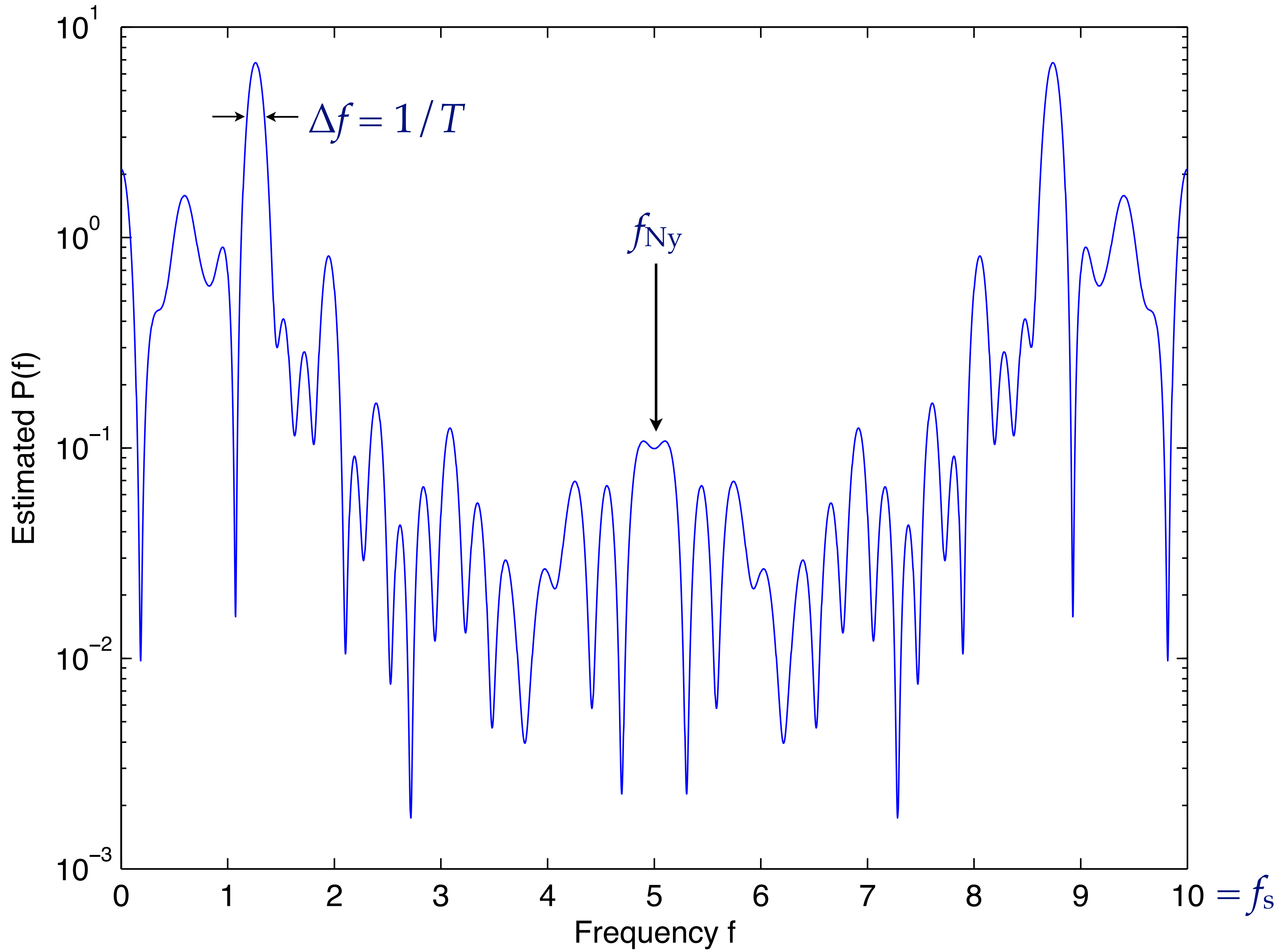
Time series example (N = 50, Δt = 0.1)



Periodogram ($f_s = 1/\Delta t = 10$): periodic with period f_s



Periodogram (only $0 < f < f_s$): mirrored at f_{Ny}



Calculating the periodogram using FFT

The periodogram for real-valued, evenly spaced data is defined by

$$\hat{P}(f) = \frac{2\Delta t}{N} \left| \sum_{j=0}^{N-1} h_j \exp(-i\,2\pi j \Delta t f) \right|^2$$

However, it should never be calculated explicitly from this formula. Mathematically equivalent but much (much!) faster is to use the Fast Fourier Transform (FFT), which is a clever ($\sim N \times \log N$) algorithm to compute the Discrete Fourier Transform

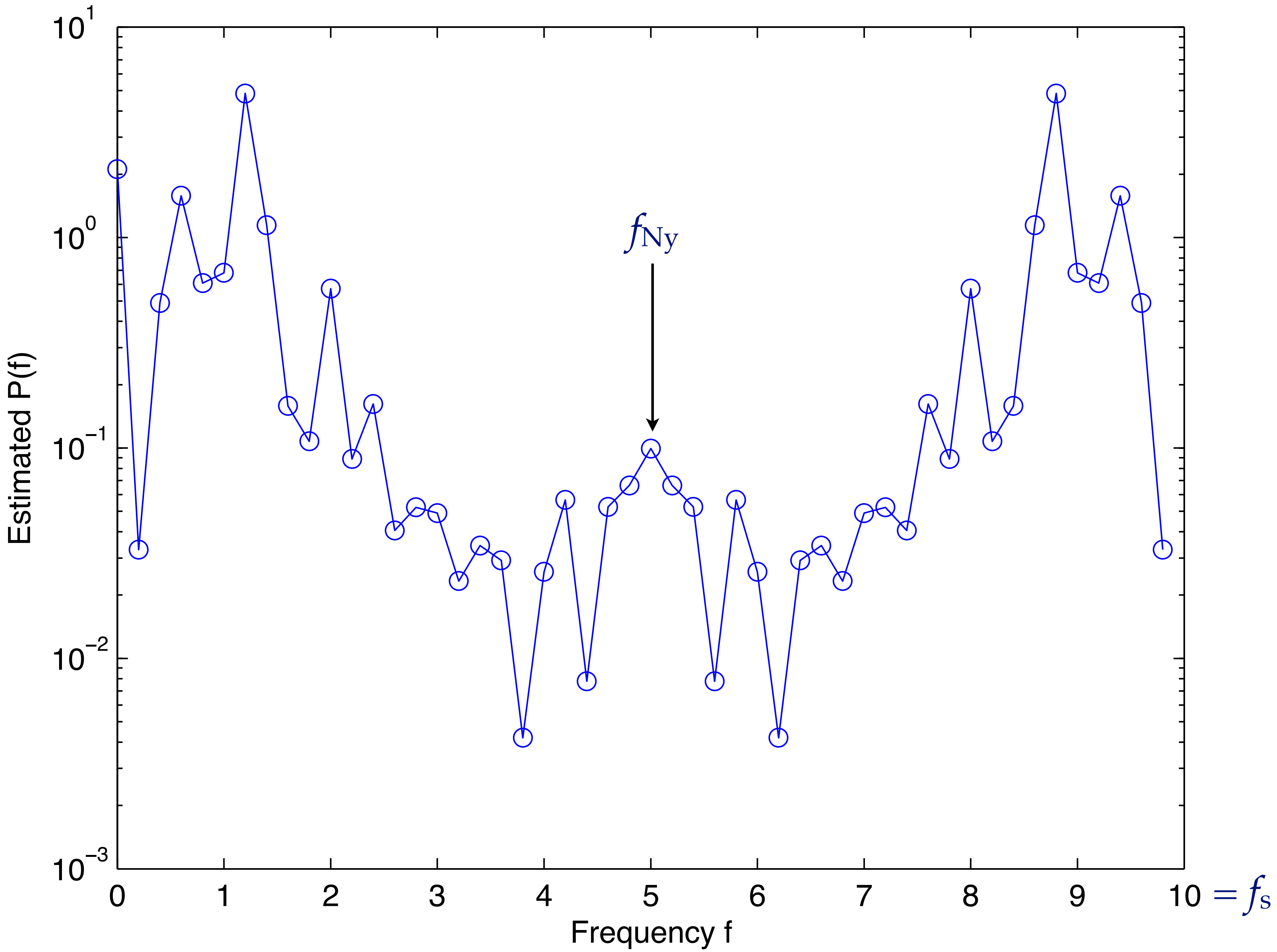
$$X_k = \sum_{j=0}^{M-1} x_j \exp(-i\,2\pi k j)\,, \quad k = 0, 1, \ldots, M-1$$

for any array of M (in general complex) values x_0, x_1, ..., x_{M-1}. The results X_k are in general complex (even if x_j are real). Putting $M = N$ and $x_j = h_j$, we clearly have

$$\hat{P}(f_k) = \frac{2\Delta t}{N} |X_k|^2\,, \quad f_k = k/\Delta t\,, \quad k = 0, 1, \ldots, N-1$$

This is shown on the next page. A disadvantage is that the periodogram is only computed for the discrete frequencies $f_k = k/\Delta t$ (circles in the diagram on next page).

Periodogram sampled at f = 0, Δf, 2Δf, ...



Interpolating the periodogram using FFT

Actually the FFT can be used to compute the periodogram for arbitrarily dense frequency points. For example, if the periodogram of the previous time series ($N = 50$, $T = 5$) needs to be sampled four times denser than in the previous plot (i.e., using $\Delta f = 0.05$ instead of 0.2), one simply makes the time series four times longer by adding zeroes:

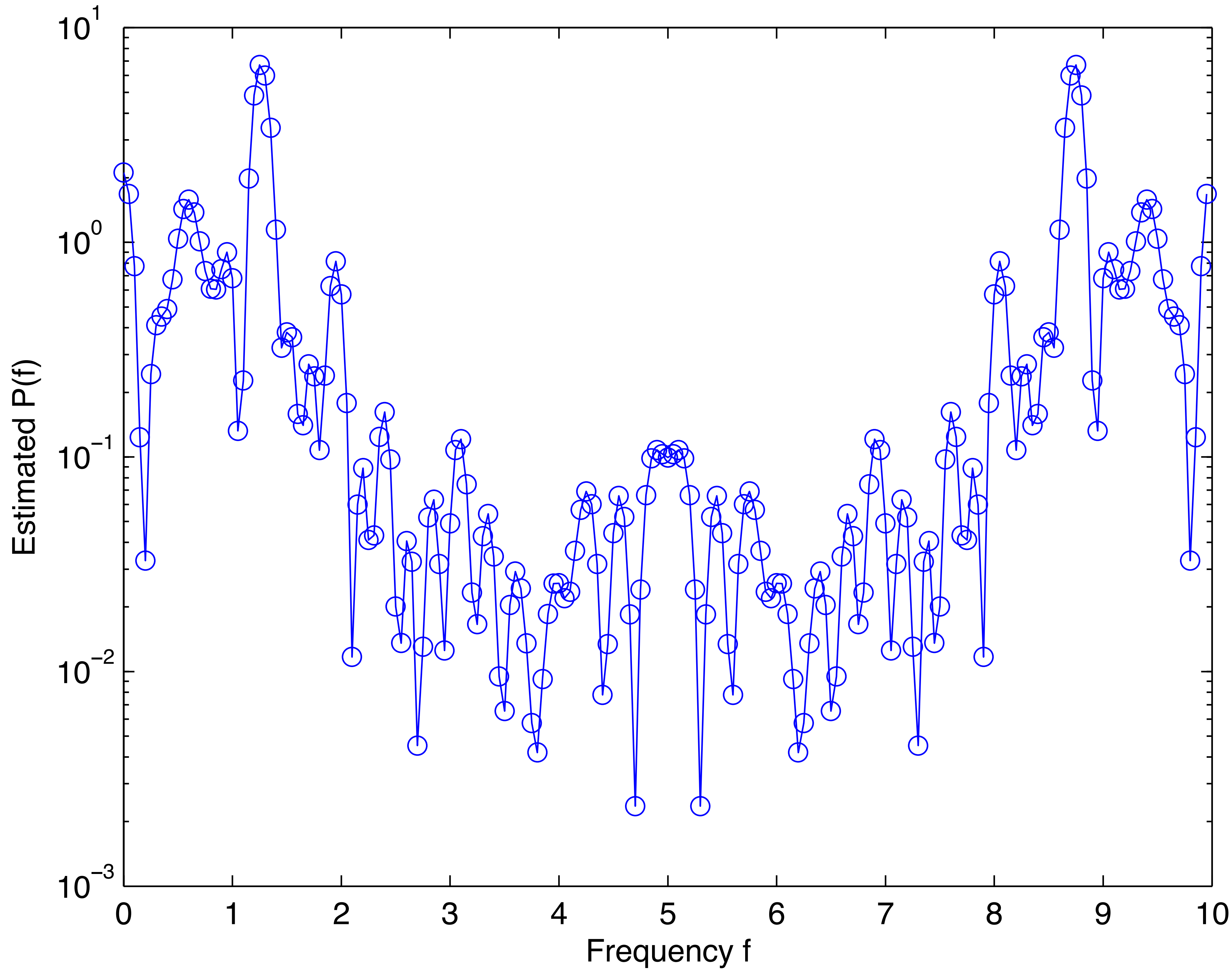
$$M = 4N\,, \quad x_j = \begin{cases} h_j & \text{if } j < N \\ 0 & \text{otherwise} \end{cases}\,, \quad \{X_k\} = \mathrm{FFT}(\{x_j\})$$

$$\Rightarrow \quad \hat{P}(f_k) = \frac{2\Delta t}{N}\,|X_k|^2\,, \quad f_k = k/(M\Delta t)\,, \quad k = 0, 1, \ldots, M-1$$

See next page, where the circles show the calculated points. (The "continuous" periodogram on p. 11 was computed this way with an 80 times oversampling.)

Note: The zero-padding just provides an interpolation of the basic periodogram at $f_k = k/(N\Delta t)$ and therefore does not add any information, only makes the plot "nicer".

Four times oversampled periodogram



Illustrating sampling / aliasing / truncation effects

$$c(t) = A\sin(2\pi f t + 0.5)$$

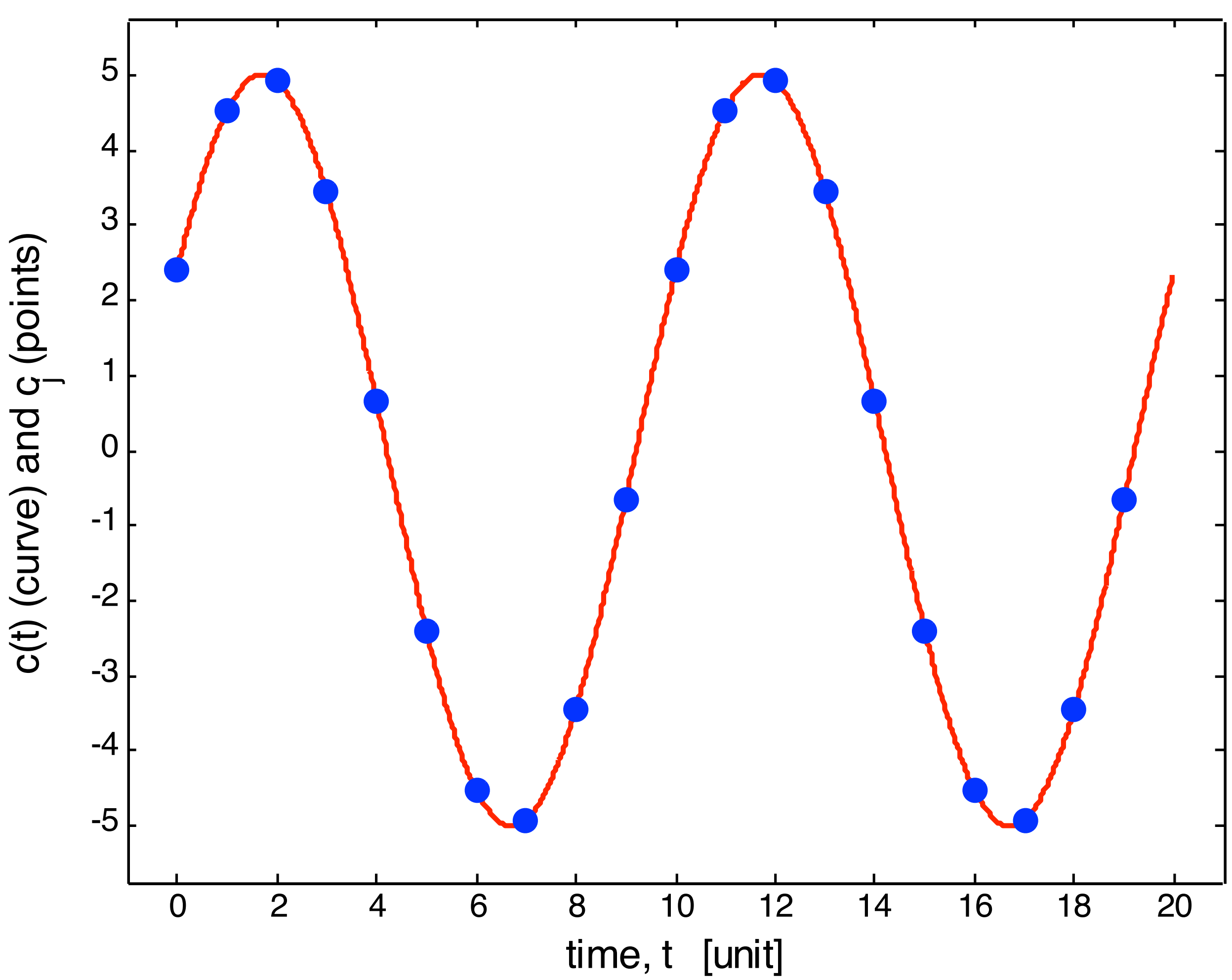
$$A = 5$$

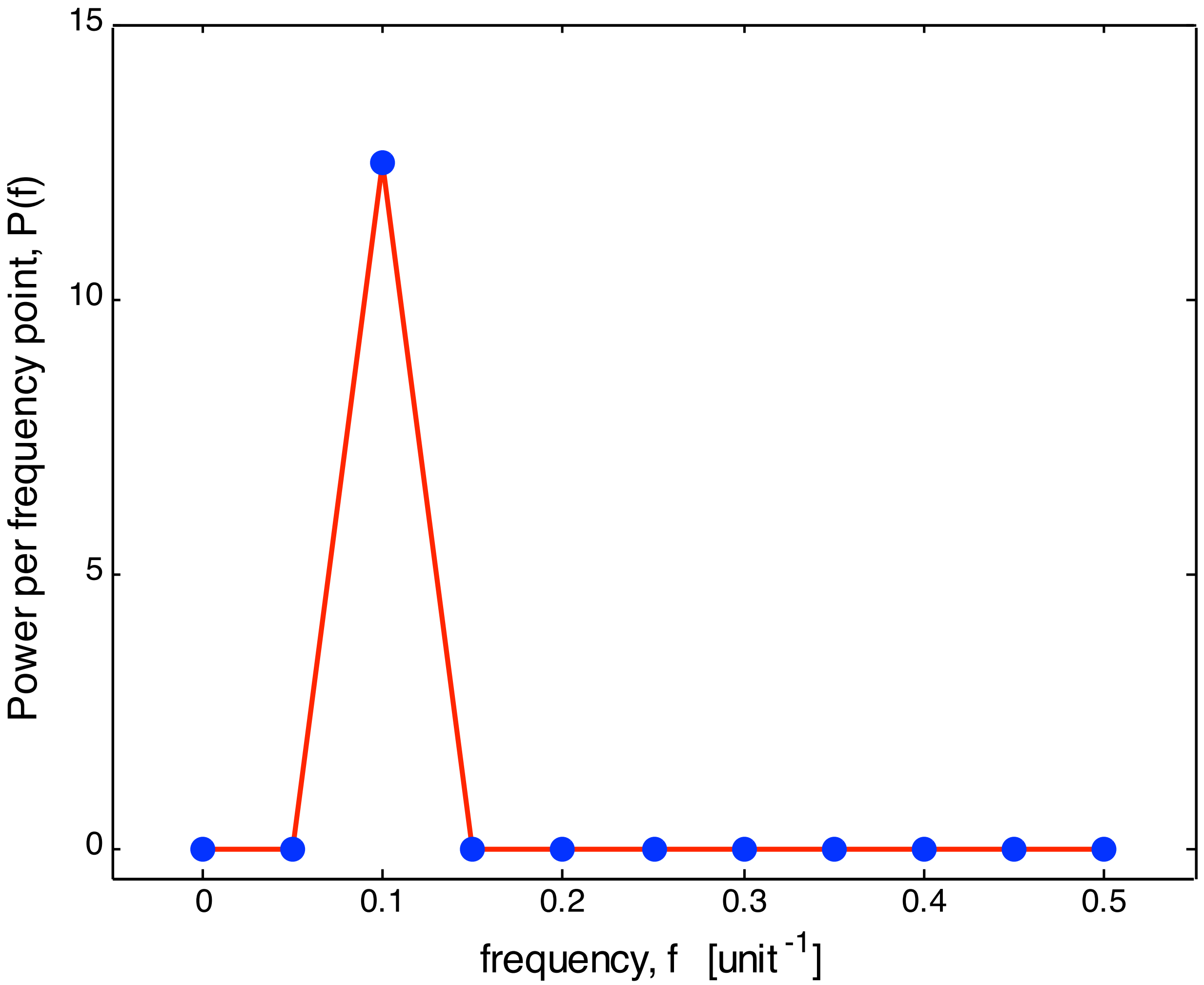
$$f = 0.1$$

$$N = 20$$

$$\Delta = 1$$

$$f_c = 0.5$$





Example periodograms, noise-free data

$$c(t) = A\sin(2\pi f t + 0.5)$$

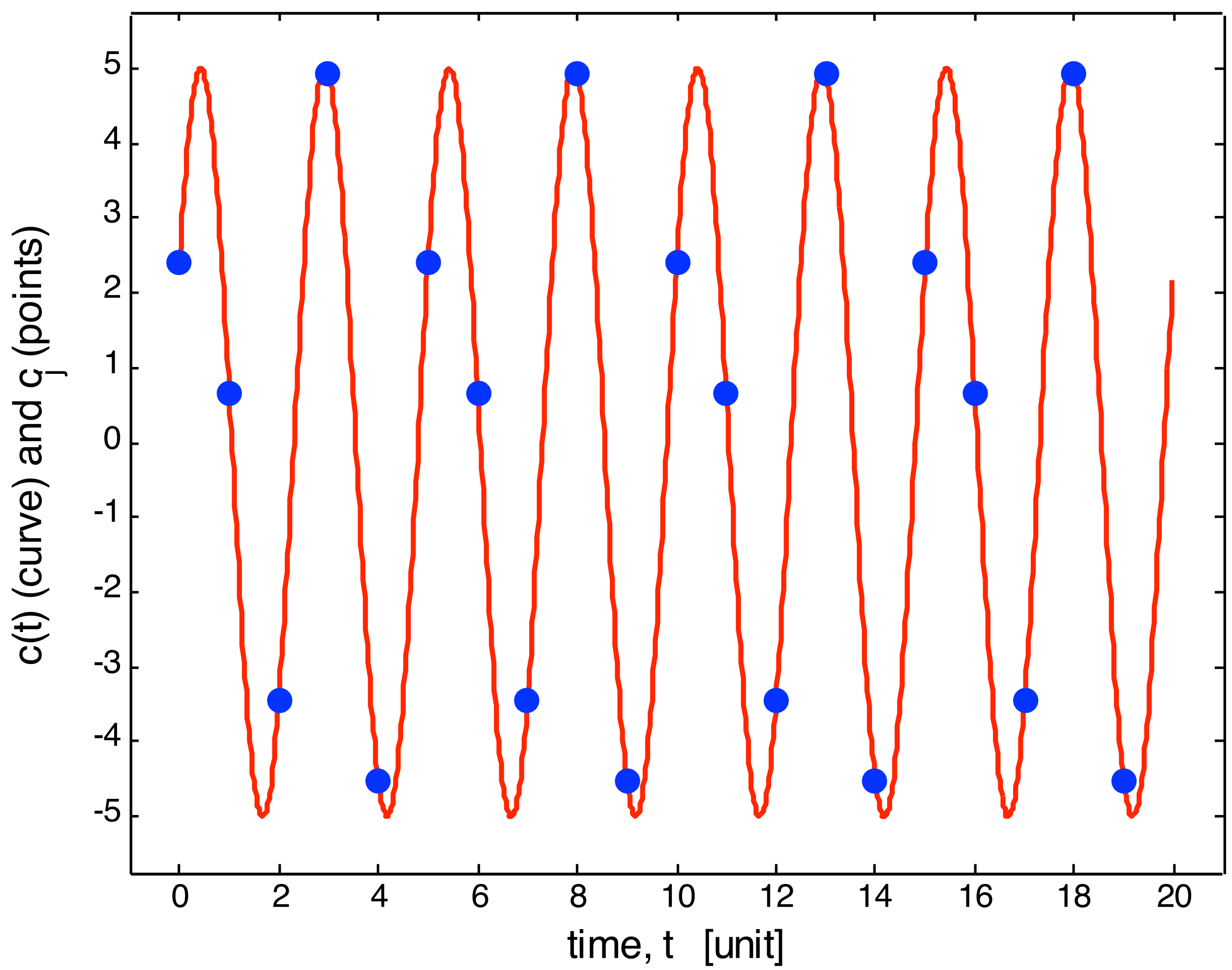
$$A = 5$$

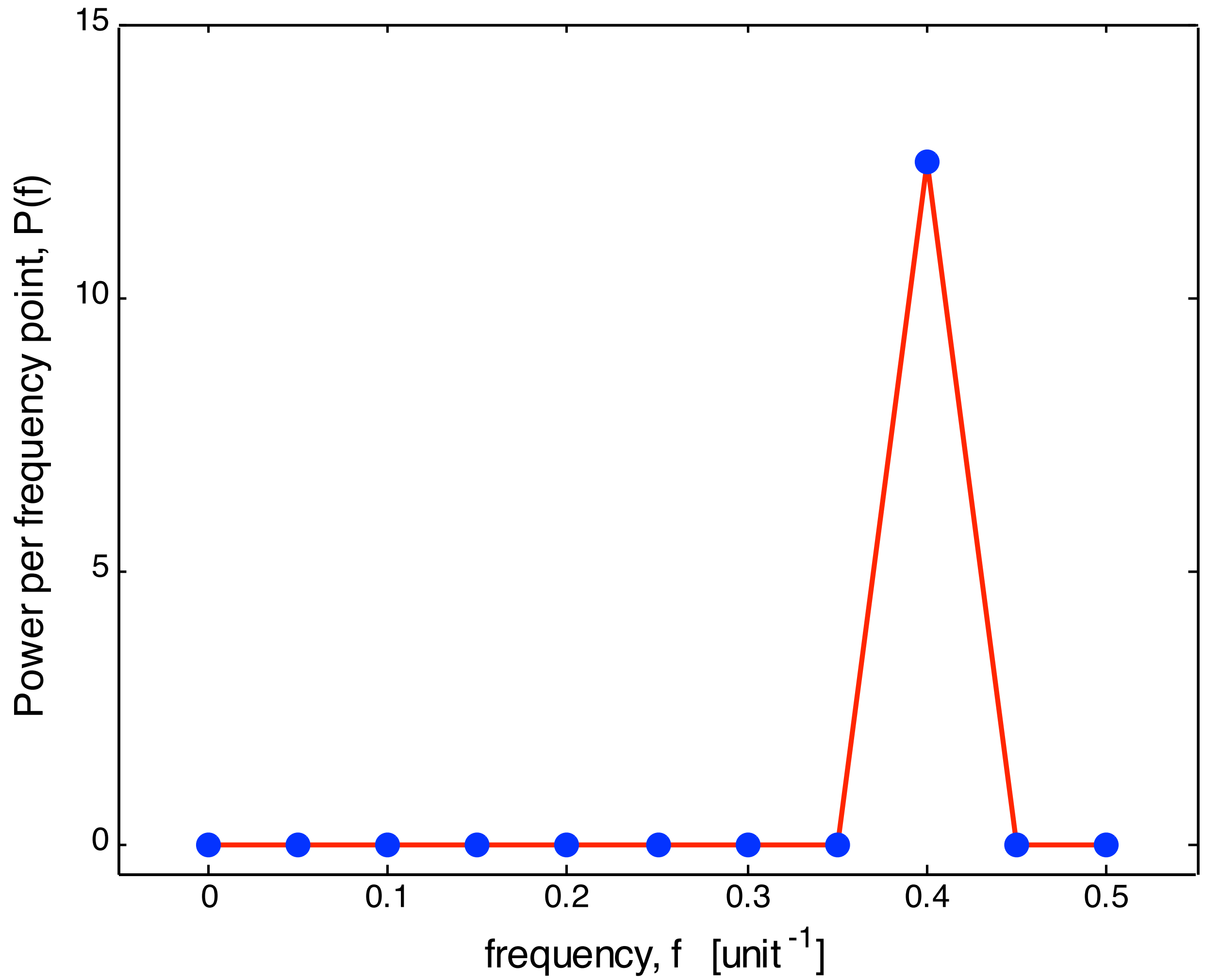
$$f = 0.4$$

$$N = 20$$

$$\Delta = 1$$

$$f_c = 0.5$$





Example periodograms, noise-free data

$$c(t) = A\sin(2\pi f t + 0.5)$$

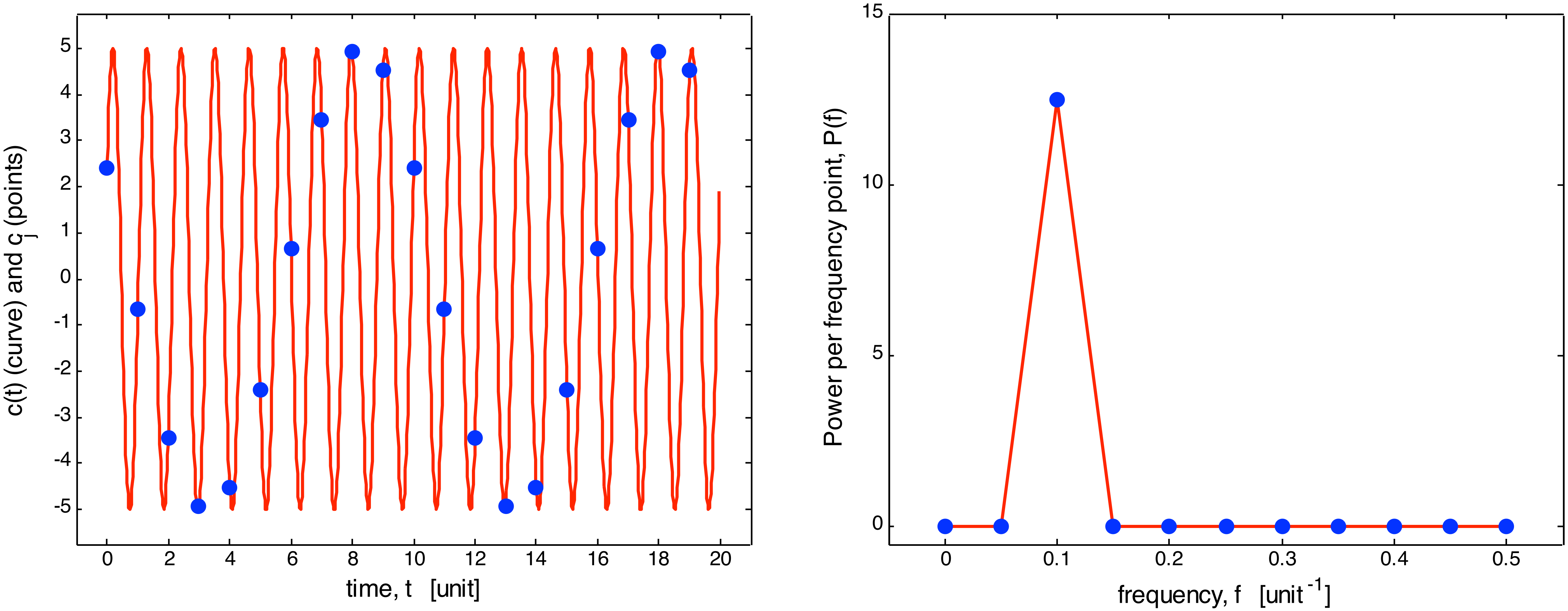
$$A = 5$$

$$f = 0.9$$

$$N = 20$$

$$\Delta = 1$$

$$f_c = 0.5$$



Aliasing: $f > f_c$ is reflected in f_c

Example periodograms, noise-free data

$$c(t) = A\sin(2\pi f t + 0.5)$$

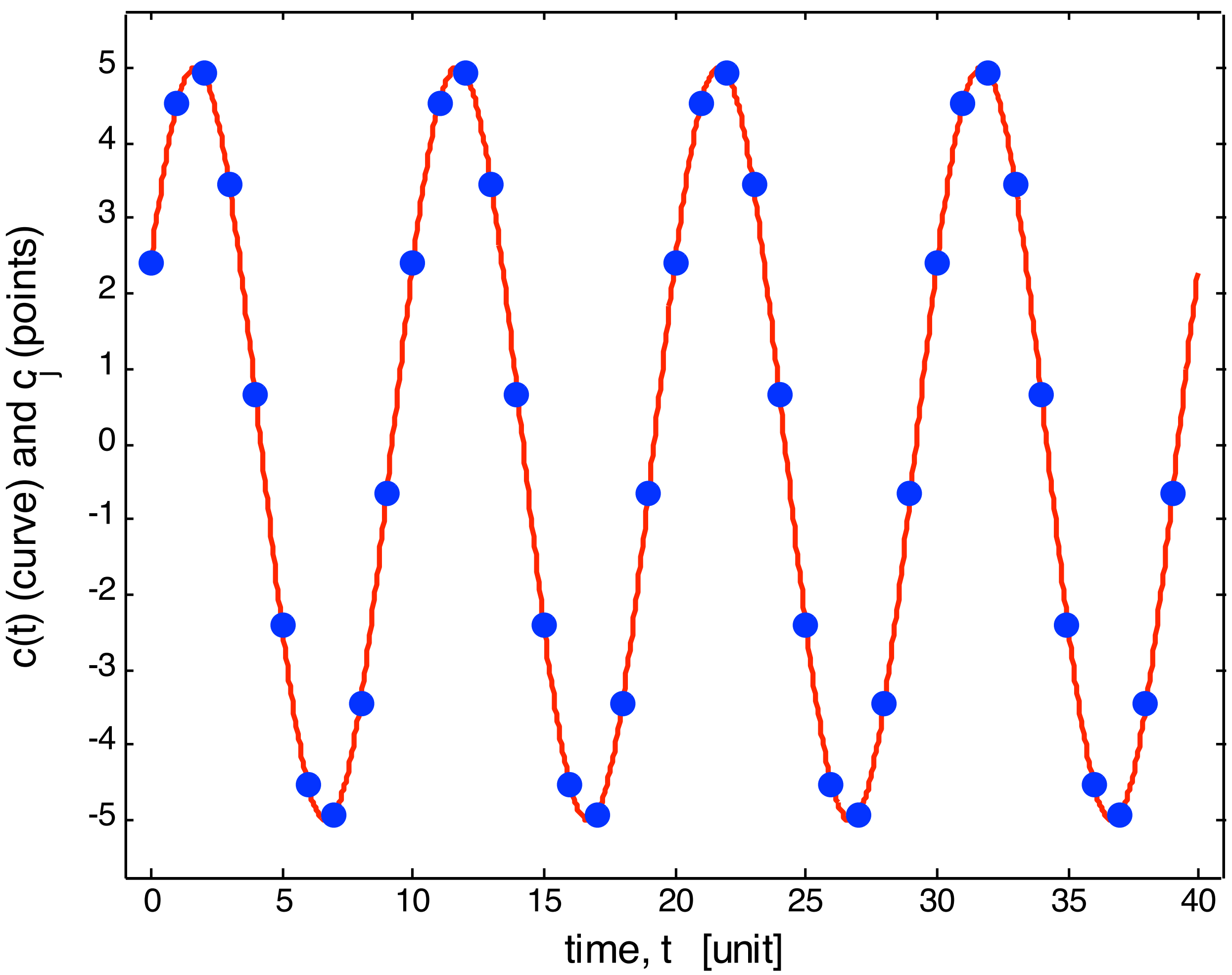
$$A = 5$$

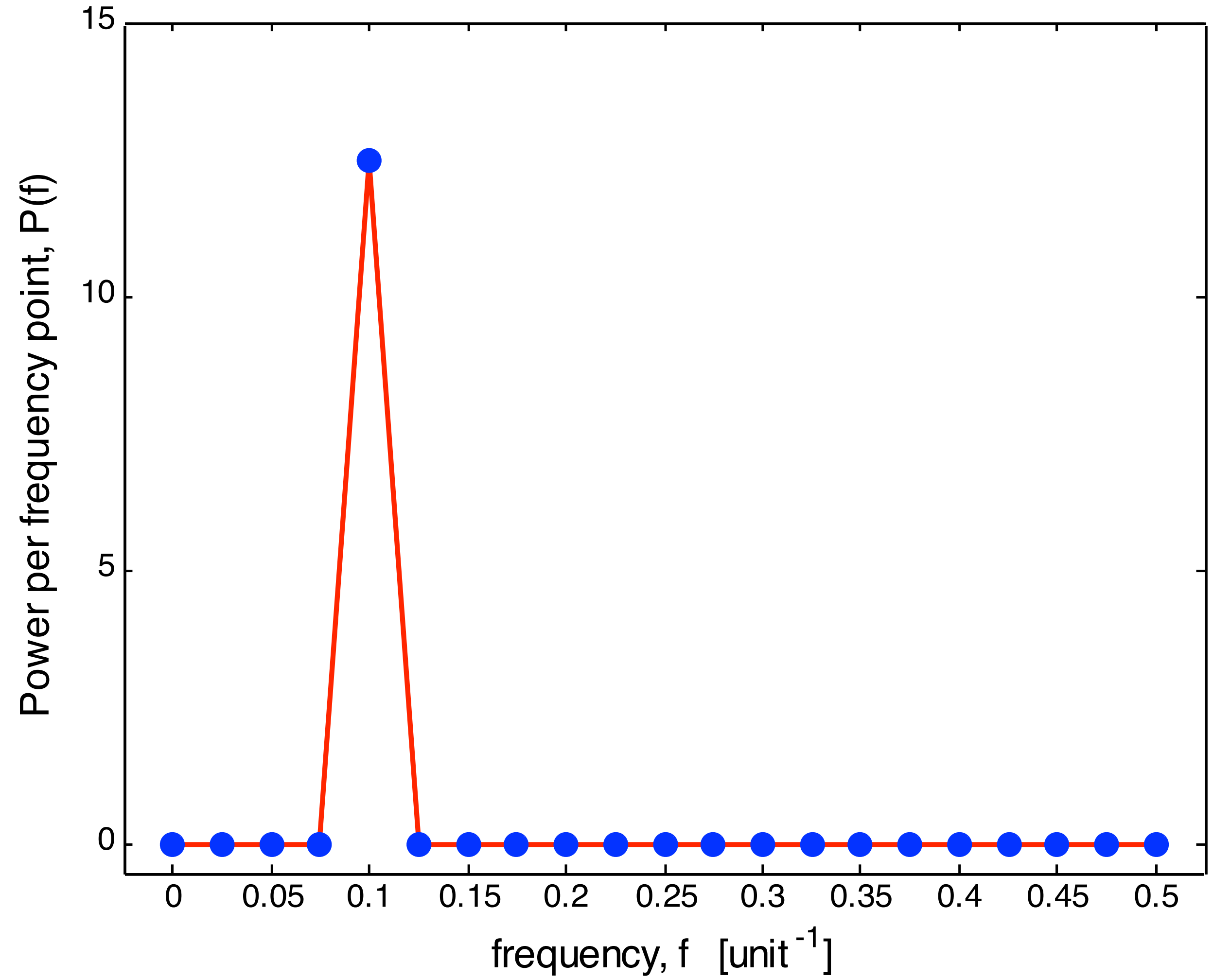
$$f = 0.1$$

$$N = 40$$

$$\Delta = 1$$

$$f_c = 0.5$$





Longer time span $\Rightarrow$ higher frequency resolution

Example periodograms, noise-free data

$$c(t) = A\sin(2\pi f t + 0.5)$$

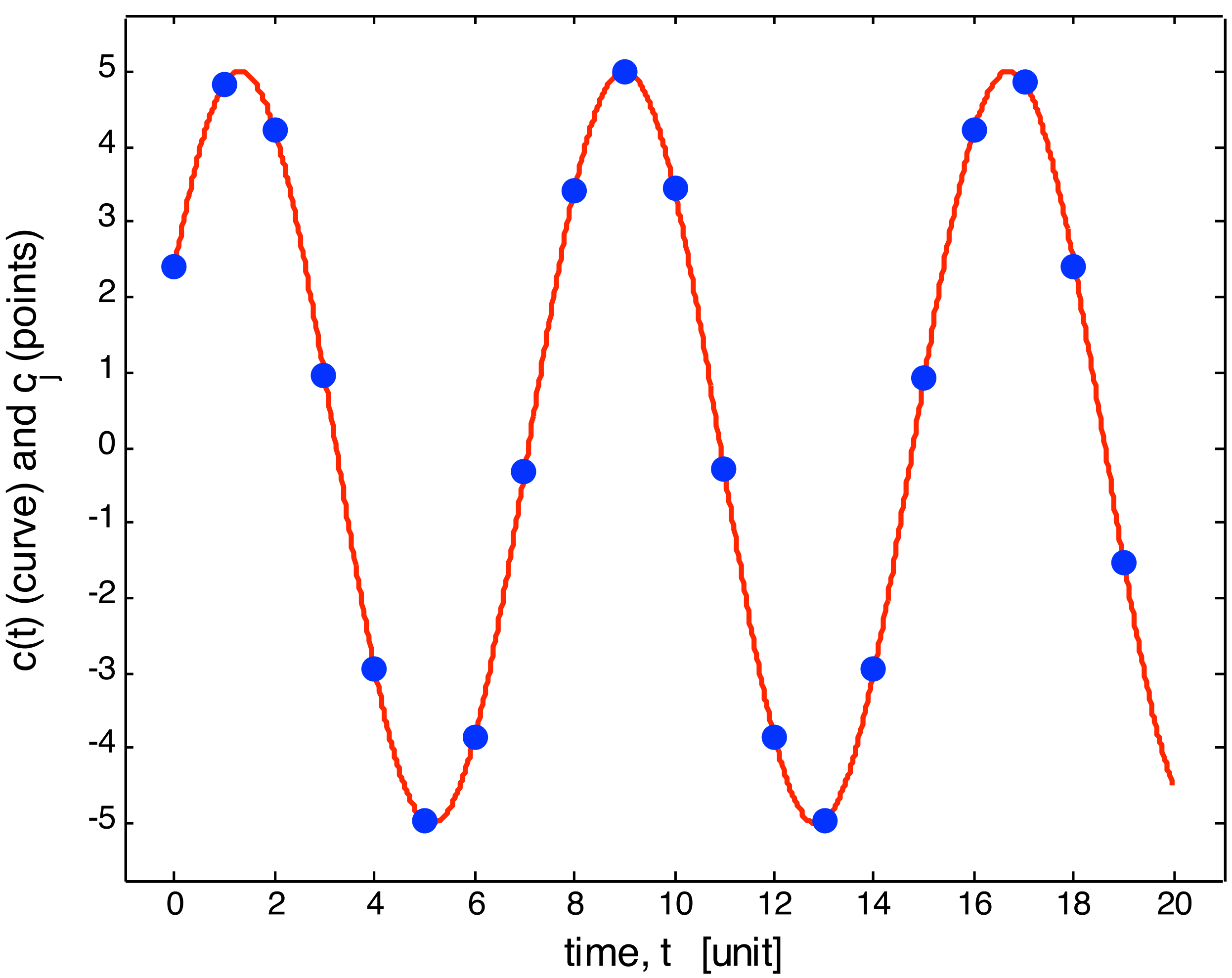
$$A = 5$$

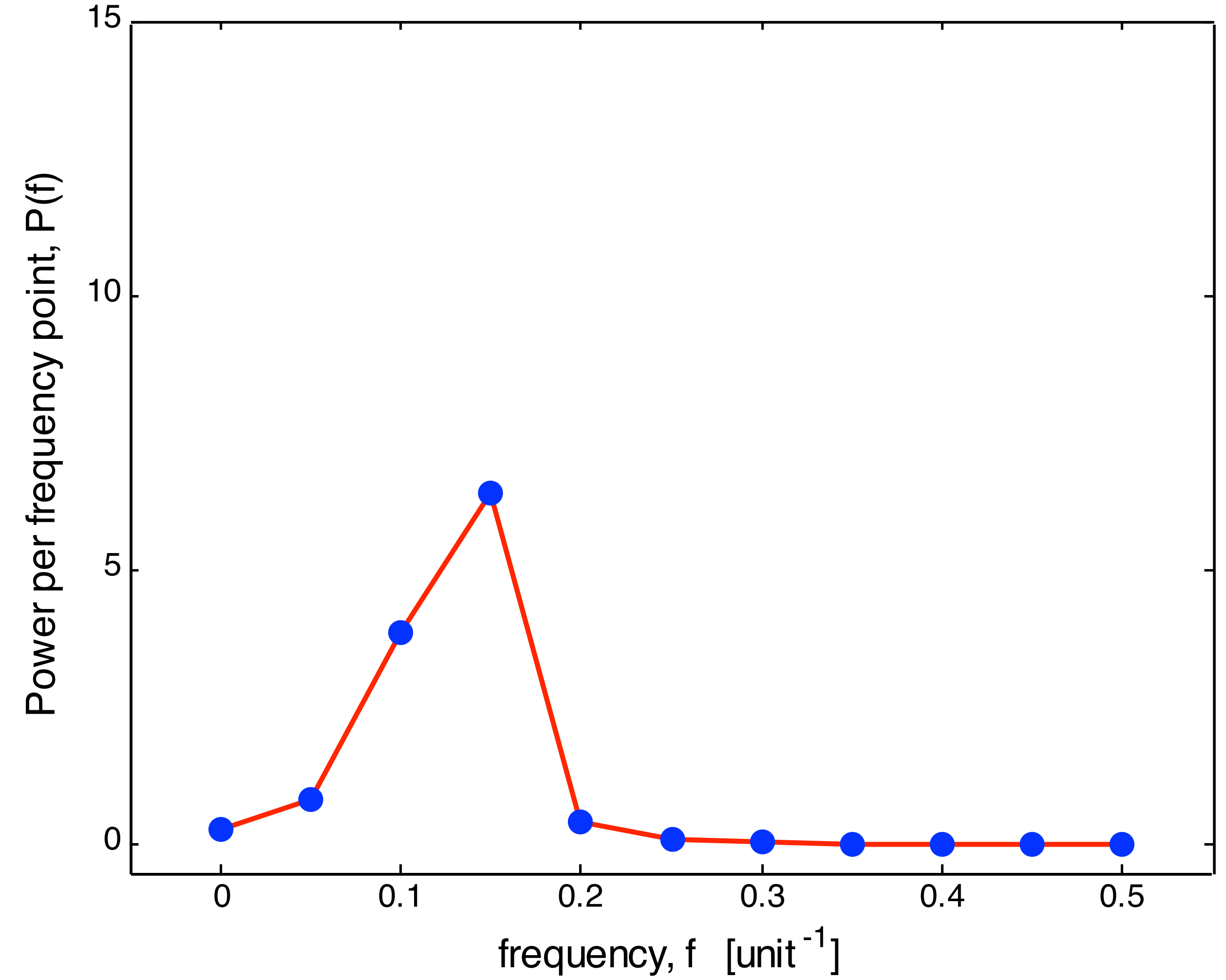
$$f = 0.13$$

$$N = 20$$

$$\Delta = 1$$

$$f_c = 0.5$$



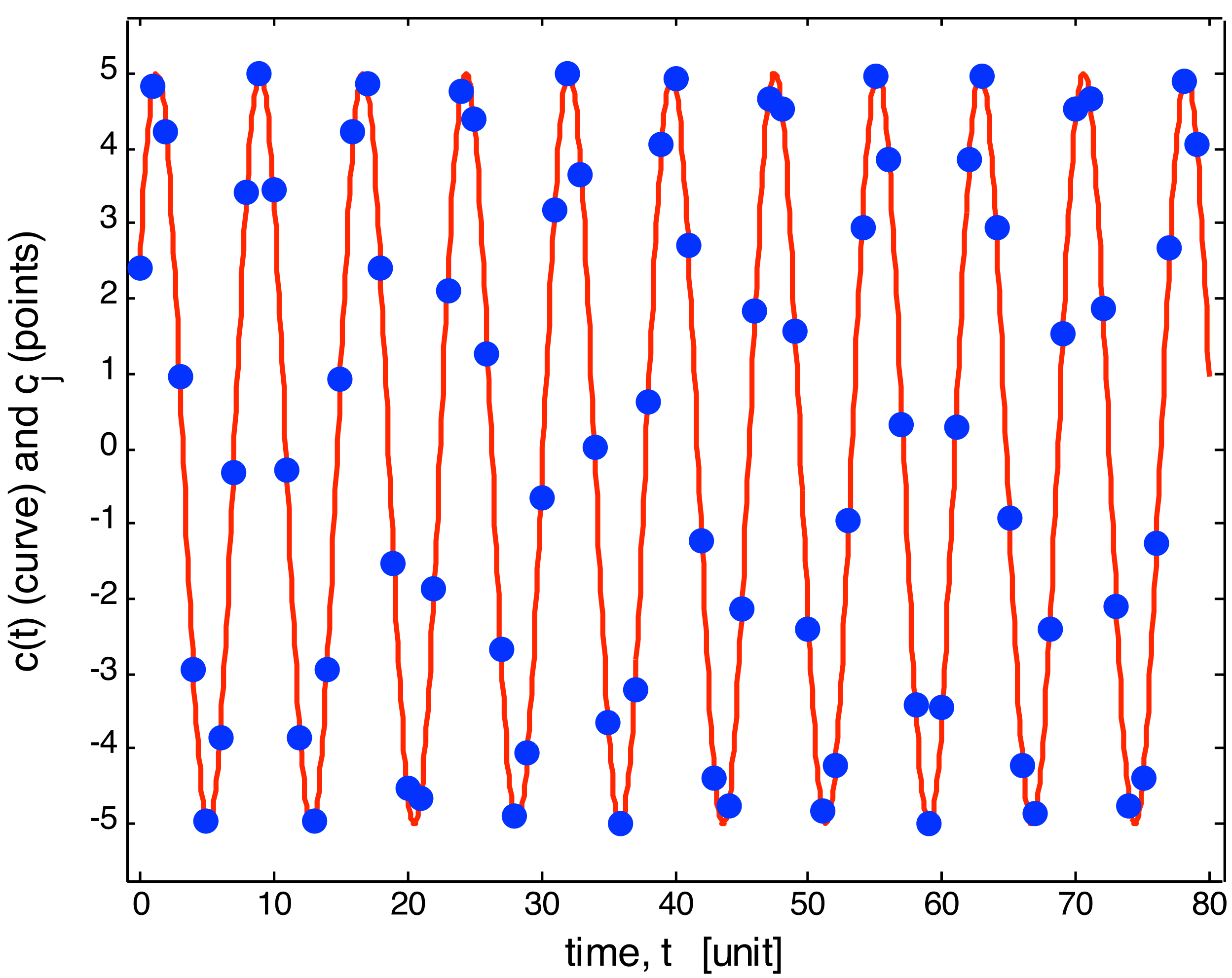


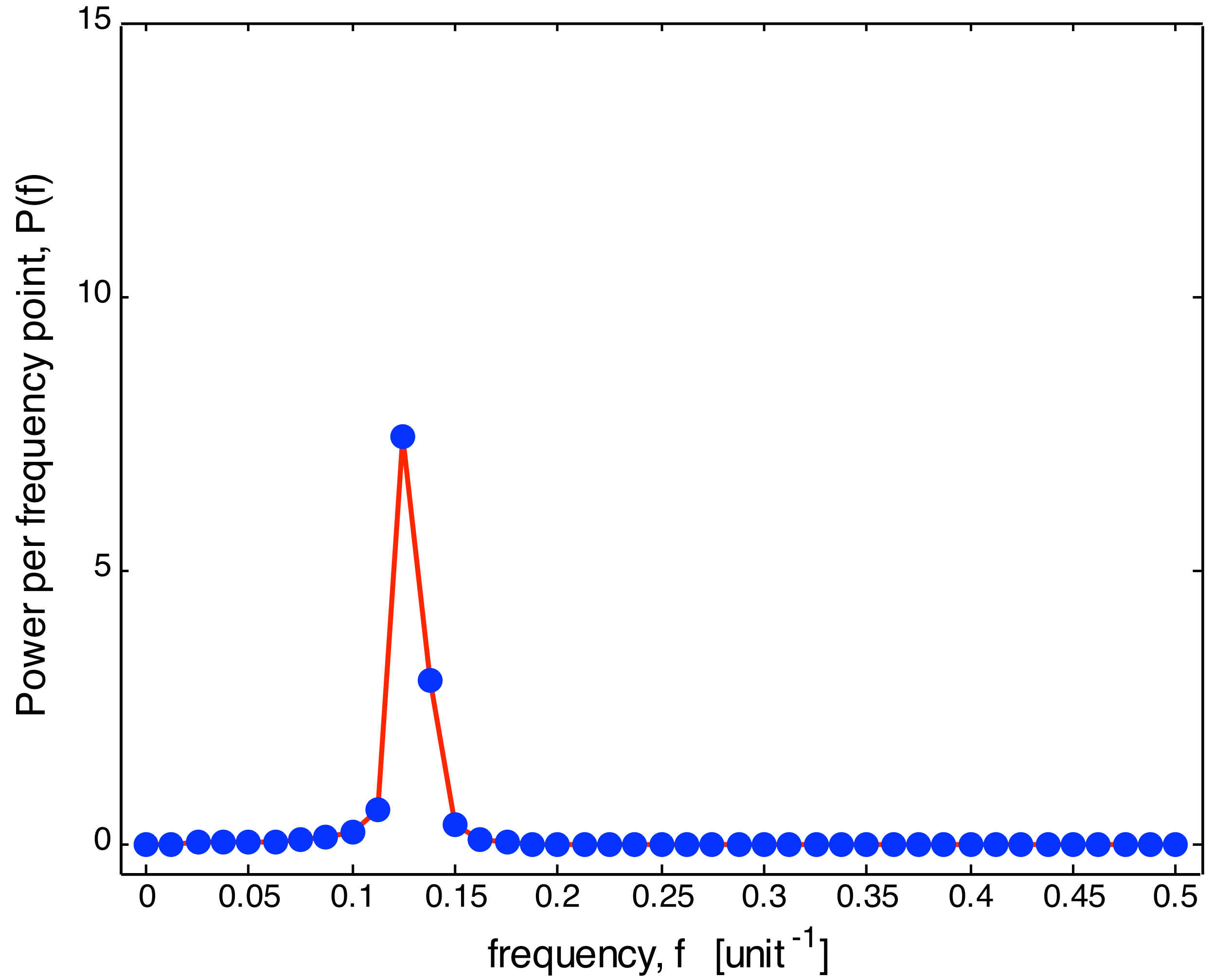
Non-integer number of periods $\Rightarrow$ frequency leakage

Example periodograms, noise-free data

$$c(t) = A\sin(2\pi f t + 0.5)$$
$$A = 5$$
$$f = 0.13$$

$$N = 80$$
$$\Delta = 1$$
$$f_c = 0.5$$





Longer time span ⇒ higher resolution ⇒ less frequency leakage

Example periodograms, noisy data

$$c(t) = A\sin(2\pi ft + 0.5) + \text{noise}$$

$$A = 0$$

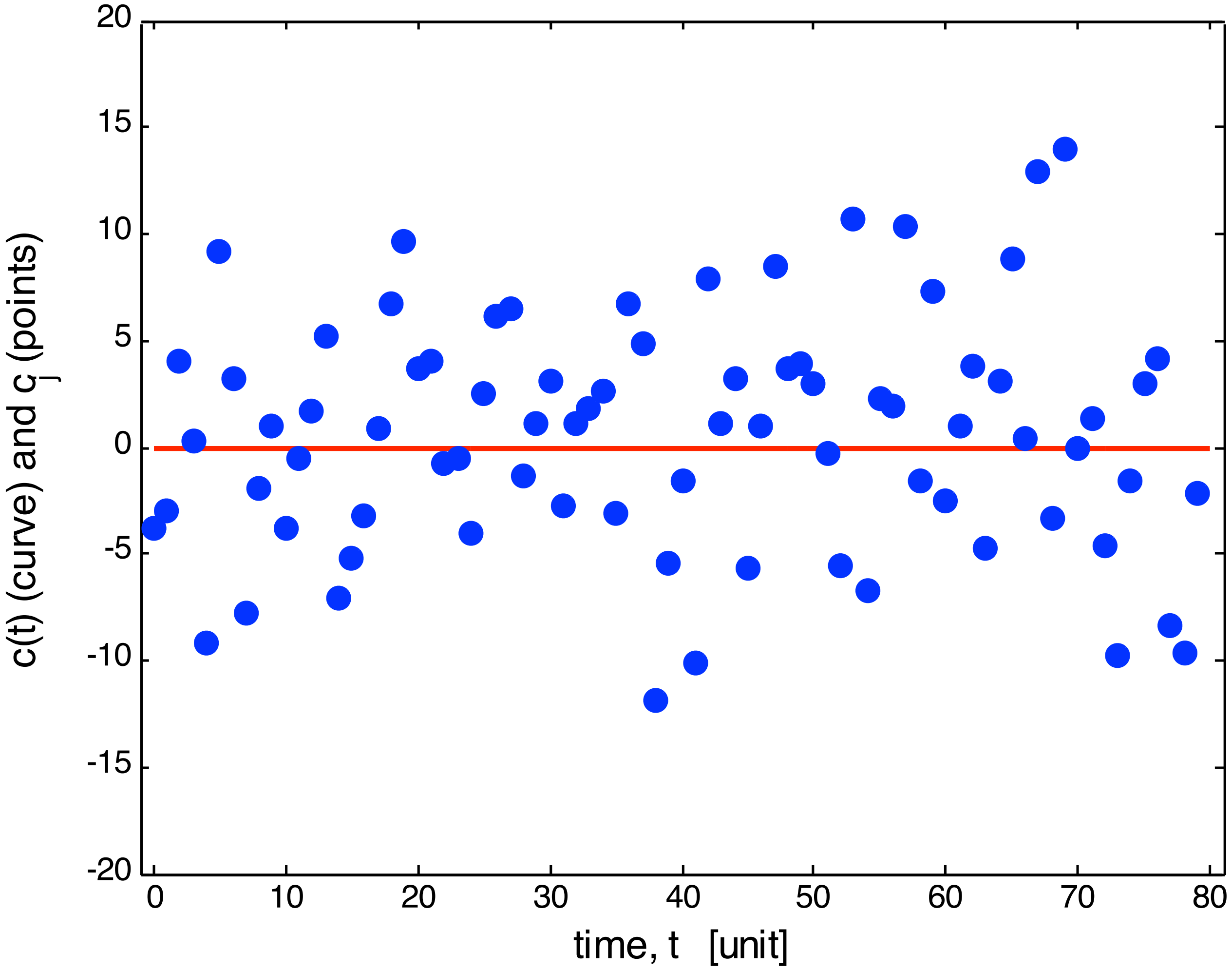
$$f =$$

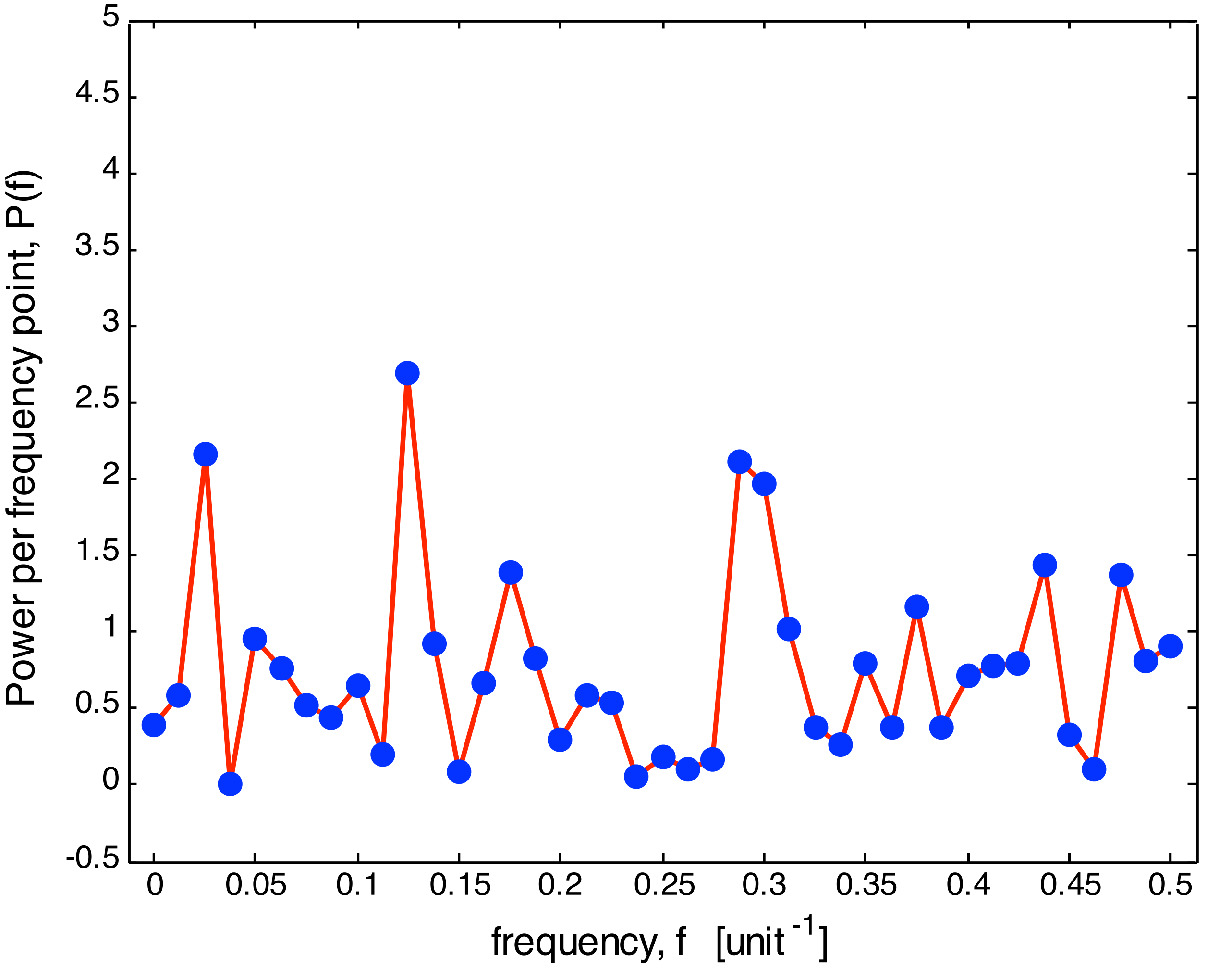
$$\sigma = 5$$

$$N = 80$$

$$\Delta = 1$$

$$f_c = 0.5$$





White noise ⇒ "constant" power (values follow exponential distribution)

Example periodograms, noisy data

$$c(t) = A\sin(2\pi ft + 0.5) + \text{noise}$$

$$A = 5$$

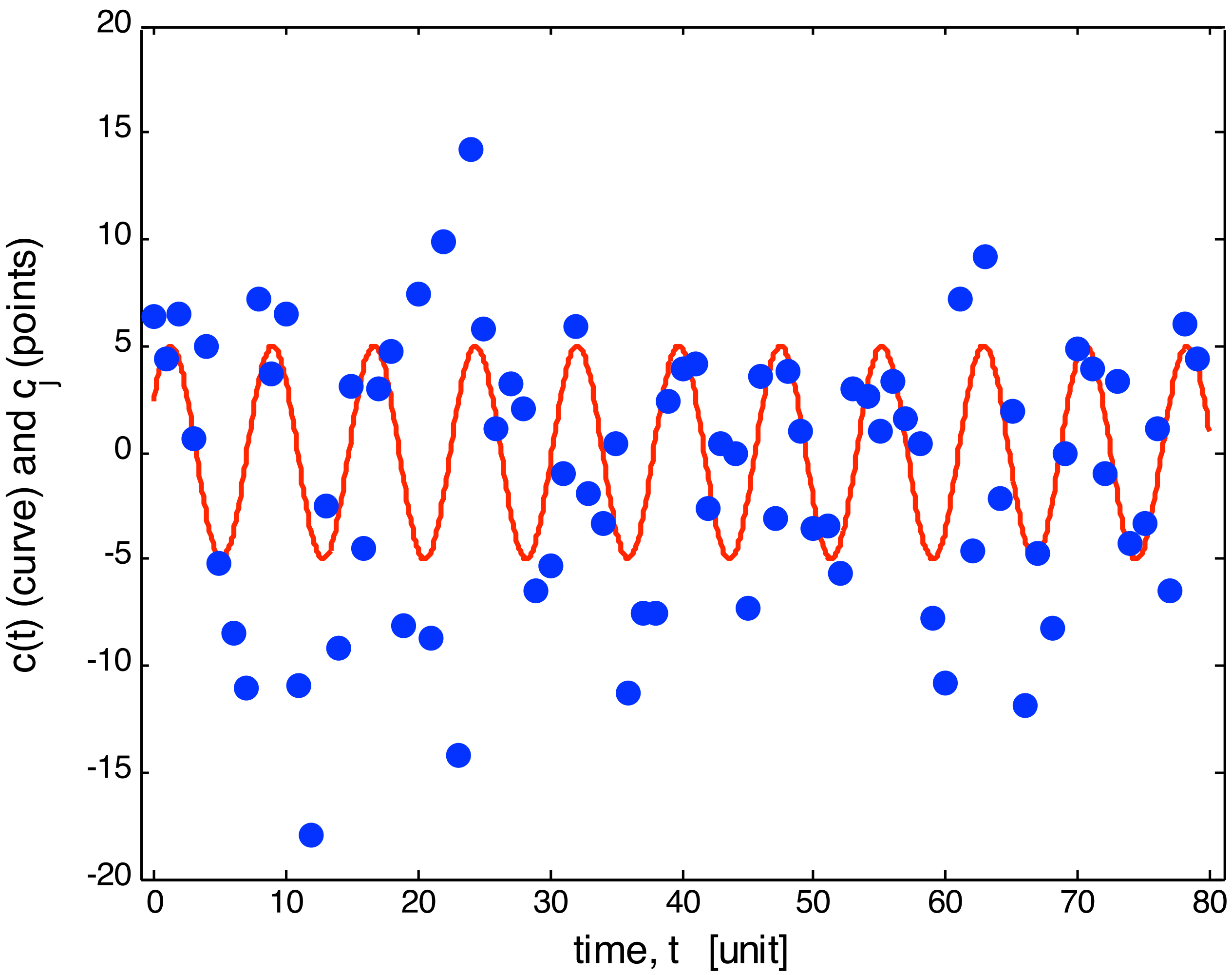
$$f = 0.13$$

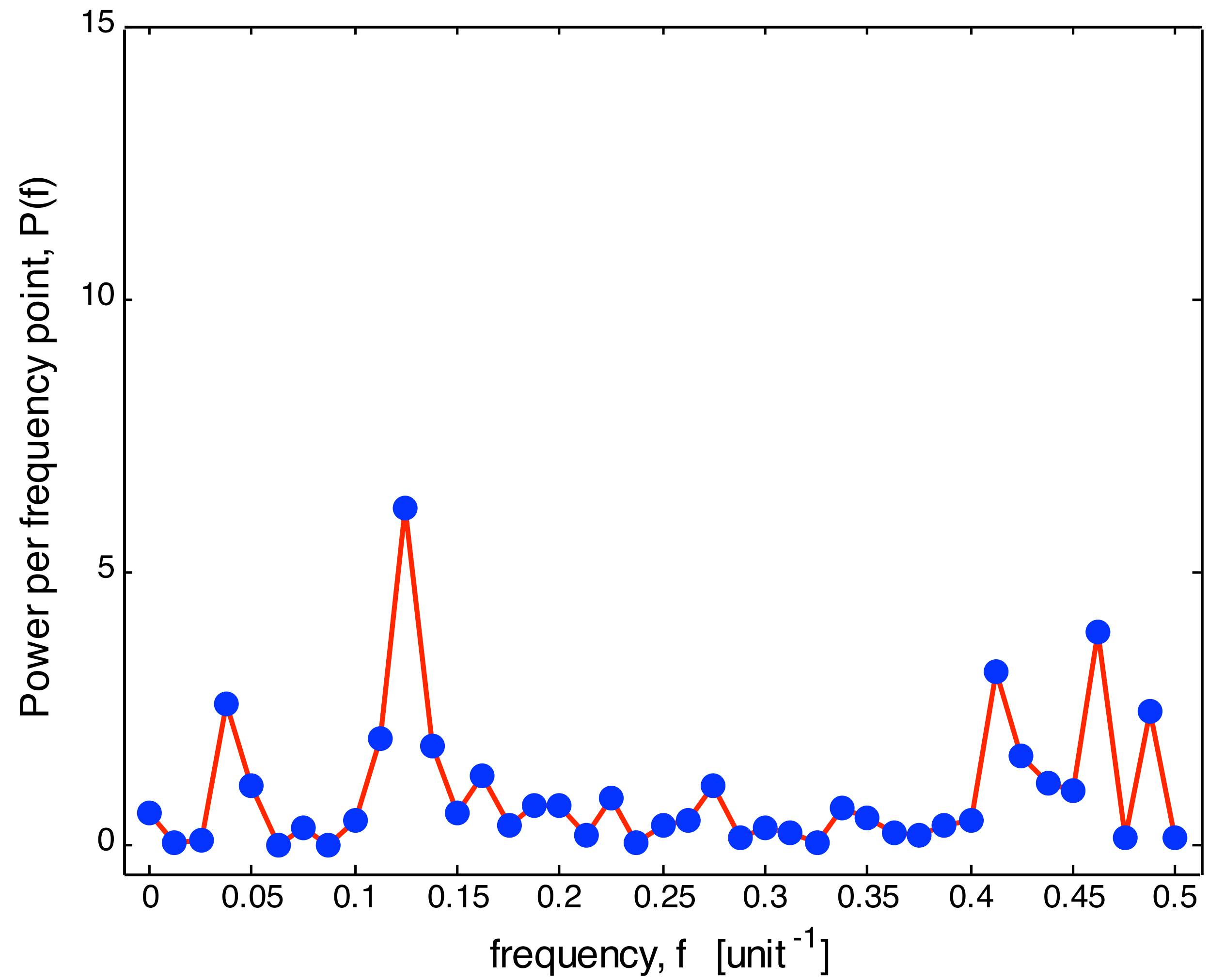
$$\sigma = 5$$

$$N = 80$$

$$\Delta = 1$$

$$f_c = 0.5$$





Periodic signal + white noise ⇒ "constant" power + peak

Example periodograms, noisy data

$$c(t) = A\sin(2\pi f t + 0.5) + \text{noise}$$

$$A = 5$$

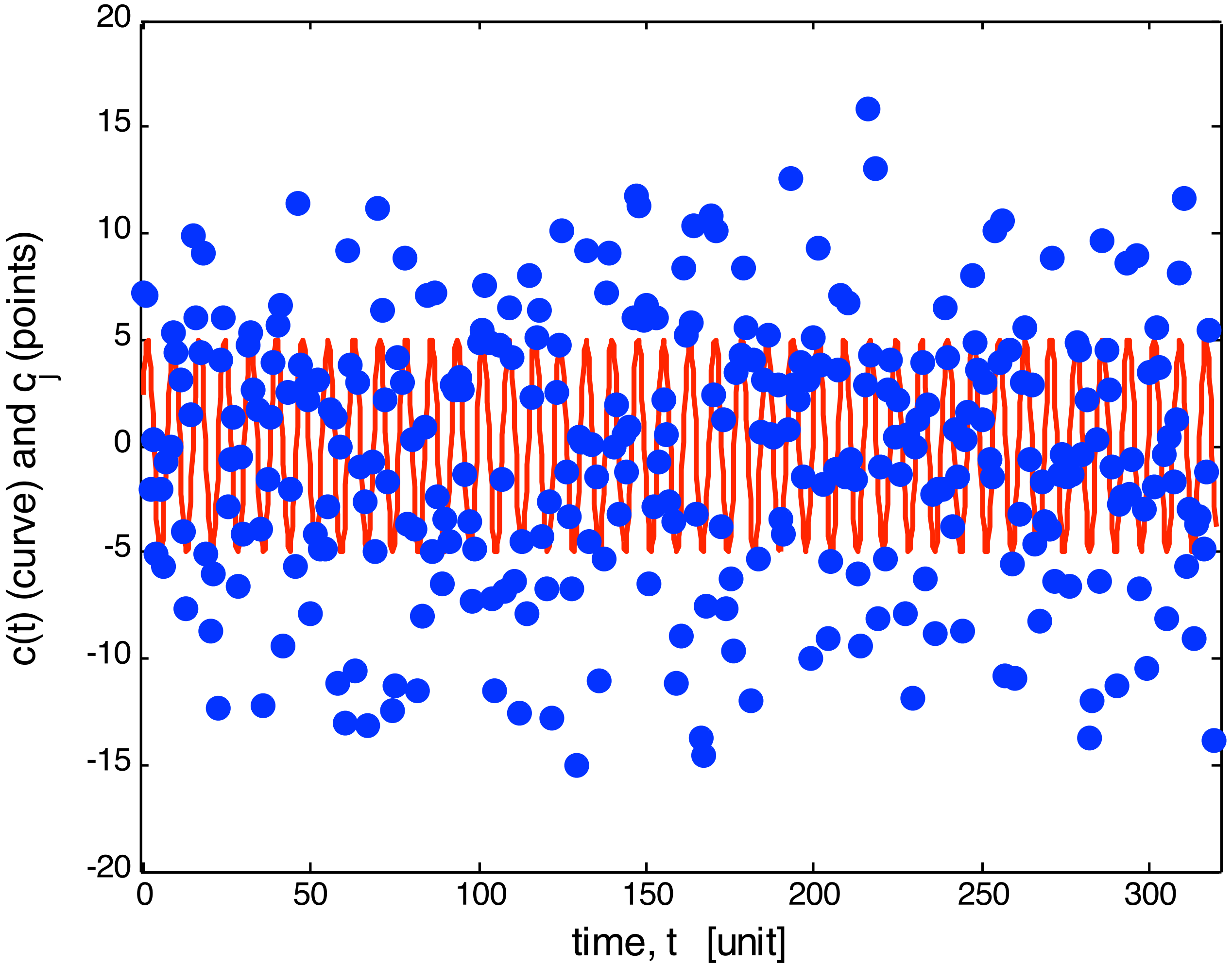
$$f = 0.13$$

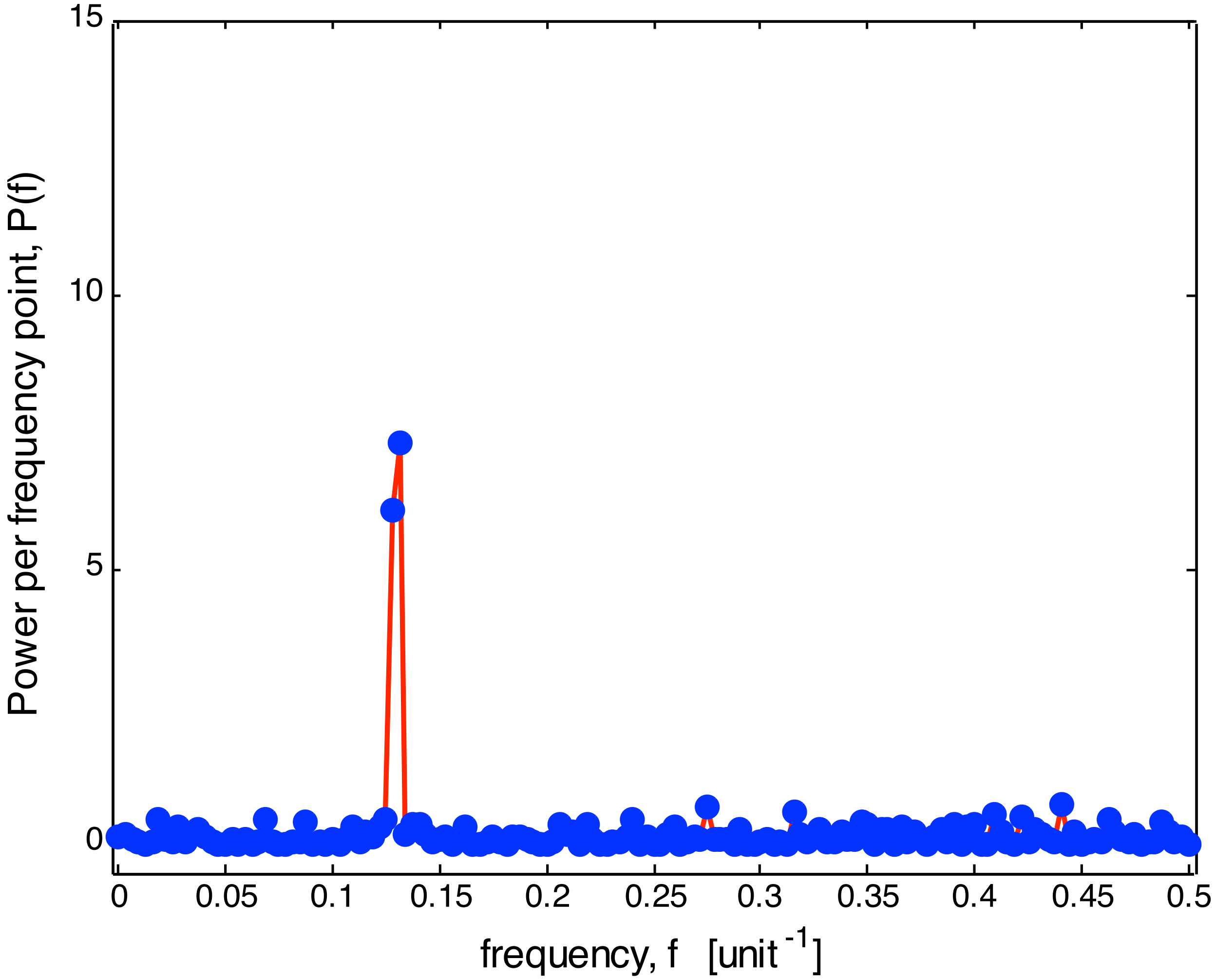
$$\sigma = 5$$

$$N = 320$$

$$\Delta = 1$$

$$f_c = 0.5$$





More data points $\Rightarrow$ less noise power per frequency point

Unevenly sampled data (1/5)

To search for a periodic signal among unevenly sampled data, we fit elementary waves of different frequencies to the data and examine the improvement of the fit as function of frequency.

Let h_j be the data values sampled at times t_j, where $j = 1,\ 2,\ \ldots N$. Assume that $\langle h_j \rangle \equiv N^{-1} \sum_j h_j = 0$ (otherwise, subtract the mean from each point). Each data point has an associated standard error σ_j.

The *null hypothesis* (H_0) is that there is no signal in the data,

$$H_0 : \qquad h_j = 0 + e_j\,, \qquad e_j \sim N(0, \sigma_j^2)$$

where $N(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2.

For trial frequency f, the alternative hypothesis (H_1) is

$$\begin{aligned} H_0 : \qquad h_j &= A\cos[\omega(t_j - \tau)] + B\sin[\omega(t_j - \tau)] + e_j \\ &= Ac_j + Bs_j + e_j\,, \qquad e_j \sim N(0, \sigma_j^2) \end{aligned}$$

Here, $\omega = 2\pi f$ and τ is some (arbitrary) time offset. For brevity we introduce $c_j = \cos[\omega(t_j - \tau)]$ and $s_j = \sin[\omega(t_j - \tau)]$.

Unevenly sampled data (2/5)

We use the chi-square (χ^2) as a measure of the goodness-of-fit for the various models. Thus, under the null hypothesis (H_0) we have

$$\chi_0^2 = \sum_j (h_j/\sigma_j)^2$$

while under H_1 the free parameters A and B are adjusted to minimize

$$\chi_1^2 = \sum_j \left(\frac{Ac_j + Bs_j - h_j}{\sigma_j} \right)^2$$

The 'power' at frequency f is taken to be half the reduction in chi-square,

$$P(f) = \frac{1}{2} \left(\chi_0^2 - \chi_{1,\min}^2 \right)$$

When calculated for a number of frequencies, this gives the so-called *Lomb–Scargle normalized periodogram* $P(f)$. If a periodic signal is present, the periodogram is expected to have a peak near the correct frequency. If there is no signal (H_0 is true), then the value $P(f)$ at any frequency follows an exponential distribution,

$$\text{Probability}(P > z) = e^{-z}$$

(this follows from the properties of the chi-square distribution).

Unevenly sampled data (3/5)

We have

$$\frac{\partial \chi_1^2}{\partial A} = \sum_j 2\left(\frac{Ac_j + Bs_j - h_j}{\sigma_j}\right)\frac{c_j}{\sigma_j}, \qquad \frac{\partial \chi_1^2}{\partial B} = \sum_j 2\left(\frac{Ac_j + Bs_j - h_j}{\sigma_j}\right)\frac{s_j}{\sigma_j}$$

so that the condition for minimum χ_1^2 is

$$\left.\begin{array}{l} \dfrac{\partial \chi_1^2}{\partial A} = 0 \\[2ex] \dfrac{\partial \chi_1^2}{\partial B} = 0 \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{l} A\displaystyle\sum_j \frac{c_j^2}{\sigma_j^2} + B\sum_j \frac{s_j c_j}{\sigma_j^2} = \sum_j \frac{h_j c_j}{\sigma_j^2} \\[3ex] A\displaystyle\sum_j \frac{c_j s_j}{\sigma_j^2} + B\sum_j \frac{s_j^2}{\sigma_j^2} = \sum_j \frac{h_j s_j}{\sigma_j^2} \end{array}\right.$$

This is a linear system of equations for the unknowns A, B.

Unevenly sampled data (4/5)

The solution is simplified if τ is chosen so that the off-diagonal terms of the equations matrix vanish:

$$\begin{aligned}
\sum_j \frac{s_j c_j}{\sigma_j^2} &= \sum_j \sigma_j^{-2} \sin(\omega t_j - \omega\tau)\cos(\omega t_j - \omega\tau) \\
&= \frac{1}{2}\sum_j \sigma_j^{-2} \sin 2(\omega t_j - \omega\tau) \\
&= \frac{1}{2}\left(\sum_j \sigma_j^{-2}\sin(2\omega t_j)\cos(2\omega\tau) - \sum_j \sigma_j^{-2}\cos(2\omega t_j)\sin(2\omega\tau)\right) \\
&= 0
\end{aligned}$$

$$\Longrightarrow \qquad \tan(2\omega\tau) = \frac{\sum_j \sigma_j^{-2}\sin(2\omega t_j)}{\sum_j \sigma_j^{-2}\cos(2\omega t_j)}$$

Unevenly sampled data (5/5)

With this τ, the solution is

$$A = \frac{\sum_j h_j c_j/\sigma_j^2}{\sum_j c_j^2/\sigma_j^2}\,, \qquad B = \frac{\sum_j h_j s_j/\sigma_j^2}{\sum_j s_j^2/\sigma_j^2}$$

After some algebra, we find

$$\chi^2_{1,\min} = \sum_j h_j^2/\sigma_j^2 - \frac{\left(\sum_j h_j c_j/\sigma_j^2\right)^2}{\sum_j c_j^2/\sigma_j^2} - \frac{\left(\sum_j h_j s_j/\sigma_j^2\right)^2}{\sum_j s_j^2/\sigma_j^2}$$

The periodogram is therefore

$$P(f) = \frac{1}{2}\left[\frac{\left(\sum_j h_j c_j/\sigma_j^2\right)^2}{\sum_j c_j^2/\sigma_j^2} + \frac{\left(\sum_j h_j s_j/\sigma_j^2\right)^2}{\sum_j s_j^2/\sigma_j^2}\right]$$

For homoscedastic data (all $\sigma_j = \sigma$) the factor σ^{-2} can be taken out of the sums, which gives the expression in NR (13.8.4).

Example periodograms, noisy data, unevenly sampled

$$c(t) = A\sin(2\pi ft + 0.5) + \text{noise}$$

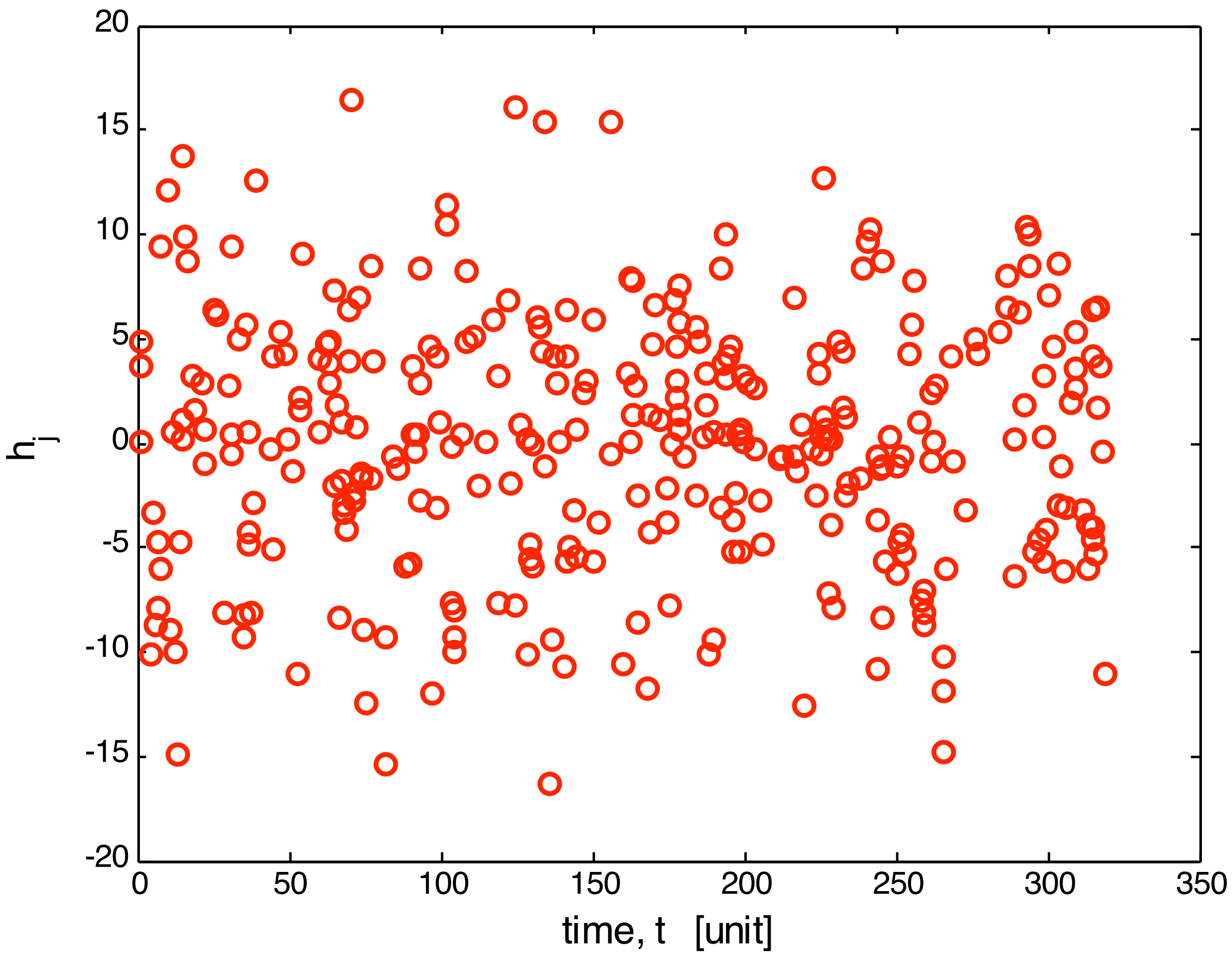
$$A = 5$$

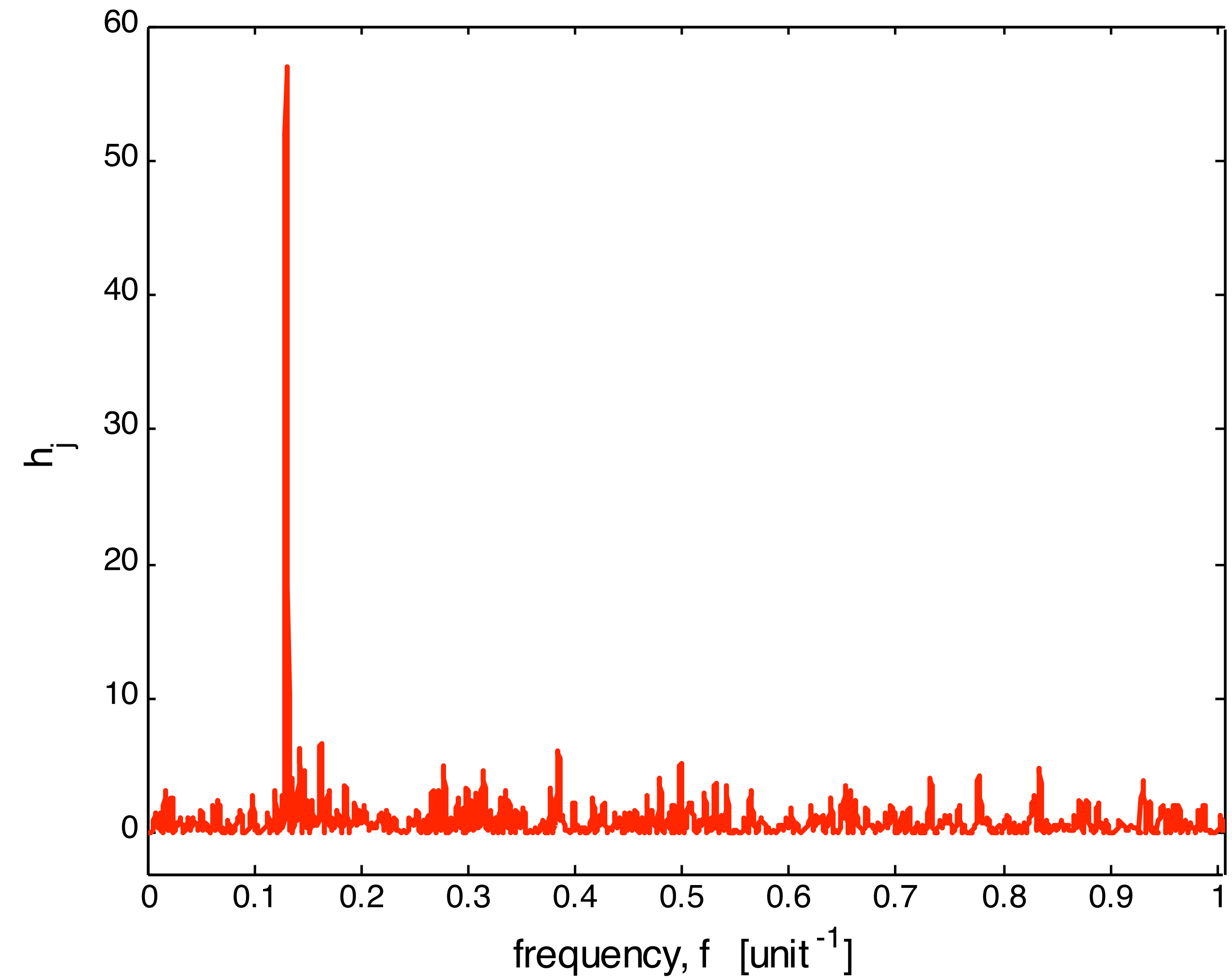
$$f = 0.13$$

$$\sigma = 5$$

$$\langle \Delta \rangle = 1$$

```
n = 320;
mjd = rand([n 1])*320;
```

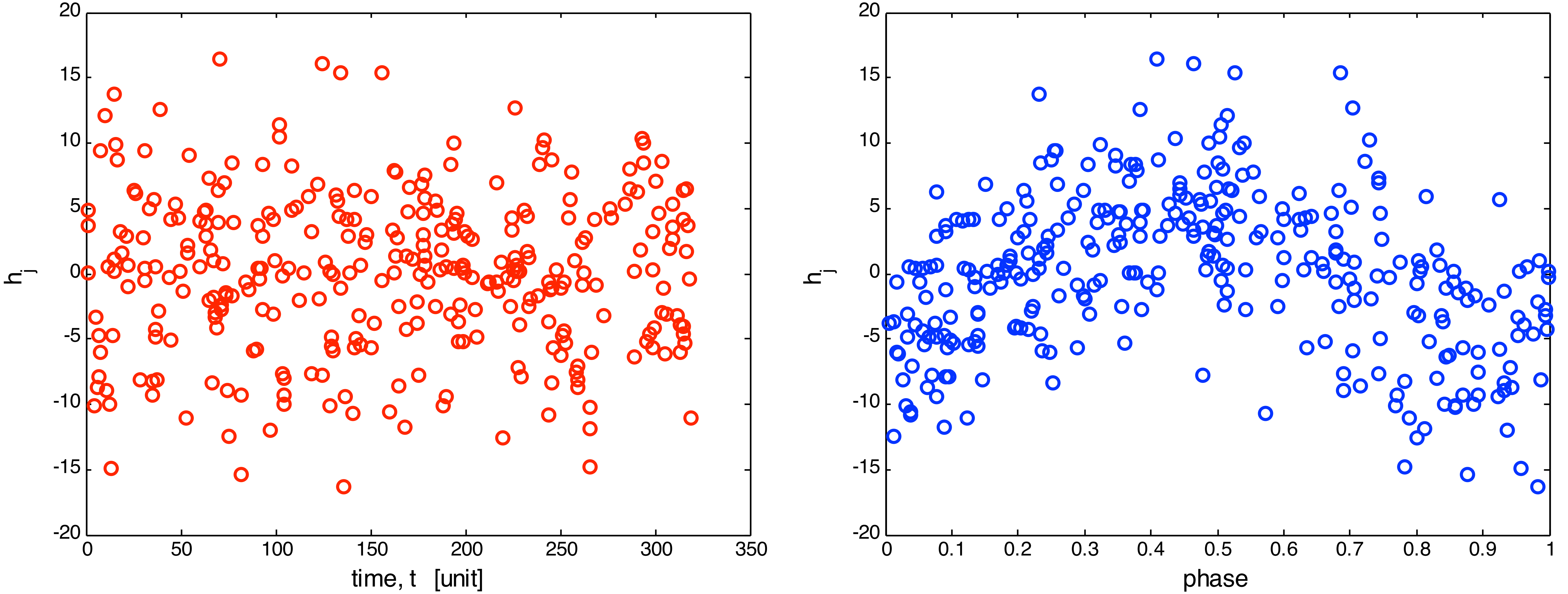




Random sampling ⇒ good frequency coverage (no aliasing!)

Checking found period (p) by folding the data

Phase (ϕ) of sample j is $$\phi_j = \frac{\mathrm{mod}(t_j, p)}{p}$$



Is scatter about the mean curve consistent with errors?

Example periodograms, noisy data, unevenly sampled

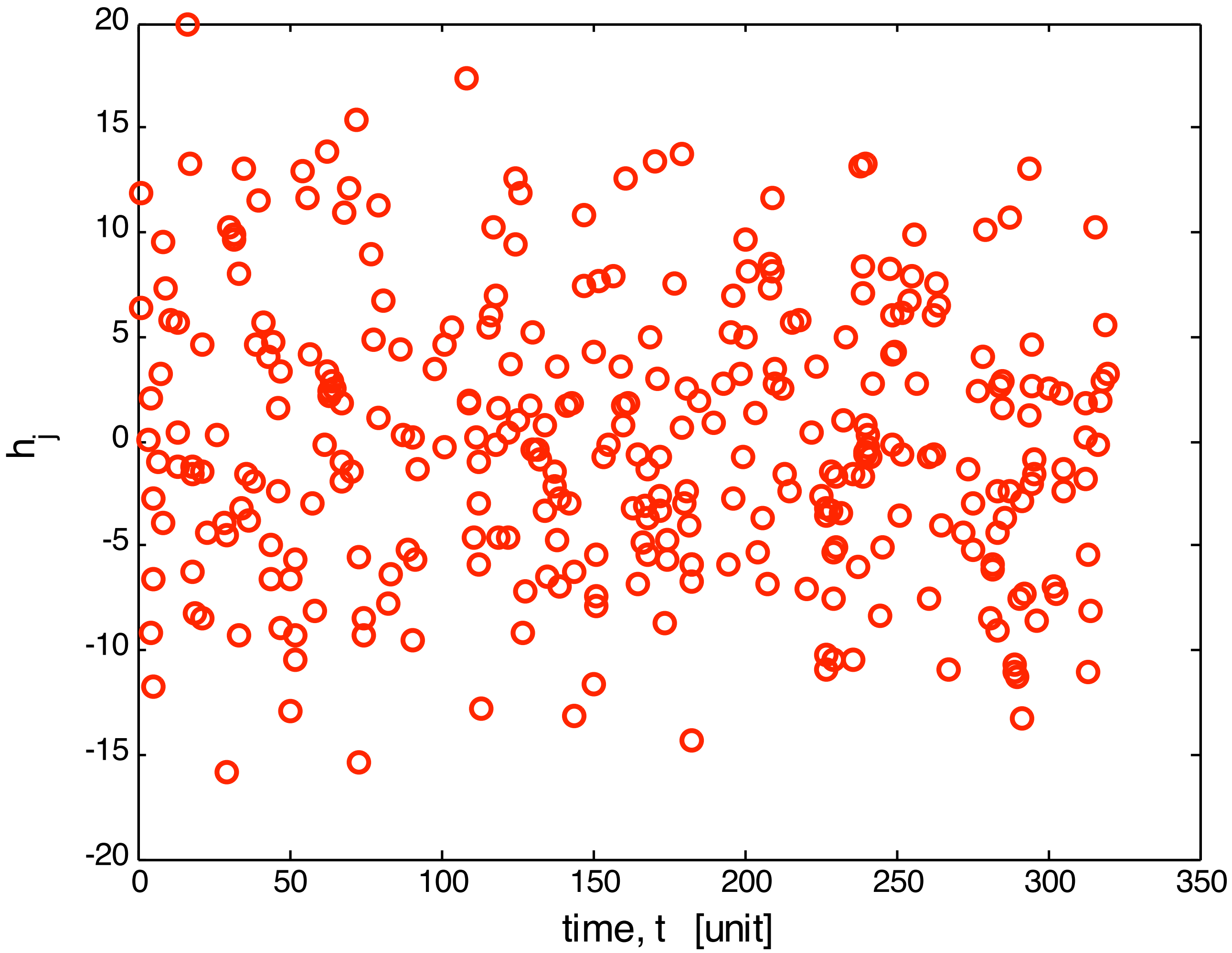
$$c(t) = A\sin(2\pi ft + 0.5) + \text{noise}$$

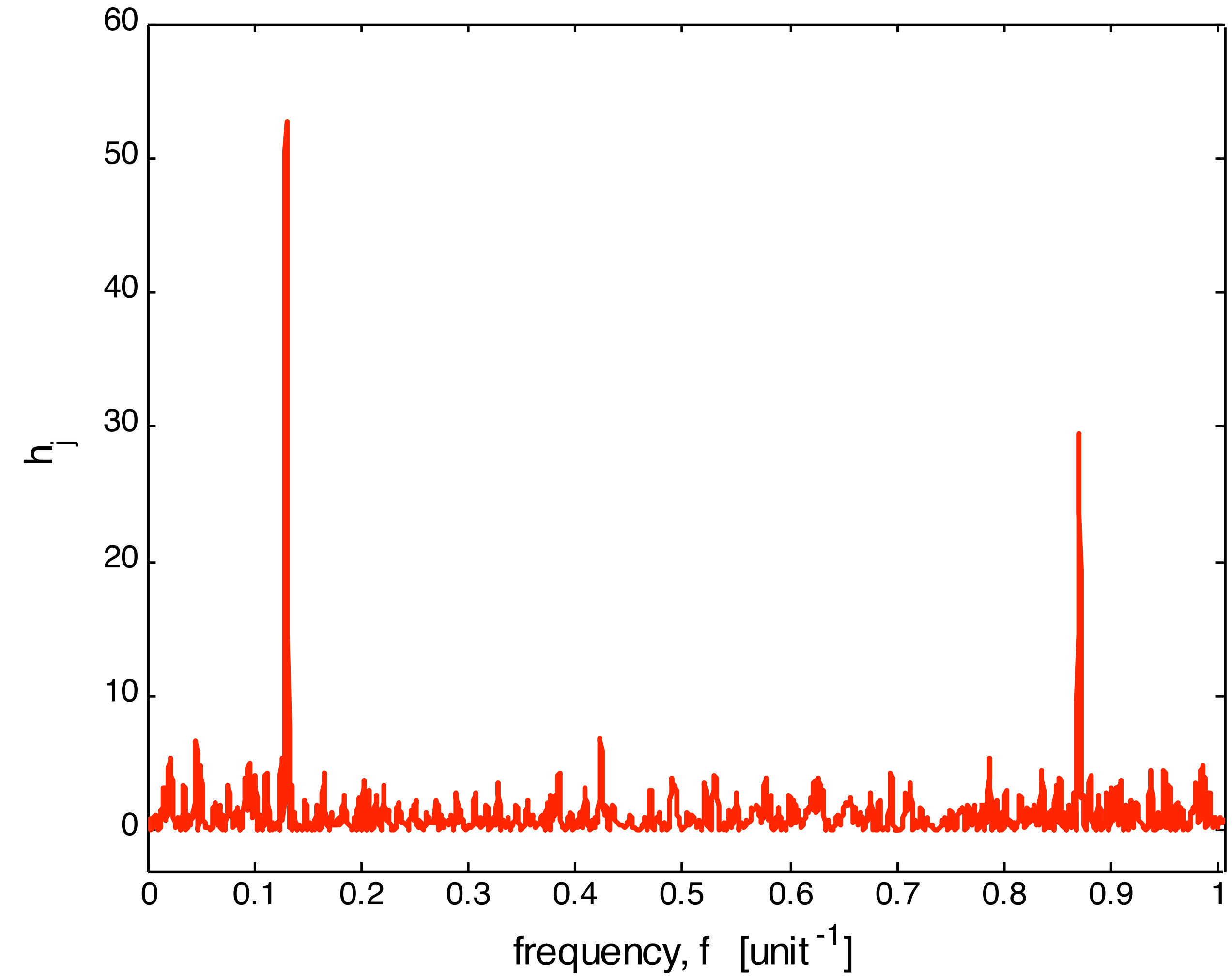
$$A = 5$$

$$f = 0.13$$

$$\sigma = 5$$

```
n = 320;
mjd = rand([n 1])*320;
for j = 1:n
    if (mod(mjd(j),1)>0.5)
        mjd(j) = mjd(j) - 0.5;
    end
end
```





Non-random sampling ⇒ look out for frequency window effects!