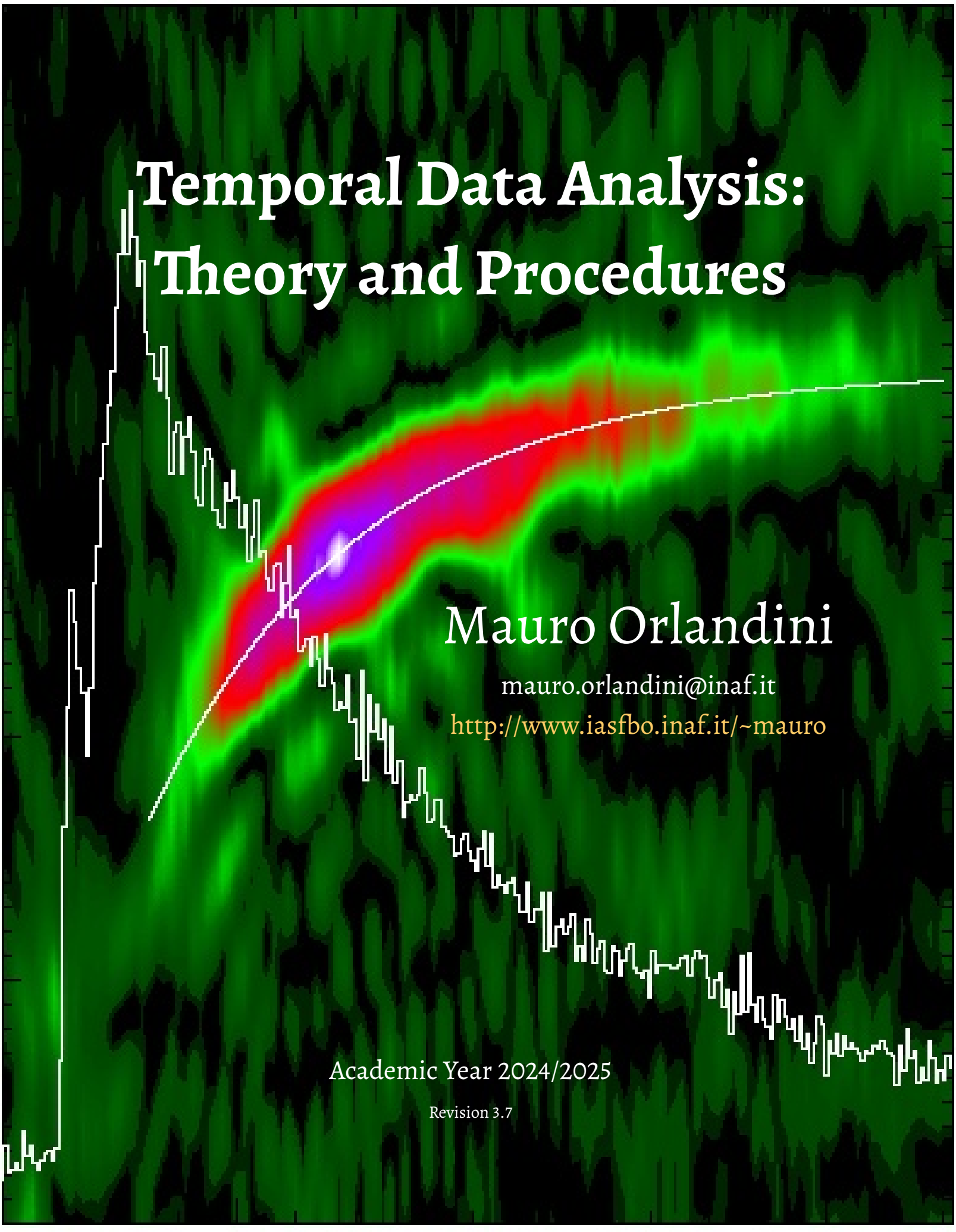
Temporal Data Analysis: Theory and Procedures



Mauro Orlandini

mauro.orlandini@inaf.it

http://www.iasfbo.inaf.it/~mauro

Academic Year 2024/2025

Revision 3.7

Please note that this document contains (a lot of!) copy-righted material that I obtained *without* the necessary authorizations. This is because I intend these notes for **INTERNAL USE ONLY!**
So, please, distribute it with care.

The material presented in Appendix B comes from a course on timing analysis given by Biswajit Paul during the first ASTROSAT Workshop. I really thank Biswajit for letting me use his material and passing me the data files.

Document information:

Author:	Mauro Orlandini
Revision:	3.7
Last Changed on:	2025/02/28 10:22
LaTeX-ed on:	2025/02/28 11:25

Course Outline

List of Figures

List of Definitions

List of Theorems

Part I

Theory

Chapter 1

Classification of Physical Data

Any observed data representing a physical phenomenon can be broadly classified as being either *deterministic* or *nondeterministic*. Deterministic data are those that can be described by an explicit mathematical relationship. There are many physical phenomena in practice which produce data that can be represented with reasonable accuracy by explicit mathematical relationships. For example, the motion of a satellite in orbit about the Earth, the potential across a condenser as it discharges through the resistor, the vibration response of an unbalanced rotating machine, or the temperature of water as heat is applied, are all basically deterministic. However, there are many other physical phenomena which produce data that are not deterministic. For example, the height of waves in a confused sea, the acoustic pressure generated by air rushing through a pipe, or the electrical output of a noise generator represent data which cannot be described by explicit mathematical relationships. There is no way to predict an exact value at a future instant of time. These data are random in character and must be described in terms of probability statements and statistical averages rather than explicit equations.
Various special classifications of deterministic and random data will now be discussed.

1.1 Deterministic Data

Data representing deterministic phenomena can be categorized as being either *periodic* or *non periodic*. Periodic data can be further categorized as being either *sinusoidal* or *complex* periodic. Non periodic data can be further categorized as being either *almost-periodic* or *transient*. These various classifications of deterministic data are schematically illustrated in Figure 1.1. Of course, any combination of these forms may also occur. For purposes of review, each of these types of deterministic data, along with physical examples, will be briefly discussed.

1.1.1 Sinusoidal Periodic Data

Sinusoidal data are those types of periodic data which can be defined mathematically by a time-varying function of the form

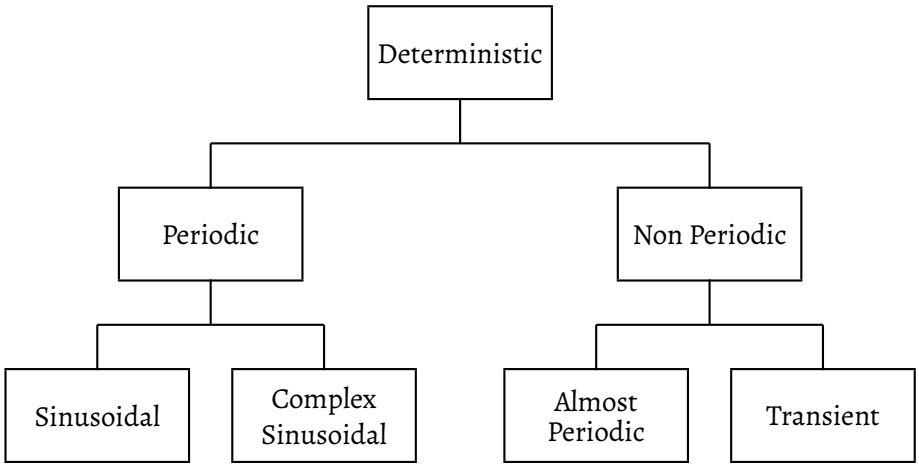


Figure 1.1: Classification of deterministic data

$$x(t) = X \sin(\omega_0 t + \varphi) \tag{1.1}$$

where X is the amplitude, ω_0 is the angular frequency, in units of radians per unit time[1], φ is the initial phase angle (in radians) with respect to the time origin, and $x(t)$ is the instantaneous value at time t. The sinusoidal time history described by (1.1) is usually referred as a sine wave. When analyzing sinusoidal data in practice, the phase angle φ is often ignored.
The time interval required for one full fluctuation or cycle of sinusoidal data is called the period T. The number of cycles per unit time is called the frequency ν.
There are many example of physical phenomena which produce approximately sinusoidal data in practice. The voltage output of an electrical alternator is one example; the vibratory motion of an unbalanced rotating weight is another. Sinusoidal data represent one of the simplest forms of time-varying data from the analysis viewpoint.

1.1.2 Complex Periodic Data

Complex periodic data are those type of periodic data which can be defined mathematically by a time-varying function whose waveform exactly repeats itself at regular intervals such that

$$x(t) = x(t \pm nT) \qquad n = 1, 2, 3, \ldots \tag{1.2}$$

As for sinusoidal data, the time interval required for one full fluctuation is called the *period* T. The angular frequency is called the *fundamental frequency* ω. With few exceptions in practice, complex periodic data may be expanded into a Fourier series according to the following formula (we will return later in greater detail on that)

[1] Not to be confused with the frequency ν, measured in Hz. The two are related by $\omega = 2\pi\nu$.

$$x(t) = \sum_{k=0}^{\infty} (A_k \cos \omega_k t + B_k \sin \omega_k t) \tag{1.3}$$

with $\omega_k = 2\pi k/T$ and $B_0 = 0$. An alternative way to express the Fourier series is

$$x(t) = X_0 + \sum_{k=1}^{\infty} X_k \cos(\omega_k t + \varphi_k) \tag{1.4}$$

In other words, (1.4) says that complex periodic data consists of a static component and an infinite number of sinusoidal components called *harmonics*, which have amplitudes X_k and phases φ_k. The frequencies of the harmonic components are all integral multiples of ω_1.
Physical phenomena which produce complex periodic data are far more common than those which produce simple sinusoidal data. In fact, the classification of data as being sinusoidal is often only an approximation for data which are actually complex. For example, the voltage output from an electrical alternator may actually display, under careful inspection, some small contributions at higher harmonic frequencies. In other cases, intense harmonic components may be present in periodic physical data.

1.1.3 Almost-Periodic Data

We have seen that periodic data can be generally reduced to a series of sine waves with commensurately related frequencies. Conversely, the data formed by summing two or more commensurately related sine waves will be periodic. However, the data formed by summing two or more sine waves with arbitrary frequencies will not be periodic.
More specifically, the sum of two or more sine waves will be periodic only when the ratios of all possible pairs of frequencies form rational numbers. This indicates that a fundamental period exists which will satisfy the requirements of (1.2). Hence,

$$x(t) = X_1 \sin(2t + \varphi_1) + X_2 \sin(3t + \varphi_2) + X_3 \sin(7t + \varphi_3)$$

is periodic since $2/3$, $2/7$ and $3/7$ are rational numbers (the fundamental period is $T = 1$). On the other hand,

$$x(t) = X_1 \sin(2t + \varphi_1) + X_2 \sin(3t + \varphi_2) + X_3 \sin(\sqrt{50}t + \varphi_3)$$

is not periodic since $2/\sqrt{50}$ and $3/\sqrt{50}$ are not rational numbers (in this case the fundamental period is infinitely long). The resulting time history in this case will have an *"almost periodic"* character, but the requirement of (1.2) will not be satisfied for any finite value of T.
Based on these discussions, almost periodic data are those types of non periodic data which can be defined mathematically by a time-varying function of the form

$$x(t) = \sum_{k=1}^{\infty} X_k \sin(\omega_k t + \varphi_k) \tag{1.5}$$

with $\omega_j/\omega_k \neq$ rational numbers in all cases. Physical phenomena producing almost periodic data frequently occur in practice when the effects of two or more unrelated periodic phenomena are mixed. A good example is the vibration response in a multiple engine propeller airplane when the engines are out of synchronization.

1.1.4 Transient Non periodic Data

Transient data are defined as all non periodic data other the almost-periodic discussed above. In other words, transient data include all data not previously discusses which can be described by some suitable time-varying function.
Physical phenomena which produce transient data are numerous and diverse. For example, the behavior of the temperature of water in a kettle (relative to room temperature) after the flame is turned off.

1.2 Random Data

Data representing a random physical phenomenon cannot be described by an explicit mathematical relationship because each observation of the phenomenon will be unique. In other words, any given observation will represent only one of the many possible results which might have occurred. For example, assume the output voltage from a thermal noise generator is recordered as a function of time. A specific voltage time history record will be obtained. However, if a second thermal noise generator of identical construction and assembly is operated simultaneously, a different voltage time history record would result. In fact, every thermal noise generator which might be constructed would produce a different voltage time history record. Hence the voltage time history for any one generator is merely one example of an infinitely large number of time histories which might be occurred.
A single time history representing a random phenomenon is called a *sample function* (or a *sample record* when observed over a finite time interval). The collection of all possible sample functions which the random phenomenon might have produced is called a *random process* or a *stochastic process*. Hence a sample record of data for a random physical phenomenon may be though of as one physical realization of a random process.
Random processes might be categorized as being either *stationary* or *non stationary*. Stationary random processes may be further categorized as being either *ergodic* or *non ergodic*. Non stationary random processes may be further categorized in terms of specific types of non stationary properties. These various classifications of random processes are schematically illustrated in Figure 1.2. The meaning and physical significance of these various types of random processes will now be discussed in broad terms.

1.2.1 Stationary Random Processes

When a physical phenomenon is considered in terms of a random process, the properties of the phenomenon can hypothetically be described at any instant of time by computing average values

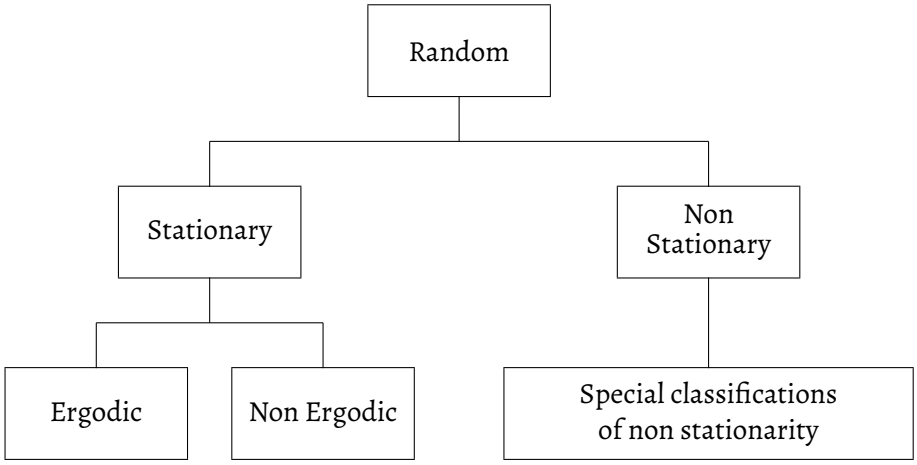


Figure 1.2: Classifications of random data

over the collection of sample functions which describe the random process. For example, consider the collection of sample functions (also called the *ensemble*) which form the random process illustrated in Figure 1.3. The *mean value* (first moment) of the random process at some time t_1 can be computed by taking the instantaneous value of each sample function of the ensemble at time t_1, summing the values, and dividing by the number of sample functions. In a similar manner, a correlation (joint moment) between the values of the random process at two different times (called *autocorrelation function*) can be computed by taking the ensemble average of the product of instantaneous values at two times, t_1 and $t_1 + \tau$. That is, for the random process $\{x(t)\}$, where the symbol $\{\ \}$ is used to denote an ensemble of sample functions, the mean value $\mu_x(t_1)$ and the autocorrelation function $R_x(t_1, t_1 + \tau)$ are given by

$$\mu_x(t_1) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1) \tag{1.6a}$$

$$R_x(t_1, t_1 + \tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1)\, x_k(t_1 + \tau) \tag{1.6b}$$

where the final summation assumes each sample function is equally likely.
For the general case where $\mu_x(t_1)$ and $R_x(t_1, t_1+\tau)$ defined in (1.6) vary as time t_1 varies, the random process $\{x(t)\}$ is said to be *non stationary*. For the special case where $\mu_x(t_1)$ and $R_x(t_1, t_1+\tau)$ do not vary as time t_1 varies, the random process $\{x(t)\}$ is said to be *weakly stationary* or stationary in the wide sense. For the weakly stationary processes, the mean value is a constant and the autocorrelation function is dependent only upon the time of displacement τ. That is, $\mu_x(t_1) = \mu_x$ and $R_x(t_1, t_1 + \tau) = R_x(\tau)$.
An infinite collection of higher order moments and joint moments of the random process $\{x(t)\}$ could also be computed to establish a complete family of probability distribution functions describing the process. For the special case where all possible moments and joint moments are

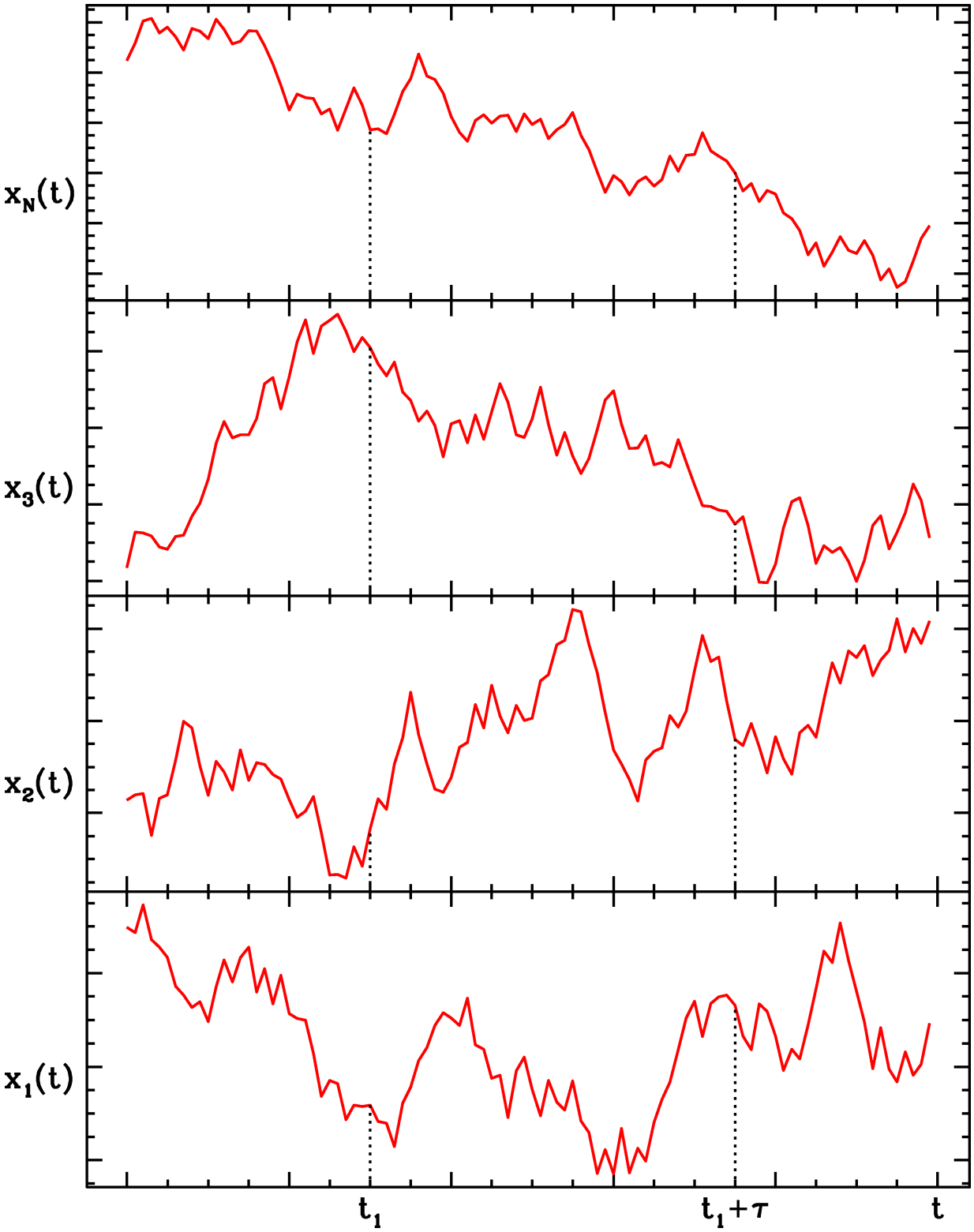


Figure 1.3: Ensemble of sample functions forming a random process

time invariant, the random process $\{x(t)\}$ is said to be *strongly stationary* or stationary in the strict sense. For many practical applications, verification of weak stationarity will justify an assumption of strong stationarity.

1.2.2 Ergodic Random Processes

The previous section discusses how the properties of a random process can be determined by computing ensamble averages at specific instants of time. In most cases, however, it is also possible to describe the properties of a stationary random process by computing time averages over specific sample functions in the ensemble. For example, consider the k-th sample function of the random process illustrated in Figure 1.3. The mean value $\mu_x(k)$ and the autocorrelation function $R_x(\tau, k)$ of the k-th sample function are given by

$$\mu_x(k) = \lim_{N\to\infty} \frac{1}{N} \int_0^T x_k(t)dt \tag{1.7a}$$

$$R_x(\tau, k) = \lim_{N\to\infty} \frac{1}{N} \int_0^T x_k(t)x_k(t+\tau)dt \tag{1.7b}$$

If the random process $\{x(t)\}$ is stationary, and $\mu_x(k)$ and $R_x(\tau, k)$ defined in (1.7) do not differ when computed over different sample functions, the random process is said to be *ergodic*. For ergodic random processes, the time averaged mean value and autocorrelation function (as well as all other time-averaged properties) are equal to the corresponding ensemble averaged value. That is, $\mu_x(k) = \mu_x$ and $R_x(\tau, k) = R_x(\tau)$. Note that only stationary random process can be ergodic.
Ergodic random processes are clearly an important class of random processes since all processes of ergodic random processes can be determined by performing time averages over a single sample function. Fortunately, in practice, random data representing stationary physical phenomena are generally ergodic. It is for this reason that the properties of stationary random phenomena can be measured properly, in most cases, from a single observed time history record.

1.2.3 Non stationary Random Processes

Non stationary random processes include all random processes which do not meet the requirements for stationarity defined in the previous section. Unless further restrictions are imposed, the properties of non stationary random processes are generally time-varying functions which can be determined only by performing instantaneous averages over the ensemble of sample functions forming the process. In practice, it is often not feasible to obtain a sufficient number of sample records to permit the accurate measurement of properties by ensemble averaging. This fact has tended to impede the development of practical techniques for measuring and analyzing non stationary random data.
In many cases, the non stationary random data produced by actual physical phenomena can be classified into special categories of non stationarity which simplify the measurement and analysis problem. For example, some type of random data might be described by a non stationary

random process $\{y(t)\}$ where each sample function is given by $y(t) = A(t)\,x(t)$. Here $x(t)$ is a sample function from a stationary random process $\{x(t)\}$ and $A(t)$ is a deterministic multiplication factor. In other words, the data might be represented by a non stationary random process consisting of a sample functions with a common deterministic time trend. If non stationary random data fit a specific model of this type, ensemble averaging is not always needed to describe the data. The various desired properties can sometimes be estimated from a single record, as is true for ergodic stationary data.

Chapter 2

Harmonic Analysis

Few preliminary remarks are in order: First, we will use the angular frequency ω when we refer to the frequency domain. The unit of the angular frequency is radians/second (or simpler s^{-1}). It is easily converted to the frequency ν (unit in Hz) using the following equation:

$$\omega = 2\pi\nu$$

Second: just let us remember the definition of even and odd functions.

Definition 2.1 (Even and odd functions). *A function is said to be even if*

$$f(-t) = f(t) \qquad \textit{even function}$$

while a function is said to be odd if

$$f(-t) = -f(t) \qquad \textit{odd function}$$

Any function can be described in terms of a mixture of even and odd functions, by means of the following decomposition (see Figure 2.1):

$$\begin{aligned} f_{\text{even}} &= \frac{f(t) + f(-t)}{2} \\ f_{\text{odd}} &= \frac{f(t) - f(-t)}{2} \end{aligned}$$

2.1 Fourier Series

This Section will deal with the mapping of *periodic* functions to a series based on the trigonometric functions sine (and odd function) and cosine (even function).

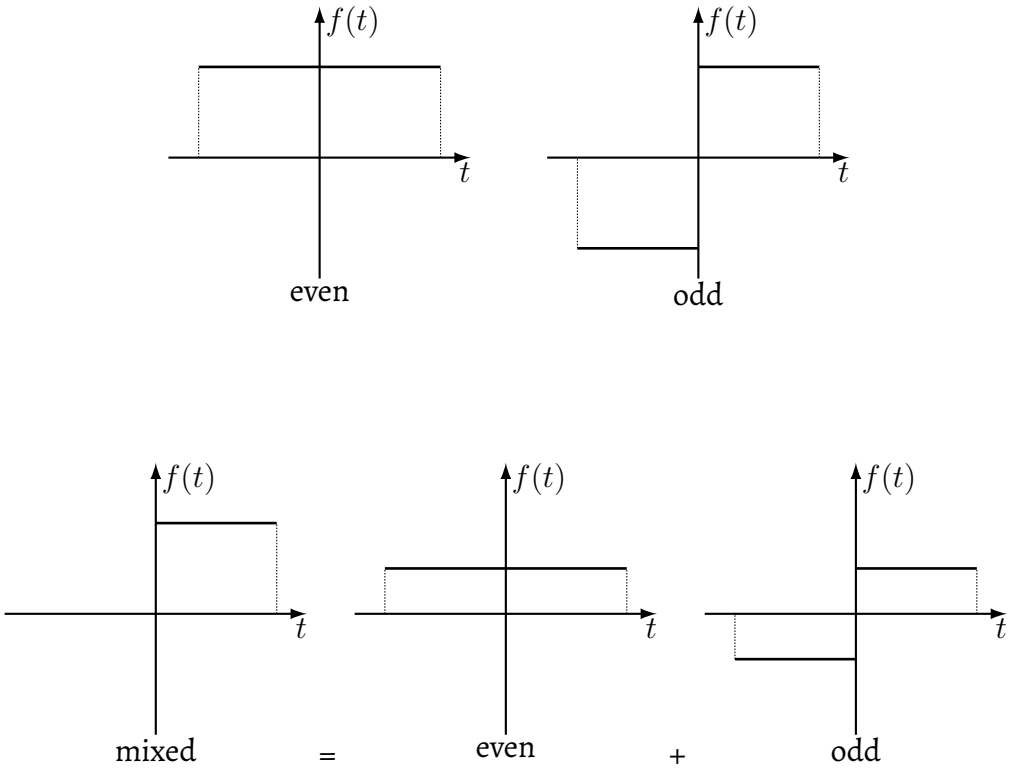


Figure 2.1: Examples of even, odd and mixed functions

2.1.1 Definition

Any periodic function $f(t)$ can be expanded into a series of trigonometric function, the so called Fourier series, as follows

Definition 2.2 (Fourier series).

$$f(t) = \sum_{k=0}^{\infty} (A_k \cos \omega_k t + B_k \sin \omega_k t) \qquad \textit{with } \omega_k = \frac{2\pi k}{T},\ B_0 = 0 \tag{2.1}$$

T is the period of the function $f(t)$. The amplitudes or *Fourier coefficients* A_k and B_k are determined in such a way that the infinite series is identical with the function $f(t)$. Equation (2.1) therefore tells us that any periodic function can be represented as a superposition of sine-function and cosine-function with appropriate amplitudes – with an infinite number of terms, if need be – yet using only precisely determined frequencies:

$$\omega = 0, \frac{2\pi}{T}, \frac{4\pi}{T}, \frac{6\pi}{T}, \dots$$

2.1.2 Calculation of the Fourier Coefficients

Before we compute the expressions of the Fourier coefficients, we need some tools. In all following integrals we integrate from $-T/2$ to $+T/2$, meaning over an interval with the period T that is symmetrical to $t = 0$. We could also pick any other interval, as long as the integrand is periodic with period T and gets integrated over a whole period. The letters n and m in the formulas below are natural numbers $0, 1, 2, \dots$ Let's have a look at the following

$$\int_{-T/2}^{+T/2} \cos \omega_n t \, dt = \begin{cases} 0 & \text{for } n \neq 0 \\ T & \text{for } n = 0 \end{cases} \tag{2.2}$$

$$\int_{-T/2}^{+T/2} \sin \omega_n t \, dt = 0 \tag{2.3}$$

This results from the fact that the areas on the positive half-plane and the ones on the negative one cancel out each other, provided we integrate over a whole number of periods. Cosine integral for $n = 0$ requires special treatment, as it lacks oscillations and therefore areas can't cancel out each other: there the integrand is 1, and the area under the horizontal line is equal to the width of the interval T. Furthermore, we need the following trigonometric identities:

$$\begin{aligned}\cos\alpha\cos\beta &= \frac{1}{2}\left[\cos(\alpha-\beta)+\cos(\alpha+\beta)\right]\\ \sin\alpha\sin\beta &= \frac{1}{2}\left[\cos(\alpha-\beta)-\cos(\alpha+\beta)\right]\\ \sin\alpha\cos\beta &= \frac{1}{2}\left[\sin(\alpha-\beta)+\sin(\alpha+\beta)\right]\end{aligned} \tag{2.4}$$

Using these identities we can demonstrate that the system of basis functions consisting of $(\sin\omega_k t,\ \cos\omega_k t)$ with $k = 0, 1, 2, \ldots$ is an *orthogonal system*. This means that

$$\int_{-T/2}^{+T/2} \cos\omega_n t\cos\omega_m t\,dt = \begin{cases} 0 & \text{for } n \neq m \\ T/2 & \text{for } n = m \neq 0 \\ T & \text{for } n = m = 0 \end{cases} \tag{2.5}$$

$$\int_{-T/2}^{+T/2} \sin\omega_n t\sin\omega_m t\,dt = \begin{cases} 0 & \text{for } n \neq m,\ n = 0,\ m = 0 \\ T/2 & \text{for } n = m \neq 0 \end{cases} \tag{2.6}$$

$$\int_{-T/2}^{+T/2} \sin\omega_n t\cos\omega_m t\,dt = 0 \tag{2.7}$$

Please note that our basis system is not an *orthonormal system*, i.e. the integrals for $n = m$ are not normalized to 1. What's even worse, the special case of $n = m = 0$ in (2.5) is a nuisance, and will keep bugging us again and again.
Using the above orthogonality relations, we are able to calculate the Fourier coefficients straight away. We need to multiply both sides of (2.1) by $\cos\omega_k t$ and integrate from $-T/2$ to $+T/2$. Due to the orthogonality, only terms with $k = k'$ will remain; the second integral will always disappear. This gives us:

$$A_k = \frac{2}{T}\int_{-T/2}^{+T/2} f(t)\cos\omega_k t\,dt \qquad \text{for } k \neq 0 \tag{2.8}$$

$$A_0 = \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\,dt \tag{2.9}$$

Please note the prefactors $2/T$ or $1/T$, respectively, in (2.8) and (2.9). Equation (2.9) simply is the average of the function $f(t)$. Now let's multiply both sides of (2.1) by $\sin\omega_k t$ and integrate from $-T/2$ to $+T/2$. We now have:

$$B_k = \frac{2}{T}\int_{-T/2}^{+T/2} f(t)\sin\omega_k t\,dt \qquad \text{for all } k \tag{2.10}$$

Equations (2.8) and (2.10) may also be interpreted like: by weighting the function $f(t)$ with $\cos\omega_k t$ or $\sin\omega_k t$, respectively, we "pick" the spectral components from $f(t)$, when integrating, corresponding to the even or odd components, respectively, of the frequency ω_k.

Ex. 2.1 *Calculation of Fourier coefficients: Constant and triangular functions*

2.1.3 Fourier Series and Music

While listening to music, we are able to clearly distinguish the sound produced by different instruments. The sound coming from a flute is quite different from the sound coming from a violin, even if they play the same note[1].
In musical terms, this difference is called ***timbre*** and it was von Helmholtz, in the second half of the XIX century, who understood that (von Helmholtz H. 1885. *"On the sensations of tone as a physiological basis for the theory of music"*)

> *"Each vibratory motion of the air in the ear canal, corresponding to a musical sound, can always be uniquely regarded as the sum of a number of vibratory movements."*

or, in mathematical terms, the timbre can be easily explained in terms of Fourier decomposition of the signal.
Indeed, if we apply Eq. (2.1) and extract the Fourier coefficients for various instruments we will obtain something like shown in Figure 2.3: it is evident that the harmonic content of different instruments is quite different.
For example, for the violin we have that the first Fourier frequencies are quite intense (and the brilliance of the violin sound is due the fact that these harmonics peak in the region where our ear is more sensitive). On the other hand, for the clarinet (here in the *chalumeau* registry) the even harmonics are quite faint, giving raise to its characteristic "hollow" sound. The typical metallic sound of the trumpet is due to the presence of very high harmonics, beyond the 21st.
It is interesting to observe that the harmonic content is different not only for different instruments playing the same note, but also for the same note played by the same instrument (the *La* played by a violin on the *La* string (not fingered) and the *La* played on the *Re* string (fingered)), or the same note played in different octaves. As an example, in Figure 2.4 we show the harmonic content of all the *Do*'s in the piano.

[1] In general terms, we call ***pitch*** of a sound the "perceived" frequency of a musical note, and it is related to the amount of Fourier frequencies we (that is, our ears) are able to distinguish.

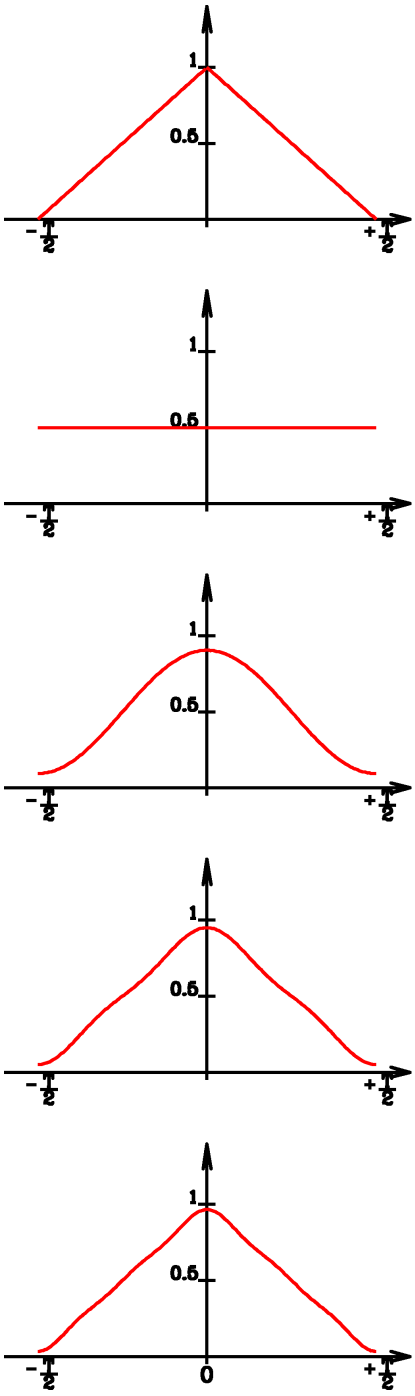


Figure 2.2: The triangular function and consecutive approximations by a Fourier series with more and more terms

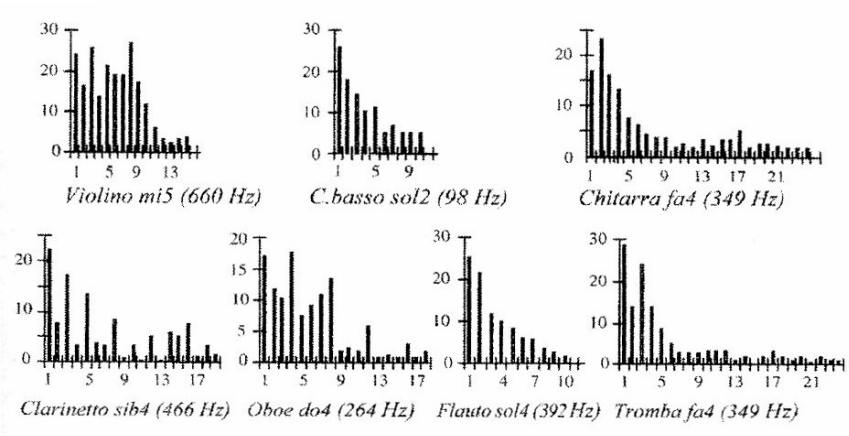


Figure 2.3: Comparison of a Fourier decomposition of a musical signal played by different instruments. On the abscissa we list the Fourier frequency index k, while the Y axis shows the power, in dB, emitted in the single harmonics (from Olson H.F., *"Music, Physics and Engineering"*).

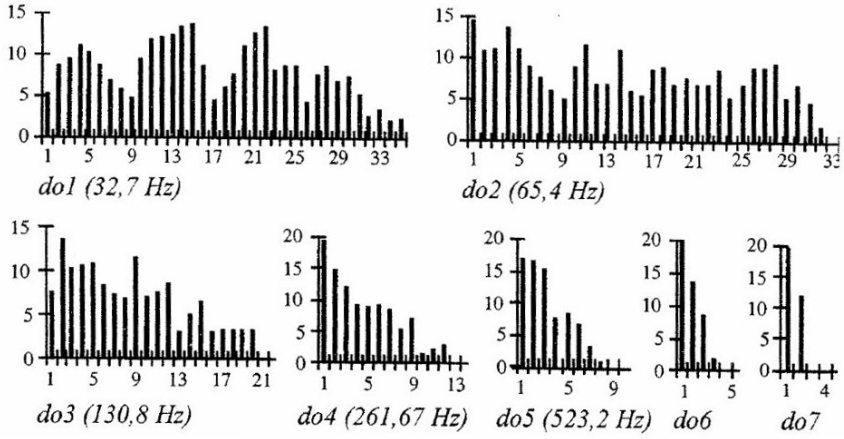


Figure 2.4: Comparison of Fourier decomposition of different *Do*'s played by a piano, from the left to the right of the keyboard. Note that for the *Do1* the fundamental and the lower harmonics are fainter than the higher ones, therefore the pitch is somewhat "virtual", while the lack of harmonics in *Do6* and *Do7* makes them an almost pure sound.

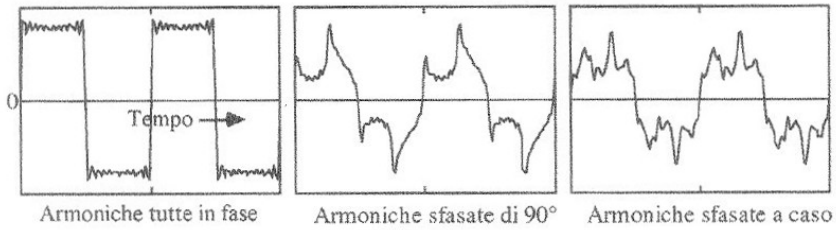


Figure 2.5: Effects of phase shifts among harmonics in a complex signal: *Left:* a square wave obtained by summing the first 21 harmonics all in phase among each other. *Central:* harmonics shifted by $\pi/2$. *Right:* random shift.

Note how the lack of harmonics in the higher *Do*'s, like *Do6* and *Do7*, makes them an almost "pure sound". Furthermore, for the *Do1* we can notice that the fundamental frequency and the lower harmonics are fainter than the harmonics between the 10th and the 15th. Despite this, our ears recognize the sound as a *Do1*. This phenomenon is called *"virtual pitch"*, and it is the demonstration that our brain is a "Fourier analyzer".
Our brain is therefore able to decompose an acoustic signal in its Fourier components, and is able to perceive each of them, independently of their phase relationships. In this perspective, the sentence that Leibniz' wrote in a letter to Christian Goldbach on April 17, 1712 was prophetic: *"Musica est exercitium arithmeticæ occultum nescientis se numerare animi"*[2].
The fact that the human brain is not able to perceive phase differences among harmonics is very important: indeed, the timbre of an instrument would change during the emission of a sound because of the different velocities of the harmonics along a string. To illustrate this phenomenon, in Figure 2.5 we show a square wave obtained by summing the first 21 harmonics, all in phase among each other. In the central panel we show the shape of the wave obtained by introducing a phase shift of $\pi/2$, while in the right panel the shift is random. While the wave shapes are completely different, if the signals are sent to a loudspeaker they are indistinguishable to the human ear.

2.1.4 Fourier Series in Complex Notation

In (2.1) the index k starts from 0, meaning that we will rule out *negative* frequencies in our Fourier series. The cosine terms didn't have a problem with negative frequencies. The sign of the cosine argument doesn't matter anyway, so we would be able to go halves as far as the spectral intensity at the positive frequency $k\omega$ was concerned: $-k\omega$ and $k\omega$ would get equal parts, as shown in Figure 2.6. As frequency $\omega = 0$ (a frequency as good as any other frequency $\omega \neq 0$) has no "brother", it will not have to go halves. A change of sign for the sine-terms arguments would result in a change of sign for the corresponding series term. The splitting of spectral intensity like "between brothers" (equal parts of $-\omega_k$ and $+\omega_k$ now will have to be like "between sisters": the sister for $-\omega_k$ also gets 50%, but hers is *minus* 50%!

[2] *Music is a hidden arithmetic exercise of the soul, which does not know that it is counting.*

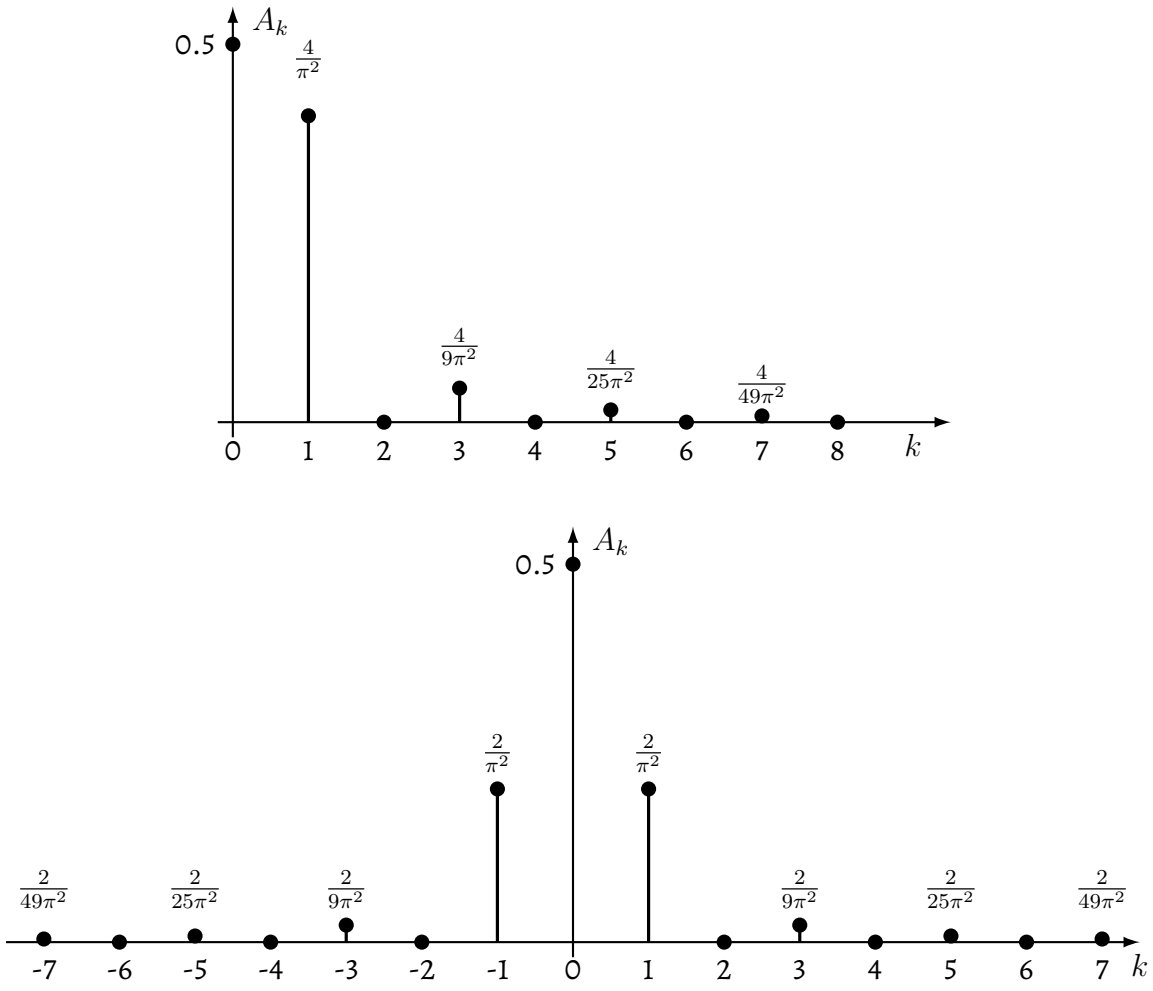


Figure 2.6: Plot of the "triangular function" Fourier frequencies: *Top:* Only positive frequencies; *Bottom:* Positive and negative frequencies.

Instead of using (2.1) we might as well use:

$$f(t) = \sum_{k=-\infty}^{+\infty} (A'_k \cos \omega_k t + B'_k \sin \omega_k t) \tag{2.11}$$

where, of course, the following is true: $A'_{-k} = A'_k$, $B'_{-k} = -B'_k$. The formulas for the computation of A'_k and B'_k for $k > 0$ are identical to (2.8) and (2.10), though the lack the extra factor 2. Equation (2.9) for A_0 stays unaffected by this. This helps us avoid to provide a special treatment for the constant term.
Now we're set and ready for the introduction of complex notation. In the following we'll always assume that $f(t)$ is a real function. Generalizing this for complex $f(t)$ is no problem. Our most important tool is Euler identity:

$$e^{i\alpha t} = \cos \alpha t + i \sin \alpha t \tag{2.12}$$

where i is the imaginary unit ($i^2 = -1$). This allows us to rewrite the trigonometric functions as

$$\begin{aligned} \cos \alpha t &= \frac{1}{2}(e^{i\alpha t} + e^{-i\alpha t}) \\ \sin \alpha t &= \frac{1}{2i}(e^{i\alpha t} - e^{-i\alpha t}) \end{aligned} \tag{2.13}$$

Inserting these relations into (2.1) we obtain

$$f(t) = A_0 + \sum_{k=1}^{\infty} \left(\frac{A_k - iB_k}{2} e^{i\omega_k t} + \frac{A_k + iB_k}{2} e^{-i\omega_k t} \right) \tag{2.14}$$

If we define

$$\begin{aligned} C_0 &= A_0 \\ C_k &= \frac{A_k - iB_k}{2} \\ C_{-k} &= \frac{A_k + iB_k}{2}, \qquad k = 1, 2, 3, \ldots \end{aligned} \tag{2.15}$$

we finally get

$$f(t) = \sum_{k=-\infty}^{\infty} C_k e^{i\omega_k t} \qquad \omega_k = \frac{2\pi k}{T} \tag{2.16}$$

Now C_k can be formulated in general terms as

$$C_k = \frac{1}{T} \int_{-T/2}^{+T/2} f(t)\, e^{-i\omega_k t}\, dt \qquad \text{for } k = 0, \pm 1, \pm 2, \ldots \tag{2.17}$$

Please note that there is a negative sign in the exponent. Please also note that the index k runs from $-\infty$ to $+\infty$ for C_k, whereas it runs from 0 to $+\infty$ for A_k and B_k.

2.1.5 Theorems and Rules

2.1.5.1 Linearity Theorem

Expanding a periodic function into a Fourier series is a linear operation. This means that we may use the two Fourier pairs:

$$\begin{aligned} f(t) &\leftrightarrow \{C_k; \omega_k\} \\ g(t) &\leftrightarrow \{C'_k; \omega_k\} \end{aligned} \tag{2.18}$$

to form the following combination

$$h(t) = a \times f(t) + b \times g(t) \leftrightarrow \{a\, C_k + b\, C'_k; \omega_k\} \tag{2.19}$$

Thus, we may easily determine the Fourier series of a function by splitting it into items whose Fourier series we already know.

2.1.5.2 The Shifting Rules

Often, we want to know how the Fourier series changes if we shift the function $f(t)$ along the time axis. This, for example, happens on a regular basis if we use a different interval, e.g. from 0 to T, instead of the symmetrical one from $-T/2$ to $+T/2$ we have used so far. In this situation, the *First Shifting Rule* comes in very handy:

$$\begin{aligned} f(t) &\leftrightarrow \{C_k; \omega_k\} \\ f(t-a) &\leftrightarrow \{C_k e^{-i\omega_k a}; \omega_k\} \end{aligned} \tag{2.20}$$

Proof.

$$\begin{aligned} C_k^{\text{new}} &= \frac{1}{T}\int_{-T/2}^{+T/2} f(t-a)\, e^{-i\omega_k t}\, dt \overset{t'=t-a}{=} \frac{1}{T}\int_{-(T/2)-a}^{+(T/2)-a} f(t')\, e^{-i\omega_k t'}\, e^{-i\omega_k a}\, dt' \\ &= e^{-i\omega_k a}\, C_k^{\text{old}} \end{aligned}$$

We integrate over a full period, that's why shifting the limits of the interval by a does not make any difference. The proof is trivial, the result of the shifting along the time axis not! The new Fourier coefficient results from the old coefficient C_k by multiplying it with the phase factor $e^{-i\omega_k a}$. As C_k generally is complex, shifting "shuffles" real and imaginary parts.

Ex. 2.2 *Shifting rules: Triangular function with average equal to zero. Quarter period shifted triangular function. Half period shifted triangular function.*

The *First Shifting Rule* showed us that shifting within the time domain leads to a multiplication by a phase factor in the frequency domain. Reversing this statement gives us the *Second Shifting Rule*:

$$\begin{aligned} f(t) &\leftrightarrow \{C_k; \omega_k\} \\ f(t)\, e^{\frac{i\,2\pi a\, t}{T}} &\leftrightarrow \{C_{k-a}; \omega_k\} \end{aligned} \tag{2.21}$$

In other words, a multiplication of the function $f(t)$ by the phase factor $e^{i\,2\pi a t/T}$ results in frequency ω_k now being related to "shifted" coefficient C_{k-a} instead of the former coefficient C_k. A comparison between (2.21) and (2.20) demonstrates the two-sided character of the two Shifting Rules. If a is an integer, there won't be any problem if we simply take the coefficient shifted by a. But what if a is not an integer?
Strangely enough nothing serious will happen. Simply shifting like we did before won't work any more, but who is to keep us from inserting $(k-a)$ into the expression for old C_k, whenever k occurs.
Before we present examples, two more ways of writing down the Second Shifting Rule are in order:

$$\begin{aligned} f(t) \leftrightarrow & \{A_k; B_k; \omega_k\} \\ f(t)\, e^{\frac{i\,2\pi a\, t}{T}} \leftrightarrow & \left\{ \frac{1}{2}\left[A_{k+a} + A_{k-a} + i\left(B_{k+a} - B_{k-a}\right)\right]; \right. \\ & \left. \frac{1}{2}\left[B_{k+a} - B_{k-a} + i\left(A_{k-a} - A_{k+a}\right)\right]; \omega_k \right\} \end{aligned} \tag{2.22}$$

Caution! This is valid for $k \neq 0$. Note that old A_0 becomes $A_a/2 + iB_a/2$. The formulas becomes a lot simpler in case $f(t)$ is real. In this case we get:

$$f(t)\, \cos\frac{2\pi a t}{T} \leftrightarrow \left\{ \frac{A_{k+a} + A_{k-a}}{2}; \frac{B_{k+a} + B_{k-a}}{2}; \omega_k \right\} \tag{2.23}$$

old A_0 becomes $A_a/2$ and

$$f(t)\, \sin\frac{2\pi a t}{T} \leftrightarrow \left\{ \frac{B_{k+a} - B_{k-a}}{2}; \frac{A_{k+a} - A_{k-a}}{2}; \omega_k \right\} \tag{2.24}$$

old A_0 becomes $B_a/2$.

Ex. 2.3 *Second Shifting Rule: constant function and triangular function*

2.1.5.3 Scaling Theorem

Sometimes we happen to want to scale the time axis. In this case, there is no need to re-calculate the Fourier coefficients. From:

$$\begin{aligned} f(t) &\leftrightarrow \{C_k; \omega_k\} \\ \text{we get: } f(at) &\leftrightarrow \{C_k; \frac{\omega_k}{a}\} \end{aligned} \tag{2.25}$$

Here, a must be real! For $a > 1$ the time axis will be stretched and, hence, the frequency axis will be compressed. For $a < 1$ the opposite is true. The proof for (2.25) is easy and follows from (2.17):

$$\begin{aligned} C_k^{\text{new}} &= \frac{a}{T} \int_{-T/2a}^{+T/2a} f(at)\, e^{-i\omega_k t}\, dt \overset{t'=at}{=} \frac{a}{T} \int_{-T/2}^{+T/2} f(t')\, e^{-i\omega_k t'/a}\, \frac{1}{a}\, dt' \\ &= C_k^{\text{old}} \qquad \text{with } \omega_k^{\text{new}} = \frac{\omega_k^{\text{old}}}{a} \end{aligned}$$

Please note that we also have to stretch or compress the interval limits because of the requirement of periodicity. Here, we have tacitly assumed $a > 0$. For $a < 0$, we would only reverse the time axis and, hence, also the frequency axis. For the special case $a = -1$ we have:

$$\begin{aligned} f(t) &\leftrightarrow \{C_k; \omega_k\} \\ f(-t) &\leftrightarrow \{C_k; -\omega_k\} \end{aligned} \tag{2.26}$$

2.1.6 Partial Sums, Parseval Equation

For practical work, infinite Fourier series have to get terminated at some stage, regardless. Therefore, we only use a partial sum, say until we reach $k_{\text{max}} = N$. This Nth partial sum then is:

$$S_N = \sum_{k=0}^{N} (A_k \cos \omega_k t + B_k \sin \omega_k t) \tag{2.27}$$

Terminating the series results in the following squared error:

$$\delta_N^2 = \frac{1}{T} \int_T [f(t) - S_N(t)]^2\, dt \tag{2.28}$$

The T below the integral symbol means integration over a full period. This definition will become plausible in a second if we look at the discrete version:

$$\delta_N^2 = \frac{1}{T} \sum_{j=0}^{N} (f_j - s_j)^2 \tag{2.29}$$

Please note that we divide by the length of the interval, to compensate for integrating over the interval T. Now we know that the following is correct for the infinite series:

$$\lim_{N\to\infty} S_N = \sum_{k=0}^{\infty}(A_k \cos\omega_k t + B_k \sin\omega_k t) \tag{2.30}$$

provided the A_k and B_k happen to be the Fourier coefficients. Does this also have to be true for the Nth partial sum? Isn't there a chance the mean squared error would get smaller, if we used other coefficients instead of Fourier coefficients? That's not the case! To prove it, we'll now insert (2.27) and (2.28) in (2.30), leave out $\lim_{N\to\infty}$ and get:

$$\begin{aligned}
\delta_N^2 &= \frac{1}{T}\left\{\int_T f^2(t)\,dt - 2\int_T f(t)S_N(t)\,dt + \int_T S_N^2(t)\,dt\right\} \\
&= \frac{1}{T}\left\{\int_T f^2(t)\,dt \right. \\
&\quad -2\int_T \sum_{k=0}^{\infty}(A_k\cos\omega_k t + B_k\sin\omega_k t)\sum_{k=0}^{N}(A_k\cos\omega_k t + B_k\sin\omega_k t)\,dt \\
&\quad \left. + \int_T \sum_{k=0}^{N}(A_k\cos\omega_k t + B_k\sin\omega_k t)\sum_{j=0}^{N}(A_j\cos\omega_j t + B_j\sin\omega_j t)\,dt\right\} \\
&= \frac{1}{T}\left\{\int_T f^2(t)\,dt - 2TA_0^2 - 2\frac{T}{2}\sum_{k=1}^{N}(A_k^2+B_k^2) + TA_0^2 + \frac{T}{2}\sum_{k=1}^{N}(A_k^2+B_k^2)\right\} \\
&= \frac{1}{T}\int_T f^2(t)\,dt - A_0^2 - \frac{1}{2}\sum_{k=1}^{N}(A_k^2+B_k^2)
\end{aligned} \tag{2.31}$$

Here, we made use of the somewhat cumbersome orthogonality properties (2.5), (2.6) and (2.7). As the A_k^2 and B_k^2 always are positive, the mean squared error will drop monotonically while N increases.

Ex. 2.4 *Approximating the triangular function*

As δ_N^2 is always positive, we finally arrive from (2.31) at the Bessel inequality

$$\frac{1}{T}\int_T f^2(t)\,dt \geq A_0^2 + \frac{1}{2}\sum_{k=1}^{N}(A_k^2+B_k^2) \tag{2.32}$$

For the border-line case of $N\to\infty$ we get the Parseval equation:

$$\frac{1}{T}\int_T f^2(t)\,dt = A_0^2 + \frac{1}{2}\sum_{k=1}^{\infty}(A_k^2 + B_k^2) \tag{2.33}$$

Parseval equation may be interpreted as follows: $1/T \int f^2(t)\,dt$ is the mean squared "signal" within the time domain, or – more colloquially – the *information content.* Fourier series don't lose this information content: it's in the squared Fourier coefficients.

2.2 Continuous Fourier Transformation

Contrary to Section 2.1, here we won't limit things to periodic $f(t)$. The integration interval is the entire real axis $(-\infty, +\infty)$. For this purpose we'll look at what happens at the transition from a series-representation to an integral-representation:

$$\text{Series:} \quad C_k = \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\,e^{-i\omega_k t}\,dt$$

$$\text{Continuous:} \quad \lim_{T\to\infty}(TC_k) = \int_{-\infty}^{+\infty} f(t)\,e^{-i\omega t}\,dt$$

2.2.1 Definition

Let us define the *Forward Fourier Transformation* and the *Inverse Fourier Transformation* as follows:

Definition 2.3 (Forward Fourier transformation).

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}\,dt \tag{2.34}$$

Definition 2.4 (Inverse Fourier transformation).

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)e^{+i\omega t}\,d\omega \tag{2.35}$$

Please note that in the case of the forward transformation, there is a minus sign in the exponent (cf. (2.17)), in the case of the inverse transformation, this is a plus sign. In the case of the inverse transformation, $1/2\pi$ is in front of the integral, contrary to the forward transformation.
The asymmetric aspect of the formulas has tempted many scientists to introduce other definitions, for example to write a factor $1/\sqrt{2\pi}$ for forward as well as inverse transformation. That's no good, as the definition of the average $F(0) = \int_{-\infty}^{+\infty} f(t)\,dt$ would be affected.
Now let us demonstrate that the inverse transformation returns us to the original function. For the forward transformation, we often will use $\text{FT}(f(t))$, and for the inverse transformation we will use $\text{FT}^{-1}(F(\omega))$. We will begin with the inverse transformation and insert:

Proof.

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)\, e^{i\omega t}\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} f(t')\, e^{-i\omega t'}\, e^{i\omega t}\, dt' \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(t')\, dt' \int_{-\infty}^{+\infty} e^{i(t-t')\omega}\, d\omega \\ &= \int_{-\infty}^{+\infty} f(t')\, \delta(t-t')\, dt' = f(t) \end{aligned} \tag{2.36}$$

where $\delta(t)$ is the Dirac δ-function[3].

2.2.2 Transformation of relevant functions

2.2.2.1 The δ-function

From (2.36), by putting $f(t) = 1$ we have

$$\begin{aligned} \text{FT}(\delta(t)) &= 1 \\ \text{FT}^{-1}(1) &= 2\pi\delta(\omega) \end{aligned} \tag{2.37}$$

We realize the dual character of the forward and inverse transformations: a very slowly varying function $f(t)$ will have a very high spectral density for very small frequencies; the spectral density will go down quickly and rapidly approaches zero. Conversely, a quickly varying function $f(t)$ will show spectral density over a very wide frequency range (we will discuss about this issue in more detail in Section 2.2.6).

2.2.2.2 The Dirac comb

A Dirac comb (also called "sampling function", see (5.9)), is an infinite sequence of Dirac δ-functions placed at even intervals of size T:

$$\text{III}_T(t) \equiv \sum_{-\infty}^{+\infty} \delta(t - nT). \tag{2.38}$$

The Fourier transform of a Dirac comb spaced with period T is a Dirac comb spaced with period $1/T$ (see Figure 2.7, fourth panel)

$$F\text{III}_T(t)) = \frac{1}{T}\,\text{III}_{\frac{1}{T}}(\omega). \tag{2.39}$$

[3] The δ-function is actually a distribution. Its value is zero anywhere except when its argument is equal to zero. In this case it is ∞.

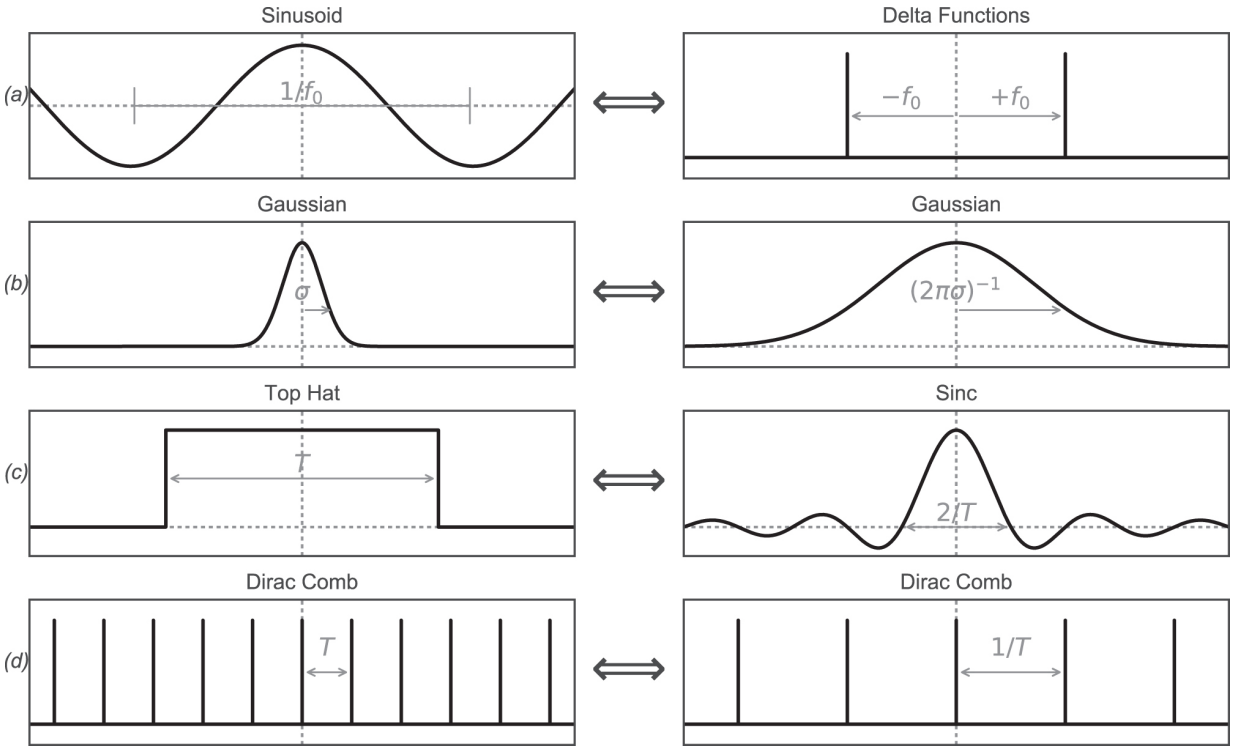


Figure 2.7: Fourier transforms of relevant functions. On the right the function, on the left the corresponding Fourier transformation. From VanderPlas 2018.

2.2.2.3 A sinusoid with frequency ω_0

From the definition of the Fourier transformation we have

$$F\left(e^{\omega_0 t}\right) = \int_{-\infty}^{+\infty} e^{-i(\omega-\omega_0)t}\, dt = \delta(\omega - \omega_0). \tag{2.40}$$

Using the Euler identity (2.12) we can write

$$\cos(\omega t) = \frac{e^{i\omega t} + e^{-i\omega t}}{2}; \qquad \sin(\omega t) = \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \tag{2.41}$$

Combining (2.40) and (2.41), along with the linearity of the Fourier transform, we obtain

$$\begin{aligned} \mathrm{FT}\left(\cos(\omega_0 t)\right) &= \frac{1}{2}\left[\delta(\omega-\omega_0) + \delta(\omega+\omega_0)\right] \\ \mathrm{FT}\left(\sin(\omega_0 t)\right) &= \frac{1}{2i}\left[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)\right] \end{aligned} \tag{2.42}$$

In other words, a sinusoidal signal with frequency ω_0 has a Fourier transform consisting of a weighted sum of δ-functions at $\pm\omega_0$ (see Figure 2.7, first panel).

2.2.2.4 The Gaussian function

The prefactor is chosen in such a way that the area under the function is normalized to unity.

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{t^2}{\sigma^2}}$$

Its Fourier transform is

$$\begin{aligned} F(\omega) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{t^2}{\sigma^2}} e^{-i\omega t}\, dt \\ &= \frac{2}{\sigma\sqrt{2\pi}} \int_{0}^{+\infty} e^{-\frac{1}{2}\frac{t^2}{\sigma^2}} \cos\omega t\, dt \\ &= \exp\left(-\frac{1}{2}\sigma^2\omega^2\right) \end{aligned}$$

Again, the imaginary part is null because the function is even. The Fourier transform of a Gaussian results to be another Gaussian. Note that the Fourier transform is *not* normalized to unit area.
$f(t)$ has σ in the exponent denominator, while $F(\omega)$ has it in the exponent numerator: the slimmer $f(t)$, the wider $F(\omega)$ and vice versa, as shown in Figure 2.7, second panel.

2.2.2.5 The "rectangular" function

Now let us discuss an important example: the Fourier transform of the "rectangular" normalized function (see Section 2.3.3 for a detailed discussion)

$$f(t) = \begin{cases} 1/T & \text{for } -T/2 \le t \le +T/2 \\ 0 & \text{else} \end{cases}$$

Its Fourier transform is

$$F(\omega) = \frac{2}{T}\int_0^{+T/2} \cos\omega t\, dt = \frac{\sin\omega T/2}{\omega T/2} \tag{2.43}$$

The imaginary part is 0, as $f(t)$ is even. The Fourier transformation of a rectangular function, therefore, is of the type $\sin x/x$. Some authors use the expression $\mathrm{sinc}(x)$ for this case. The "c" stands for *cardinal* and we will discuss about its importance in signal analysis in Section 3.4. The functions $f(t)$ and $F(\omega)$ are shown in Figure 2.7, third panel.

Ex. 2.5 *Fourier transformation of relevant functions: bilateral exponential, unilateral exponential*

2.2.3 Theorems and Rules

2.2.3.1 Linearity Theorem

For completeness' sake, once again:

$$\begin{aligned} f(t) &\leftrightarrow F(\omega); \\ g(t) &\leftrightarrow G(\omega); \\ a \times f(t) + b \times g(t) &\leftrightarrow a \times F(\omega) + b \times G(\omega) \end{aligned} \tag{2.44}$$

2.2.3.2 Shifting Rules

We already know: shifting in the time domain means modulation in the frequency domain, and a modulation in the time domain results in a shift in the frequency domain

$$\begin{aligned} f(t) &\leftrightarrow F(\omega); \\ f(t-a) &\leftrightarrow F(\omega)\, e^{-i\omega a} \\ f(t)\, e^{-i\omega_0 t} &\leftrightarrow F(\omega - \omega_0) \end{aligned} \tag{2.45}$$

2.2.3.3 Scaling Theorem

$$\begin{aligned} f(t) &\leftrightarrow F(\omega); \\ f(at) &\leftrightarrow \frac{1}{|a|} F\left(\frac{\omega}{a}\right) \end{aligned} \tag{2.46}$$

Proof. Analogously to (2.25) with the difference that here we cannot stretch or compress the interval limits $\pm\infty$:

$$\begin{aligned} F(\omega)^{\text{new}} \quad &= \quad \frac{1}{T} \int_{-\infty}^{+\infty} f(at)\, e^{-i\omega t}\, dt \\ &\overset{t'=at}{=} \quad \frac{1}{T} \int_{-\infty}^{+\infty} f(t')\, e^{-i\omega t'/a}\, \frac{1}{a}\, dt \\ \frac{1}{|a|} F(\omega)^{\text{old}} \quad &\text{with} \quad \omega = \frac{\omega^{\text{old}}}{a} \end{aligned}$$

Here, we tacitly assumed $a > 0$. For $a < 0$ we would get a minus sign in the prefactor; however, we would also have to interchange the integration limits and thus get together the factor $1/|a|$. This means: stretching (compressing) the time-axis results in the compression (stretching) of the frequency-axis. For the special case $a = -1$ we get:

$$\begin{aligned} f(t) &\rightarrow F(\omega); \\ f(-t) &\rightarrow F(-\omega); \end{aligned} \tag{2.47}$$

Therefore, turning around the time axis ("looking into the past") results in turning around the frequency axis.

2.2.4 Convolution, Parseval Theorem

2.2.4.1 Convolution

The convolution of a function $f(t)$ with another function $g(t)$ is defined as:

Definition 2.5 (Convolution).

$$h(t) = \int_{-\infty}^{+\infty} f(\xi)\, g(t-\xi)\, d\xi \equiv f(t) \otimes g(t) \tag{2.48}$$

Please note that there is a minus sign in the argument of $g(t)$. The convolution is commutative, distributive, and associative. This means

$$\begin{aligned} \text{commutative:} \quad & f(t) \otimes g(t) = g(t) \otimes f(t) \\ \text{distributive:} \quad & f(t) \otimes (g(t) + h(t)) = f(t) \otimes g(t) + f(t) \otimes h(t) \\ \text{associative:} \quad & f(t) \otimes (g(t) \otimes h(t)) = (f(t) \otimes g(t)) \otimes h(t) \end{aligned}$$

Before going on with the mathematical demonstration of the convolution theorem, let us present two physical examples of convolution.
A real observation of a physical phenomenon cannot lasts forever, but will be performed for a certain amount of time. This corresponds to "convolute" a continuous signal with a rectangular function, as shown in Figure 2.8. According to the convolution theorem, the Fourier transform of the convolution is the point wise product of the individual Fourier transforms. This concept will be discussed into details in the Second Part of the course.
As another example of convolution, let us take a pulse that looks like an unilateral exponential function

$$f(t) = \begin{cases} e^{-t/\tau} & \text{for } t \geq 0 \\ 0 & \text{else} \end{cases} \tag{2.49}$$

Any device that delivers pulses as a function of time, has a finite rise-time/decay-time, which for simplicity's sake we'll assume to be a Gaussian

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{t^2}{\sigma^2}\right) \tag{2.50}$$

That is how our device would represent a δfunction – we can't get sharper than that. The function $g(t)$, therefore, is the device's resolution function, which we'll have to use for the convolution of all signals we want to record. An example would be the bandwidth of an oscilloscope. We then need:

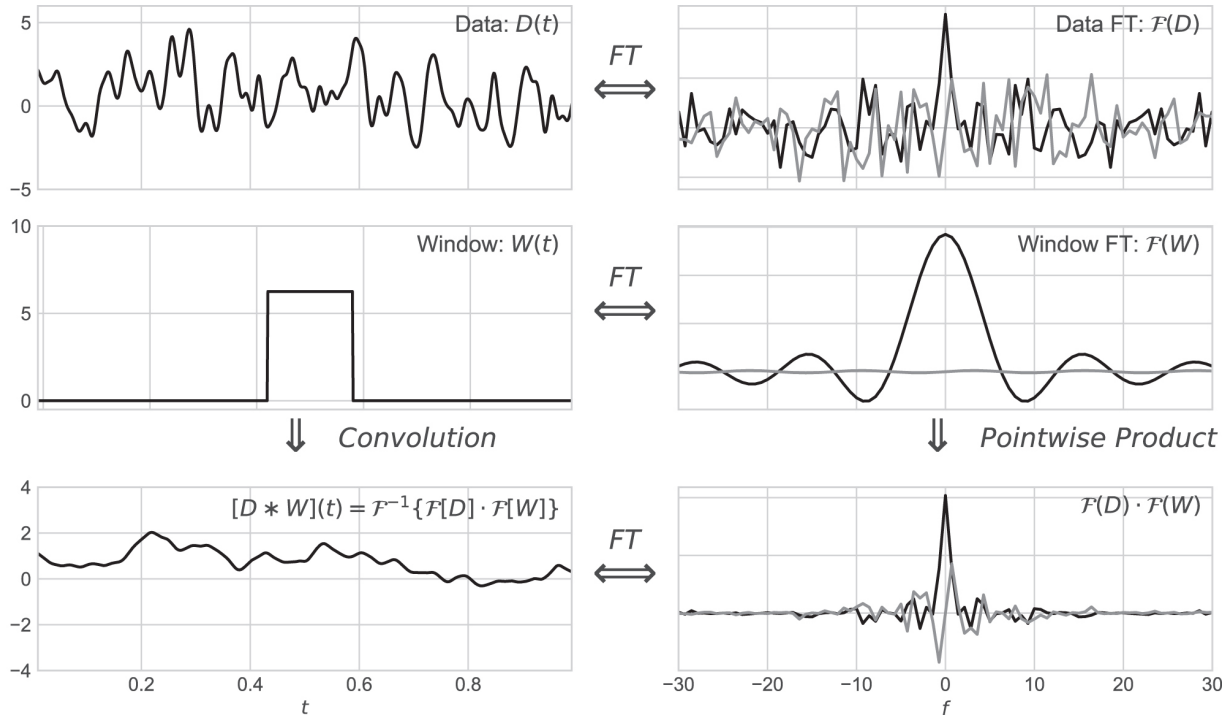


Figure 2.8: Visualization of the convolution theorem: In the right panels, the black and gray lines represent the real and imaginary parts of the transforms, respectively. From VanderPlas 2018.

$$S(t) = f(t) \otimes g(t) \tag{2.51}$$

where $S(t)$ is the experimental, *smeared* signal. It's obvious that the rise at $t = 0$ will not be as steep, and the peak of the exponential function will get "ironed out". We'll have to take a closer look:

$$
\begin{aligned}
S(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_0^{+\infty} e^{-\xi/\tau} \exp\left(-\frac{1}{2}\frac{(t-\xi)^2}{\sigma^2}\right) d\xi \\
&= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{t^2}{\sigma^2}\right) \int_0^{+\infty} \exp\left[-\frac{\xi}{\tau} + \frac{t\xi}{\sigma^2} - \frac{1}{2}\frac{\xi^2}{\sigma^2}\right] d\xi \\
&= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{t^2}{\sigma^2}\right) \exp\left(\frac{1}{2}\frac{t^2}{\sigma^2}\right) \exp\left(-\frac{t}{\tau}\right) \exp\left(\frac{\sigma^2}{2\tau^2}\right) \times \\
&\qquad \int_0^{+\infty} \exp\left\{-\frac{1}{2\sigma^2}\left[\xi - \left(t - \frac{\sigma^2}{\tau}\right)\right]^2\right\} d\xi \qquad (2.52) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t}{\tau}\right) \exp\left(\frac{\sigma^2}{2\tau^2}\right) \times \\
&\qquad \int_{-(t-\sigma^2/\tau)}^{+\infty} \exp\left(-\frac{1}{2}\frac{\xi'^2}{\sigma^2}\right) d\xi' \qquad \text{with } \xi' = \xi - \left(t - \frac{\sigma^2}{\tau}\right) \\
&= \frac{1}{2} \exp\left(-\frac{t}{\tau}\right) \exp\left(\frac{\sigma^2}{2\tau^2}\right) \operatorname{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} - \frac{t}{\sigma\sqrt{2}}\right)
\end{aligned}
$$

Here, $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary error function, where

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \qquad (2.53)$$

Figure 2.9 shows the result of the convolution of the exponential function with the Gaussian. The following properties immediately stand out: (i) The finite time resolution ensures that there is also a signal at negative times, whereas it was 0 before convolution. (ii) The maximum is not at $t = 0$ any more. (iii) What can't be seen straight away, yet is easy to grasp, is the following: the center of gravity of the exponential function, which was at $t = \tau$, doesn't get shifted at all upon convolution.
Now we prove the extremely important *Convolution Theorem*:

Theorem 2.1 (Convolution theorem). *Let be*

$$
\begin{aligned}
f(t) &\leftrightarrow F(\omega) \\
g(t) &\leftrightarrow G(\omega)
\end{aligned}
$$

Then

$$h(t) = f(t) \otimes g(t) \leftrightarrow H(\omega) = F(\omega) \times G(\omega) \qquad (2.54)$$

The *convolution integral* becomes, through Fourier transformation, a product of the Fourier-transformed ones.

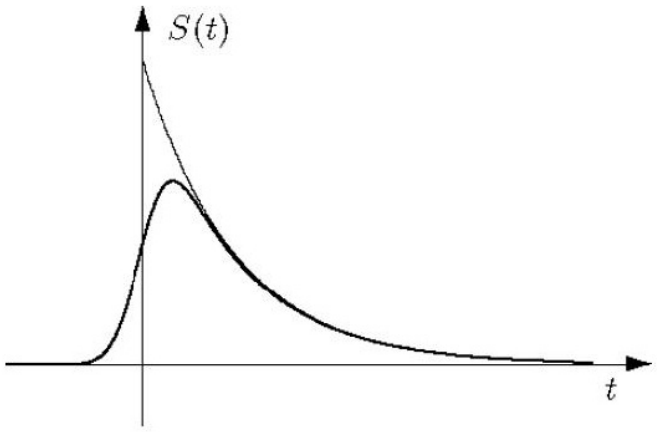


Figure 2.9: Result of the convolution of an unilateral exponential function with a Gaussian. The exponential function without convolution is indicated with the thin line

Proof.

$$\begin{aligned} H(\omega) &= \int\int f(\xi)g(t-\xi)\,d\xi\; e^{-i\omega t}\,dt \\ &= \int f(\xi)\,e^{-i\omega\xi}\left[\int g(t-\xi)\,e^{-i\omega(t-\xi)}\,dt\right]\,d\xi \\ &\overset{t'=t-\xi}{=} \int f(\xi)\,e^{-i\omega\xi}\,d\xi\,G(\omega) \\ &= F(\omega)\,G(\omega) \end{aligned}$$

The integration boundaries $\pm\infty$ did not change by doing that, and $G(\omega)$ does not depend on ξ. The *inverse* Convolution theorem is:

Theorem 2.2 (Inverse convolution theorem). *Let be*

$$\begin{aligned} f(t) &\leftrightarrow F(\omega) \\ g(t) &\leftrightarrow G(\omega) \end{aligned}$$

Then

$$h(t) = f(t) \times g(t) \leftrightarrow H(\omega) = \frac{1}{2\pi}\,F(\omega) \otimes G(\omega) \tag{2.55}$$

Proof.

$$\begin{aligned} H(\omega) &= \int f(t)\, g(t)\, e^{-i\omega t}\, dt \\ &= \int \left(\frac{1}{2\pi} \int F(\omega')\, e^{+i\omega' t}\, d\omega' \times \frac{1}{2\pi} \int G(\omega'')\, e^{+i\omega'' t}\, d\omega'' \right) e^{-i\omega t}\, dt \\ &= \frac{1}{(2\pi)^2} \int F(\omega') \int G(\omega'') \underbrace{\int e^{i(\omega'+\omega''-\omega)t}\, dt}_{=2\pi\delta(\omega'+\omega''-\omega)}\, d\omega' d\omega'' \\ &= \frac{1}{2\pi} \int F(\omega')\, G(\omega-\omega')\, d\omega' \\ &= \frac{1}{2\pi} F(\omega) \otimes G(\omega) \end{aligned}$$

Contrary to the Convolution Theorem (2.54) in (2.55) there is a factor $1/2\pi$ in front of the convolution of the Fourier transforms.
A widely popular exercise is the *unfolding* of data: the instruments' resolution function "smears out" the quickly varying functions, but we naturally want to reconstruct the data to what they would look like if the resolution function was infinitely good – provided we precisely knew the resolution function. In principle, that's a good idea – and thanks to the Convolution Theorem, not a problem: we Fourier-transform the data, divide by the Fourier-transformed resolution function and transform it back. For practical applications it doesn't quite work that way. As in real life, we can't transform from $-\infty$ to $+\infty$, we need low-pass filters, in order not to get "swamped" with oscillations resulting from cut-off errors. Therefore, the advantages of unfolding are just as quickly lost as gained. Actually, the following is obvious: whatever got "smeared" by finite resolution, can't be reconstructed unambiguously. Imagine that a very pointed peak got eroded over millions of years, so there's only gravel left at its bottom. Try reconstructing the original peak from the debris around it! The result might be impressive from an artist's point of view, an artifact, but it hasn't got much to do with the original reality.

Ex. 2.6 *Convolution: Gaussian frequency distribution. Lorentzian frequency distribution*

2.2.4.2 Cross Correlation

Sometimes, we want to know if a measured function $f(t)$ has anything in common with another measured function $g(t)$. Cross correlation is ideally suited to that.

Definition 2.6 (Cross correlation).

$$h(t) = \int_{-\infty}^{+\infty} f(\xi)\, g^*(t+\xi)\, d\xi \equiv f(t) \star g(t) \tag{2.56}$$

Important note: Here, there is a plus sign in the argument of g, therefore we don't mirror $g(t)$. For even functions $g(t)$ this doesn't matter. The asterisk $*$ means complex conjugated. We may disregard it for real functions. The symbol $\star$ means cross correlation, and is not to be confounded with $\otimes$ for folding. Cross correlation is associative and distributive, yet not commutative. That's not only because of the complex-conjugated symbol, but mainly because of the plus sign in the argument of $g(t)$. Of course, we want to convert the integral in the cross correlation to a product by using Fourier transformation.

Theorem 2.3 (Cross correlation). *Let be*

$$\begin{aligned} f(t) &\leftrightarrow F(\omega) \\ g(t) &\leftrightarrow G(\omega) \end{aligned}$$

Then

$$h(t) = f(t) \star g(t) \leftrightarrow H(\omega) = F(\omega) \times G^*(\omega) \tag{2.57}$$

Proof.

$$\begin{aligned} H(\omega) &= \int \int f(\xi)\, g^*(t+\xi)\, d\xi\, e^{-i\omega t}\, dt \\ &= \int f(\xi) \left[\int g^*(t+\xi)\, e^{-i\omega t}\, dt \right] d\xi \\ &= \int f(\xi)\, G^*(+\omega)\, e^{-i\omega\xi}\, d\xi \\ &= F(\omega) \times G^*(\omega) \end{aligned}$$

In the third passage we used the first shifting rule (2.45) with $\xi = -a$. In the last passage we use the following identity:

$$\begin{aligned} G(\omega) &= \int g(t)\, e^{-i\omega t}\, dt \\ G^*(\omega) &= \int g^*(t)\, e^{+i\omega t}\, dt \\ G^*(-\omega) &= \int g^*(t)\, e^{-i\omega t}\, dt \end{aligned}$$

The interpretation of (2.57) is simple: if the spectral densities of $f(t)$ and $g(t)$ are a good match, i.e. have much in common, then $H(\omega)$ will become large on average, and the cross correlation $h(t)$ will also be large, on average. Otherwise, if $F(\omega)$ would be small e.g. where $G^*(\omega)$ is large and vice versa, so that there is never much left for the product $H(\omega)$. Then also $h(t)$ would be small, i.e. there is not much in common between $f(t)$ and $g(t)$.

2.2.4.3 Autocorrelation

The autocorrelation function is the cross correlation of a function $f(t)$ with itself. We may ask, for what purpose we'd want to check for what $f(t)$ has in common with $f(t)$. Autocorrelation, however, seems to attract many people in a magical manner. We often hear the view, that a signal full of noise can be turned into something really good by using the autocorrelation function, i.e. the signal-to-noise ratio would improve a lot. Don't you believe a word of it! We'll see why shortly.

Definition 2.7 (Autocorrelation).

$$h(t) = \int f(t)\, f^*(\xi + t)\, d\xi \tag{2.58}$$

From its definition and the cross-correlation theorem (2.57) we have the so called

Theorem 2.4 (Wiener-Khinchin theorem).

$$\begin{gathered} f(t) \leftrightarrow F(\omega) \\ h(t) = f(t) \star f(t) \leftrightarrow H(\omega) = F(\omega) \times F^*(\omega) = |F(\omega)|^2 \end{gathered} \tag{2.59}$$

We may either use the Fourier transform $F(\omega)$ of a noisy function $f(t)$ and get angry about the noise in $F(\omega)$, or we first form the autocorrelation function $h(t)$ from the function $f(t)$ and are then happy about the Fourier transform $H(\omega)$ of function $h(t)$. Normally, $H(\omega)$ does look a lot less noisy, indeed. Instead of doing it the roundabout way by using the autocorrelation function, we could have used the square of the magnitude of $F(\omega)$ in the first place. We all know, that a squared representation in the ordinate always pleases the eye, if we want to do cosmetics to a noisy spectrum. Big spectral components will grow when squared, small ones will get even smaller. But isn't it rather obvious that squaring doesn't change anything to the signal-to-noise ratio? In order to make it "look good", we pay the price of losing linearity.

2.2.4.4 The Parseval Theorem

The autocorrelation function also comes in handy for something else, namely for deriving Parseval theorem. We start out with (2.58), insert especially $t = 0$, and get Parseval theorem:

Theorem 2.5 (Parseval theorem).

$$h(0) = \int |f(t)|^2\, dt = \frac{1}{2\pi} \int |F(\omega)|^2\, d\omega \tag{2.60}$$

The second equal sign is obtained by inverse transformation of $|F(\omega)|^2$ where, for $t = 0$, $e^{i\omega t}$ becomes unity.
Equation (2.60) states that the *information content* of the function $f(x)$ – defined as integral over the square of the magnitude – is just as large as the *information content* of its Fourier transform $F(\omega)$ (same definition, but with $1/(2\pi)$).

2.2.5 Fourier Transformation of Derivatives

When solving differential equations, we can make life easier using Fourier transformation. The derivative simply becomes a product:

$$\begin{aligned} f(t) &\leftrightarrow F(\omega) \\ f'(t) &\leftrightarrow i\omega F(\omega) \end{aligned} \tag{2.61}$$

The proof is straight-forward:

Proof.

$$\begin{aligned} \mathrm{FT}(f'(t)) &= \int_{-\infty}^{+\infty} f'(t)\, e^{-i\omega t}\, dt = f(t)\, e^{-i\omega t}|_{-\infty}^{+\infty} - (-i\omega) \int_{-\infty}^{+\infty} f(t)\, e^{-i\omega t}\, dt \\ &= i\omega F(\omega) \end{aligned}$$

The first term in the partial integration is discarded, as $f(t) \to 0$ for $t \to \infty$, otherwise $f(t)$ could not be integrable. This game can go on:

$$\mathrm{FT}\left(\frac{df^n(t)}{d^n t}\right) = (i\omega)^n F(\omega) \tag{2.62}$$

For negative n we may also use the formula for integration. We can also formulate in a simple way the derivative of a Fourier transform $F(\omega)$ with respect to the frequency ω:

$$\frac{dF(\omega)}{d\omega} = -i\, \mathrm{FT}(t\, f(t)) \tag{2.63}$$

Proof.

$$\frac{dF(\omega)}{d\omega} = \int_{-\infty}^{+\infty} f(t)\, \frac{d}{d\omega} e^{-i\omega t}\, dt = -i \int_{-\infty}^{+\infty} f(t)\, t\, e^{-i\omega t}\, dt = -i\, \mathrm{FT}(t\, f(t))$$

2.2.6 Fourier Transform and Uncertainty Relation

At this point it should be clear that the behavior of a function and its Fourier transform is in a certain sense complementary: to a function which is "wide spread" in the time domain corresponds a Fourier transform which is "narrow" in the frequency domain and vice versa (see, e.g., the case for $f(t)$ a constant). This rather qualitative statement can be proven mathematically, but in order to do that we need the following

Lemma 2.6 (Cauchy-Schwarz Inequality). *For any square integrable functions $z(x)$ and $w(x)$ defined on the interval $[a, b]$,*

$$\left| \int_a^b z(x)w(x)\,dx \right|^2 \leq \int_a^b |z(x)|^2\,dx \int_a^b |w(x|^2\,dx \tag{2.64}$$

and equality holds if and only if $z(x)$ is proportional to $w^*(x)$ (almost everywhere in $[a, b]$).

Proof. Assume $z(x)$ and $w(x)$ are real (the extension to complex-valued functions is straightforward). Let

$$\begin{aligned} I(y) &= \int_a^b [z(x) - y\,w(x)]^2\,dx \\ &= \underbrace{\int_a^b z^2(x)\,dx}_{A} - 2y \underbrace{\int_a^b z(x)w(x)\,dx}_{B} + y^2 \underbrace{\int_a^b w^2(x)\,dx}_{C} \\ &= A - 2By + Cy^2 \end{aligned}$$

Clearly, $I(y) \geq 0$ for all $y \in \mathbb{R}$. But if $I(y) = A - 2By + Cy^2 \geq 0$ for all $y \in \mathbb{R}$ then $B^2 - AC \leq 0$. If $B^2 - AC = 0$, then $I(y)$ has a real double root such that $I(k) = 0$ for $y = k$. Therefore (2.64) holds and if it is an equality, then $I(y)$ has a real root which implies

$$I(k) = \int_a^b [z(x) - k\,w(x)]^2\,dx = 0$$

But this can only occur if the integrand is identically zero; thus $z(x) = k\,w(x)$ for all x.

Definition 2.8 (Energy and distances for a signal). *Suppose $f(t)$ is a finite energy signal with Fourier transform $F(\omega)$. Let*

$$\begin{aligned} E &\equiv \int_{-\infty}^{+\infty} |f(t)|^2\,dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(\omega)|^2\,d\omega \\ d^2 &\equiv \frac{1}{E} \int_{-\infty}^{+\infty} t^2\,|f(t)|^2\,dt \\ D^2 &\equiv \frac{1}{2\pi E} \int_{-\infty}^{+\infty} \omega^2 |F(\omega)|^2\,d\omega \end{aligned}$$

Theorem 2.7 (Uncertainty Principle). *If $\sqrt{|t|}\,f(t) \to 0$ as $|t| \to \infty$, then*

$$D\,d \geq \frac{1}{2} \tag{2.65}$$

and equality holds if and only if $f(t)$ has the form $f(t) = K\,e^{-\alpha t^2}$.

Proof. Assume $f(t)$ is real. Lemma 2.6 implies

$$\left| \int_{-\infty}^{+\infty} t\, f\, \frac{df}{dt}\, dt \right|^2 \leq \int_{-\infty}^{+\infty} t^2\, f^2\, dt \int_{-\infty}^{+\infty} \left| \frac{df}{dt} \right|^2 dt \tag{2.66}$$

Let us define

$$\begin{aligned} A &\equiv \int_{-\infty}^{+\infty} t\, f\, \frac{df}{dt}\, dt \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} t\, \frac{df^2}{dt}\, dt \\ &= \underbrace{\frac{1}{2} t f^2 |_{-\infty}^{+\infty}}_{\alpha} - \underbrace{\frac{1}{2} \int_{-\infty}^{+\infty} f^2\, dt}_{\beta} \end{aligned}$$

when in the last passage we integrated by parts. By assumption $\sqrt{|t|}\, f \to 0 \to |t| f^2 \to 0 \Longrightarrow t f^2 \to 0$. Thus $\alpha = 0$. Furthermore $\beta = E/2$ and so

$$A = -\frac{E}{2} \tag{2.67}$$

Recalling (2.61), the Fourier transformation of derivatives, we have $df/dt \leftrightarrow i\omega F(\omega)$. From the Parseval Theorem (2.60) we obtain

$$\int_{-\infty}^{+\infty} \left| \frac{df}{dt} \right|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega^2 |F(\omega)|^2\, d\omega \tag{2.68}$$

Substituting (2.67) and (2.68) into (2.66) we obtain

$$\left| -\frac{E}{2} \right|^2 = \left| \int_{-\infty}^{+\infty} t\, f\, \frac{df}{dt}\, dt \right|^2 \leq \underbrace{\int_{-\infty}^{+\infty} t^2 f^2\, dt}_{Ed^2} \times \underbrace{\frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega^2 |F(\omega)|^2\, d\omega}_{ED^2} \tag{2.69}$$

That is

$$d\, D \geq \frac{1}{2} \tag{2.70}$$

If (2.70) is an equality, then (2.66) must be. This is possible only if (Lemma 2.6)

$$\frac{d}{dt} f(t) = k\, t\, f(t)$$

which means

$$f(t) = K\, e^{-\alpha t^2}$$

Remember that the Fourier transform of a Gaussian is a Gaussian.

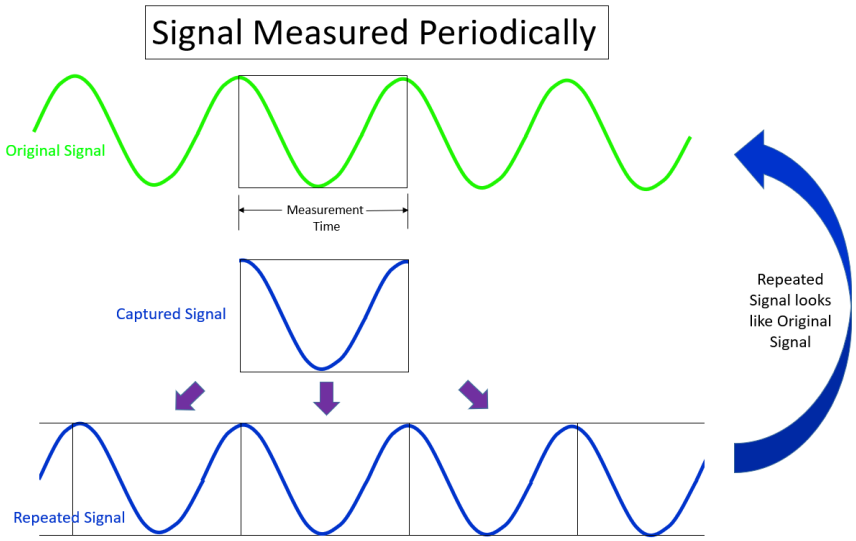
The *Uncertainty Relation* (2.70) states that one cannot jointly localize a signal in time and frequency arbitrarily well; either one has poor frequency localization or poor time localization. In signal analysis, this means that no window function can be chosen which is arbitrarily sharply concentrated both in time and frequency domain, and that the function that gives the best compromise about the "localization" is a Gaussian.
In quantum mechanics, the momentum and position wave functions are Fourier transform pairs (that is conjugate variables), to within a factor of Planck's constant. With this constant properly taken into account, the inequality (2.70) becomes the statement of the Heisenberg uncertainty principle.

2.3 Spectral Leakage

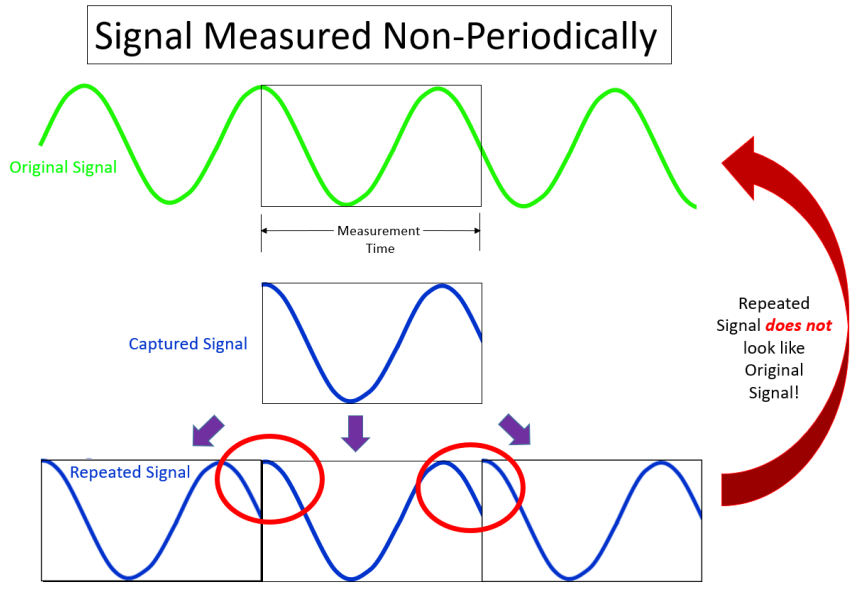
By necessity, every observed signal we process must be of finite extent. The extent may be adjustable and selectable, but it must be finite. If we observe the signal for T units of time, in order to apply our *continuous* Fourier transformations we perform a so-called *period extension* of our data, as shown in Figure 2.10a. It is evident that if the periodic extension of a signal does not commensurate with its natural period, then discontinuities at the boundaries will be present (see Figure 2.10b). These discontinuities will introduce spurious frequencies, responsible for spectral contributions over the entire set of frequencies. This effect is called **spectral leakage**.
In order to clarify why we are talking about *leakage*, let us take as an example a 3 Hz sine wave, as in the left panel of Figure 2.11. Suppose that the data are sampled at 1 Hz frequency (that is, every one second). In this case there is no problem to display the three Hertz sine wave in the frequency domain, because three Hertz is an integer multiple of the frequency resolution of 1 Hz (we are in the case shown in Figure 2.10a).
On the right panel of Figure 2.11 we show the case when we want to analyze a 2.5 Hz sine wave: it is not clear how to handle this situation because, due to the acquisition settings, displaying data at 2.5 Hz is not possible. So the 2.5 Hz signal will *leak* from zero Hertz to the full bandwidth, as shown in Figure 2.12. This might be surprising, as intuitively one might guess that the spectral leakage would be confined to the adjacent frequency lines (2 Hz and 3 Hz in this case). But remember: we are performing a *period extension* of our data, that is, we transform our data in a *continuous* periodic signal extended from $T = -\infty$ to $T = +\infty$, in order to apply continuous Fourier transformation. This means that we are spreading the not periodicty over the entire frequency range.
A signal with leakage (green in Figure 2.12) has lower amplitude and a broader frequency response than a signal with no leakage (red in Figure 2.12). This makes it difficult to quantify the signal properly in the frequency domain.

2.3.1 Window Functions

In order to reduce spectral leakage associated with finite observations intervals we apply to the data weighting functions, called *windows*. From one viewpoint, the window is applied to data (as a multiplicative weighting) to reduce the order of the discontinuity of the periodic extension



(a) The captured signal happened to be periodic, and the recreated signal matches the original.



(b) The captured signal is not periodic, causing discontinuities in the recreated signal.

Figure 2.10: Periodic extension of a sinusoidal signal periodic and not periodic in the observation interval.

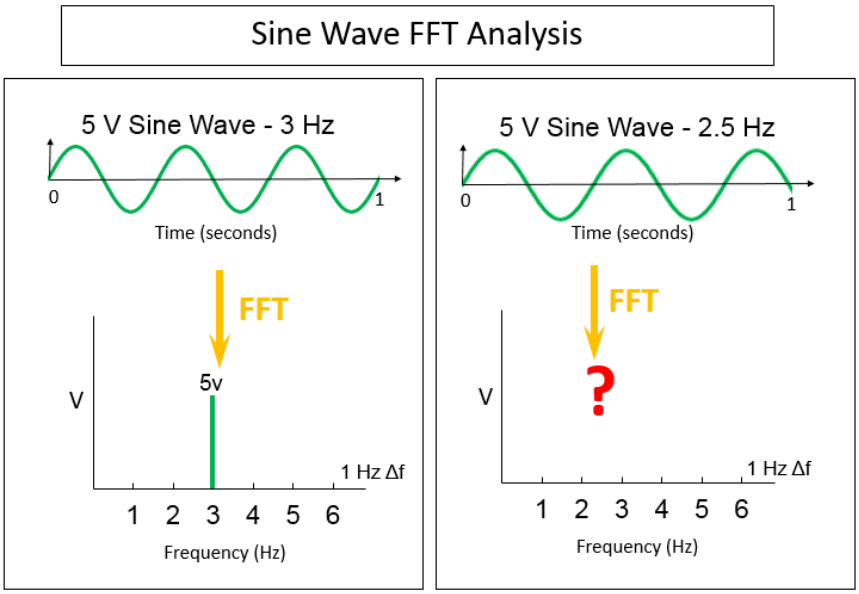


Figure 2.11: Left: A 3 Hz sine wave has the correct amplitude at a 1 Hz frequency resolution. Right: When the sine wave is not an integer multiple of the frequency resolution.

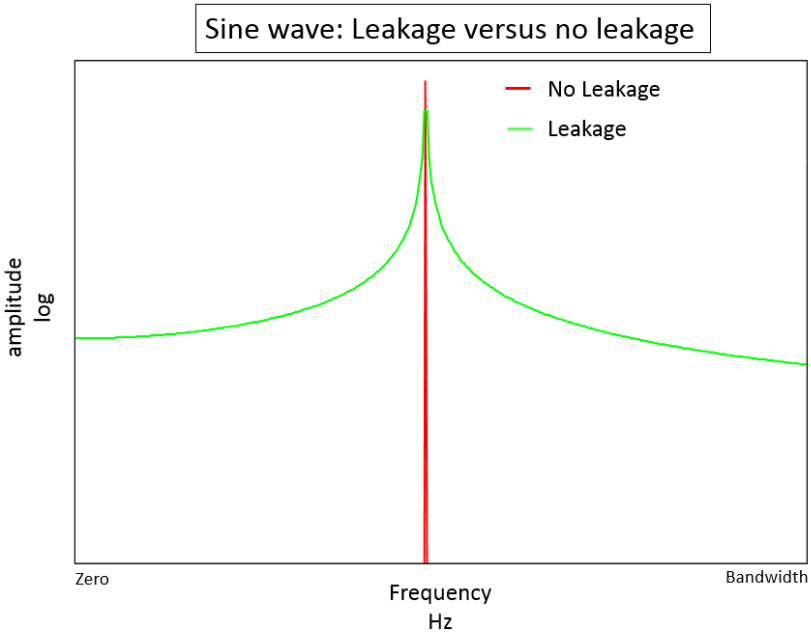


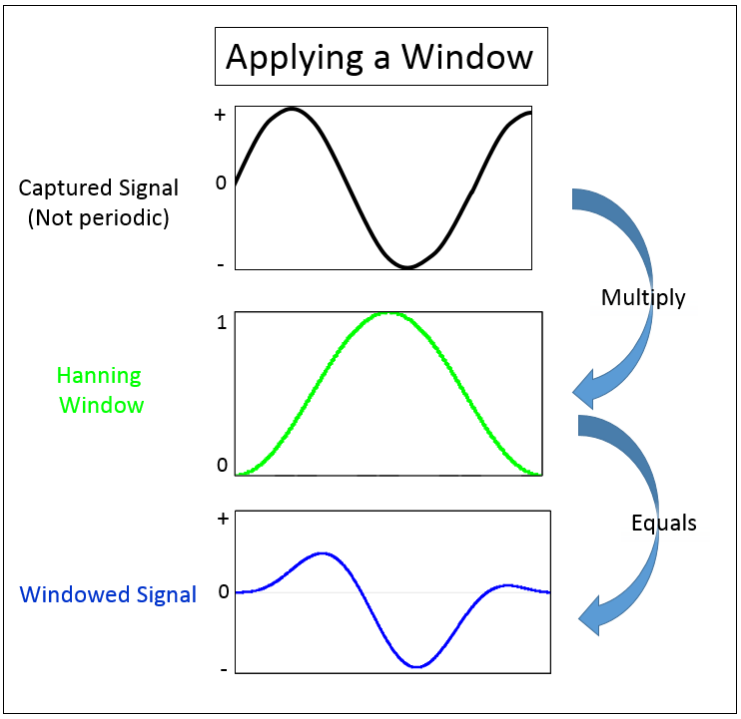
Figure 2.12: Frequency spectrum of sine wave aligning with frequency resolution (red) and sine wave not aligning with frequency resolution (green).

(Figure 2.13a). This is accomplished by matching as many orders of derivatives (of the weighted data) as possible at the boundary. The easiest way to achieve this matching is by setting the value of these derivatives to zero or near zero. Thus windowed data are smoothly brought to zero at the boundaries so that the periodic extension of the data is continuous in many orders of derivatives (Figure 2.13b).
From another viewpoint, the window is multiplicatively applied to the Fourier frequencies so that a signal of arbitrary frequency will exhibit a significant component for frequencies close to the Fourier frequencies. Of course both viewpoints lead to identical results.
In Figure 2.14 we show the effect of applying a window to our data in order to reduce the leakage. It could be argued that, by applying a window to our data, the signal is not perfectly replicated. But the main benefit is that the leakage is now confined over a smaller frequency range, instead of affecting the entire frequency bandwidth of the measurement.
All window functions are, of course, even functions. The Fourier transforms of the window function therefore don't have an imaginary part. We require a large dynamic range in order to better compare window qualities. That's why we'll use logarithmic representations covering equal ranges. And that's also the reason why we can't have negative function values. To make sure they don't occur, we'll use the power representation, i.e. $|F(\omega)|^2$.

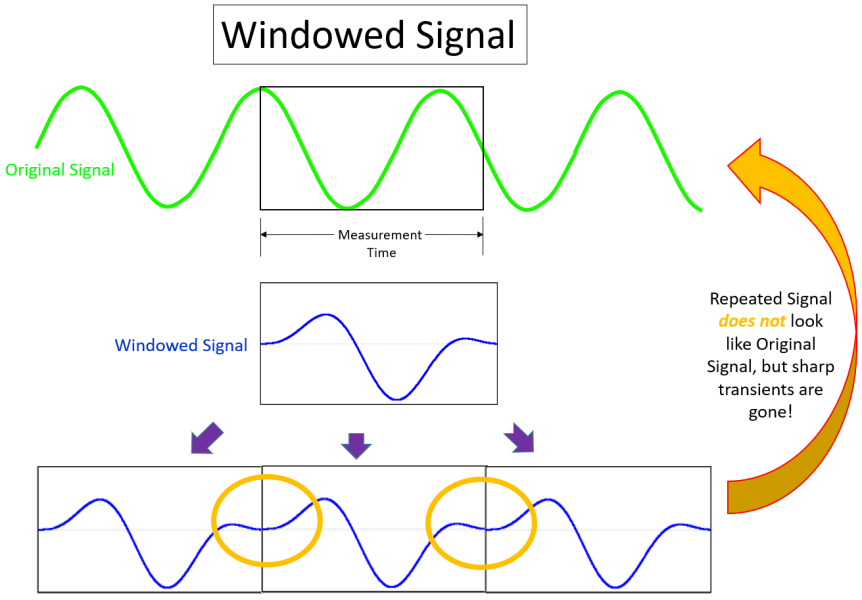
2.3.2 Types of Window Functions

There are several different types of window functions that we can apply depending on the signal. An actual plot of a window shows that the frequency characteristic of a window is a continuous spectrum with a main lobe and several side lobes. The main lobe is centered at each frequency component of the time-domain signal, and the side lobes approach zero. The height of the side lobes indicates the affect the windowing function has on frequencies around main lobes (see Figure 2.15). The side lobe response of a strong sinusoidal signal can overpower the main lobe response of a nearby weak sinusoidal signal. Typically, lower side lobes reduce leakage but increase the bandwidth of the major lobe. The side lobe roll-off rate is the asymptotic decay rate of the side lobe peaks. By increasing the side lobe roll-off rate, we can reduce spectral leakage.
Selecting a window function is not a simple task. Each window function has its own characteristics and suitability for different applications. To choose a window function, we must estimate the frequency content of the signal.

- ❏ If the signal contains strong interfering frequency components distant from the frequency of interest, choose a smoothing window with a high side lobe roll-off rate.
- ❏ If the signal contains strong interfering signals near the frequency of interest, choose a window function with a low maximum side lobe level.
- ❏ If the frequency of interest contains two or more signals very near to each other, spectral resolution is important. In this case, it is best to choose a smoothing window with a very narrow main lobe.



(a) Windows are designed to reduce the sharp transient in the recreated signal as much as possible. The captured signal is multiplied by the window.



(b) The windowed signal is period extended: the sharp transients are eliminated and smoothed out, even though the repeated signal does not match the original signal.

Figure 2.13: Effect of application of a window to a not periodic data in the observation interval.

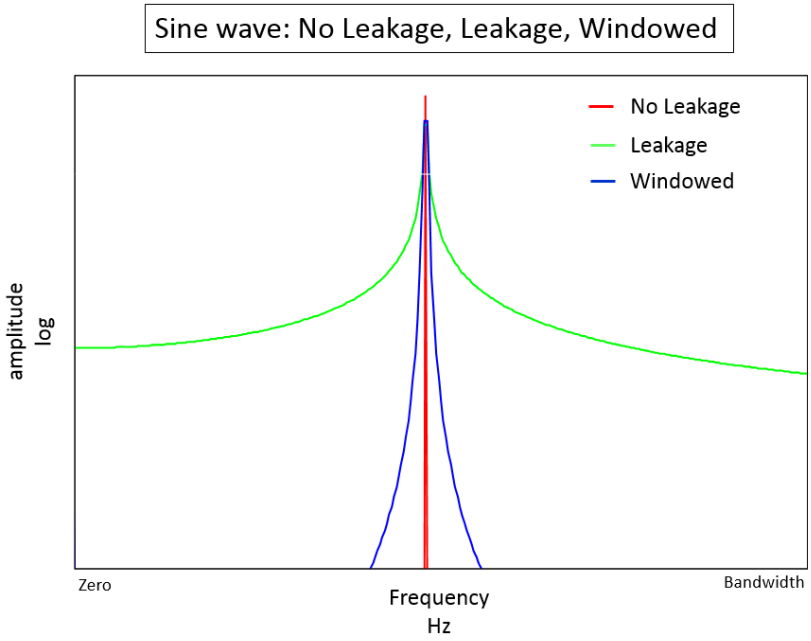


Figure 2.14: Periodic sine wave without leakage (red), non-periodic sine wave with leakage (green), and windowed non-periodic sine wave with reduced leakage (blue).

- ❏ If the amplitude accuracy of a single frequency component is more important than the exact location of the component in a given frequency bin, choose a window with a wide main lobe.

- ❏ If the signal spectrum is rather flat or broadband in frequency content, use the uniform window, or no window.

- ❏ In general, the Hanning (Hann) window is satisfactory in 95 percent of cases. It has good frequency resolution and reduced spectral leakage. If we do not know the nature of the signal but we want to apply a smoothing window, start with the Hann window.

Even if we use no window, the signal is convolved with a rectangular-shaped window of uniform height, by the nature of taking a snapshot in time of the input signal and working with a discrete signal. This convolution has a sine function characteristic spectrum. For this reason, no window is often called the uniform or rectangular window because there is still a windowing effect.
The Hamming and Hann window functions both have a sinusoidal shape. Both windows result in a wide peak but low side lobes. However, the Hann window touches zero at both ends eliminating all discontinuity. The Hamming window doesn't quite reach zero and thus still has a slight discontinuity in the signal. Because of this difference, the Hamming window does a better job of cancelling the nearest side lobe but a poorer job of canceling any others.
These window functions are useful for noise measurements where better frequency resolution than some of the other windows is wanted but moderate side lobes do not present a problem.

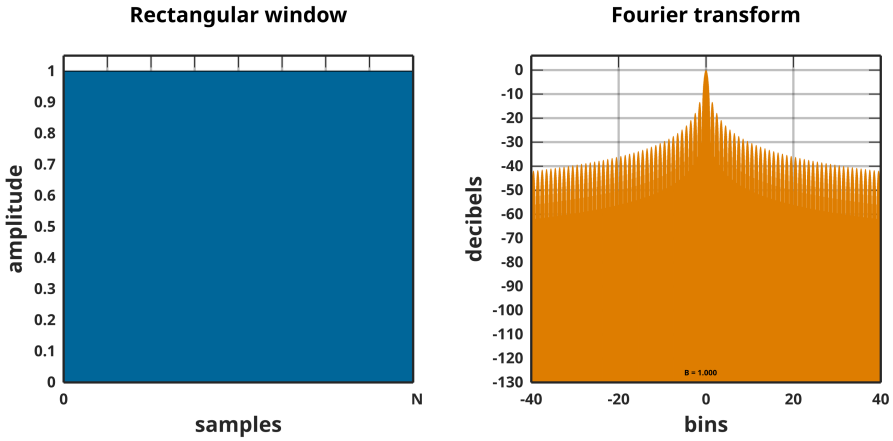


Figure 2.15: Rectangular window function and its Fourier transform in power representation.

2.3.3 The Rectangular Window

$$f(t) = \begin{cases} 1 & \text{for } -T/2 \leq t \leq T/2 \\ 0 & \text{else} \end{cases} \tag{2.71}$$

has the power representation of the Fourier transform (see (2.43)):

$$|F(\omega)|^2 = T^2 \left(\frac{\sin(\omega T/2)}{\omega T/2} \right)^2 \tag{2.72}$$

The rectangular window and this function are shown in Figure 2.15.

2.3.3.1 Zeroes

Where are the zeros of this function? We'll find them at $\omega T/2 = l\pi$ with $l = 1, 2, 3, \ldots$ and without the zero! The zeros are equidistant, the zero at $l = 0$ in the numerator gets plugged by a zero in the denominator.

2.3.3.2 Intensity at the Central Peak

Now we want to find out how much intensity is at the central peak, and how much gets lost in the sidebands (sidelobes). To get there, we need the first zero at $\omega T/2 = \pi$ or $\omega = \pm 2\pi/T$ and:

$$\int_{-2\pi/T}^{+2\pi/T} T^2 \left(\frac{\sin(\omega T/2)}{\omega T/2} \right)^2 d\omega \overset{\omega T/2 = x}{=} T^2 \frac{2}{T^2} 2 \int_0^{+2\pi} \frac{\sin^2 x}{x} dx = 4T \, \mathrm{Si}(2\pi) \tag{2.73}$$

where $\mathrm{Si}(x)$ is the sine integral, defined as

$$\mathrm{Si}(x) \equiv \int_0^x \frac{\sin y}{y} dy \tag{2.74}$$

The last passage in (2.73) may be proved as follows. We start out with

$$\int_0^\pi \frac{\sin^2 x}{x}\,dx$$

and integrate per parts with $u = \sin^2 x$ and $v = -1/x$:

$$\begin{aligned}\int_0^\pi \frac{\sin^2 x}{x}\,dx &= \left.\frac{\sin^2 x}{x}\right|_0^\pi + \int_0^\pi \frac{2\sin x\cos x}{x}\,dx \\ &= 2\int_0^\pi \frac{\sin 2x}{2x}\,dx \\ &\overset{2x=y}{=} \mathrm{Si}(2\pi)\end{aligned} \tag{2.75}$$

By means of the Parseval theorem (2.60) we get the total intensity

$$\int_{-\infty}^{+\infty} T^2 \left(\frac{\sin(\omega T/2)}{\omega T/2}\right)^2 d\omega = 2\pi \int_{-T/2}^{+T/2} 1^2\,dt = 2\pi T \tag{2.76}$$

The ratio of the intensity at the central peak to the total intensity is therefore

$$\frac{4T\,\mathrm{Si}(2\pi)}{2\pi T} = \frac{2}{\pi}\,\mathrm{Si}(2\pi) = 0.903$$

This means that $\approx 90\%$ of the intensity is in the central peak, whereas some 10% are "wasted" in the sidelobes.

2.3.3.3 Sidelobe Suppression

Now let's determine the height of the first sidelobe. To get there, we need:

$$\frac{d|F(\omega)|^2}{d\omega} = 0 \qquad \text{or also} \qquad \frac{dF(\omega)}{d\omega} = 0 \tag{2.77}$$

and this occurs when

$$\frac{d}{dx}\frac{\sin x}{x} = 0 \overset{x=\omega T/2}{=} \frac{x\cos x - \sin x}{x^2}$$

Solving this transcendental equation gives us the smallest possible solution $x = 4.4934$ or $\omega = 8.9868/T$. Inserting this value in $|F(\omega)|^2$ results in:

$$\left|F\left(\frac{8.9868}{T}\right)\right|^2 = T^2 \times 0.04719 \tag{2.78}$$

For $\omega = 0$ we get $|F(0)|^2 = T^2$, the ratio of the first sidelobe height to the central peak height is therefore 0.04719. It is customary to express ratios between two values spanning several order of magnitude in decibel (short dB). The definition of decibel is

$$\text{dB} \equiv 10\ \log_{10} x \tag{2.79}$$

Quite regularly people forget to mention what the ratio's based on, which can cause confusion. Here we're talking about intensity-ratios. If we're referring to amplitude-ratios, (that is, $F(\omega)$), this would make precisely a factor of two in logarithmic representation! Here we have a sidelobe suppression (first sidelobe) of:

$$10\ \log_{10} 0.04719 = -13.2\ \text{dB}$$

2.3.3.4 3 dB Bandwidth

As the $10\ \log_{10}(1/2) = -3.0103 \approx -3$, the 3 dB bandwidth tells us where the central peak has dropped to half its height. This is easily calculated as follows

$$T^2 \left(\frac{\sin(\omega T/2)}{\omega T/2} \right)^2 = \frac{1}{2} T^2$$

Using $x = \omega T/2$ we have

$$\sin^2 x = \frac{x}{2} \qquad \text{or } \sin x = \frac{x}{\sqrt{2}}$$

This transcendental equation has the following solution:

$$x = 1.3915, \qquad \text{thus } \omega_{3\,\text{dB}} = \frac{2.783}{T}$$

This gives the total width ($\pm\omega_{3\,\text{dB}}$):

$$\Delta\omega = \frac{5.566}{T} \tag{2.80}$$

This is the slimmest central peak we can get using Fourier transformation. Any other window function will lead to larger 3 dB-bandwidths. Admittedly, it's more than nasty to stick more than 10% of the information into the sidelobes. If we have, apart from the prominent spectral component, another spectral component, with – say – an approx. 10 dB smaller intensity, this component will be completely smothered by the main component's sidelobes. If we're lucky, it will sit on the first sidelobe and will be visible; if we're out of luck, it will fall into the gap (the zero) between central peak and first sidelobe and will get swallowed. So it pays to get rid of these sidelobes.

Warning! This 3 dB-bandwidth is valid for $|F(\omega)|^2$ and not for $F(\omega)$! Since one often uses $|F(\omega)|$ or the cosine-/sine-transformation one wants the 3 dB-bandwidth thereof, which corresponds to the 6 dB-bandwidth of $|F(\omega)|^2$. Unfortunately, we cannot simply multiply the 3 dB-bandwidth of $|F(\omega)|^2$ by $\sqrt{2}$, we have to solve a new transcendental equation. However, it's still good as a first guess because we merely interpolate linearly between the point of 3 dB-bandwidth and the point of the 6 dB-bandwidth. We'd overestimate the width by less than 5%.

2.3.4 The Triangular Window (Fejer Window)

The first real weighting function is the triangular window:

$$f(t) = \begin{cases} 1 + \dfrac{2t}{T} & \text{for } -T/2 \leq t \leq 0 \\ 1 - \dfrac{2t}{T} & \text{for } 0 \leq t \leq T/2 \\ 0 & \text{else} \end{cases} \tag{2.81}$$

Its Fourier transform is

$$F(\omega) = \frac{T}{2}\left(\frac{\sin(\omega T/4)}{\omega T/4}\right)^2 \tag{2.82}$$

The zeros are twice as far apart as in the case of the "rectangular function": $\omega T/4 = l\pi$ or $\omega = 4l\pi/T$ with $l = 1, 2, 3, \ldots$ The intensity at the central peak is 99.7%. The height of the first sidelobe is suppressed by $2 \times (-13.2\text{ dB}) \approx -26.5$ dB. The 3 dB-bandwidth is computed as follows:

$$\sin\frac{\omega T}{4} = \frac{1}{\sqrt[4]{2}}\frac{\omega T}{4} \rightarrow \Delta\omega = \frac{8.016}{T}\text{ full width}$$

that is some 1.44 times wider than in the case of the rectangular window. The asymptotic behavior of the sidelobes is -12 dB/octave.

2.3.5 The Gauss Window

A pretty obvious window function is the Gauss function.

$$f(t) = \begin{cases} \exp\left(-\dfrac{1}{2}\dfrac{t^2}{\sigma^2}\right) & \text{for } -T/2 \leq t \leq +T/2 \\ 0 & \text{else} \end{cases} \tag{2.83}$$

Its Fourier transform is

$$f(\omega) = \sigma\sqrt{\frac{\pi}{2}}\, e^{-\frac{\sigma^2\omega^2}{4}}\left[\text{erfc}\left(-i\frac{\sigma^2\omega^2}{\sqrt{2}} + \frac{T^2}{8\sigma^2}\right) + \text{erfc}\left(+i\frac{\sigma^2\omega^2}{\sqrt{2}} + \frac{T^2}{8\sigma^2}\right)\right] \tag{2.84}$$

As the error function occurs with complex arguments, though together with the conjugate complex argument, $F(\omega)$ is real. The function $f(t)$ with $\sigma = 2$ and $|F(\omega)|^2$ is shown in Figure 2.16.

2.4 Windowing or Convolution?

In principle, we have two possibilities to use window functions:

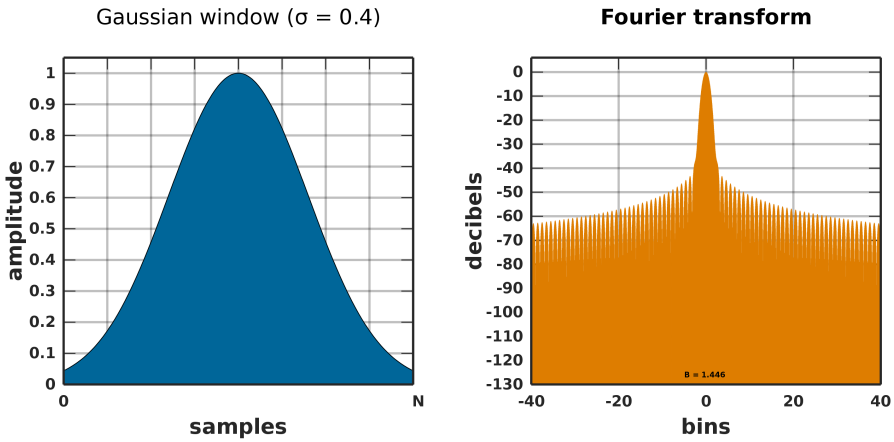


Figure 2.16: Gauss window and power representation of the Fourier transform

- ❏ Either we weight, i.e. we multiply, the input by the window function and subsequently Fourier-transform, or
- ❏ We Fourier-transform the input and convolute the result with the Fourier transform of the window function.

According to the Convolution Theorem (2.54) we get the same result. What are the pros and cons of both procedures? There is no easy answer to this question. What helps in arguing is thinking in discrete data. Take, e.g., a weighting window. Let's start with a reasonable value for the its parameter, based on considerations of the trade-off between 3 dB-bandwidth (i.e. resolution) and sidelobe suppression. In the case of windowing we have to multiply our input data, say N real or complex numbers, by the window function which we have to calculate at N points. After that we Fourier-transform. Should it turn out that we actually should require a better sidelobe suppression and could tolerate a worse resolution – or vice versa – we would have to go back to the original data, window them again and Fourier-transform again.
The situation is different for the case of convolution: we Fourier-transform without any bias concerning the eventually required sidelobe suppression and subsequently convolute the Fourier data (again N numbers, however in general complex!) with the Fourier-transformed window function, which we have to calculate for a sufficient number of points. What is a sufficient number? Of course, we drop the sidelobes for the convolution and only take the central peak! This should be calculated at least for five points, better more. The convolution then actually consists of five (or more) multiplications and a summation for each Fourier coefficient. This appears to be more work; however, it has the advantage that a further convolution with another, say broader Fourier-transformed window function, would not require to carry out a new Fourier transformation. Of course, this procedure is but an approximation because of the truncation of the sidelobes. If we included all data of the Fourier-transformed window function including the sidelobes, we had to carry out N (complex) multiplications and a summation per point, already quite a lot of computational effort, yet still less than a new Fourier transformation. This could

be relevant for large arrays, especially in two or three dimensions like in image processing and tomography.

Chapter 3

Temporal Analysis on Digital Data

As we have already mentioned, every observed signal we process must be of finite extent. Processing a finite-duration observation imposes interesting and interacting considerations on the harmonic analysis. Furthermore, for practicality the data we process are N uniformly (this condition can also be raised) spaced samples of the observed signal. For convenience, N is highly composite, and we will assume N is even. The harmonic estimates we obtain through the discrete Fourier transformation (DFT) are N uniformly spaced samples of the associated periodic spectra. This approach is elegant and attractive when the processing scheme is cast as a spectral decomposition in an N-dimensional orthogonal vector space. Unfortunately, in many practical situations, to obtain meaningful results this elegance must be compromised. One such compromise consists of applying windows to the sampled data set, or equivalently, smoothing the spectral samples.
The two operations to which we subject the data are *sampling* and *windowing*. These operations can be performed in either order. We will address the interacting considerations of window selection in harmonic analysis and examine the special considerations related to sampled windows for DFT.

3.1 Discrete Fourier Transformation

Often we do not know a function's continuous "behavior" over time, but only what happens at N discrete times:

$$t_k = k\,\Delta t, \qquad k = 0, 1, 2, \ldots, N-1$$

In other words: we've taken our "pick", that's "samples" $f(t_k) = f_k$ at certain points in time t_k. Any digital data-recording uses this technique. So the data set consists of a series $\{f_k\}$. Outside the sampled interval $T = N\Delta t$ we don't know anything about the function. The discrete Fourier transformation (DFT) automatically assumes that $\{f_k\}$ will continue periodically outside the interval's range. At first glance this limitation appears to be very annoying, maybe $f(t)$ isn't periodic at all, and even if $f(t)$ were periodic, there's a chance that our interval happens to

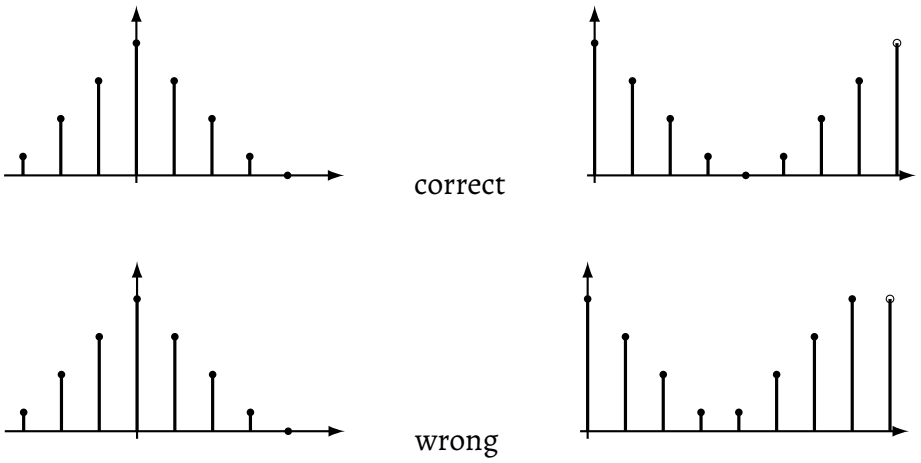


Figure 3.1: Correctly wrapped-around (top); incorrectly wrapped-around (bottom)

truncate at the wrong time (meaning: not after an integer number of periods. See Figure 2.10b). To make life easier, we'll also take for granted that N is a power of 2. We'll have to assume the latter anyway for the Fast Fourier Transformation (FFT) which we'll cover in Section 3.5.

3.1.1 Even and Odd Series and Wrap-around

A series is called even if the following is true for all f_k:

$$f_{-k} = f_k \tag{3.1}$$

A series is called odd if the following is true for all f_k:

$$f_{-k} = -f_k \tag{3.2}$$

Please note that f_0 is compulsory! Any series can be broken up into an even and an odd series. But what about negative indices? We'll extend the series periodically:

$$f_{-k} = f_{N-k} \tag{3.3}$$

This allows us, by adding N, to shift the negative indices to the right end of the interval, or using another word, "wrap them around", as shown in Figure 3.1. Please make sure f_0 doesn't get wrapped, something that often is done by mistake. The periodicity with period N, which we always assume as given for the discrete Fourier transformation, requires $f_N = f_0$. In the second example – the one with the mistake – we would get f_0 twice next to each other (and apart from that, we would have overwritten f_4).

3.1.2 The Kronecker Symbol or the Discrete δ-Function

Before we get into the definition of the discrete Fourier transformation (forward and inverse transformation), a few preliminary remarks are in order. From the continuous Fourier trans-

formation term $e^{i\omega t}$ we get the discrete times $t_k = k\,\Delta t$, $k = 0, 1, 2, \ldots, N-1$ with $T = N\Delta t$:

$$\exp(i\omega t) \to \exp\left(i\frac{2\pi t_k}{T}\right) = \exp\left(i\frac{2\pi k\,\Delta t}{N\Delta t}\right) = \exp\left(\frac{2\pi i\,k}{N}\right) \equiv W_N^k \tag{3.4}$$

We will use the abbreviation for the "kernel" W_N as

$$W_N = \exp\left(\frac{2\pi i}{N}\right) \tag{3.5}$$

Occasionally we will also use the discrete frequencies ω_j

$$\omega_j = \frac{2\pi j}{N\Delta t} \tag{3.6}$$

related to the discrete Fourier coefficients F_j (see below). The kernel W_N has the following properties:

$$\begin{gathered} W_N^{n\,N} = e^{2\pi i\,n} = 1 \qquad \text{for all integer } n \\ W_N \text{ is periodic in } j \text{ and } k \text{ with period } N \end{gathered} \tag{3.7}$$

We can define the discrete δ-function as follow:

$$\sum_{j=0}^{N-1} W_N^{(k-k')j} = N\,\delta_{k,k'} \tag{3.8}$$

where $\delta_{k,k'}$ is the Kronecker symbol with the following property:

$$\delta_{k,k'} = \begin{cases} 1 & \text{for } k = k' \\ 0 & \text{else} \end{cases} \tag{3.9}$$

This symbol (with prefactor N) accomplishes the same tasks the δ-function had when doing the continuous Fourier transformation.

3.1.3 Definition of the Discrete Fourier Transformation

Now we want to determine the spectral content $\{F_j\}$ of the series $\{f_k\}$ using discrete Fourier transformation. For this purpose, we have to make the transition in the definition of the Fourier series:

$$c_j = \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\,e^{-2\pi ij/T}\,dt \longrightarrow \frac{1}{N}\sum_{k=0}^{N-1} f_k\,e^{-2\pi ijk/N} \tag{3.10}$$

with $f(t)$ periodic of period T.

In the exponent we find $k\Delta t/N\Delta t$, meaning that Δt can be eliminated. The prefactor contains the sampling raster Δt, so the prefactor becomes $\Delta t/T = \Delta t/(N\Delta t) = 1/N$. During the transition (3.10) we tacitly shifted the limits of the interval from $-T/2$ to $+T/2$ to 0 to T, something that was okay, as we integrate over an *integer* period and $f(t)$ was assumed to be periodic of period T. The sum has to come to an end at $N-1$, as this sampling point plus Δt reaches the limit of the interval. Therefore we get, for the discrete Fourier transformations:

Definition 3.1 (Discrete Fourier transformations).

$$F_j = \frac{1}{N}\sum_{k=0}^{N-1} f_k\, W_N^{-kj} \qquad \text{with } W_N = e^{2\pi i/N} \tag{3.11a}$$

$$f_k = \sum_{j=0}^{N-1} F_j\, W_N^{+kj} \qquad \text{with } W_N = e^{2\pi i/N} \tag{3.11b}$$

Please note that the inverse Fourier transformation (3.11b) doesn't have a prefactor $1/N$.
A bit of a warning is called for here. Instead of (3.11) we also come across definition equations with positive exponents for the forward transformation and with negative exponent for the inverse transformation. For example Press *et al.* use this convention, and this is what is used in X–ray astronomy (see (5.1) at page 91). This doesn't matter as far as the real part of $\{F_j\}$ is concerned. The imaginary part of $\{F_j\}$, however, changes its sign. Because we want to be consistent with the previous definitions of Fourier series and the continuous Fourier transformation we'd rather stick with the definitions (3.11) and remember that, for example, a negative, purely imaginary Fourier coefficient F_j belongs to a positive amplitude of a sine wave (given positive frequencies), as i of the forward transformation multiplied by i of the inverse transformation results in precisely a change of sign $i^2 = -1$.
Often also the prefactor $1/N$ of the forward transformation is missing (again this is the case for Press *et al.*) This prefactor has to be there because F_0 is to be equal to the average of all samples. As we will see, also the Parseval theorem will be grateful if we stick with our definition of the forward transformation. Using (3.8) we can see straight away that the inverse transformation (3.11b) is correct:

$$\begin{aligned} f_k &= \sum_{j=0}^{N-1} F_j\, W_N^{+kj} = \sum_{j=0}^{N-1}\frac{1}{N}\sum_{k'=0}^{N-1} f_{k'} W_N^{-k'j} W_N^{+kj} \\ &= \frac{1}{N}\sum_{k'=0}^{N-1} f_{k'} \sum_{j=0}^{N-1} W_N^{(k-k')j} = \frac{1}{N}\sum_{k'=0}^{N-1} f_{k'} N\,\delta_{k,k'} = f_k \end{aligned} \tag{3.12}$$

Ex. 3.1 *Discrete Fourier Transformation: Constant function. Cosine function. Sine function.*

3.2 Theorems and Rules

3.2.1 Linearity Theorem

If we combine in a linear way $\{f_k\}$ and its series $\{F_j\}$ with $\{g_k\}$ and its series $\{G_j\}$, the we get:

$$\begin{aligned} \{f_k\} &\leftrightarrow \{F_j\} \\ \{g_k\} &\leftrightarrow \{G_j\} \\ a \times \{f_k\} + b \times \{g_k\} &\leftrightarrow a \times \{F_j\} + b \times \{G_j\} \end{aligned} \tag{3.13}$$

Please always keep in mind that the discrete Fourier transformation contains only linear operators, but that the power representation is *not* a linear operation.

3.2.2 Shifting Rules

$$\begin{aligned} \{f_k\} &\leftrightarrow \{F_j\} \\ \{f_{k-n}\} &\leftrightarrow \{F_j W_N^{-jn}\} \qquad n \text{ integer} \end{aligned} \tag{3.14}$$

A shift in the time domain by n results in a multiplication by a phase factor W_N^{-jn}. Let us proof it:

Proof.

$$\begin{aligned} F_j^{\text{shifted}} &= \frac{1}{N} \sum_{k=0}^{N-1} f_{k-n} W_N^{-kj} \\ &\overset{k-n=k'}{=} \frac{1}{N} \sum_{k'=-n}^{N-1-n} f_{k'} W_N^{-(k'+n)j} \\ &= \frac{1}{N} \sum_{k'=0}^{N-1} f_{k'} W_N^{-k'j} W_N^{-nj} \\ &= F_j^{\text{old}} W_N^{-nj} \end{aligned} \tag{3.15}$$

Because of the periodicity of f_k, we may shift the lower and the upper summation boundaries by n without a problem.

Ex. 3.2 *First Shifting Rule: Shifted cosine with $N = 2$*

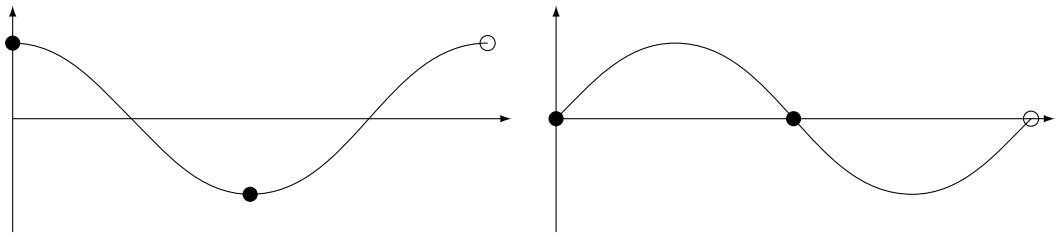


Figure 3.2: Two samples per period: cosine (left); sine (right)

$$\begin{aligned} \{f_k\} &\leftrightarrow \{F_j\} \\ \{f_k W_N^{-nk}\} &\leftrightarrow \{F_{j+n}\} \qquad n \text{ integer} \end{aligned} \tag{3.16}$$

A modulation in the time domain with W_N^{-nk} corresponds to a shift in the frequency domain. The proof is trivial.

Ex. 3.3 *Second Shifting Rule: Modulated cosine with $N = 2$*

3.2.3 Scaling Rule/Nyquist Frequency

We saw above that the highest frequency ω_{max} or also $-\omega_{\text{max}}$ corresponds to the center of the series of Fourier coefficients. This we get by inserting $j = N/2$ in definition of the discrete frequency (3.6):

$$\omega_{N/2} = \frac{2\pi}{N\Delta t}\frac{N}{2} = \frac{\pi}{\Delta t} \equiv \Omega_{\text{Nyq}} \tag{3.17}$$

This frequency is called *Nyquist frequency* or cut-off frequency. This corresponds to take *two samples* per period, as shown in Figure 3.2.
While we'll get away with this in the case of the cosine it definitely won't work for the sine! Here we grabbed the samples at the wrong moment, or maybe there was no signal after all. In fact, the imaginary part of f_k at the Nyquist frequency always is 0. The Nyquist frequency therefore is the highest possible spectral component for a cosine wave; for the sine it is only up to:

$$\omega = \frac{2\pi(N/2-1)}{N\Delta t} = \Omega_{\text{Nyq}}(1-2/N)$$

Equation (3.17) is our scaling theorem, as the choice of Δt allows us to stretch or compress the time axis, while keeping the number of samples N constant. This only has an impact on the

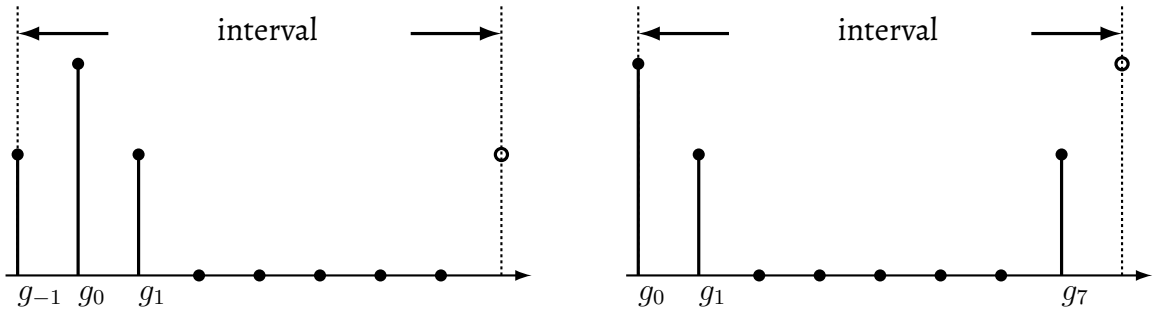


Figure 3.3: Resolution function $\{g_k\}$: without wrap-around (left); with wrap-around (right)

frequency scale running from $\omega = 0$ to $\omega = \Omega_{\text{Nyq}}$. Δt doesn't appear anywhere else! The normalization factor we came across in (2.25) and (2.46), is done away with here, as using discrete Fourier transformation we normalize to the number of samples N, regardless of the sampling raster Δt.

3.3 Convolution, Parseval Theorem

Before we're able to formulate the discrete versions of the (2.48), (2.56), (2.58), and (2.60), we have to get a handle on two problems:

- ❏ The number of samples N for the two functions $f(t)$ and $g(t)$ we want to convolute or cross-correlate, must be the same. This often is not the case, for example, if $f(t)$ is the "theoretical" signal we would get for a δ-shaped instrumental resolution function, which, however, has to be convoluted with the finite resolution function $g(t)$. There's a simple fix: we pad the series $\{g_k\}$ with zeros so we get N samples, just like in the case of series $\{f_k\}$.

- ❏ Don't forget, that $\{f_k\}$ is periodic in N and our "padded" $\{g_k\}$, too. This means that negative indices are wrapped-around to the right end of the interval. The resolution function $g(t)$ mentioned in Figure 3.3, which we assumed to be symmetrical, had three samples and got padded with five zeros to a total of $N = 8$ and is displayed in Figure 3.3.

As long as $\{f_k\}$ is periodic in N, there's nothing wrong with the fact upon convolution data from the end/beginning of the interval will be "mixed into" data from the beginning/end of the interval. If we don't like that – for whatever reasons – rather also pad $\{f_k\}$ with zeros, using precisely the correct number of zeros so $\{g_k\}$ won't create overlap between f_0 and f_{N-1} any more.

3.3.1 Convolution

We will define the discrete convolution as follows:

Definition 3.2 (Discrete convolution).

$$h_k \equiv (f \otimes g)_k = \frac{1}{N} \sum_{l=0}^{N-1} f_l \, g_{k-l} \tag{3.18}$$

The "convolution sum" is commutative, distributive and associative. The normalization factor $1/N$ in context: the convolution of $\{f_k\}$ with the "discrete δ-function" $\{g_k\} = N\delta_{k,0}$ is to leave the series $\{f_k\}$ unchanged. Following this rule, also a "normalized" resolution function $\{g_k\}$ should respect the condition $\sum_{k=0}^{N-1} = N$. Unfortunately often the convolution also gets defined without the prefactor $1/N$.
The Fourier transform of $\{h_k\}$ is

$$\begin{aligned} H_j &= \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{N} \sum_{l=0}^{N-1} f_l \, g_{k-l} W_N^{-kj} \\ &= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f_l W_N^{-lj} \, g_{k-l} W_N^{-kj} \, W_N^{+lj} \\ &\overset{k'=k-l}{=} \frac{1}{N^2} \sum_{l=0}^{N-1} f_l W_N^{-lj} \sum_{k'=-l}^{N-1-l} g_{k'} W_N^{-k'j} \\ &= F_j \, G_j \end{aligned} \tag{3.19}$$

In our last step we took advantage of the fact that, due to the periodicity in N, the second sum may also run from 0 to $N-1$ instead of $-l$ to $N-1-l$. This, however, makes sure that the current index l has been totally eliminated from the second sum, and we get the product of the Fourier transform F_j and G_j . So we arrive at the discrete Convolution Theorem:

Theorem 3.1 (Discrete convolution theorem). *Let be*

$$\begin{aligned} \{f_k\} &\leftrightarrow \{F_j\} \\ \{g_k\} &\leftrightarrow \{G_j\} \end{aligned}$$

Then

$$\{h_k\} = \{(f_{\otimes} g)_k\} \leftrightarrow \{H_j\} = \{F_j \times G_j\} \tag{3.20}$$

The convolution of the series $\{f_k\}$ and $\{g_k\}$ results in a product in the Fourier space.

Theorem 3.2 (Inverse discrete convolution theorem). *Let be*

$$\begin{aligned} \{f_k\} &\leftrightarrow \{F_j\} \\ \{g_k\} &\leftrightarrow \{G_j\} \end{aligned}$$

Then

$$\{h_k\} = \{f_k\} \times \{g_k\} \leftrightarrow \{H_j\} = \{N(F \otimes G)_j\} \tag{3.21}$$

Proof.

$$\begin{aligned} H_j &= \frac{1}{N}\sum_{k=0}^{N-1} f_k g_k W_N^{-kj} = \frac{1}{N}\sum_{k=0}^{N-1} f_k g_k \underbrace{\sum_{k'=0}^{N-1} W_N^{-k'j}\,\delta_{k,k'}}_{k'\text{-sum "artificially" introduced}} \\ &= \frac{1}{N^2}\sum_{k=0}^{N-1} f_k \sum_{k'=0}^{N-1} g_{k'} W_N^{-k'j} \underbrace{\sum_{l=0}^{N-1} W_N^{-l(k-k')}}_{l-\text{sum yields } N\delta_{k,k'}} \\ &= \sum_{l=0}^{N-1} \frac{1}{N}\sum_{k=0}^{N-1} f_k W_N^{-lk} \frac{1}{N}\sum_{k'=0}^{N-1} g_{k'} W_N^{-k'(j-l)} \\ &= \sum_{l=0}^{N-1} F_l\, G_{j-l} = N(F \otimes G)_j \end{aligned}$$

Ex. 3.4 *Discrete Fourier Transform: Nyquist frequency with $N = 8$*

3.3.2 Cross Correlation

We define for the discrete cross correlation between $\{f_k\}$ and $\{g_k\}$, similar to what we did in (2.56):

Definition 3.3 (Discrete cross correlation).

$$h_k \equiv (f \star g)_k = \frac{1}{N}\sum_{l=0}^{N-1} f_l \times g^*_{l+k} \tag{3.22}$$

If the indices at g_k go beyond $N-1$, then we'll simply subtract N (periodicity). The cross correlation between $\{f_k\}$ and $\{g_k\}$, of course, results in a product of their Fourier transforms:

$$\begin{gathered} \{f_k\} \leftrightarrow \{F_j\} \\ \{g_k\} \leftrightarrow \{G_j\} \\ \{h_k\} = \{(f_\star g)_k\} \leftrightarrow \{H_j\} = \{F_j \times G_j^*\} \end{gathered} \tag{3.23}$$

Proof.

$$\begin{aligned} H_j &= \frac{1}{N}\sum_{k=0}^{N-1}\frac{1}{N}\sum_{l=0}^{N-1} f_l\, g^*_{l+k} W_N^{-kj} \\ &= \frac{1}{N}\sum_{l=0}^{N-1} f_l \frac{1}{N}\sum_{k=0}^{N-1} g^*_{l+k} W_N^{-kj} \\ &= \frac{1}{N}\sum_{l=0}^{N-1} f_l\, G^*_j\, W_N^{-jl} = F_j\, G^*_j \end{aligned}$$

3.3.3 Autocorrelation

Here we have $\{g_k\} = \{f_k\}$, which leads to

$$h_k \equiv (f \star f)_k = \frac{1}{N}\sum_{l=0}^{N-1} f_l \times f^*_{l+k} \tag{3.24}$$

and

$$\begin{gathered} \{f_k\} \leftrightarrow \{F_j\} \\ \{h_k\} = \{(f_\star f)_k\} \leftrightarrow \{H_j\} = \{F_j \times F_j^*\} = \{|F_j|^2\} \end{gathered} \tag{3.25}$$

In other words: the Fourier transform of the autocorrelation of $\{f_k\}$ is the modulus squared of the Fourier series $\{F_j\}$ or its power representation.

3.3.4 Parseval Theorem

We use (3.24) for $k = 0$, that is h_0, and get on the one side:

$$h_0 = \frac{1}{N}\sum_{l=0}^{N-1} |f_l|^2 \tag{3.26}$$

On the other hand, the inverse transformation of $\{H_j\}$, especially for $k = 0$, results in (see (3.11b))

$$h_0 = \sum_{j=0}^{N-1} |F_j|^2 \tag{3.27}$$

Put together, this gives us the discrete version of Parseval theorem:

$$\frac{1}{N}\sum_{l=0}^{N-1} |f_l|^2 = \sum_{j=0}^{N-1} |F_j|^2 \tag{3.28}$$

3.4 The Sampling Theorem

When discussing the Nyquist frequency, we already mentioned that we need at least two samples per period to show cosine oscillations at the Nyquist frequency. Now we'll turn the tables and claim that as a matter of principle we won't be looking at anything but functions $f(t)$ that are "bandwidth-limited", meaning that outside the interval $[-\Omega_{\text{Nyq}}, \Omega_{\text{Nyq}}]$ their Fourier transforms $F(\omega)$ are 0. In other words: we'll refine our sampling to a degree where we just manage to capture all the spectral components of $f(t)$. Now we'll skillfully use formulas we've learned when dealing with the Fourier series expansion and the continuous Fourier transformation with each other, and then pull the sampling theorem out of the hat. For this purpose we will recall (2.16) and (2.17) which show that a periodic function $f(t)$ can be expanded into an (infinite) Fourier series:

$$\begin{aligned} f(t) &= \sum_{k=-\infty}^{\infty} C_k e^{i\omega_k t} \qquad \omega_k = \frac{2\pi k}{T} \\ C_k &= \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\, e^{-i\omega_k t}\, dt \qquad \text{for } k = 0, \pm 1, \pm 2, \ldots \end{aligned}$$

Since $F(\omega)$ is 0 outside $[-\Omega_{\text{Nyq}}, \Omega_{\text{Nyq}}]$, we can continue this function periodically and expand it into an infinite Fourier series. So we replace: $f(t) \to F(\omega), t \to \omega, T/2 \to \Omega_{\text{Nyq}}$, and get

$$\begin{aligned} F(\omega) &= \sum_{k=-\infty}^{\infty} C_k e^{i\pi k\omega/\Omega_{\text{Nyq}}} \\ C_k &= \frac{1}{2\Omega_{\text{Nyq}}}\int_{-\Omega_{\text{Nyq}}}^{+\Omega_{\text{Nyq}}} F(\omega)\, e^{-i\pi k\omega/\Omega_{\text{Nyq}}}\, d\omega \end{aligned} \tag{3.29}$$

A similar integral also occurs in the defining equation for the inverse continuous Fourier transformation (see (2.35)):

$$f(t) = \frac{1}{2\pi}\int_{-\Omega_{\text{Nyq}}}^{+\Omega_{\text{Nyq}}} F(\omega)\, e^{i\omega t}\, d\omega \tag{3.30}$$

The integrations boundaries are $\pm\Omega_{\text{Nyq}}$, as $F(\omega)$ is band-limited. By comparing (3.30) with (3.29) we have

$$\begin{aligned} \int_{-\Omega_{\text{Nyq}}}^{+\Omega_{\text{Nyq}}} F(\omega) e^{-i\pi k\omega/\Omega_{\text{Nyq}}}\, d\omega &= 2\, C_k\, \Omega_{\text{Nyq}} \\ \int_{-\Omega_{\text{Nyq}}}^{+\Omega_{\text{Nyq}}} F(\omega)\, e^{i\omega t}\, d\omega &= 2\pi\, f(t) \end{aligned}$$

and the two integrals are the same if

$$-i\pi k\omega/\Omega_{\text{Nyq}} = i\omega t \rightarrow t = -\frac{\pi\, k}{\Omega_{\text{Nyq}}}$$

therefore

$$2\Omega_{\text{Nyq}} C_k = 2\pi f(-\pi k/\Omega_{\text{Nyq}}) \tag{3.31}$$

Once we have inserted this in (3.29) we get:

$$F(\omega) = \frac{\pi}{\Omega_{\text{Nyq}}} \sum_{k=-\infty}^{+\infty} f(-\pi k/\Omega_{\text{Nyq}})\, e^{i\pi k\omega/\Omega_{\text{Nyq}}} \tag{3.32}$$

When we finally insert this expression in the defining equation (3.30), we get:

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\Omega_{\text{Nyq}}}^{+\Omega_{\text{Nyq}}} \frac{\pi}{\Omega_{\text{Nyq}}} \sum_{k=-\infty}^{+\infty} f(-\pi k/\Omega_{\text{Nyq}})\, e^{i\pi k\omega/\Omega_{\text{Nyq}}}\, e^{i\omega t}\, d\omega \\ &= \frac{1}{2\Omega_{\text{Nyq}}} \sum_{k=-\infty}^{+\infty} f(-k\,\Delta t)\, 2 \int_0^{+\Omega_{\text{Nyq}}} \cos\omega(t + k\,\Delta t)\, d\omega \\ &= \frac{1}{\Omega_{\text{Nyq}}} \sum_{k=-\infty}^{+\infty} f(-k\,\Delta t)\, \frac{\sin\Omega_{\text{Nyq}}(t + k\,\Delta t)}{(t + k\,\Delta t)} \end{aligned} \tag{3.33}$$

where we have defined $\Delta t = \pi/\Omega_{\text{Nyq}}$. By replacing $k \rightarrow -k$ (it's not important in which order the sums are calculated) we get the Sampling Theorem:

Theorem 3.3 (Sampling theorem). *If a function $f(t)$ contains no frequencies higher that W cps (that is is "bandwidth-limited"), then it is completely determined by giving its ordinates at a series of points spaced $1/2W$ apart. By choosing $W = \Omega_{Nyq}$ we obtain*

$$f(t) = \sum_{k=-\infty}^{+\infty} f(k\,\Delta t)\, \frac{\sin\Omega_{Nyq}(t - k\,\Delta t)}{\Omega_{Nyq}(t - k\,\Delta t)} \tag{3.34}$$

In other words, we can reconstruct the function $f(t)$ for all times t from the samples at the times $k\,\Delta t$, provided the function $f(t)$ is bandwidth-limited. To achieve this, we only need to multiply $f(k\Delta t)$ with the function $(\sin x/x)$ (with $x = \Omega_{\text{Nyq}}(t - k\,\Delta t)$) and sum up over all samples. The factor $(\sin x/x)$ naturally is equal to 1 for $t = k\,\Delta t$, for other times, $(\sin x/x)$ decays and slowly oscillates towards zero, which means, that $f(t)$ is a composite of plenty of $(\sin x/x)$-functions at the location $t = k\,\Delta t$ with the amplitude $f(k\Delta t)$. Note that for adequate sampling with $\Delta t = \pi/\Omega_{\text{Nyq}}$, each k-term in the sum in (3.34) contributes $f(k\,\Delta t)$ at the sampling points $t = k\,\Delta t$ and zero at all other sampling points whereas all terms contribute to the interpolation between sampling points.

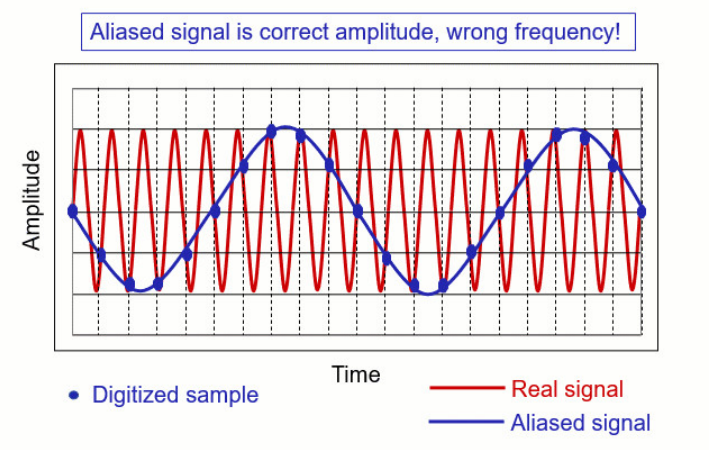


Figure 3.4: Aliasing is caused when the digital sampling rate is not adequate to capture the fluctuations in an analog signal, and results in the wrong frequency being identified. The red sine wave is the original signal. The blue dots represent how often the signal is being sampled. The blue line is how the signal will appear at the wrong frequency due to the low sampling rate.

Ex. 3.5 *Sampling theorem with $N = 2$*

3.4.1 Aliasing

What happens if, for some reason or other, our sampling happens to be too coarse and $F(\omega)$ above Ω_{Nyq} was unequal to 0? In Figure 3.4 we show a case, in which the original signal (in red) is sampled at a frequency such that only one data point is tracked per period (the blue dots). The blue line represents the reconstructed signal: the amplitude is correct but the frequency is not. In order to correctly reconstruct not only the amplitude but also the frequency content of the signal, we need *at least* two points per period. And the reason is evident from Figure 3.5: on the left we show the sampling of a sine wave when only one point is sampled per period, $f_s = f_{\text{sine}}$. In this case the amplitude of the reconstructed signal is zero. On the other hand, when we have at least two points sampled per period ($f_s = 2f_{\text{sine}}$), we are able to extract the true frequency of the sine wave.

Therefore, to properly sample all the desired frequency content of an incoming signal, one must sample at (or above) the Nyquist rate. In data acquisition, the sampling frequency is twice as high as the specified bandwidth. So, all frequency content below the specified bandwidth will be sampled at a rate sufficient to accurately capture the frequency content.

When the incoming signal contains frequency content above the specified bandwidth, the sampling frequency (2 times the bandwidth) will violate the Nyquist theorem (Eq. 3.34) for this higher frequency content.

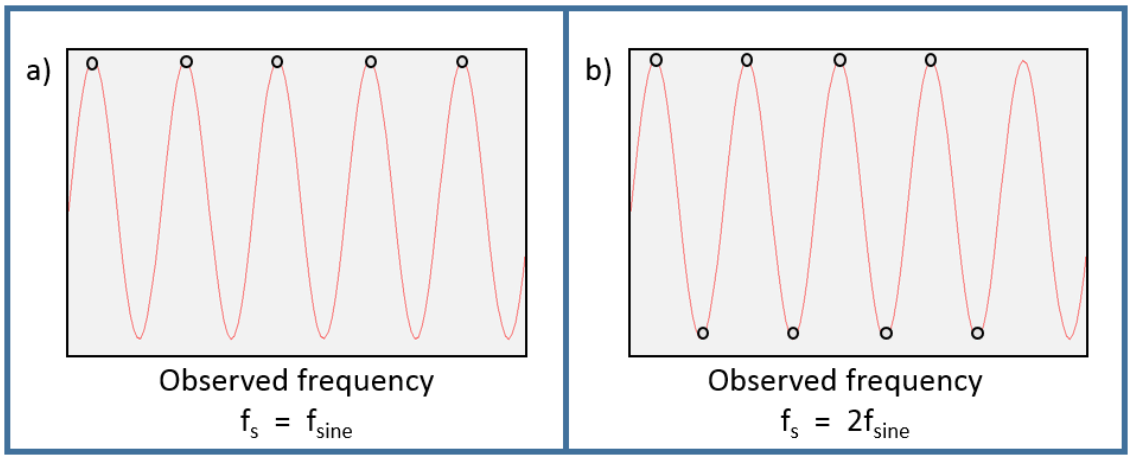


Figure 3.5: f_s represents the sampling frequency, f_{sine} represents the frequency of the sine wave. a) When sampling at the same frequency as the incoming signal, the observed frequency is zero Hertz. b) When sampling at twice the frequency of the sine wave, the observed frequency is f_{sine}, the true frequency of the sine wave.

When the Nyquist theorem is violated, spectral content above the bandwidth is mirrored about the bandwidth frequency in an accordion-plated fashion. Specifically, for any frequency ω in the range $0 \leq \omega \leq \Omega_{\text{Nyq}}$, the higher frequencies which are aliased with ω are defined by

$$(2\Omega_{\text{Nyq}} \pm \omega), (4\Omega_{\text{Nyq}} \pm \omega), \ldots, (2k\Omega_{\text{Nyq}} \pm \omega), \ldots \tag{3.35}$$

To prove this, observe that, for $t = (2\pi)/(2\Omega_{\text{Nyq}}) = \pi/\Omega_{\text{Nyq}}$

$$\cos \omega t = \cos\left[\left(2k\Omega_{\text{Nyq}} \pm \omega\right) \frac{\pi}{\Omega_{\text{Nyq}}}\right] = \cos\left(2k\,\pi \pm \pi\frac{\omega}{\Omega_{\text{Nyq}}}\right) = \cos\left(\pi\frac{\omega}{\Omega_{\text{Nyq}}}\right) \tag{3.36}$$

Thus all data at frequencies $2k\Omega_{\text{Nyq}} \pm \omega$ have the same cosine function as data at frequency ω when sampled at points π/Ω_{Nyq} apart. The same for the sine function. For example, if $\Omega_{\text{Nyq}} = 100$ cps, then data at 30 cps would be aliased with data at frequencies 170 cps, 230 cps, 370 cps, 430 cps, and so forth. Similarly, the power at these confounding frequencies is aliased with the power in the lower frequencies, because the power quantities depends on terms containing $\sin^2$ and $\cos^2$ functions.

Thus, higher frequency content appears to be at a lower frequency, or an “alias" frequency, as shown in Figure 3.6. In other words: spectral density that would appear at $\approx 2\Omega_{\text{Nyq}}$, appears at $\omega \approx 0$! This “corruption” of the spectral density through insufficient sampling is called **aliasing**, similar to someone acting under an assumed name. In a nutshell: when sampling, rather err on the fine side than the coarse one! Coarser rasters can always be achieved later on by compressing data sets, but it will never work the other way, round!

Let us see how the choice of the frequency resolution affects the signal processing. Because the “golden equation” of digital processing $\Delta f = 1/T$, we have

- The *finer* the desired frequency resolution, the longer the acquisition time;

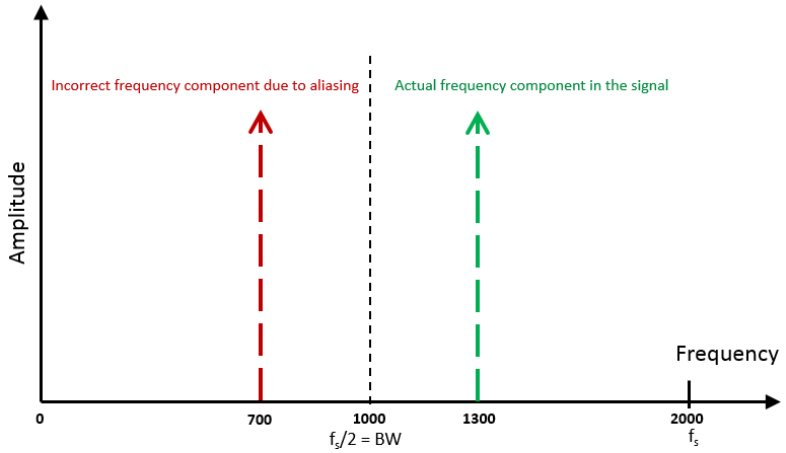


Figure 3.6: Aliasing causes frequency above the bandwidth to be mirrored across the bandwidth.

- The shorter the acquisition time, or frame size, the *coarser* the frequency resolution.

In Figure 3.7 two sine tones (100 Hertz and 101 Hertz) have been digitized, and a Fourier Transform performed. This was done with two different frequency resolutions: 1.0 Hertz and 0.5 Hertz.
With the finer frequency resolution of 0.5 Hertz, rather than 1.0 Hertz, the spectrum shows two separate and distinct peaks. The benefit of a finer frequency resolution is very obvious. This might beg the question, why not use the finest frequency resolution possible in all cases?
There is a tradeoff. Per the "golden equation" the amount of time data per frame is higher as the frequency resolution is made finer. This can cause requirements for long time data acquisition:
- 10 Hz frequency resolution is desired, only 0.1 seconds of data is required
- 1 Hertz frequency resolution requires 1 second of data
- 0.1 Hertz frequency resolution requires 10 seconds of data
- 0.01 Hertz frequency resolution requires 100 seconds of data!

In some situations, these long time acquisition requirements are not practical. For example, a sports car may go from idle to full speed in just 4 seconds, making a 100 second acquisition, and the corresponding 0.01 frequency resolution, impossible.
Rather than using the sine formulation of the Fourier Transform, a wavelet formulation can be used instead (see Section 3.6 for a brief introduction on the wavelet transform). This can address some of the time-frequency tradeoffs.
As we will see into detail in Part II, two practical methods exist for handling this aliasing problem. The first method is to choose the Nyquist frequency sufficiently large so that it is physically unreasonable for data to exist above Ω_{Nyq}. In general, it is a good rule to select Ω_{Nyq} to be one-and-a-half or two times greater that the maximum anticipated frequency. The choosing Ω_{Nyq} equal to the maximum frequency of interest will give accurate results for frequencies below Ω_{Nyq}.
The second method is to filter the data *prior* to sampling so that the information above a maximum frequency of interest is no longer contained in the filtered data.
An anti-aliasing filter is a low-pass filter that removes spectral content that violates the Nyquist

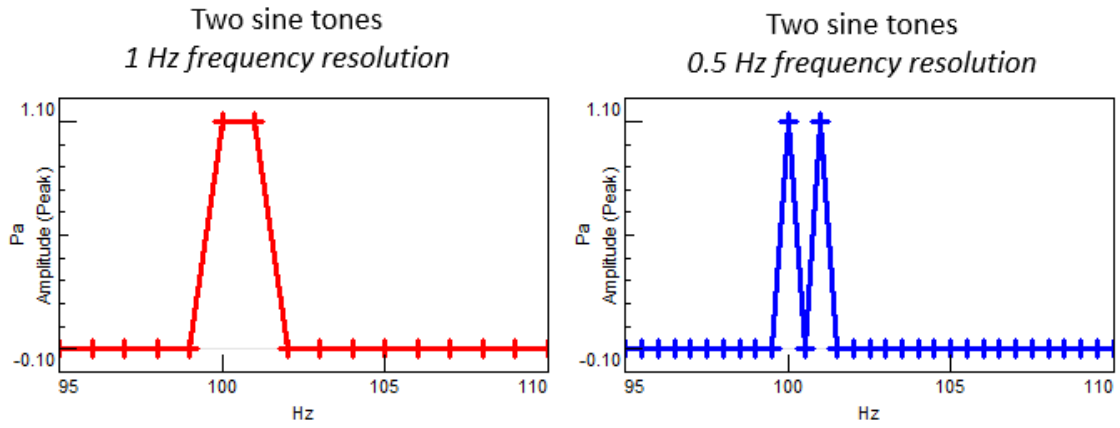


Figure 3.7: Left: Spectrum with 1.0 Hertz frequency resolution makes two separate tones appear as one peak. Right: Spectrum with 0.5 Hertz frequency resolution makes two separate tones appear as two different peaks.

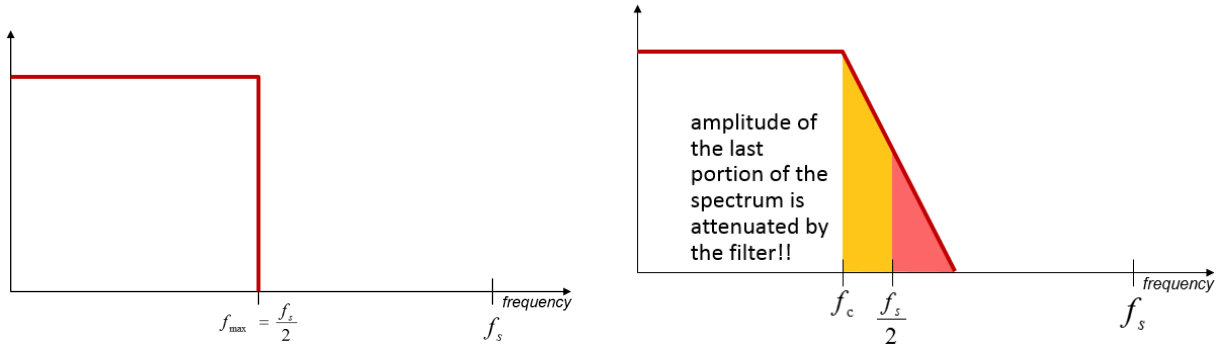


Figure 3.8: Right: The ideal anti-aliasing filter would be shaped like a wall: cutting off all frequencies beyond the specified bandwidth ($f_s/2$). Right: The anti-aliasing filter has a -3dB roll off point at the bandwidth.

criteria. This makes it so a 125 Hertz sine wave does not show up as 75 Hertz. The ideal anti-aliasing filter would be shaped like a "brick wall", completely attenuating all signals beyond the specified bandwidth, as shown in the left panel of Figure 3.8.

In the real world, it is impossible to have this "wall shaped" filter. Instead, a very sharp analog filter is used that has a -3dB roll off at the bandwidth and attenuates all frequencies 20% beyond the bandwidth to zero as shown in in the right panel of Figure 3.8. This is why the "trustable", alias-free region of the spectrum is from zero Hz to 80% of the bandwidth. This alias-free range is called the frequency span.

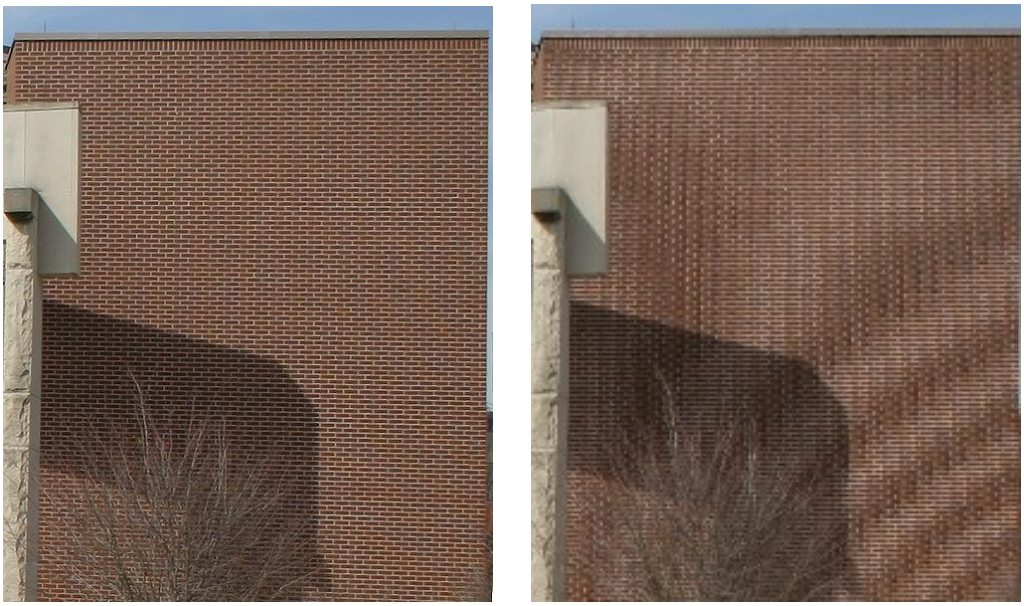


Figure 3.9: Aliasing in images: On the left a properly sampled image of a brick wall. On the right a spatial aliasing creates a Moire pattern

3.4.2 Application to Multi-Variable Signals and Images

The sampling theorem is usually formulated for functions of a single variable. Consequently, the theorem is directly applicable to time-dependent signals and is normally formulated in that context. However, the sampling theorem can be extended in a straightforward way to functions of arbitrarily many variables. Gray-scale images, for example, are often represented as two-dimensional arrays (or matrices) of real numbers representing the relative intensities of pixels (picture elements) located at the intersections of row and column sample locations. As a result, images require two independent variables, or indices, to specify each pixel uniquely – one for the row, and one for the column.
Similar to one-dimensional discrete-time signals, images can also suffer from aliasing if the sampling resolution, or pixel density, is inadequate. For example, a digital photograph of a striped shirt with high frequencies (in other words, the distance between the stripes is small), can cause aliasing of the shirt when it is sampled by the camera's image sensor. The aliasing appears as a *Moire pattern* (see Figure 3.9). The "solution" to higher sampling in the spatial domain for this case would be to move closer to the shirt, use a higher resolution sensor, or to optically blur the image before acquiring it with the sensor.

3.4.3 Geometrical Representation of the Signal

Let us discuss on the sampling theorem in a different way. The $2TW$ evenly spaced samples of a signal $f(t)$ can be thought of as coordinates of a point in a space of $2TW$ dimensions. Each particular selection of these numbers corresponds to a particular point in this space. Thus there

is exactly one point corresponding to each signal in the band and with duration T.
The number of dimensions $2TW$ will be, in general, very high, and needless to say, such a space cannot be visualized. It is possible, however, to study analytically the properties of an n-dimensional space. To a considerable extent, these properties are a simple generalization of the properties of two- and three-dimensional space, and can often be arrived at by inductive reasoning from these cases.
The advantage of this geometrical representation of the signals is that we can use the vocabulary and the results of geometry. If we imagine the $2TW$ coordinate axes to be at right angles to each other, then distances in the space have a simple interpretation. The distance from the origin to a point is analogous to the two- and three-dimensional cases

$$d = \sqrt{\sum_{k=1}^{2TW} x_k^2} \tag{3.37}$$

where x_k is the kth sample. From the sampling theorem (3.34) we have

$$f(t) = \sum_{k=1}^{2TW} x_k \frac{\sin \pi(2Wt - k)}{\pi(2Wt - k)} \tag{3.38}$$

therefore

$$\int_{-\infty}^{+\infty} f(t)^2 \, dt = \frac{1}{2W} \sum_{k=1}^{2TW} x_k^2 \tag{3.39}$$

where we used the property

$$\int_{-\infty}^{+\infty} \frac{\sin \pi(2Wt - k)}{\pi(2Wt - k)} \frac{\sin \pi(2Wt - l)}{\pi(2Wt - l)} \, dt = \begin{cases} 0 & k \neq l \\ \dfrac{1}{2W} & k = l \end{cases}$$

Hence the square of the distance to a point is $2W$ times the energy of the corresponding signal

$$d^2 = 2W \, E = 2W \, (TP) \tag{3.40}$$

where P is the average power over the time T. If we consider only signals whose average power is less than P, these will correspond to points within a sphere of radius $r = \sqrt{2W \, TP}$. If noise is added to the signal in transmission, it means that the point corresponding to the signal has been moved a certain distance in the space proportional to the rms value of the noise. Thus noise produces a small region of uncertainty about each point in the space.

3.5 Fast Fourier Transform

Cooley and Tukey started out from the simple question: what is the Fourier transform of a series of numbers with only **one** real number $(N = 1)$? There are at least 3 answers:

- From (3.11a) with $N = 1$ follows:

$$F_0 = \frac{1}{1} f_0 W_1^{-0} = f_0 \tag{3.41}$$

- From the Parseval theorem (2.60) follows:

$$|F_0|^2 = \frac{1}{1} |f_0|^2 \tag{3.42}$$

  Because f_0 is real and even, this leads to $F_0 = \pm f_0$. Furthermore, F_0 is also to be equal to the average of the series of numbers, so there is no chance to get the minus sign.

- We know that the Fourier transform of a δ-function results in a constant and vice versa. How do we represent a constant in the world of 1-term series? By using the number f_0. How do we represent in this world a δ-function? By using this number f_0. So in this world there's no difference any more between a constant and a δ-function. Result: f_0 is its own Fourier transform.

This finding, together with the trick to achieve $N = 1$ by smartly halving the input again and again (that's why we have to stipulate: $N = 2^p$, p integer), (almost) saves us the Fourier transformation. For this purpose, let's first have a look at the first subdivision. We'll assume as given: $\{f_k\}$ with $N = 2^p$. This series will get cut up in a way that one sub-series will only contain the even elements and the other sub-series only the odd elements of $\{f_k\}$:

$$\begin{aligned} \{f_{1,k}\} &= \{f_{2k}\} \qquad k = 0, 1, 2, \ldots, M-1 \\ \{f_{2,k}\} &= \{f_{2k+1}\} \qquad M = N/2 \end{aligned} \tag{3.43}$$

Proof.

$$f_{1,k+M} = f_{2k+2M} = f_{2k} = f_{1,k} \tag{3.44}$$

because of $2M = N$ and f periodic in N. Analogously for $f_{2,k}$.

The respective Fourier transforms are:

$$\begin{aligned} F_{1,j} &= \frac{1}{M} \sum_{k=0}^{M-1} f_{1,k} W_M^{-kj} \\ F_{2,j} &= \frac{1}{M} \sum_{k=0}^{M-1} f_{2,k} W_M^{-kj} \end{aligned} \tag{3.45}$$

The Fourier transform of the original series is:

$$
\begin{aligned}
F_j &= \frac{1}{N}\sum_{k=0}^{N-1} f_k\, W_N^{-kj} \\
&= \frac{1}{N}\sum_{k=0}^{M-1} f_{2k}\, W_N^{-2kj} + \frac{1}{N}\sum_{k=0}^{M-1} f_{2k+1}\, W_N^{-(2k+1)j} \\
&= \frac{1}{N}\sum_{k=0}^{M-1} f_{1,k}\, W_M^{-kj} + \frac{W_N^{-j}}{N}\sum_{k=0}^{M-1} f_{2,k}\, W_M^{-kj} \qquad j = 0, 1, 2, \ldots, N-1
\end{aligned} \tag{3.46}
$$

In our last step we used:

$$
\begin{aligned}
W_N^{-2kj} &= e^{-2\times 2\pi ikj/N} = e^{-2\pi ikj/(N/2)} = W_M^{-kj} \\
W_N^{-(2k+1)j} &= e^{-2\pi i(2k+1)j/N} = W_M^{-kj} W_N^{-j}
\end{aligned}
$$

Together we get (remember that $N = 2M$):

$$
F_j = \frac{1}{2}F_{1,j} + \frac{1}{2}W_N^{-j}F_{2,j} \qquad j = 0, 1, 2, \ldots, N-1
$$

or better

$$
\begin{aligned}
F_j &= \frac{1}{2}(F_{1,j} + W_N^{-j}F_{2,j}) \\
F_{j+M} &= \frac{1}{2}(F_{1,j} - W_N^{-j}F_{2,j})
\end{aligned} \tag{3.47}
$$

Please note that in (3.47) we allowed j to run from 0 to $M-1$ only. In the second line in front of $F_{2,j}$ there really should be the factor:

$$
\begin{aligned}
W_N^{-(j+M)} = W_N^{-j}W_N^{-M} &= W_N^{-j}W_N^{-N/2} = W_N^{-j}e^{-2\pi i\frac{N}{2}/N} \\
&= W_N^{-j}e^{-i\pi} = -W_N^{-j}
\end{aligned} \tag{3.48}
$$

This "decimation in time" can be repeated until we finally end up with 1-term series whose Fourier transforms are identical to the input number, as we know. The normal Fourier transformation requires N^2 calculations, whereas here we only need $pN = N\ln N$.

Ex. 3.6 *FFT: Saw-tooth with $N = 2$ and $N = 4$*

3.6 The Wavelet Transform

As we have discussed in Section 3.5, the traditional FFT decomposes a time signal into its component sine functions of various frequency, amplitude, and phase.
From there, a spectrum (a plot of amplitude vs frequency) is generated. It is possible to calculate many spectra for a single signal to see how the amplitudes and frequencies change with time.
When calculating a spectrum, there is an inverse relationship between observation time and frequency resolution (the so called "golden equation" discussed in Section 3.4.1, page 66). Essentially, the longer the time-chunk that is analyzed, the finer the frequency resolution that can be obtained in the spectrum.
This means that when analyzing very short-time events (transients) with a short time window, the frequency resolution is forced to be rather coarse. If the frequency resolution is refined, the time block will be much greater than the transient event.
Therefore, when doing an FFT on short time duration events, there is a fair bit of smearing in the frequency domain due to coarse frequency resolution. When attempting to dial in a finer frequency resolution, the time domain resolution will suffer (Figure 3.10). This is a disadvantage when using the FFT on short-time events.
On the other hand, the signal can be decomposed into wavelets (instead of sine functions). A wavelet is a function that rapidly increases, oscillates about a zero mean, and rapidly decays (see Figure 3.11).
In order to understand how wavelets correspond to frequency and time, let's take a look at *scaling* and *shifting*.
Scaling is straight forward, the wavelet is simply stretched or compressed in time.
A stretched wavelet (Figure 3.11, left) helps quantify the slow changing portion of a signal (low frequency) while a compressed wavelet (Figure 3.11, right) helps quantify the abruptly changing (high frequency) content of the signal.
The wavelet in the frequency domain has a band-pass characteristic. By stretching and compressing the wavelet, the center frequency of the band-pass filter is shifted higher or lower.
The output of the wavelet analysis is frequency (scale) vs time (shift). The wavelet is both shifted and scaled to determine how it aligns with various features of the signal.
For the purposes of simplicity, imagine that the wavelet is shifted though the time data. At each location, the shape of the wavelet is "compared" to the shape of the time data. Similarities between the time data and wavelet indicate that the frequency content that the wavelet represents is present (Figure 3.13).
Wavelets of different scales and shifts are convolved with the original signal to determine if the original signal has similar frequency content.

- Each wavelet has a corresponding "frequency", and the result of the convolution will determine if the original signal at that particular shift (time) also contains that same frequency.
- Therefore, it can be determined what frequency content is present at what time via wavelet analysis.

Essentially, wavelets can be thought of as a discrete-time filter-bank of band-pass filters.
So, what are the main differences between FFT and Wavelet? By nature of the processing type, the traditional FFT has a fixed relationship between time and frequency. Conversely, the wavelet

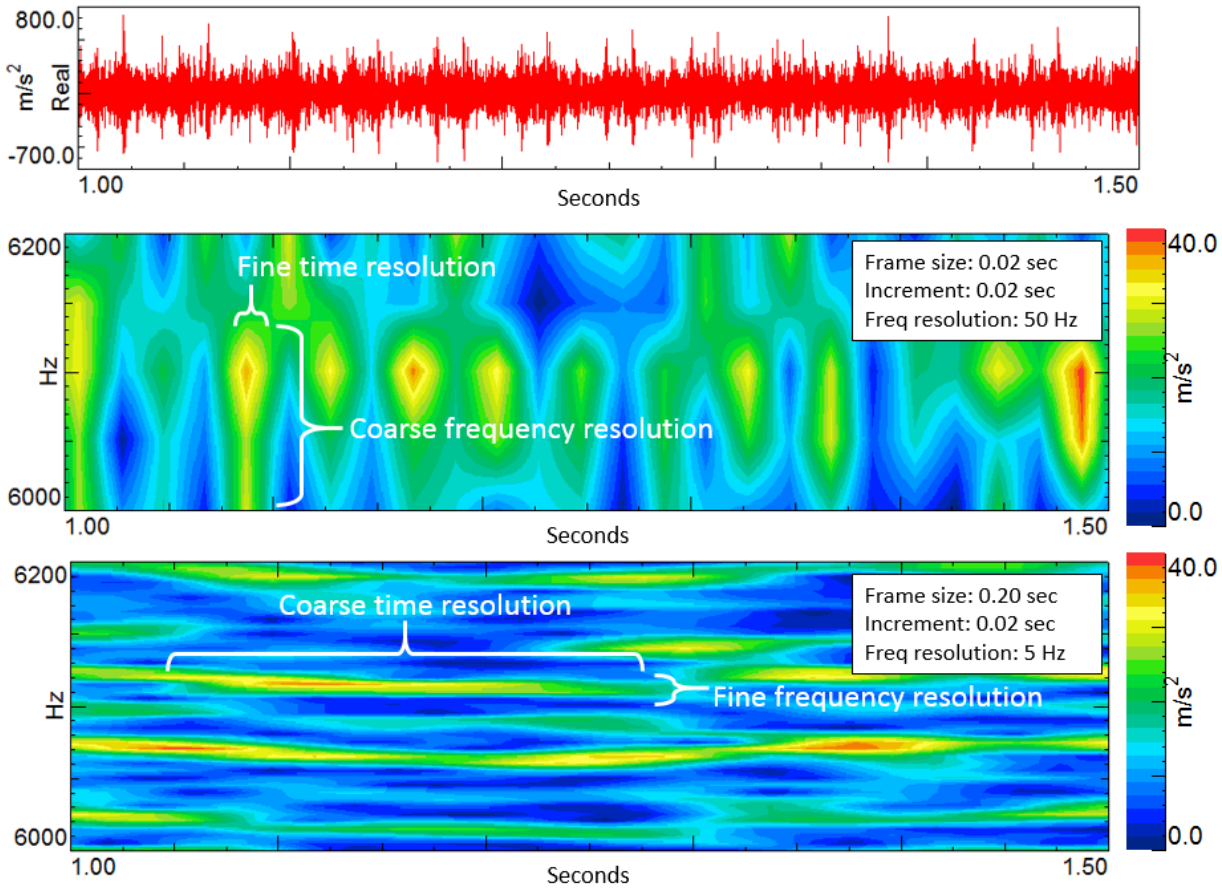


Figure 3.10: Top: Time signal of transient event. Middle: FFT with a 0.02 second frame size resulting in 50Hz frequency resolution. Fine time resolution, coarse frequency resolution. Bottom: FFT with a 0.20 second frame size resulting in 5Hz frequency resolution. Finer frequency resolution, coarser time resolution.

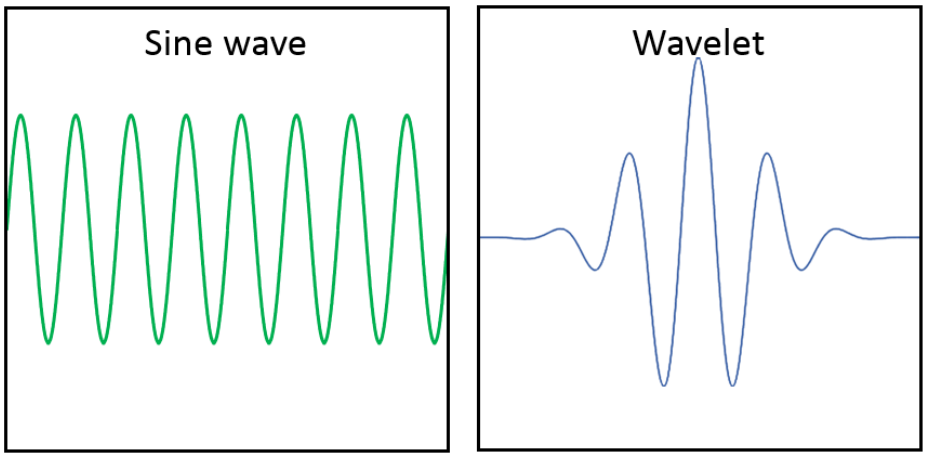


Figure 3.11: Comparison of a sine wave vs a wavelet function. A sine wave oscillates in time from negative infinity to infinity. Contrarily, a wavelet oscillates for a short time duration.

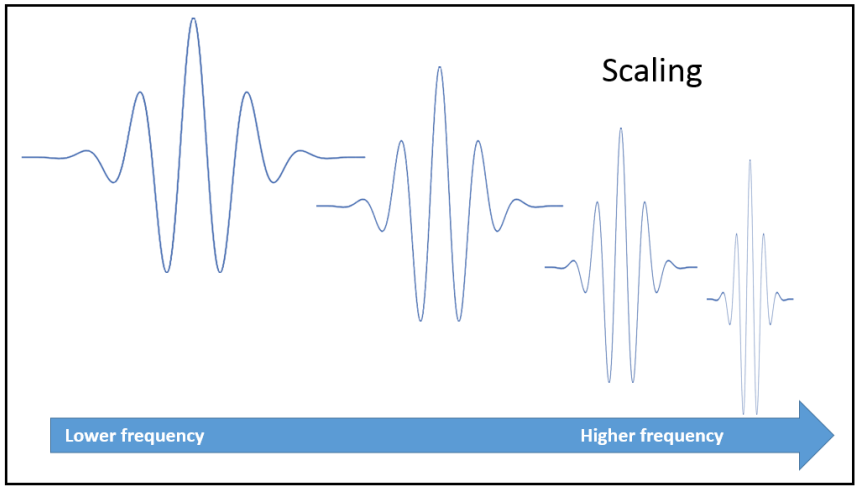


Figure 3.12: Stretched wavelets (left) represent lower frequencies, while compressed wavelets (right) represent higher frequencies.

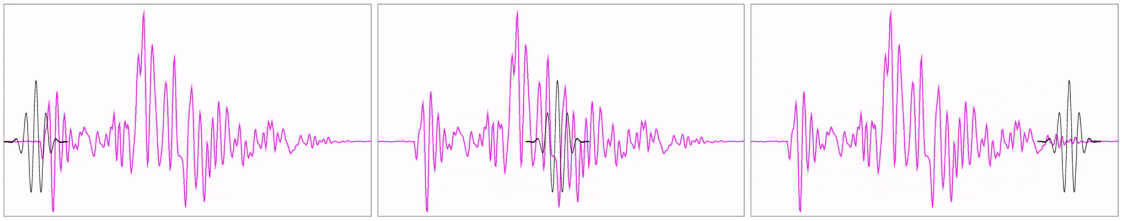


Figure 3.13: Shifting of a wavelet though time data: from left to right the wavelet is shifted in time relative to the signal being analyzed.

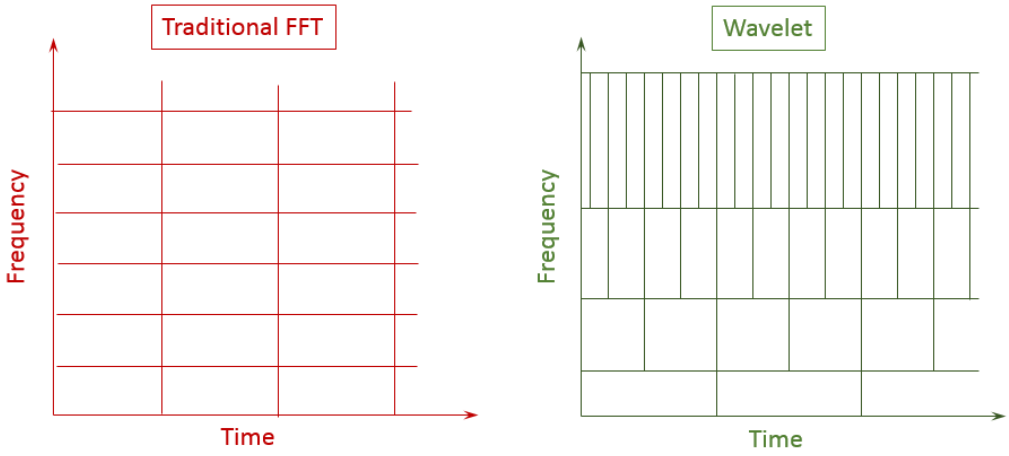


Figure 3.14: Change of frequency resolution in FFT and wavelet analysis. Left: The FFT smear and resolution is equal at all frequencies. Right: Conversely, wavelets have a variable relationship between time and frequency.

does not have a fixed relationship between time and frequency.
As shown in Figure 3.14, wavelets have different behavior at different frequencies:

- ❏ At lower frequencies, the data will be finer in the frequency domain, and more smeared in the time domain. At lower frequencies, octave bands are narrower, resulting in less smearing. At high frequencies, octave bands are broader resulting in more smearing.
- ❏ At higher frequencies, the data will be finer resolution in the time domain, and more smeared in the frequency domain The change in time resolution is due to the stretching of the wavelets at low frequency and the shrinking of wavelets at high frequency.

The change in the frequency resolution is due to the fact that the frequency scale for the wavelet processing is based on octaves.
Figure 3.15 shows the wavelet result of a transient event. Look at the right side of Figure 3.14 to better understand how data is smeared in a wavelet map as shown in Figure 3.15.
An impulse in the time domain is represented by broadband frequency response in the frequency domain. Most transient events are “impulsive" in nature and therefore have a rather broadband signature in the frequency domain. Therefore, a fine frequency resolution at high frequencies is typically not necessary. However, the improved time resolution the wavelet has to offer can be hugely beneficial when analyzing transients.

3.6.1 Fourier Transform vs Wavelet

Let’s take a closer look at the results to better understand the differences between FFT and wavelet analysis. The colormap results for both the FFT and wavelet transformations are created by stacking a series of tracked results together.

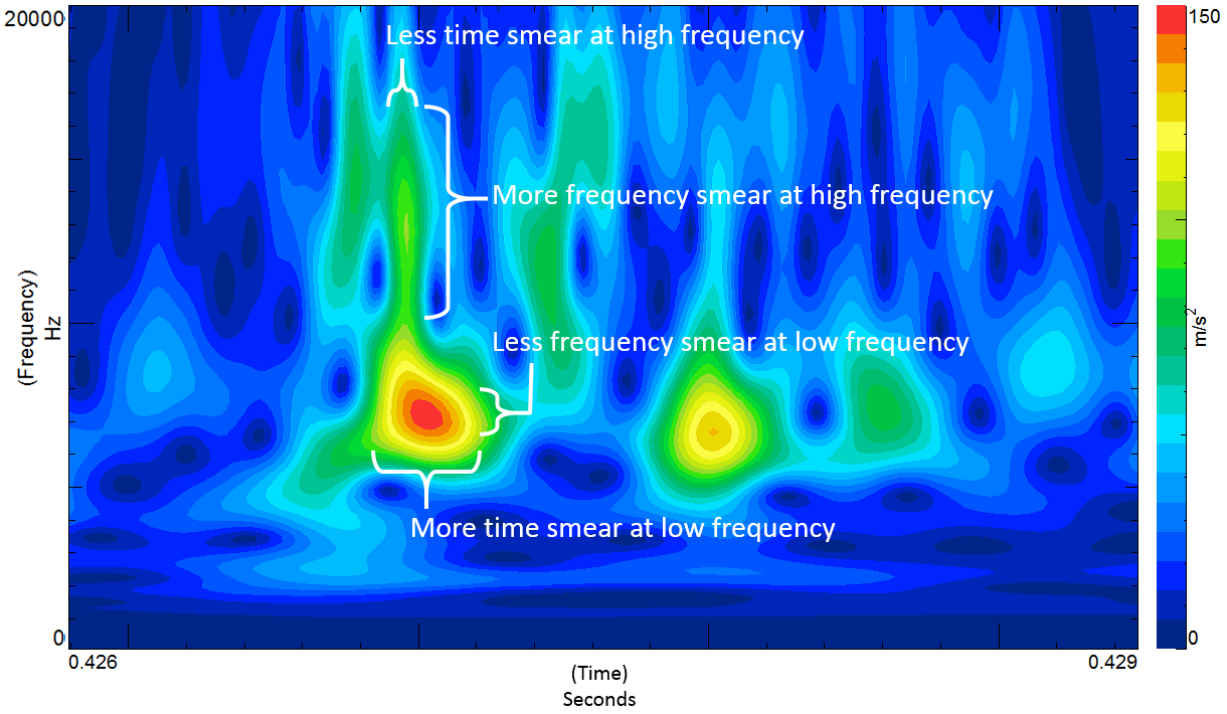
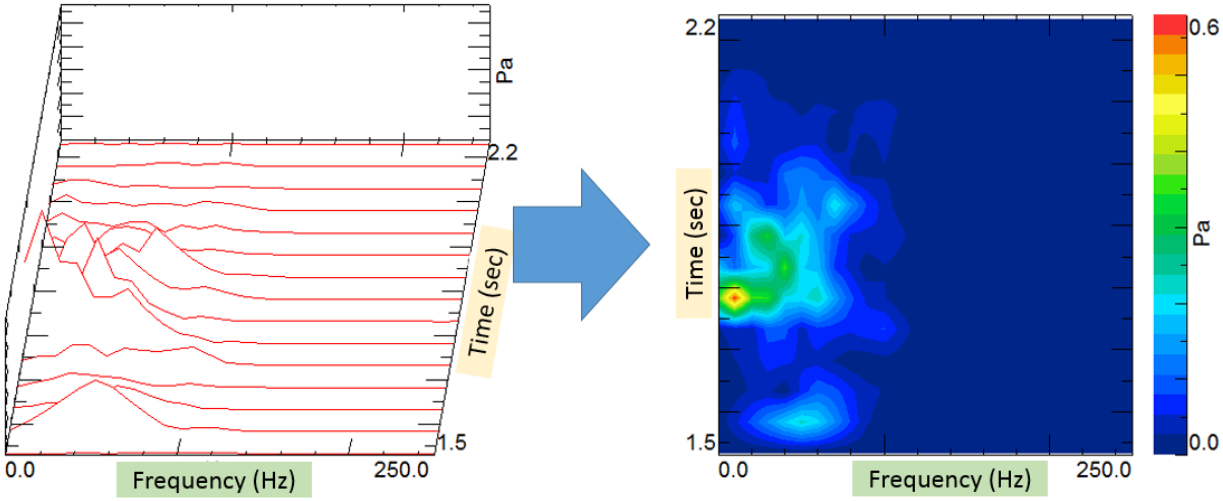
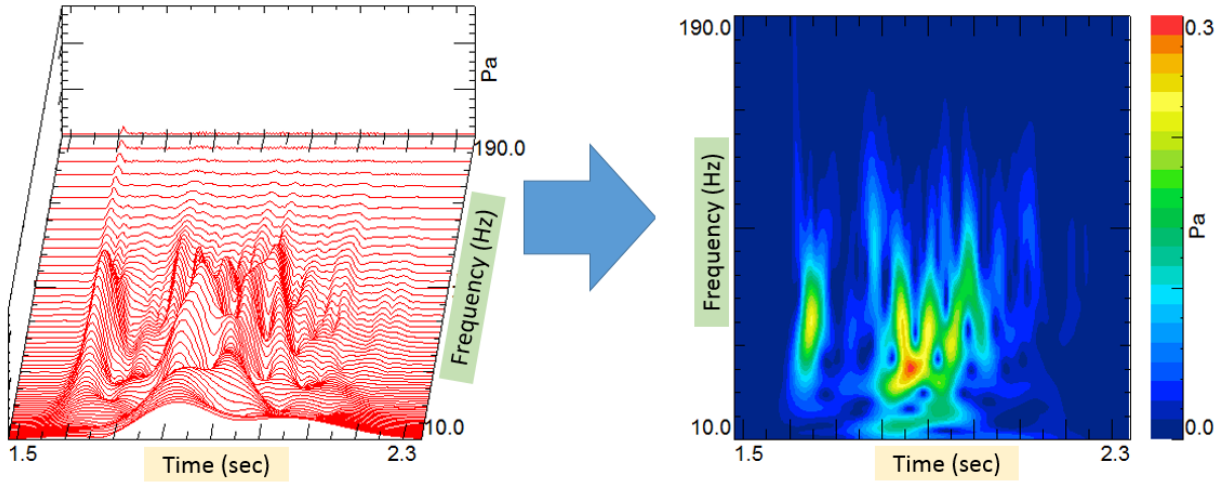


Figure 3.15: In a wavelet analysis, at low frequencies the frequency resolution is finer. At high frequencies, the time resolution is finer.

In the case of the FFT, for each time increment an “amplitude vs frequency" result is created. These results are stacked together to create the colormap.
In Figure 3.16a, each individual calculation of “amplitude vs frequency" in the “waterfall" map display (left) is shown. These individual calculations are smoothed together to create the “colormap" display on the right.
Alternatively, the wavelet analysis (3.16b) will create an “amplitude vs time" result for each frequency increment (as specified by the wavelets per octave setting). These individual calculations (as seen in the waterfall display, left) are smoothed together to create the “colormap" display on the right.
To conclude the comparison between Fourier and wavelet transform, let us analyze a real case: piston slap noise. Piston slap occurs when the piston inside of a cylinder hits the cylinder wall during the operating cycle. This causes an audible transient noise.
It may be desired to know the timing of the piston slap. Traditional FFT methods do not make it obvious when the piston slap occurs. Alternatively, the wavelet analysis highlights the timing and the frequency content of the piston slap.
In Figure 3.17, time vs pressure data from a microphone near an engine block is displayed (top). Some of the transient events are highlighted. This data is analyzed in two way: FFT (middle) and wavelet (bottom).
The wavelet has much better time resolution. The FFT is smeared both in the time domain and the frequency domain.



(a) The Fourier transform results in amplitude vs frequency spectra for each increment in time.



(b) The wavelet analysis results in time versus amplitude results for each frequency increment.

Figure 3.16: Comparison of Fourier and wavelet transform analysis. These results are smoothed together to create the colormap on the right.

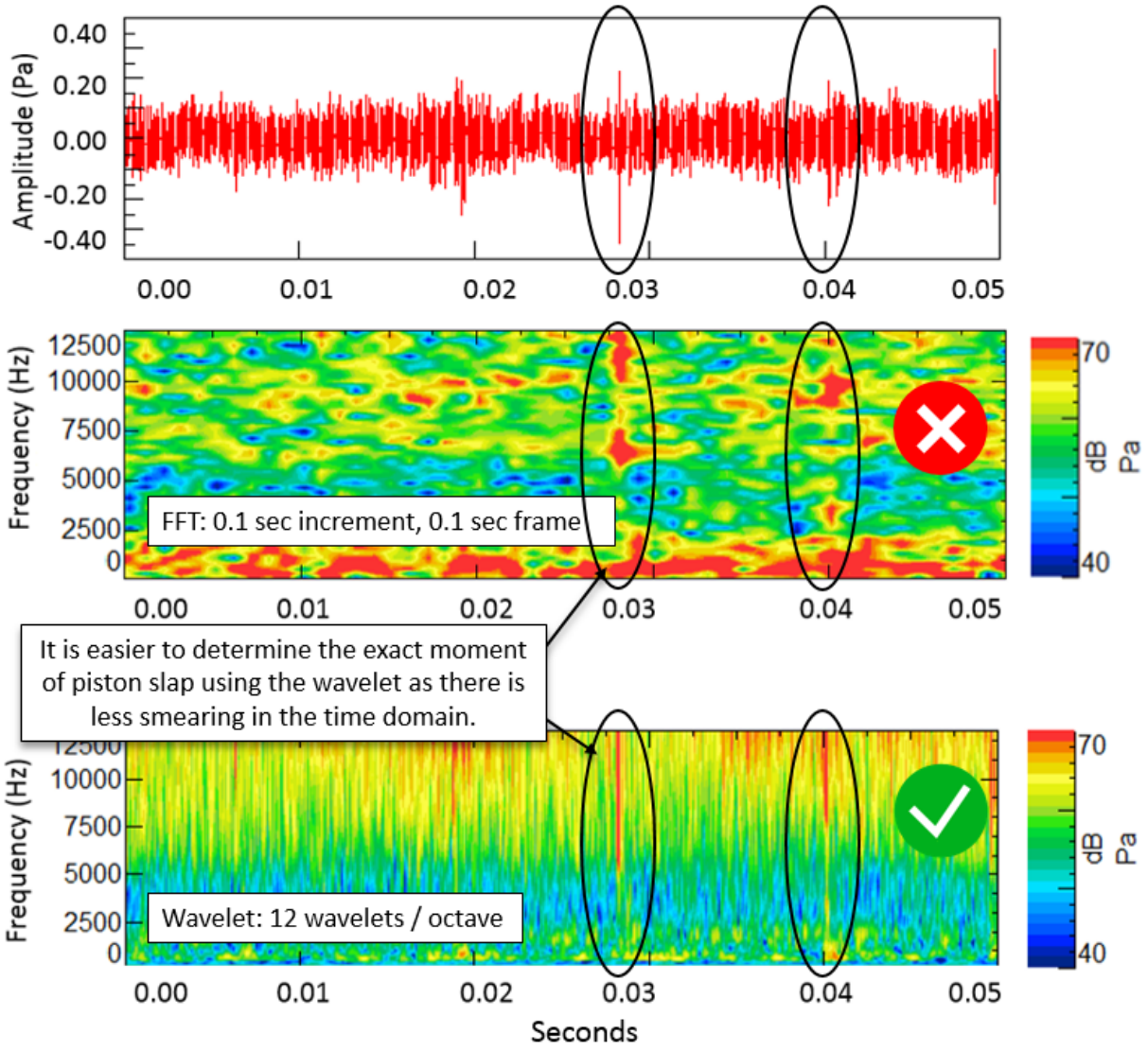


Figure 3.17: Fourier and wavelet transform analysis on a piston slap noise. Top: Time history with multiple transient events. Middle: FFT versus Time analysis does not contain clear indication of the exact timing and frequency content of the transient events. Bottom: Wavelet analysis shows both time and frequency content of transients accurately.

Part II

Procedures

Chapter 4

Procedures for Analyzing Random Data

The procedures for analyzing the properties of random data may be divided logically into two categories: the procedures for analyzing individual sample records, and the procedures for analyzing a collection of sample records given the properties of the individual records. Applicable data analysis procedures for these two categories will be now outlined.

4.1 Procedures for Analyzing Individual Records

An overall procedure for analyzing pertinent statistical properties of individual sample time history records is presented in Figure 4.1. Note that many of the suggested steps in the procedure might be omitted for some applications while additional steps would be required for other applications. Each block in Figure 4.1 will now be discussed.

4.1.1 Mean and Mean Square Value Analysis

The first step indicated by Block A is a mean and mean square value (or variance) measurement (also called ANOVA, ANalysis Of VAriance). This step is almost universally performed for one or more of three sound reasons:

① Since the mean and the mean square values are the basic measures of central tendency and dispersion, their calculation is generally required for even the most rudimentary applications.

② The calculation of a short time averaged mean and mean square value estimates provides a basis for evaluating the stationarity of the data.

③ Mean and mean square value estimates can be extracted from other descriptive properties (probability density plots, correlograms, and/or power spectra) which might be measured later. The comparison of directly measured mean and mean square values estimates to the corresponding estimates extracted from other analyses provides an excellent method for checking the data analysis equipment or computer programs for correct operation.

Figure 4.1: General procedure for analyzing individual sample records

4.1.2 Autocorrelation Analysis

The next suggested analysis is autocorrelation, as indicated by Block B. The autocorrelation function of stationary data is the inverse Fourier transform of the power spectral density function (see (3.25)). Thus the determination of the autocorrelation function will technically not yield any new information over the power spectrum. There might be applications, however, when an autocorrelogram would present the desired information in a more convenient format. Autocorrelation functions can be a useful tool for detecting periodicities in otherwise random data. Furthermore, autocorrelation functions might be computed as an intermediate step in the calculation of power spectral estimates.

4.1.3 Power Spectral Density Analysis

Perhaps the most important single descriptive characteristic of stationary random data is the power spectral density function, which defines the frequency composition of the data. For constant parameter linear physical systems, the output power spectrum is equal to the input power spectrum multiplied by the square of the gain factor of the system. Thus power spectra measurements can yield information concerning the dynamic characteristics of the system. The total area under a power spectrum (that is $\int |F(\omega)|^2 \, d\omega$) is equal to the mean square value. To be more general, the mean square value of the data in any frequency range of concern is determined by the area under the power spectrum bounded by the limits of that frequency range. Obviously, the measurement of power spectra data, as indicated in Block C, will be valuable for many analysis objectives, like detection of periodicities and as an intermediate step in the calculation of autocorrelation functions.

4.1.4 Probability Density Analysis

The last fundamental analysis included in the procedure is probability density analysis, as indicated by Block D. Probability density analysis is often omitted from a data analysis procedure because of the tendency to assume that all random phenomena are normally distributed (this analysis is performed in the so-called Exploratory Data Analysis – EDA). In some cases, however, random data may deviate substantially from the Gaussian form. If such deviations are detected by a test for normality, then the probability density function of the data must be measured to establish the actual probabilistic characteristics of the data. Furthermore, a probability density function estimate is sometimes used as a basis for a normality test.

4.1.5 Nonstationary and Transient Data Analysis

All of the analysis techniques discussed so far apply only to sample records of stationary data. If the data are determined to be nonstationary during the qualification phase of the processing, then special analysis techniques will be required as indicated by Block E. Note that certain classes of nonstationary data can be sometimes be analyzed using the same equipment or com-

puter programs employed for stationary data analysis. However, the results of such analyses must be interpreted with caution.

4.1.6 Periodic and Almost-Periodic Data Analysis

If sinusoids due to periodic or almost-periodic contributions are detected in the data during the qualification phase, then special attention is warranted. Specifically, one of two approaches should be followed. First, the sinusoidal components might be isolated from the random portion of the data by filtering operations and analyzed separately, as illustrated by Block F. Second, the sinusoidal components might be analyzed along with the random portion of the data, and simply accounted for in the results. For example, if a power spectrum is computed for data which include sinusoidal components, a delta function symbol might be superimposed on each spectral peak at the frequency of an identified sinusoid, and labeled with the mean square value of the sinusoid. The mean square value can be estimated from the spectral plot by multiplying the maximum indicated spectral density of the peak by the resolution bandwidth used for the analysis. If this is not done, the physical significance of such spectral peaks might be misinterpreted.

4.1.7 Specialized Data Analysis

Various other analyses of individual time history records are often required, depending upon the specific goals of the data processing. For example, studies of fatigue damage in mechanical systems usually involve the calculation of peak probability density functions of strain data. Spectral descriptions other than power spectral density functions are sometimes desired: for example, the spectra for acoustic noise levels are commonly presented in terms of rms values in 1/1 or 1/3 octave frequency bands. Such specialized analyses, as indicated by Block G, must be established in the context of specific engineering problem of concern.

4.2 Procedures for Analyzing a Collection of Records

The preceding Section presented methods for analyzing each individual sample record from an experiment. A procedure for analyzing further pertinent statistical properties of a collection of sample records is presented in Figure 4.2. As for the analysis of individual sample records outlined in Figure 4.1, many of the suggested steps in Figure 4.2 might be omitted for some applications while additional steps would be required for others. Furthermore, the suggested steps assume the individual records are stationary.

4.2.1 Analysis of Individual Records

The first step is to analyze the pertinent statistical properties of the individual sample records, as outlined in Figure 4.1. Hence the applicable portions of Figure 4.1 constitute Block A in Figure 4.2.

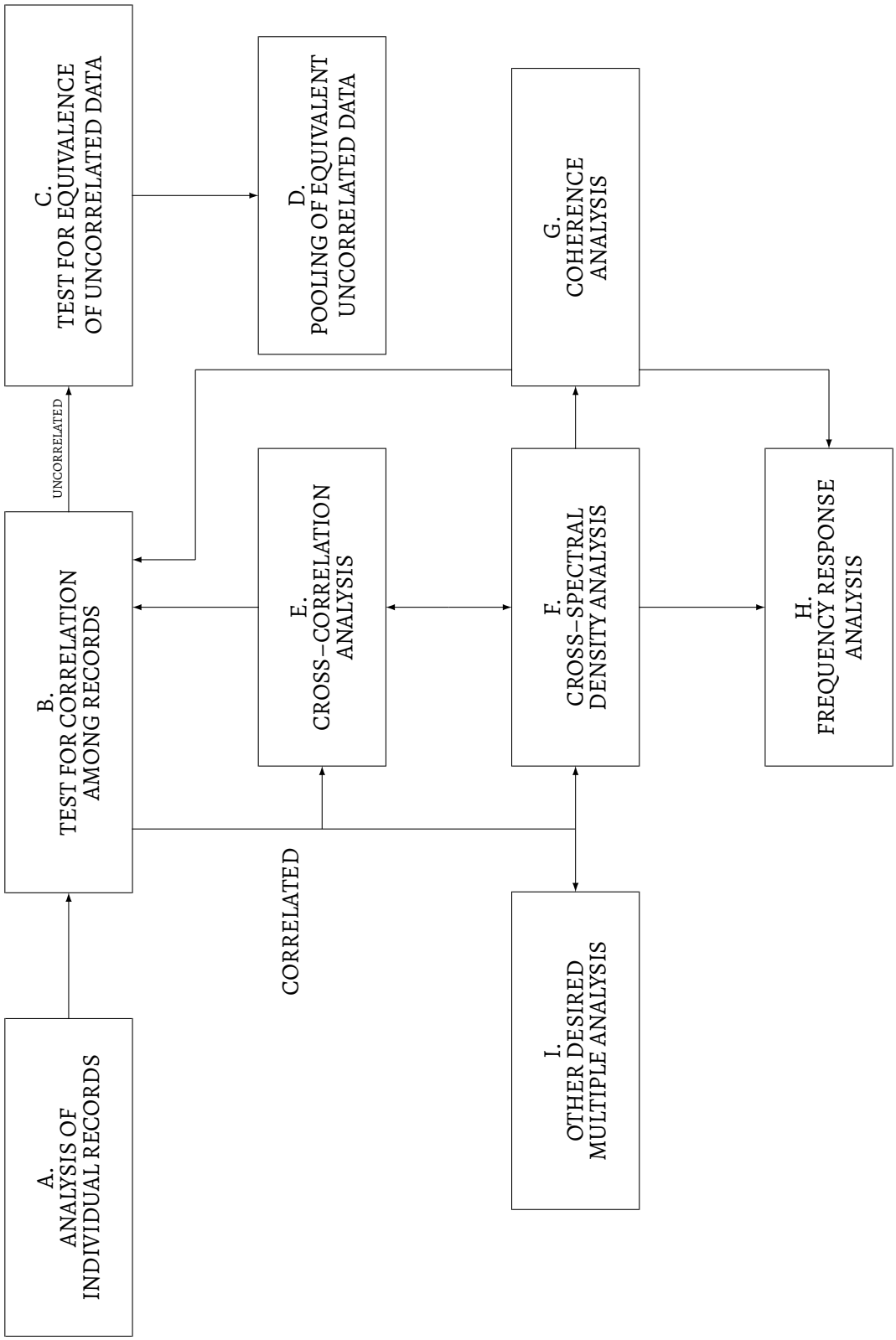


Figure 4.2: General procedure for analyzing a collection of sample records

4.2.2 Test for Correlation

The next step indicated by Block B is to determine whether or not the individual sample records are correlated. In many cases, this decision involves little more than a cursory evaluation of pertinent physical considerations. For example, if the collection of sample records represent measurements of a physical phenomenon over widely separated time intervals, then usually the individual records can be accepted as uncorrelated without further study. On the other hand, if the collection represents simultaneous measurements of the excitation and response of a physical system, the correlation would be anticipated. For those cases where a lack of correlation is not obvious from basic considerations, a test for correlation among the sample records should be performed using cross correlation functions or coherence functions.

4.2.3 Test for Equivalence of Uncorrelated Data

If sample records are found to be uncorrelated in Block B, then these records should be tested for equivalent statistical properties as indicated by Block C. This is an important but often overlooked step in the analysis of random data. Far too often the analyzed results for a large number of sample records are presented as individual plots when in fact the results differ only by the amounts which fall within the acceptable limits of random error. The formal presentation of such redundant data is usually of no value, and can be detrimental in several ways. First, large quantities of analyzed data will sometimes tend to overwhelm the user and unnecessarily complicate the interpretation of the results. Second, the unsophisticated user might interpret the statistical scatter in individual results as physically meaningful differences. Third, more accurate results could be presented for the equivalent data if they were pooled prior of plotting, as will be discussed in the next Section. Note that for most applications, an equivalence of power spectra is a sufficient criterion for equivalence of sampled data.

4.2.4 Pooling of Equivalent Uncorrelated Data

The analyzed results for individual sample records which are found to represent equivalent data should be pooled together as indicated by Block D. This is dome by computing appropriately weighted averages of the results for the individual records. For example, assume two power spectral density function estimates were computed from two uncorrelated sample records which now we found to represent equivalent data. If $G_1(\omega)$ and $G_2(\omega)$ were the original power spectral estimates with n_1 and n_2 sampled points, respectively, a new pooled estimate for the power spectrum is given by

$$G_p(\omega) = \frac{n_1 G_1(\omega) + n_2 G_2(\omega)}{n_1 + n_2} \tag{4.1}$$

where $G_p(\omega)$ has $n_p = n_1 + n_2$ sampled points. Equation (4.1) may be generalized for q estimates from uncorrelated but equivalent samples as follows

$$G_p(\omega) = \frac{\sum_{i=1}^{q} n_i G_i(\omega)}{\sum_{i=1}^{q} n_i} \tag{4.2}$$

It is clear that the pooling operation produces a power spectrum estimate with a reduced random error. However, it should also be noted that the pooling operation generally will not suppress the systematic error (bias) in the power spectra estimates. This fact often leads data analyst to re-process sample records with equivalent statistical properties in a manner designed to reduce bias errors. For the case of power spectra estimates, the reprocessing might consist of a re-computation of power spectral density estimates from the original sample records using a greatly reduced resolution bandwidth to suppress the bias error at the expense of increase random errors. The random errors in the individual estimates are then suppresses by the pooling operation.
Another approach which is sometimes employed with analog data analysis equipment is to splice together the original sample records in order to obtain one long sample record for reprocessing. This procedure can produce acceptable results, but it must be remembered that certain inherent limitations imposed by the length of the original records still apply. Specifically, if q sample records each of length T are spliced together to form a record of length qT, the lowest frequency which can be defined in the data is still $\omega = 2\pi/T$ and not $\omega = 2\pi/(qT)$ – see discussion in Section 5.2.1.

4.2.5 Cross Correlation Analysis

As for the case of autocorrelation and power spectral density functions, the cross correlation and cross spectral density functions are Fourier transform pairs. Hence the measurement of a cross correlogram will technically not yield any new information over the cross spectrum. However, it may present desired information in a more convenient format. An example is the measurement of a simple time delay between two measurement points. Therefore, the cross correlation analysis is included in the procedure as a separate step indicated by Block E. Note that a cross correlation estimate can be used as a test for correlation between two individual records, and as an intermediate step in the calculation of a cross spectral density estimate.

4.2.6 Cross Correlation Spectral Analysis

The most important join measurement for a collection of correlated sample records is the cross spectral density analysis indicated by Block F. Cross spectral density functions provide information concerning the linear relationships which might exist among the collection of sample records.

4.2.7 Coherence Function Analysis

Block G indicates the calculation of coherence functions based upon power and cross spectral density estimates. Coherence functions of various types (ordinary, multiple, and partial) are valuable in several ways. First, they can be used to test for correlation among the collection of sample records. Second, they constitute a vital parameter in assessing the accuracy of frequency response function estimates.

4.2.8 Frequency Response Function Analysis

The ultimate goal in the analysis of a collection of sample records is often to establish linear relationships among the data represented by the various records. The existence of such linear relationships can be detected from cross correlation, cross spectral density, or coherence function estimates. However, a meaningful description of the linear relationship is best provided by computing the frequency response functions of the relationships, as indicated by Block H.

4.2.9 Other Desired Multiple Analysis

Block I indicates other joint analyses of a collection of sample records needed to satisfy special data processing goals. Included might be joint probability density and distribution functions.

Chapter 5

Temporal Analysis in X–ray Astronomy

Now we will apply all the mathematical tools developed in Part I to real data. In particular we will explore the techniques that are commonly used in timing studies of X–ray sources. The regime we will be referring to is that of equidistantly binned timing data, the background noise of which is dominated by counting statistics. If there are gaps in the data, they are far apart and the data are not "sparse" in the sense that nearly all time bins are empty. This kind of data are eminently suited to analysis with FFT techniques, and the discussed methods will be based on these techniques.

5.1 Power Spectra in X–ray Astronomy

If we indicate with x_k, $k = 0, 1, 2, \ldots, N-1$, the number of photons detected in bin k by our instrument, then the discrete Fourier transform a_j, with $j = -N/2, \ldots, N/2-1$, decomposes this signal into N sine waves. The following expressions describe the signal transform pair:

Definition 5.1 (Discrete Fourier transform in X–ray astronomy).

$$a_j = \sum_{k-0}^{N-1} x_k \, e^{2\pi i j k/N} \qquad j = -\frac{N}{2}, \ldots, \frac{N}{2} - 1 \tag{5.1a}$$

$$x_k = \frac{1}{N} \sum_{j=-N/2}^{N/2-1} a_j \, e^{-2\pi i j k/N} \qquad k = 0, 1, \ldots, N-1 \tag{5.1b}$$

Important Note: please note the difference between this definition and the definition (3.11). The signs in the exponents in (5.1) are reversed with respect to the ones in (3.11). In X–ray astronomy it is customary the use of the convention as in Press *et al.* Accordingly, the prefactor $1/N$ is present in the *inverse* discrete Fourier transform and not in the *direct* one. The consequence is that a_0 will not be anymore the average, but the total number of counts $N_{\rm ph} = \sum_k x_k$. As we said before, it is only a question of convention.

If the signal is an equidistant time series of length T, so that x_k refers to a time $t_k = k(T/N)$, then the transform is an equidistant "frequency series", and a_j refers to a frequency $\omega_j = 2\pi\nu_j = 2\pi j/T$. The time step is $\delta t = T/N$; the frequency step is $\delta\nu = 1/T$.
Note that $a_{-N/2} = \sum_k x_k\, e^{-\pi i k} = \sum_k x_k\,(-1)^k = a_{N/2}$, and that a_0 is nothing else that the total number of detected photons $a_0 = \sum_k x_k \equiv N_{\rm ph}$.
We have already seen that the Parseval theorem relates the a_j and x_k:

$$\sum_{k=0}^{N-1} |x_k|^2 = \frac{1}{N} \sum_{j=-N/2}^{N/2-1} |a_j|^2 \tag{5.2}$$

This implies that there is a relation between the summed squared modulus of the Fourier amplitudes and the total variance of the data:

$$\begin{aligned}
\mathrm{Var}(x_k) &= \sum_k (x_k - \bar{x})^2 = \sum_k x_k^2 - 2\bar{x} \underbrace{\sum_k x_k}_{N\bar{x}} + N(\bar{x})^2 \\
&= \sum_k x_k^2 - N(\bar{x})^2 = \sum_k x_k^2 - \frac{1}{N}\left(\sum_k x_k\right)^2 \\
&\overset{(5.2)}{=} \frac{1}{N}\sum_j |a_j| - \frac{1}{N}|a_0|^2
\end{aligned}$$

Therefore we have

$$\mathrm{Var}(x_k) = \frac{1}{N} \sum_{\substack{j=-N/2 \\ j\neq 0}}^{N/2-1} |a_j|^2 \tag{5.3}$$

Adopting the normalization used by Leahy *et al.* (1983), we will define

Definition 5.2 (Power spectrum).

$$P_j \equiv \frac{2}{N_{ph}} |a_j|^2 \qquad j = 0, 1, 2, \ldots, \frac{N}{2} \tag{5.4}$$

where N_{ph} is the total number of photons.

Taking into account that for real data $|a_j| = |a_{-j}|$ and that the term at the Nyquist frequency occurs only once in (5.3), we find the expression for the total variance in terms of P_j:

$$\mathrm{Var}(x_k) = \frac{N_{\rm ph}}{N} \left(\sum_{j=1}^{N/2-1} P_j + \frac{1}{2} P_{N/2} \right) \tag{5.5}$$

Note the difference in the indexing of a_j and P_j. Often the variance is expressed in terms of the fractional root-mean-square (rms) variation in the x_k:

$$\text{rms} = \frac{\sqrt{\frac{1}{N}\,\text{Var}(x_k)}}{\bar{x}} = \sqrt{\frac{1}{N_{\text{ph}}}\left(\sum_{j=1}^{N/2-1} P_j + \frac{1}{2}\,P_{N/2}\right)} \tag{5.6}$$

Sometimes rms is expressed in terms of percentage, and is then called the "percentage rms variation". A sinusoidal signal $x_k = A\sin(2\pi\nu_j t_k)$ at the Fourier frequency ν_j will cause a spike at ν_j in the power spectrum with

$$P_{j,\text{sine}} = \frac{1}{2}\,\frac{N^2}{N_{\text{ph}}}\,A^2 \tag{5.7}$$

The reason for choosing this apparently rather awkward normalization for the powers lies in the statistical properties of the noise power spectrum, to be described later.
Finally, let us discuss on relation between the *sampled* x_k (with Fourier transform a_j) and the *continuous* function $x(t)$ (with Fourier transform $a(\nu)$). It is easy to understand that x_k is given by a **double multiplication** by two functions:

Window Function

$$w(t) = \begin{cases} 1 & 0 \le t < T \\ 0 & \text{else} \end{cases} \tag{5.8}$$

Sampling Function

$$i(t) = \sum_{k=-\infty}^{+\infty} \delta\left(t - k\frac{T}{N}\right) \tag{5.9}$$

Therefore in order to obtain the power spectrum of x_k we must perform a double convolution with both the window and the sampling functions. The power spectrum of the shifted window function is (see Section 2.3.3):

$$|W(\nu)|^2 = \left|\frac{\sin \pi\nu T}{\pi\nu}\right|^2 \tag{5.10}$$

The Fourier transform on an infinitely extended periodic series of δ-functions is

$$I(\nu) = \frac{N}{T}\sum_{k=-\infty}^{+\infty} \delta\left(\nu - k\frac{N}{T}\right) \tag{5.11}$$

The functions $w(t)$ and $i(t)$, together with the corresponding power spectra $W(\nu)$ and $I(\nu)$, are shown in Figure 5.1.
The convolution of $a(\nu)$ with $W(\nu)$ causes all features in the power spectrum to become wider. We have already seen that the convolution with a δ-function at ν_0 causes a shift of the function by ν_0: $f(\nu) * \delta(\nu - \nu_0) = f(\nu - \nu_0)$ Therefore the convolution of $a(\nu)$ with $I(\nu)$, which is a series

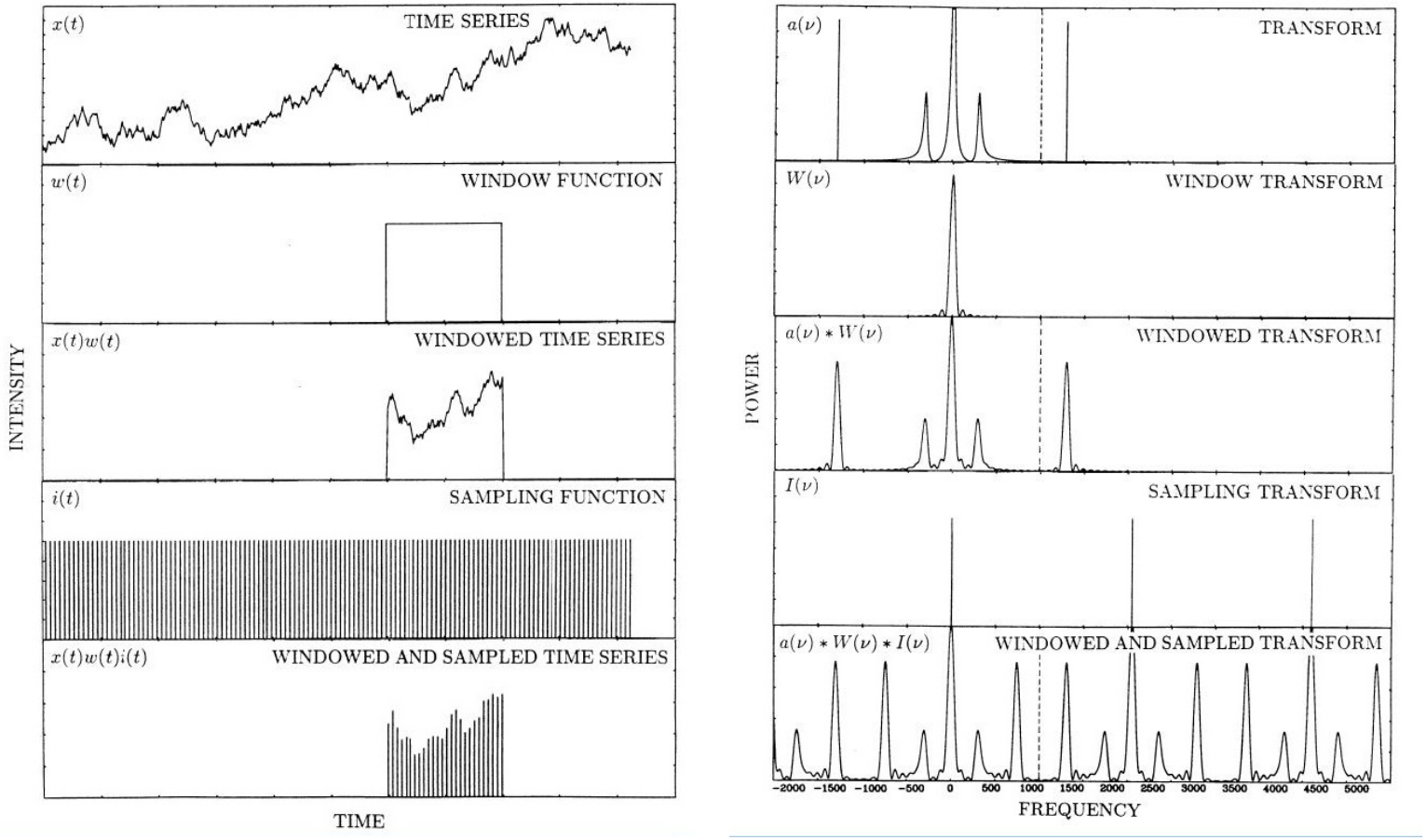


Figure 5.1: *Left:* Obtaining the discrete time series x_k involves the application of the two function $w(t)$ (window function) and $i(t)$ (sampling function). The bottom panel show the final results. *Right:* The discrete Fourier transform a_j of x_k is obtained out of the continuous Fourier transform by a double convolution. These are the power spectra corresponding to the various Fourier transforms. Vertical dashed lines indicate the Nyquist frequency

of δ-functions with spacing N/T results in a convolved function $a(\nu) * I(\nu)$ that repeats every N/T frequency units.
For a real signal $x(t)$ we have, as before, $a(-\nu) = a^*(\nu)$, so that $|a(\nu)|^2 = |a(-\nu)|^2$: the power spectrum is symmetric with respect to $v = 0$. The final result is that the power spectrum of the convolved function $|a(\nu) * I(\nu)|^2$ is reflected around the Nyquist frequency $\nu_{N/2} = \frac{1}{2}N/T$. This causes features with a frequency exceeding the Nyquist frequency by ν_x to appear also at a frequency $\nu_{N/2} - \nu_x$, a phenomenon we have already seen and known as *aliasing*.
From their definitions, is is straightforward to show that the discrete Fourier amplitudes a_j are the values at the Fourier frequencies $\nu_j = j/T$ of the windowed and aliased continuous Fourier transform $a_{\text{WI}}(\nu)$

$$\begin{aligned} a_{\text{WI}}(\nu) = a(\nu) * W(\nu) * I(\nu) = \int_{-\infty}^{+\infty} x(t)\, w(t)\, i(t)\, e^{2\pi i \nu t}\, dt = \\ \int_{-\infty}^{+\infty} x(t) \sum_{k=0}^{N-1} \delta\left(t - k\frac{T}{N}\right) e^{2\pi i \nu t}\, dt = \sum_{k=0}^{N-1} x\left(k\frac{T}{N}\right) e^{2\pi i \nu k T/N} \end{aligned} \tag{5.12}$$

so that $a_{\text{WI}}(j/T) = a_j$. Explicitly performing the convolution of $a(\nu)$ with $I(\nu)$ we finally have

$$a_j = a_{\text{WI}}(j/T) = a_{\text{W}}(j/T) * a_I(j/T) = \frac{N}{T} \sum_{k=-\infty}^{+\infty} a_{\text{W}}\left(v_k - k\frac{N}{T}\right) \tag{5.13}$$

were we used (5.11) and where $\nu_j = j/T$ and $a_{\text{W}} = a(\nu) * W(\nu)$.
To summarize: the transition from the continuous Fourier transform to the discrete Fourier transform involves two operations: windowing, a convolution with the function $W(\nu)$ which is essentially a peak with a width $\delta\nu = 1/T$ plus sidelobes, and aliasing, a reflection of features above the Nyquist frequency back into the range $(0, \nu_{N/2})$. Windowing is caused by the finite extent, aliasing by the discrete sampling of the data.
In practice, aliasing is not so much of a problem as one might fear, as the data are not really discretely sampled at intervals $\delta t = T/N$, but rather binned into time bins with a width δt. This is equivalent of convolving with the "binning window"

$$b(t) = \begin{cases} N/T & -\dfrac{T}{2N} < t < \dfrac{T}{2N} \\ 0 & \text{else} \end{cases} \tag{5.14}$$

before the discrete sampling. Applying the inverse convolution theorem, we can see that the effect of this on the Fourier transform will be that $a(\nu)$ is multiplied with the transform of $b(t)$:

$$B(\nu) = \frac{\sin \pi\nu T/N}{\pi\nu T/N} \tag{5.15}$$

This function drops from a value of 1 at $\nu = 0$ to 0 at $\nu = N/T$; halfway, at the Nyquist frequency it has the value $2/\pi$. The effect of this multiplication is a considerable repression of the

high-frequency features that could be aliased back into the frequency range $(0, \nu_{N/2})$. This is understandable: the effect of the binning is nothing else than averaging the time series over the bin width T/N so that variations with a frequency close to N/T are largely averaged out.
The problem caused by the windowing can be more serious: the "leakage" cause by the finite width of the central peak of $W(\nu)$ and its sidelobes can strongly distort steep power spectra (they becomes less steeper) and it can spread out δ-functions over the entire power spectrum.

5.2 Power Spectral Statistics

In general, signal processing is devoted to *detection* and *estimation*. Detection is the task of determining if a specific signal set is present in the observation, while estimation is the task of obtaining the values of the parameters describing the signal. The process of detecting something in a power spectrum against a background of noise has several steps. To quantify the power of the source signal, that is to determine what the power signal $P_{j,\text{signal}}$ would have been in the absence of noise, we must consider the *interaction* between the noise and the signal.
As our starting point we will make the assumption that our signal will be due to the *sum* of two *independent* processes: signal and noise. This corresponds to assume $x_k = x_{k,\text{signal}} + x_{k,\text{noise}}$. For the linearity of the Fourier transform if b_j and c_j are the Fourier transforms of $x_{k,\text{signal}}$ and $x_{k,\text{noise}}$, then $a_j = b_j + c_j$. This means that a similar properties **does not apply** to power spectra:

$$|a_j|^2 = |b_j + c_j|^2 = |b_j|^2 + |c_j|^2 + \text{cross terms} \tag{5.16}$$

If the noise is random and uncorrelated, and if many powers are averaged, then the cross terms will tend to average out to zero, and we can write down

$$P_j = P_{j,\text{signal}} + P_{j,\text{noise}} \tag{5.17}$$

5.2.1 The Probability Distribution of the Noise Powers

For a wide range of type of noise, the noise powers $P_{j,\text{noise}}$ follow a χ^2 distribution with 2 degrees of freedom (dof). Indeed, if A_j and B_j are the Fourier coefficient of the noise signal, then the Parseval theorem says that $P_{j,\text{noise}} = A_j^2 + B_j^2$. But A_j and B_j are linear combinations of the x_k, therefore if x_k are normally distributed, then the A_j and B_j do as well, so that $P_{j,\text{noise}}$, by definition, is distributed according to the χ^2 distribution with 2 dof.
If the x_k follow some other probability distribution, for example the Poisson distribution, then it follows from the central limit theorem that for "certain" conditions on this other distribution (i.e. for large N), the A_j and B_j will be approximately normally distributed.
In practice, one finds out that noise powers are nearly always χ^2 distributed, not only for Poisson noise, but also for many other type of noise.
The power spectrum normalization defined in (5.4) is chosen in such a way that if the noise in the photon counting data x_k is pure Poissonian counting noise, then the distribution of the P_j, noise is exactly given by a χ^2 distribution with 2 dof. Therefore the probability to exceed a certain threshold power level P_{thr} is given by

$$\text{Prob}(P_{j,\text{noise}} > P_{\text{thr}}) = Q(P_{\text{thr}}|2) \qquad j = 1, 2, \ldots, N/2 - 1 \tag{5.18}$$

where the integral probability of the χ^2 is defined as

$$Q(\chi^2|n) = \frac{1}{2^{n/2}\,\Gamma\left(\frac{n}{2}\right)} \int_{\chi^2}^{\infty} t^{\frac{n}{2}-1}\, e^{-\frac{t}{2}}\, dt \tag{5.19}$$

where n is the number of dof.
Because the $P_{j,\text{noise}}$ follow this distribution, the power spectrum is very noisy; the standard deviation of the noise powers is equal to their mean value:

$$\sigma_{P_j} = \langle P_j \rangle = 2 \tag{5.20}$$

Two more or less equivalent methods are often used to decrease the large variance of the $P_{j,\text{noise}}$ (see discussion in Section 4.2.4)

- ❏ Rebin the power spectrum, averaging b consecutive frequency bins;
- ❏ Divide the data up into M equal segments, transform these segments each individually and then average the resulting M power spectra, each normalized according to (5.4). The N_{ph} is now the number of photons in *each* transform.

These two methods, of course, degrade the frequency resolution.
Because the time required to computed the Fourier transform of N data point using an FFT algorithm is proportional to $N \log N$, there is a computational advantage in the second method; the time saving factor is about $1 + \log M / \log N$.
For a variable source, a further advantage of the second method is the possibility to follow the variations of the power spectra as a function of time and/or intensity (see Figure 5.2).
The first method, on the other hand, has the advantage of producing a power spectrum that extends to lower frequencies. It is of course possible to combine both methods: each power in the final spectrum will be the average of Mb original powers.
Because of the additive properties of the χ^2 distribution, the sum of Mb powers is distributed according to the χ^2 distribution with $2\,Mb$ dof, so that the probability for a given power $P_{j,\text{noise}}$ in the average spectrum to exceed a P_{thr} will be

$$\text{Prob}(P_{j,\text{noise}} > P_{\text{thr}}) = Q(Mb\,P_{\text{thr}}|2\,Mb) \tag{5.21}$$

For large Mb this distribution tends asymptotically to a normal distribution with a mean of 2 and a standard deviation of $2/\sqrt{Mb}$:

$$\lim_{Mb\to\infty} \text{Prob}(P_{j,\text{noise}} > P_{\text{thr}}) = Q_{\text{Gauss}}\left(\frac{P_{\text{thr}} - 2}{2/\sqrt{Mb}}\right) \tag{5.22}$$

where the integral probability of the normal distribution is

$$Q_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-t^2/2}\, dt \tag{5.23}$$

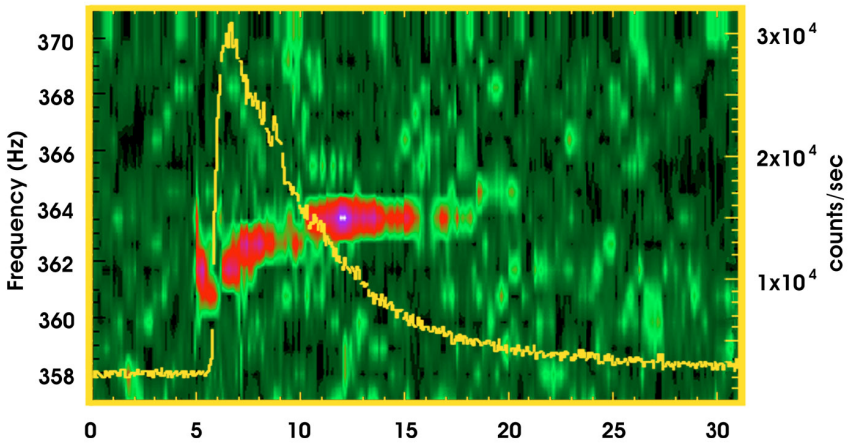


Figure 5.2: Dynamic power spectrum of the low mass X–ray binary 4U 1728–34. The color map shows increasing power in the order green, red, blue, and white. The time increases along the horizontal axis with a resolution of 1 sec. The total time shown is 32 sec. During the burst (the yellow line) the source exhibits pulsation at ~ 363 Hz (Strohmayer *et al.* 1996. *ApJ* 469, L9)

5.2.2 The Detection Level: The Number of Trials

Assuming the χ^2 properties of the noise powers (5.21) we can now determine how large a power must be to constitute a significant excess above the noise.

Definition 5.3 (Detection level). *Let us define $(1-\varepsilon)$ the confidence detection at level P_{det} as the power level that has only the small probability ε to be exceeded by a noise power.*

So, if there is a power P_j that exceeds $P_{\rm det}$, then there is a large probability $(1-\varepsilon)$ that P_j is not purely due to noise, but also contains signal power $P_{j,\rm signal}$ (this is because of (5.17)).
Because of the way we have built our power spectrum, each P_j will be the sum of Mb power spectra (or less if only certain frequency range in the spectrum is considered). We call the number of different P_j values the **number of trials** $N_{\rm trial}$. The probability to exceed $P_{\rm det}$ by noise should have the small value ε for all the powers in the frequency range of interest *together*, so that the chance *per trial* should have the much smaller value of

$$(1-\varepsilon)^{1/N_{\rm trial}} \simeq \frac{\varepsilon}{N_{\rm trial}} \qquad \text{for } \varepsilon \ll 1 \tag{5.24}$$

So the detection level $P_{\rm det}$ is

$$\frac{\varepsilon}{N_{\rm trial}} = Q(Mb\, P_{\rm det}|2\, Mb) \tag{5.25}$$

As an example, if in a power spectrum normalized according (5.4) we find a feature at a level of 44, the probability of finding a $\chi^2 > 44$ for 2 dof by chance is $Q(44|2) = 3\cdot 10^{-10}$. Taking into account $N_{\rm trial} = 65000$ we obtain that the probability of finding our feature by chance is $3\cdot 10^{-10}\times 65000 = 2\cdot 10^{-5}$: so our feature is quite significant!

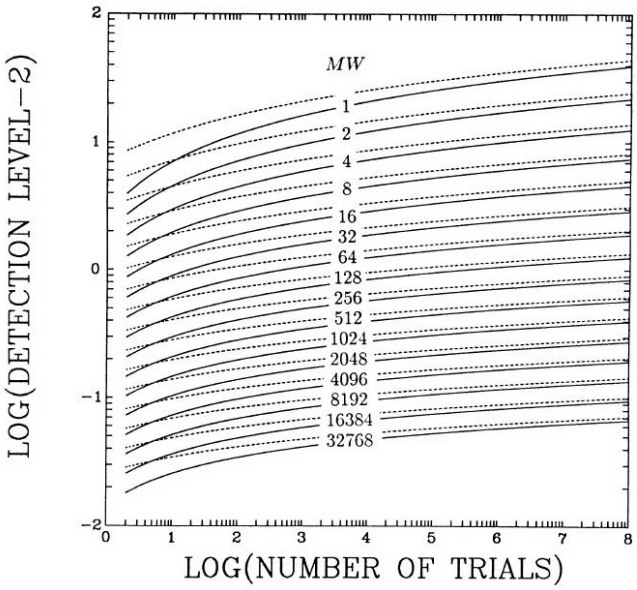


Figure 5.3: Confidence detection level at the 90% (continuous) and 99% (dashed) as a function of the number of trials. The number of independent powers, Mb, averaged together due to rebinning of the power spectra by a factor b and averaging M different power spectra increases by a factor 2 in consecutive curves. The trials are assumed to be independent, so no overlaps between the b-bins averages are allowed. As an example, for a power spectrum produced by averaging together 2 "raw" spectra of 4096 bins each and binning up the resulting spectrum by a factor 4 to produce a 1024-bin average spectrum, the 90% confidence detection level can be read from the curve $Mb = 2 \times 4 = 8$ at $N_{\text{trial}} = 1024$ to be 5.8.

In Figure 5.3 P_{det} is plotted as a function of N_{trial} for various values of Mb and for a confidence level of 90% ($\varepsilon = 0.1$), and 99% ($\varepsilon = 0.01$). Note that although P_{det} increases with the number of trials N_{trial}, the increase is relatively slow.

5.3 The Signal Power

Any quantitative statement one can make about the signal power $P_{j,\text{signal}}$ will be a statement of a probability based on the probability distribution of the noise powers $P_{j,\text{noise}}$ because, from (5.17),

$$P_{j,\text{signal}} = P_j - P_{j,\text{noise}} \tag{5.26}$$

Therefore, supposing we have a detection (i.e. for a given j it is true that $P_j > P_{\text{thr}}$), then what is the probable value of the signal power $P_{j,\text{signal}}$ at j?
If we define a "limiting noise power level" P_{NL} that has only a small probability ε' to be exceeded in one trial

$$\varepsilon' = Q(Mb\, P_{\text{NL}} | 2\, Mb) \tag{5.27}$$

then, with confidence $(1-\varepsilon')$, we can say that, for a given j, $P_{j,\text{noise}} < P_{\text{NL}}$. Therefore, from (5.26)

$$P_{j,\text{signal}} > P_j - P_{\text{NL}} \qquad (1-\varepsilon') \text{ confidence} \tag{5.28}$$

If no significant power level has been attained by any of the P_j, then it is useful to determine an upper limit to the signal power. The $(1-\delta)$ confidence upper limit P_{UL} to the signal power is defined as the power level for which $P_{j,\text{signal}} < P_{\text{UL}}$ at $(1-\delta)$ confidence, irrespective of where this signal power may have occurred.
To determine P_{UL} we define a power level P_{exce} that has the large probability $(1-\delta)$ to be exceeded by a given individual $P_{j,\text{noise}}$:

$$(1-\delta) = Q(Mb\, P_{\text{exce}} | 2\, Mb) \tag{5.29}$$

So a fraction of approximately $(1-\delta)$ of all powers considered will exceed P_{exce} in absence of source signal. We now find the largest *observed* power P_{max} in the given frequency interval, and write

$$P_{\text{UL}} = P_{\text{max}} - P_{\text{exce}} \tag{5.30}$$

In Figure 5.4 we show the relations between the different quantities defined so far.

5.3.1 Sensitivity to Signal Power

It is sometimes useful to predict the capabilities of a planned experiment in terms of its sensitivity to signal power. The sensitivity level P_{sens} can be calculated on the basis of the expected probability distribution of the noise power as

$$P_{\text{sens}} = P_{\text{det}} - P_{\text{exce}} \tag{5.31}$$

If there occurs a $P_{j,\text{signal}}$ somewhere in the spectrum that exceeds P_{sens} then it will be detected with $(1-\delta)$ confidence (see Figure 5.4).

5.3.2 The rms Variation in the Source Signal

Assuming that the signal power spectrum has been properly separated from the total power spectrum, we can convert the signal power into the rms variation of the source signal x_k using the expression

$$\text{rms} = \sqrt{\frac{b \sum_j P_{j,\text{signal}}}{N_{\text{ph}}}} \tag{5.32}$$

where P_j is an Mb times averaged power and where N_{ph} is the number of photons *per transform*.

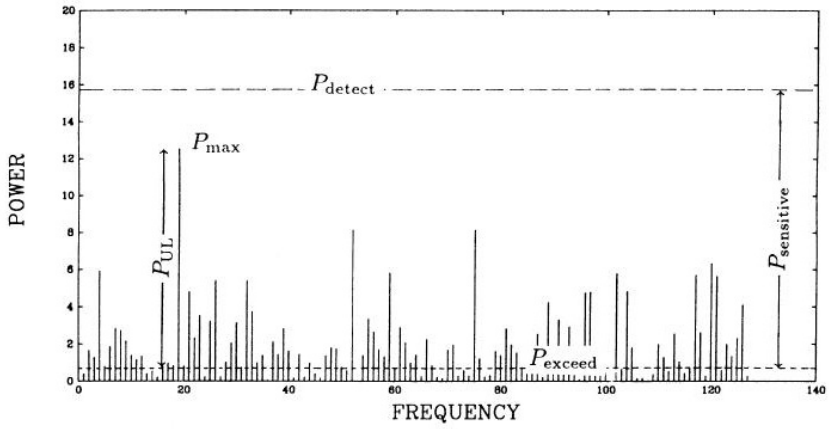


Figure 5.4: Relations between the detection level P_{det}, the exceeded level P_{exce}, the maximum observed power $P_{\max}$, the upper limit P_{UL} and the sensitivity level P_{sens}.

5.4 Detection of Features

If the signal power of a narrow feature in a power spectrum is P_{signal}, then it will drop to P_{signal}/Mb after the frequency resolution has been degraded by a factor Mb by one of the methods described above. Also the detection level degrades, both because the probability distribution of the noise powers in the average power spectrum becomes narrower and because the number of trials decreases by a factor Mb.
However, in the final analysis the sensitivity level always drops more slowly that $1/Mb$, so that the conclusion is that *for detecting a narrow feature in the power spectrum the highest sensitivity is reached for the maximum possible frequency resolution, that is* $Mb = 1$.
Similar reasoning shows that for a feature of finite width $\Delta\nu$ the signal power summed over all frequency bins in the feature will drop proportionally to $1/Mb$ when the frequency resolution of the power spectrum is degraded. However, as long as the width of the feature exceeds the frequency resolution $\Delta\nu > Mb/T_{\text{obs}}$, where $T_{\text{obs}} = MT$ is the total duration of the observation, the signal power in one frequency bin within the feature will remain constant.

5.5 Power Spectral Searches Made Easy

In this section we collect all previous results into a "how-to" recipe for testing the power spectrum for a weak signal using equal statistically independent trials

① Determine the M and b. The optimal choice for Mb is that which approximately matches the expected width of the power spectral feature one desires to detect, $\Delta\nu \geq Mb/T_{\text{obs}}$ (see Figure 5.5 for the effects of choosing the right b). Note that gaps in the data or the desire to observe the time evolution of the power spectrum may dictate M.

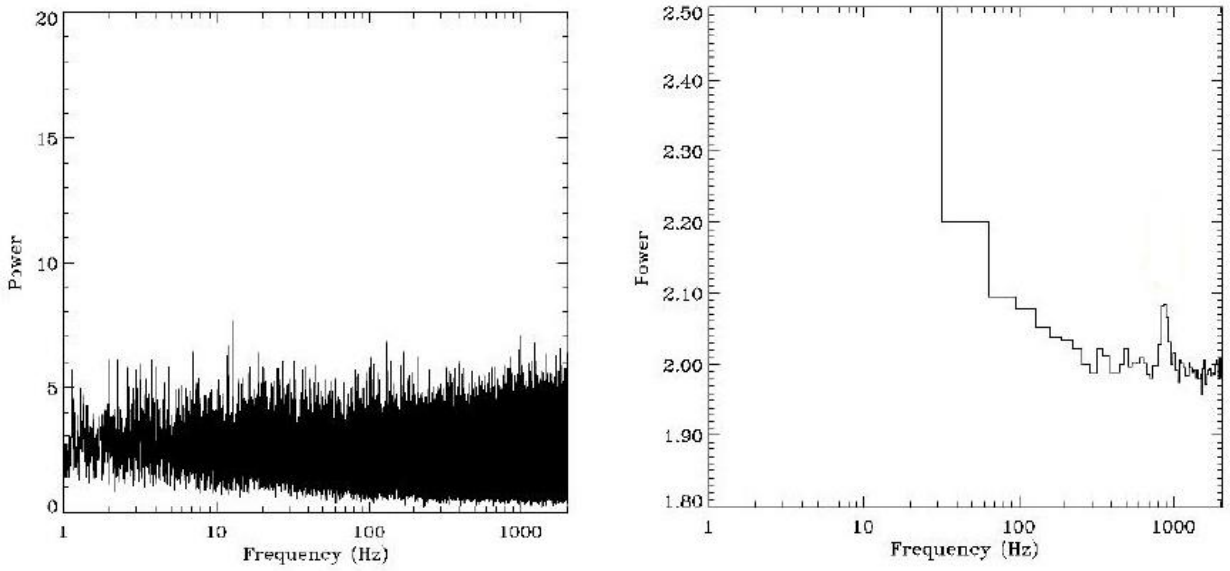


Figure 5.5: Effect of choosing the binning size in detecting weak features: the case of the kHz QPO in 4U 1728−34. The same data are shown in both the two panels, but the right bin size reveals the QPO at ~ 800 Hz

② Calculate the M power spectra normalized according to (5.4). Note that x_k is the number of photons in bin k and $N_{\rm ph}$ is the number of photons in *one* power spectrum.

③ Average the M power spectra.

④ Observe the noise power distribution. Is the noise power spectrum flat? Is its mean level equal to 2? If so, the noise is probably dominated by Poissonian counting statistics. If not, it is necessary to find out why.

⑤ Determine the detection level.

⑥ Check the average spectrum for powers exceeding the detection level.

⑦ Quantify the signal power in terms of a detection or an upper limit.

⑧ Convert the signal power into the relative rms variation of the source signal, defined as

$$\text{rms} = \sqrt{\frac{1}{N}\sum_k (\text{RATE}_k - \langle \text{RATE} \rangle)^2} \tag{5.33}$$

and compute the *excess* variance

$$\text{Excess Variance} = \sqrt{\text{rms}^2 - \frac{1}{N}\sum_k \text{ERROR}_k^2} \tag{5.34}$$

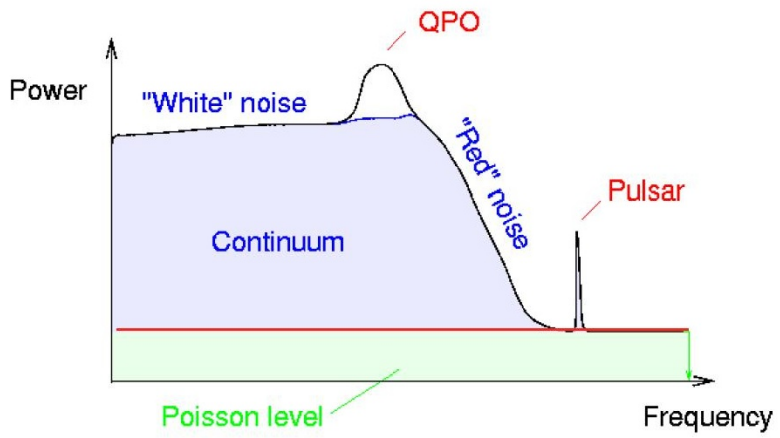


Figure 5.6: Noise classification in astronomical power spectra

⑨ To say more about the signal, we need to model its power spectrum.

5.6 Type of Variability

In the previous Section we were left with the last point in our "how-to" with the problem of modeling a power spectrum. In this section we will deal with the problem of linking the *shape* of a power spectrum with the statistical processes that originated the timing variability. In Figure 5.6 we show a schematic power spectrum of an X–ray source displaying characteristic features: a continuum described in terms of $1/f$ noise, a Quasi-Periodic Oscillation (QPO) and a sharp peak due to a coherent signal (in this case the rotation period of the object). We have already discussed on the Poissonian level; now we will now analyze in details the other components.

5.6.1 $1/f$ Noise

Definition 5.4 ($1/f$ noise). *$1/f$ refers to the phenomenon of the spectral density, $S(f)$, having the form*

$$S(f) = K\, f^{-\alpha} \tag{5.35}$$

where f is the frequency.

$1/f$ noise is an intermediate between the well understood white noise with no correlation in time and random walk (Brownian motion) noise with no correlation between increments (see Figure 5.7). Brownian motion is the integral of white noise, and integration of a signal increases the exponent α by 2 whereas the inverse operation of differentiation decreases it by 2. Therefore, $1/f$ noise can not be obtained by the simple procedure of integration or of differentiation of such convenient signals. Moreover, there are no simple, even linear stochastic differential equations generating signals with $1/f$ noise.

The widespread occurrence of signals exhibiting such behavior suggests that a generic mathematical explanation might exist. Except for some formal mathematical descriptions like fractional Brownian motion (half-integral of a white noise signal), however, no generally recognized physical explanation of $1/f$ noise has been proposed. Consequently, the ubiquity of $1/f$ noise is one of the oldest puzzles of contemporary physics and science in general.
The case of $\alpha = 1$, or *pink noise*, is both the canonical case, and the one of most interest, but the more general form, where $0 < \alpha \leq 3$, is sometimes referred to simply as $1/f$. $1/f^{\alpha}$ noise is of interest because it occurs in many different systems, both in the natural world and in man-made processes (see Figure 5.8) from physics, biology, neuroscience, and psychology.
Although 1/f noise appears in many natural systems and has been intensively studied for decades with many attempts to describe the phenomenon mathematically, researchers have not yet been able to agree on a unified explanation. Thus, there exist at present several formulations of systems that give rise to $S(f) = K/f^{\alpha}$.

5.6.2 Shot Noise Process

First, let t_k be a Poisson point process. A shot noise process is obtained by attaching to each t_k a relaxation function (unilateral exponential function)

$$N(t) = N_0 e^{-\lambda t}, \qquad t \geq 0$$

and summing on k (see Figure 5.9). The Fourier transform of the shot noise process is (see Page 28)

$$S(f) = \lim_{T\to\infty} \frac{1}{T} \langle \mid F(f) \mid^2 \rangle = \frac{N_0^2 n}{\lambda^2 + f^2}$$

where n is the average rate at which t_k occur, and T is the interval over which the process is observed. As we have already seen, the power spectrum of an unilateral exponential function is a Lorentzian function. For an aggregation of shot noise processes with λ uniformly distributed on $[\lambda_1, \lambda_2]$, the power spectrum is

$$S(f) = \begin{cases} N_0^2 n & \text{if } 0 \ll f \ll \lambda_1 \ll \lambda_2 \\ \dfrac{N_0^2 n \pi}{2f(\lambda_2 - \lambda_1)} \cdot \dfrac{1}{f} & \text{if } \lambda_1 \ll f \ll \lambda_2 \\ N_0^2 n \cdot \dfrac{1}{f^2} & \text{if } 0 \ll \lambda_1 \ll \lambda_2 \ll f \end{cases}$$

If the impulse response function is a power law, $N_0 x^{-\beta}$, the process is called *fractal* shot noise, and the power spectrum is of the form

$$S(f) \approx \frac{k}{f^{2(1-\beta)}}$$

When $\beta = 1/2$, we obtain $S(f) \approx 1/f$.

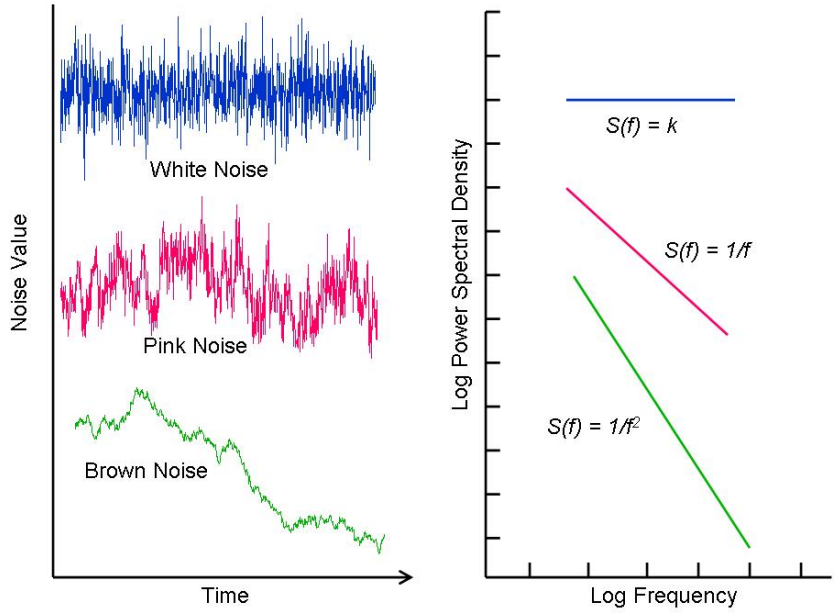


Figure 5.7: Examples of $1/f$ noise: on the left the time series and on the right the corresponding power spectrum

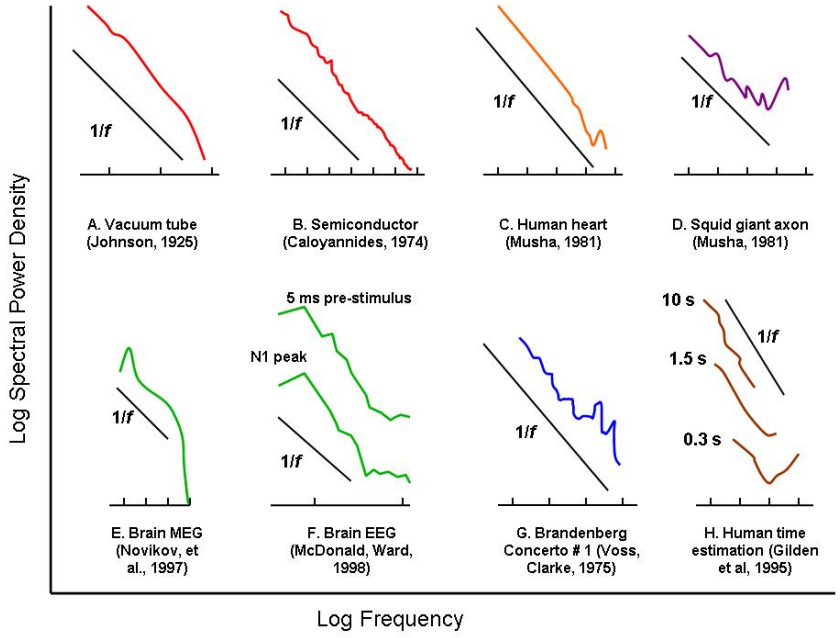


Figure 5.8: Examples of $1/f$ noise observed in both in the natural world and in man-made processes from physics, biology, neuroscience, and psychology

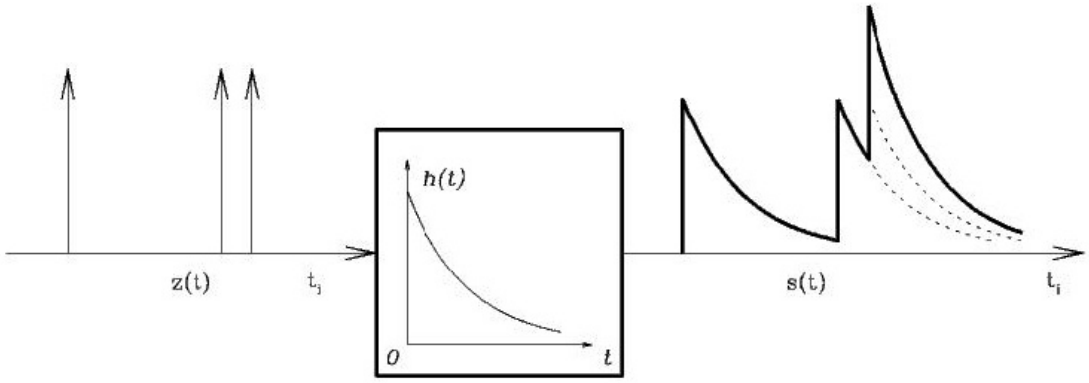


Figure 5.9: The shot noise process

5.6.3 A Clustering Poisson Point Process

Another example based on a Poisson point process is the clustering Poisson. To each Poisson point, t_k, is attached an additional set of points, called a cluster, that occur after it; clusters can overlap each other. The number of points in each cluster, m, is a random variable whose distribution, p_m, is concentrated on a finite set of integers. The points in the cluster are spaced at independent and identically distributed intervals with an arbitrary inter-point distribution. The power spectral density turns out to be a sum of Lorentzian-like functions. When p_m is proportional to $1/m^2$ we obtain $S(f) \propto 1/f$.
A slightly different formulation is a gating process in which clusters do not overlap. Here, a Poisson point process is multiplied by a gating process that is 1 on a random interval and then 0 on a random interval and so on. To obtain a $1/f$ noise let the intervals of 1 be exponentially distributed and the intervals of 0 be geometrically distributed, or vice versa. Then roughly the same computations as just summarized yield the $1/f$ approximation. Notice that for the shot noise processes, the cluster and gating processes, and the AR(1) aggregation (see below), the power spectral density computation yields a sum of Lorentzian or Lorentzian-like functions.

5.6.4 Recurrence Models

In these models the signal consists of pulses or events

$$x(t) = a \sum_k \delta(t - t_k).$$

Here $\delta(t)$ is the Dirac delta function, $\{t_k\}$ is a set of the occurrence times at which the particles or pulses cross the section of observation, and a is the contribution to the signal of one pulse or particle. The inter-pulse, inter-event, inter-arrival, recurrence or waiting times $\tau_k = t_{k+1} - t_k$ of the signal are described by the general Langevin equation with multiplicative noise, which is also stochastically diffuse in some interval, resulting in the power-law distribution.

Another recurrence time process generating a power-law probability distribution is a multiplicative stochastic process

$$\tau_{k+1} = \tau_k + \gamma\,\tau_k^{2\mu-1} + \sigma\,\tau_k^{\mu}\varepsilon_k,$$

where the ϵ_k are independent and identically distributed Gaussian noise, γ is very small and σ, the standard deviation of the noise, is also small, while μ represents the degree of multiplicativity of the process. A particular form of the model is the autoregressive[1] AR(1) process

$$(\tau_k - \bar{\tau}) = (1-\gamma)(\tau_k - \bar{\tau}) + \sigma\epsilon_k,$$

where $\bar{\tau}$ is the mean of the inter-event intervals.
Notice that the power spectrum of this AR(1) time series process, composed of successive values of $(\tau_k - \bar{\tau})$, is proportional to $1/f^2$ on a long interval when γ is small, and thus this power spectrum is not the same as that of the point process (whose points t_k generate the time series) on that interval.
In such point process models the intrinsic origin of $1/f$ noise is in Brownian fluctuations of the mean inter-event time of the (Poisson-like) signal pulses, similar to Brownian fluctuations of signal amplitude that result in $1/f^2$ noise. The random walk of the inter-event time on the time axis is a property of randomly perturbed or complex systems that display self-organization.

Without enter into the details, other ways of obtaining $1/f$ noise are through models of a particular class of *stochastic differential equation* and through a *reversible Markov chain model*.

5.7 Fitting Power Spectra Continuum with Lorentzians

Instead of describing the observed power spectrum continua in terms of $1/f$ noise, recently it has become quite popular a different approach. The power spectrum from X–ray sources like low-mass X–ray binaries (LMXB) can be described in terms of a flat-top continuum at low frequencies that becomes steeper at high frequencies, with bumps and wiggles. This continuum can be fit *without* the need of power-law components, but as a sum of Lorentzians, some of which are broad (Belloni *et al.* 2002).
The power spectra are described as the sum of Lorentzian components $L(\nu)$ of the form

$$L(\nu) = \frac{r^2\Delta}{\pi}\,\frac{1}{\Delta^2 + (\nu-\nu_0)^2} \tag{5.36}$$

[1] The notation AR(p) refers to the autoregressive model of order p. The AR(p) model is defined as

$$X_t = c + \sum_{i=1}^{p}\varphi_i X_{t-i} + \varepsilon_t$$

where $\varphi_1, \ldots, \varphi_p$ are the parameters of the model, c is a constant and ε_t is white noise. The constant term is omitted by many authors for simplicity. An autoregressive model can thus be viewed as the output of an all-pole infinite impulse response filter whose input is white noise.

where r is the integrated fractional rms (over $-\infty$ to $+\infty$) of each Lorentzian and Δ is its Half-Width at Half Maximum (HWHM=FWHM/2). The power spectra are then displayed in a $\nu\, P_\nu$ plot. The frequency $\nu_{\rm max}$ at which the $\nu\, P_\nu$ attains its maximum is

$$\nu_{\rm max} = \sqrt{\nu_0^2 + \Delta^2} = \nu_0 \sqrt{1 + \frac{1}{4Q^2}} \tag{5.37}$$

where $Q \equiv \nu_0/2\Delta$ is called quality factor. Note that $\nu_{\rm max} \geq \nu_0$: the difference is small for narrow features but becomes large in the case of broad ones. In Figure 5.10 we show an example of such a fit for two LMXB: XTE 1118+480 and 1E 1724−3045.
With this phenomenological modelization it is possible to use a limited number of fit components and compare the power spectra of different sources. But what is the *physical mechanism* responsible for the observed shape of the power spectra is still an open issue.

5.8 Quasi-Periodic Oscillations (QPO)

Quasi-Periodic Oscillations (QPO) are broad features observed in the power spectra of many X–ray sources. They are described in terms both of a Gaussian or a Lorentzian shape. As discussed above, the Lorentzian shape has a physical basis as due to a shot noise process. The QPO can be therefore characterized by its centroid frequency LC, its width LW, and its normalization LN. Instead of LN t is customary to give the QPO percentage rms, defined as

$$\text{percentage rms} = 100 \sqrt{\frac{I}{\langle \text{RATE} \rangle}} \tag{5.38}$$

where I is the Lorentzian integral, defined as

$$I = \frac{\pi}{2}\, \text{LC} \times \text{LW}$$

and $\langle \text{RATE} \rangle$ is the source average count rate. Sometimes, for a QPO is given the quality factor Q, defined as

$$Q\text{-factor} = \frac{\text{LC}}{\text{LW}} \tag{5.39}$$

In Figure 5.11 we show a Lorentzian fit to a QPO observed in the low-mass X-ray binary and atoll source 4U 1735-44.

5.9 Analysis of Unevenly Sampled Data

Thus far, we have been dealing exclusively with evenly sampled data. There are situations, however, where evenly sampled data cannot be obtained (for example, for astronomical data, where the observer cannot completely control the time of observations).

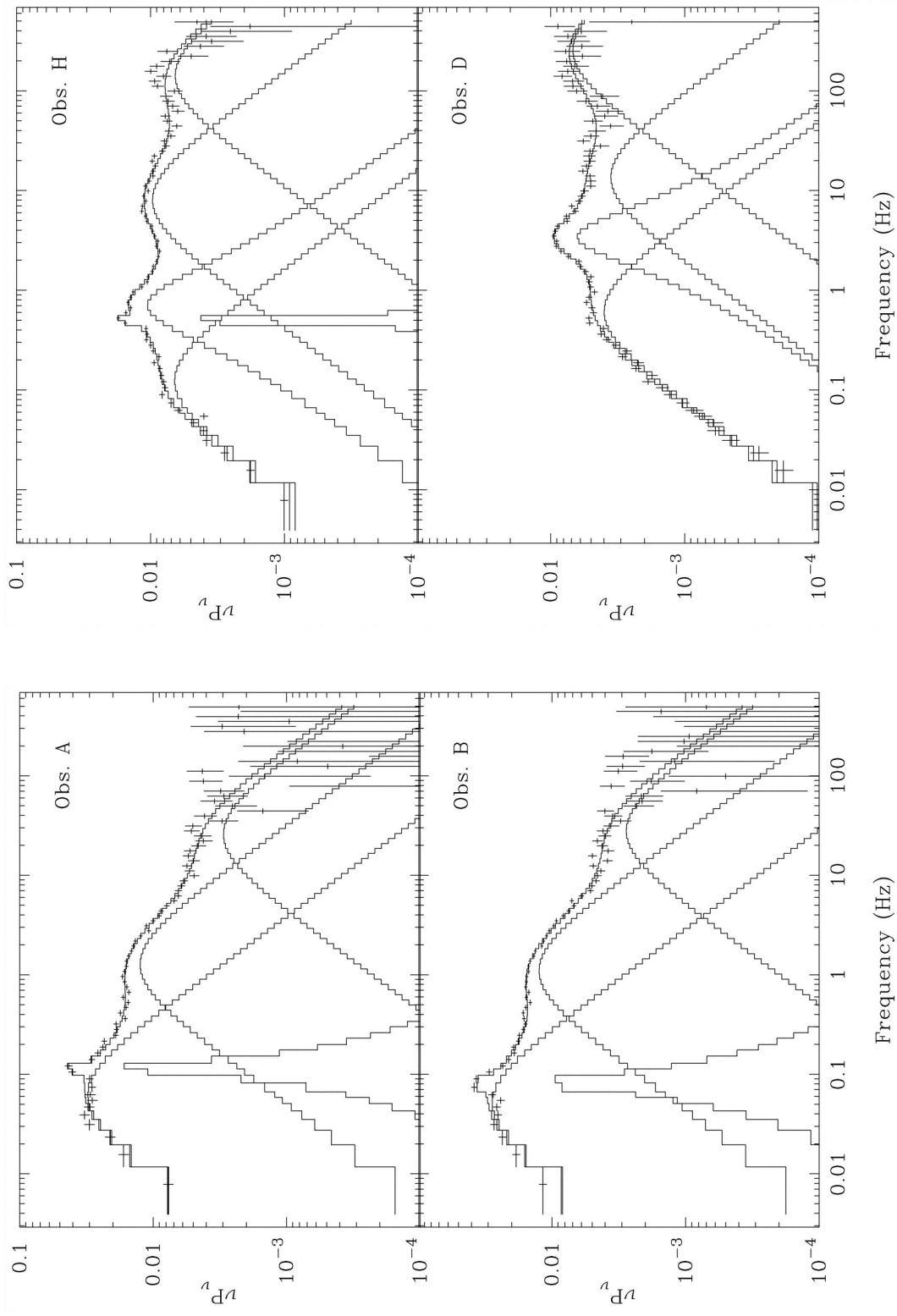


Figure 5.10: Power spectra in $\nu\, P_\nu$ form for two LMXB XTE 1118+480 (left) and 1E 1724−3045 (right). The presence of more power at high frequency for 1E 1724−3045 is interpreted as due to the presence of a neutron star in the system (while XTE 1118+480 should hosts a black hole)

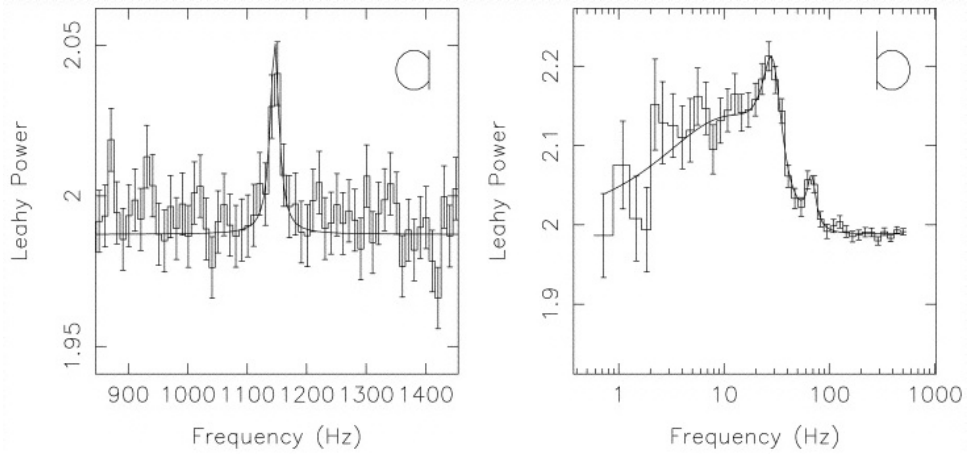


Figure 5.11: Typical Leahy normalized power spectra in the energy range of 2–18 keV. (a) The kHz QPO at 1150 Hz; (b) the complex high-frequency noise and the 67 Hz QPO. From Wijnands *et al.* 1998. *ApJ* 495, L39

There are some obvious ways to get from unevenly spaced t_k to evenly spaced ones. Interpolation is one way: lay down a grid of evenly spaced times on our data and interpolate values onto that grid; then use FFT methods. If a lot of consecutive points are missing, we might set them to zero, or might be fix them at the value of the last measured point. Unfortunately, these techniques perform poorly. Long gaps in the data, for example, often produce a spurious bulge of power at low frequencies.
A completely different method of spectral analysis of unevenly sampled data was developed by Lomb and additionally elaborated by Scargle. The Lomb-Scargle method evaluates data, and sines and cosines, only at times t_k that are actually measured.
Suppose that there are N data points $h_k \equiv h(t_k)$ with $k = 0, 1, \ldots, N-1$. Then first find the mean and variance of the data in the usual way

$$\bar{h} \equiv \frac{1}{N}\sum_{k=1}^{N-1} h_k \qquad\qquad \sigma \equiv \frac{1}{N-1}\sum_{k=1}^{N-1}(h_k - \bar{h})^2 \tag{5.40}$$

Definition 5.5 (Lomb-Scargle normalized periodogram).

$$P_N(\omega) \equiv \frac{1}{2\sigma^2}\left\{\frac{\left[\sum_k (h_k - \bar{h})\,\cos\omega(t_k - \tau)\right]^2}{\sum_k \cos^2\omega(t_k - \tau)} + \frac{\left[\sum_k (h_k - \bar{h})\,\sin\omega(t_k - \tau)\right]^2}{\sum_k \sin^2\omega(t_k - \tau)}\right\} \tag{5.41}$$

Here τ is defined by the relation

$$\tan(2\omega\tau) = \frac{\sum_k \sin 2\omega t_k}{\sum_k \cos 2\omega t_k} \tag{5.42}$$

The constant τ is a kind of offset that makes $P_N(\omega)$ completely independent of shifting all the t_k by any constant. This particular choice of τ has another, deeper effect. It makes (5.41) identical

to the equation that one would obtain if one estimated the harmonic content of a data set, at a given frequency ω, by linear least-squares fitting to the model

$$h(t) = A\ \cos\omega t + B\ \sin\omega t$$

This fact gives some insight into why the method can give results superior to FFT methods: it weight the data on a "per-time interval" basis, when uneven sampling can render the latter seriously in error.

As we have seen in Section 5.3, the assessment of the signal significance in the case of FFT methods is not easy. On the other hand, the significance of a peak in the $P_N(\omega)$ spectrum can be assessed rigorously.

The word "normalized" refers to the σ^2 factor in the denominator of (5.41). Scargle shows that with this normalization, at any particular ω and in the case of the *null hypothesis* that our data are *independent Gaussian random values*, then $P_N(\omega)$ has an exponential probability distribution with unit mean. In other words, the probability that $P_N(\omega)$ will be between some positive z and $z + dz$ is $\exp(-z)\, dz$. It readily follows that, if we scan some M *independent* frequencies, the probability that none give values larger than z is $(1 - e^{-z})^M$. So

$$\text{Prob}(> z) = 1 - (1 - e^{-z})^M \tag{5.43}$$

is the *false-alarm probability* of the null hypothesis, that is, the *significance level* of any peak in $P_N(\omega)$ that we do see. A small value for the false-alarm probability indicates a highly significant periodic signal.

To evaluate the significance, we need to know M. A typical procedure will be to plot $P_N(\omega)$ as a function of many closely spaced frequencies in some large frequency range. How many of them are independent?

Before answering, let us first see how accurately we need to know M. The interesting region is where the significance is a small (significant) number, $\ll 1$. There (5.43) can be expanded in series to give

$$\text{Prob}(> z) \approx M\, e^{-z} \tag{5.44}$$

We see that the significance scale linearly with M. Practical significance levels are numbers like 0.05, 0.01, 0.001, etc, therefore our estimate of M need not to be accurate.

The results of Monte Carlo experiments aimed to determine M show that M depends on the number of frequencies sampled, the number of data points N, and their detailed spacing. It turns out that M is very nearly equal to N when the data points are approximately equally spaced and when the sampled frequencies "fill" (oversample) the frequency range $[0, \Omega_{\text{Nyq}}]$. Figure 5.12 shows the results of applying a Lomb-Scargle periodogram to a set of $N = 100$ data points, Poissonian distributed in time. There is certainly no sinusoidal signal evident to the eye. The lower plot shows $P_N(\omega)$ against the frequency $\nu = \omega/2\pi$. The Nyquist frequency that would obtain if the points were evenly spaced is at $\nu_{\text{Nyq}} = 0.5$ Hz. Since we have searched up to about twice that frequency, and oversampled the ν's to the point where successive values of $P_N(\omega)$ vary smoothly, we take $M = 2N$. One see a highly significant peak at a frequency of 0.81 Hz. This is indeed the frequency of the sine wave used to create the data (we have to take our word for this!).

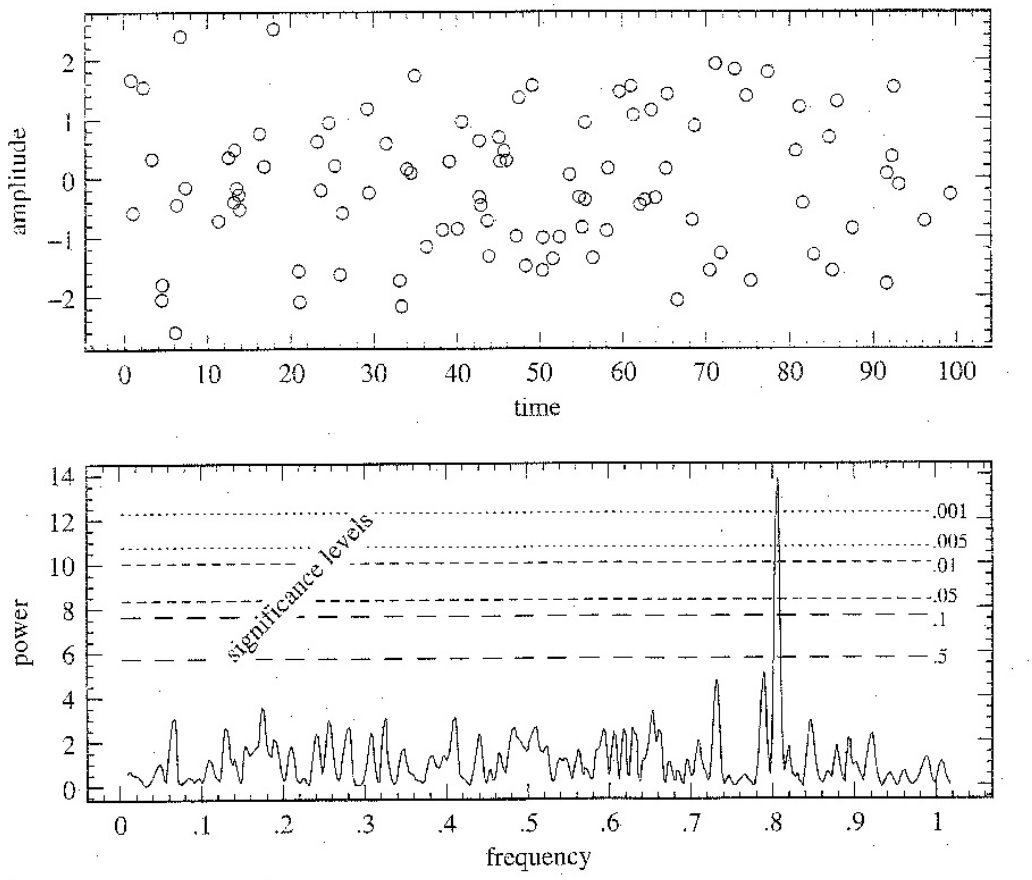


Figure 5.12: Example of the Lomb-Scargle algorithm in action. The 100 data points *(upper panel)* are at random times between 0 and 100. Their sinusoidal component is readily uncovered *(lower panel)* by the algorithm, at a significance level $p < 0.001$. If the 100 data points had been evenly spaced at unit interval, the Nyquist frequency would have been 0.5 Hz. Note that, for these unevenly spaced points, there is no visible aliasing into the Nyquist range

Note that two other peaks approach but not exceed the 50% significance level: that is about what one might expect by chance. It is also worth commenting on the fact that the significant peak was found *above the Nyquist frequency* and without any significant aliasing down into the Nyquist interval. That would not be possible for evenly spaced data.

5.10 Analysis of a Coherent Signal

In X–ray astronomy (and in astronomy in general) the detection of coherent signal is quite common: for example, we detect periodic signal as due to star pulsations, pulse periods in pulsars, orbital modulations and eclipses, precession.
Two methods of analysis are used to examine data for evidence for periodic signals: FFT and epoch folding. In general, both techniques have certain advantages and disadvantages in their application. There latter are worsened both by the presence of gaps in the data and the large number of statistically independent frequencies which could, in principle, be examined.
Epoch folding is more sensitive to non sinusoidal pulse shapes encountered in X–ray astronomy. Furthermore, the technique is relatively insensitive to randomly occurring gaps in the data so long as the net pulse phase coverage is reasonably uniform. Epoch folding is, however, extremely computer time-consuming (even if now the increased CPU power of the current computers makes this issue less important).
The FFT, on the other hand, is extremely efficient. However, the FFT is difficult to interpret in the presence of gaps in the data (and in this case is better to use the Lomb-Scargle periodogram technique, as discussed in Section 5.9).
The epoch folding consists in folding the data modulo a trial period and then grouping the observations according to phase, in order to obtain a high signal-to-noise profile. The χ^2 statistics is then used to test the high signal-to-noise profile for uniformity. This statistic is χ^2_{n-1} distributed, where n is the number of phase bins. By varying the trial period we can build a χ^2 vs period diagram and find out the one that gives the maximum χ^2 (that is, the rejection of the uniformity hypothesis). Because the χ^2 distribution resembles a triangular distribution (see Figure 5.13), it can be often well be fit by a Gaussian function, the mean of which may be considered the *"best"* coherent period P present in the data. The FWHM of the χ^2 distribution should be of the order of $\sim P^2/T$, where T is the total elapsed time of the observation.
Of course this method works if there are not intrinsic period variations, like the one due to orbital motions. In this case it is necessary to perform a time transformation that makes the signal coherent. This transformation is called a *timing model*. The timing model predicts a model profile, or *template*, that is correlated to the *average* profile so that a phase offset can be determined. When multiplied by the instantaneous pulse period, that phase yields a time offset that can be added to a high-precision reference point on the profile (for example, the edge of the profile) to create the time-of-arrival or TOA, as shown in Figure 5.14. The general procedure to derive information on the source from the measured TOAs is depicted in Figure 5.15.
The TOA of the pulse number n is, by definition,

$$t_n = t_0 + n\,P \tag{5.45}$$

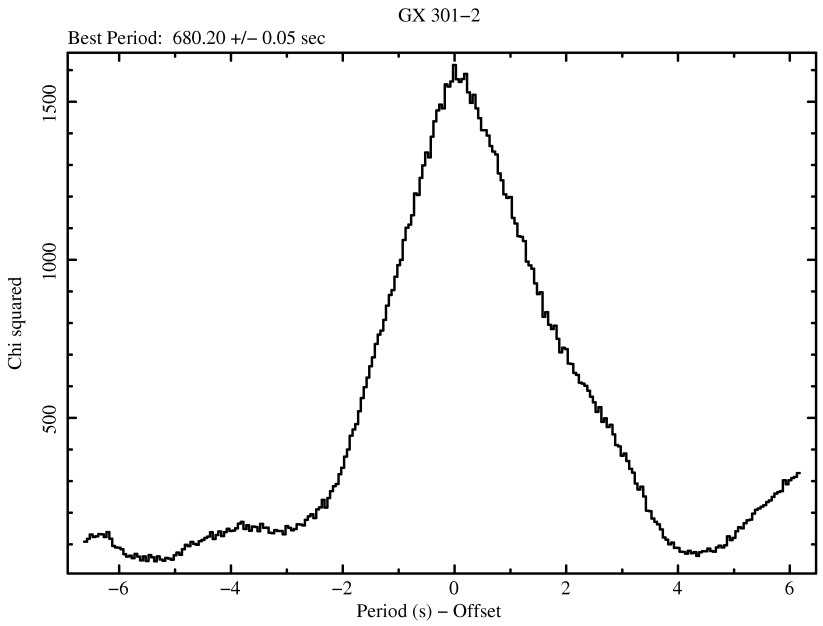


Figure 5.13: Pulse period of the X–ray binary pulsar GX 301–2 obtained by means of the epoch folding technique

where t_0 is a reference time (usually the start of the observation). When including intrinsic period variations we can perform a Taylor expansion in P and write down

$$t_n = t_0 + n\,P + \frac{1}{2}n^2\,P\dot{P} + \frac{1}{6}n^3\,P^2\ddot{P} + \cdots \tag{5.46}$$

Equation (5.46) can be inverted and expressed in terms of the pulse phase φ at time $(t - t_0)$

$$\varphi = \varphi_0 + f_0(t - t_0) + \frac{1}{2}\dot{f}(t - t_0)^2 + \frac{1}{6}\ddot{f}(t - t_0)^3 + \cdots \tag{5.47}$$

where $f_0 = 1/P$ and f is the frequency. The precision with which a TOA can be determined is approximately equal to the duration of a sharp pulse feature (e.g., the leading edge) divided by the signal-to-noise ratio of the average profile. It is usually expressed in terms of the width of the pulse features W_f in units of the period P, the pulse period P, and the signal-to-noise ratio SNR such that $\sigma_\mathrm{TOA} \propto W_\mathrm{f}P/\mathrm{SNR}$. Therefore strong, fast pulsars with narrow pulse profiles provide the best arrival times.

Before proceeding, it is better to eliminate the very first cause of period variations: the motion of the Earth and/or of the spacecraft. This is done by referencing the TOAs to a nearly inertial reference frame: the Solar System barycenter (see Figure 5.16).

The variation of the TOAs due to the orbital motion of the pulsar in the binary system can be written as an addition term in Equation (5.46) of the form

Figure 5.15: General procedure to derive information on the source from the measured TOAs

Figure 5.14: Measuring the phase shift with respect to a template profile

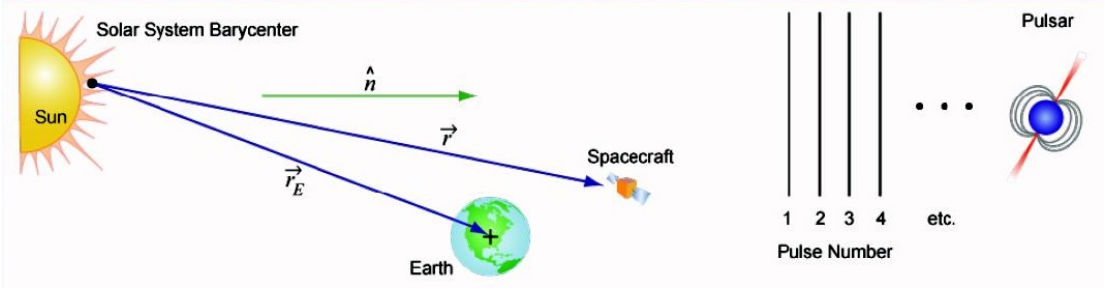


Figure 5.16: The time of arrivals of pulses from celestial objects are referenced to the nearly inertial reference frame of the Solar System barycenter

$$\cdots + \frac{a_x \sin i}{c} F(e, \omega, \tau, \theta) \tag{5.48}$$

where $a_x \sin i$ is the projected semi-major axis of the pulsating star orbiting with an inclination angle i in the binary system. The function $F(e, \omega, \tau, \theta)$ represents the eccentric orbit of the pulsar around the center of mass of the binary system, where e is the eccentricity, ω the longitude of the periastron, τ the time of periastron passage, and $\theta \equiv 2\pi(t-\tau)/P_{\text{orb}}$ is the mean anomaly (see Figure 5.17 for the definition of the orbital parameters). In particular, we have that

$$F(e, \omega, \tau, \theta) = (1 - e^2) \frac{\sin(\upsilon + \omega)}{1 + e \cos \upsilon} \tag{5.49}$$

where the true anomaly υ can be calculated from the observed mean anomaly θ using the relations

$$\tan \frac{\upsilon}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2}; \qquad E - e \sin E = \theta \tag{5.50}$$

The last relation is the so-called *Kepler Equation*. By fitting the series of TOAs with Equation (5.46) plus (5.48), we are able to obtain the pulse period P at time t_0, its time derivatives $\dot{P}$ and $\ddot{P}$, together with the orbital parameters of the pulsar $a \sin i$, e, ω, τ and P_{orb}. An example of the Doppler curve, representing the pulse delay times due to the light transit across the binary orbit, is shown in Figure 5.18 for the X–ray binary pulsar Hercules X–1.

5.10.1 Determination of Neutron Star Masses

X–ray binary pulsars are an important laboratory for measurements of astrophysical quantities. One of the most important is the determination of the mass of the neutron star orbiting its companion. From the orbital parameters determined in the previous section, the *mass function* can be expressed as

$$f(M) = \frac{4\pi^2 (a_x \sin i)^3}{G\, P_{\text{orb}}^2} = 1.1 \times 10^{-3} \left(\frac{a_x \sin i}{1 \text{ lt-s}} \right)^3 \left(\frac{P_{\text{orb}}}{1 \text{ d}} \right)^{-2} M_\odot \tag{5.51}$$

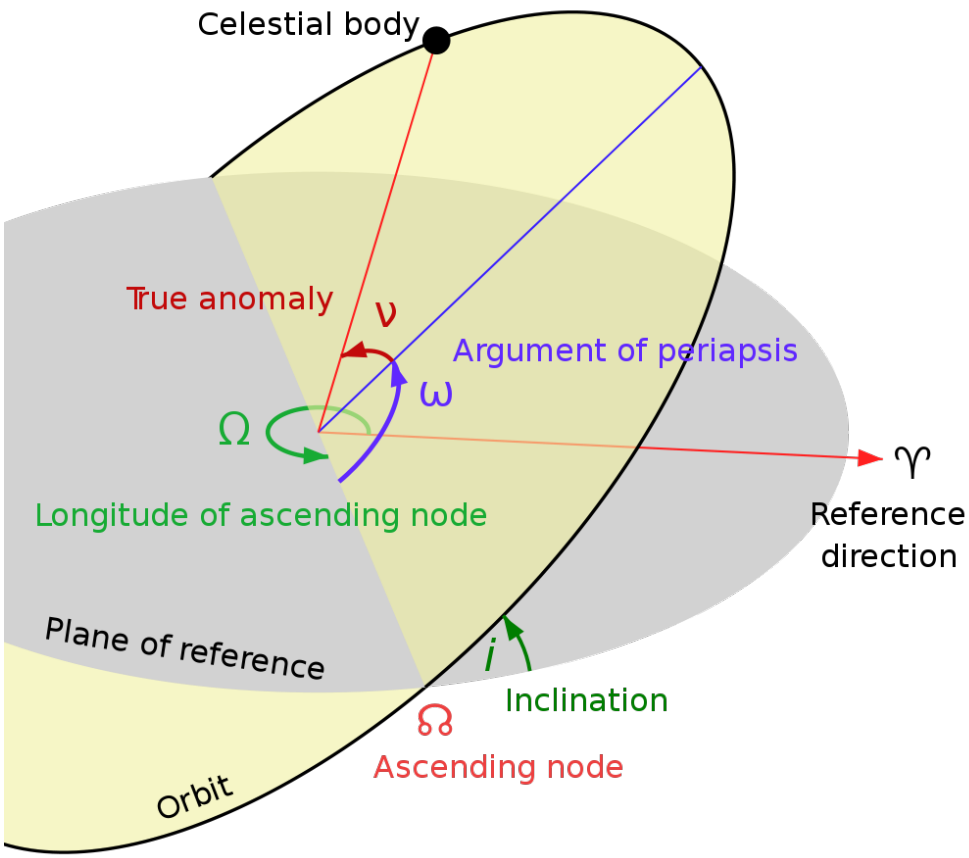


Figure 5.17: In this diagram, the orbital plane (yellow) intersects a reference plane (gray). The intersection is called the line of nodes, as it connects the center of mass with the ascending and descending nodes

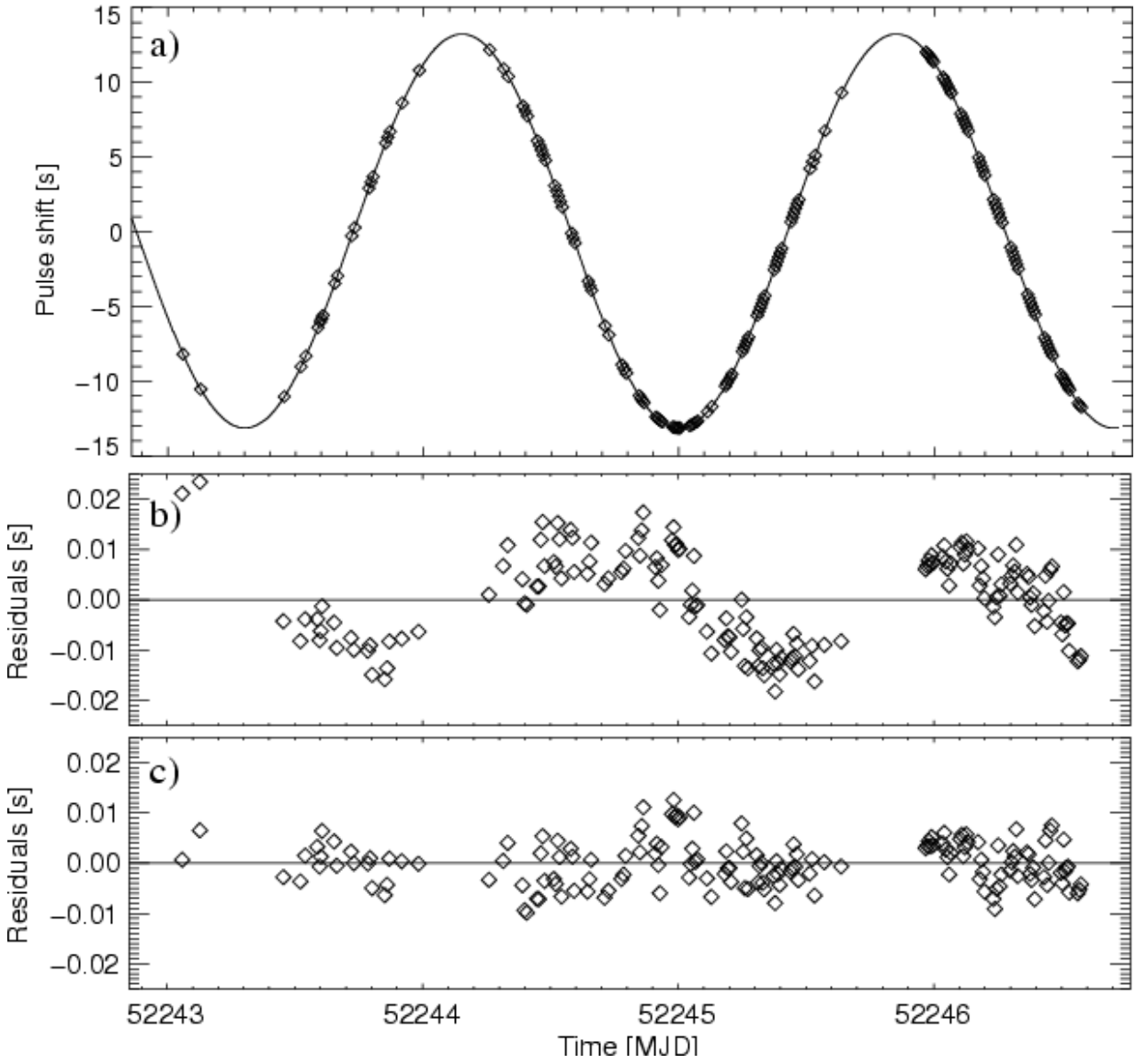


Figure 5.18: Delays of the TOA in Her X–1 due to its orbital motion and theoretical sine curve for the 35 day orbital period. The residuals refer to different orbital parameters solutions. From Staubert *et al.* 2009

where G is the gravitational constant. The ratio of the neutron star mass M_x to that of the companion star M_c is obtained as

$$\begin{aligned} q &\equiv \frac{M_x}{M_c} = \frac{K_c\, P_{\rm orb}\sqrt{1-e^2}}{2\pi a_x \sin i} \\ &= 4.6\times 10^{-2}\sqrt{1-e^2}\left(\frac{K_c}{1\ \mathrm{Km\ s^{-1}}}\right)\left(\frac{P_{\rm orb}}{1\ \mathrm{d}}\right)\left(\frac{a_x \sin i}{1\ \mathrm{lt\text{-}s}}\right)^{-1} \end{aligned}$$

where K_c is the semi-amplitude of the Doppler velocity curve of the companion star as measured from optical observations.
Because in the Doppler fitting we are able to determine only the product $a_x \sin i$, we need an independent estimate of the inclination angle i. This can be obtained if the source shows X–ray eclipses. Indeed, the average radius of the companion star R_c is related to the eclipse half-angle Θ_c as

$$R_c = a\ \left(\cos^2 i + \sin^2 i \sin^2 \Theta_c\right)^{1/2} \tag{5.52}$$

where $a = a_x + a_c$ is the separation of the centers of mass of the two stars in the binary system. Putting the things together, we can estimate the inclination angle in terms of the critical Roche-lobe radius R_L as

$$\sin i = \frac{\left[1-\beta^2\left(\frac{R_L}{a}\right)^2\right]^{1/2}}{\cos\Theta_c} \tag{5.53}$$

where $\beta \equiv R_c/R_L$. With all this information it is possible to *resolve* the binary system and derive an estimate of the mass of the neutron star. All measured masses are consistent at 4σ with a maximum mass of 1.5–1.65 $M_\odot$.

Chapter 6

Bibliography

Reference Textbooks

- **Random Data: Analysis and Measurement Procedures, 4th Edition**, Bendat J.S. and Piersol A.G., Wiley-Interscience (2010)
- **Fourier Transformation for Pedestrian**, Butz T., Springer-Verlag (2006)
- **Handbook of Pulsar Astronomy**, Lorimer D.R. and Kramer M., Cambridge University Press (2004)
- **The Fourier Transform and its Applications**, Osgood Brad G., 30 video lectures (about 50 min each) available online from https://see.stanford.edu/Course/EE261
- **Numerical Recipes: The Art of Scientific Computing**, Press W.H., Flannery B.P., Teukolsky S.A. and Vetterling W.T. Cambridge University Press (2007) available online from http://www.numerical.recipes/book
- **Fourier Techniques in X-ray Timing**, van der Klis M., in *Timing neutron Stars*, H. Ögelman and E.P.J. van den Heuvel (eds), Kluwer Academic Pub., p.27–69 (1989) available online from http://dare.uva.nl/record/218477

In-depth Readings

Belloni T, Psaltis D., van der Klis M. 2002. *A unified description of the timing features of accreting X–ray binaries*. ApJ 572, 392

Davies S.R. 1990. *An improved test for periodicity.* MNRAS 244, 93

Harris F.J. 1978. *On the use of windows for harmonic analysis with the discrete Fourier transform*. Proc. IEEE 66, 51

Jerri A.J. 1977. *The Shannon sampling theorem – Its various extensions and applications: A tutorial review.* Proc. IEEE 65, 1565

Leahy D.A., Darbro W., Elsner R.F., Weisskopf M.C., Sutherland P.G., Kahn S., Grindlay J.E. 1983. *On searches for pulsed emission with application to four globular clusters X–ray sources: NGC 1851, 6441, 6624 and 6712.* ApJ 266, 160

Lomb N.R. 1976. *Least-squares frequency analysis of unequally spaced data.* Astrophys. Space Sci. 39, 447

Nowak M.A., Vaughan B.A., Wilms J., Dove J.B., Begelman M.C. 1999. *RXTE observation of Cygnus X–1: II. Timing analysis.* ApJ 510, 874

Scargle J.D. 1982. *Studies in astronomical time series analysis. II. Statistical aspect of spectral analysis of unevenly spaced data.* ApJ 263, 835

Shannon C.E. 1998. *Communication in the presence of noise.* Proc. IEEE 86, 447

Staubert R., Klochkov D., Wilms J. 2009. *Updating the orbital ephemeris of Hercules X-1: Rate of decay and eccentricity of the orbit.* A&A 500, 883

VanderPlas J.T. 2018. *Understanding the Lomb-Scargle Periodogram.* ApJ Suppl. Ser. 236, 16

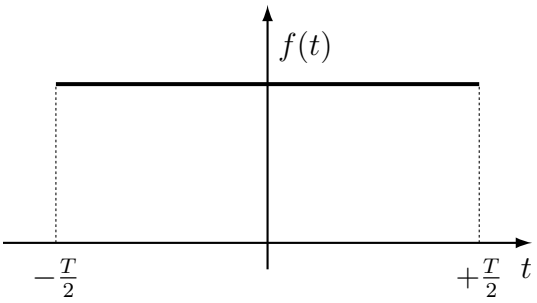
Vaughan B.A., van der Klis M., Wood K.S., Norris J.P., Hertz P., Michelson P.F., van Paradijs J., Lewin W.H.G., Mitsuda K., Penninx W. 1994. *Searches for millisecond pulsations in low-mass X-ray binaries. II.* ApJ 435, 362

Appendix A

Examples Shown at the Blackboard

2.1 Calculation of Fourier Coefficients — Page 15

Constant Function



$$f(t) = 1 \qquad \text{for } -T/2 \le t \le +T/2$$

From the definitions of the Fourier coefficients A_k (2.8) and A_0 (2.9) we have

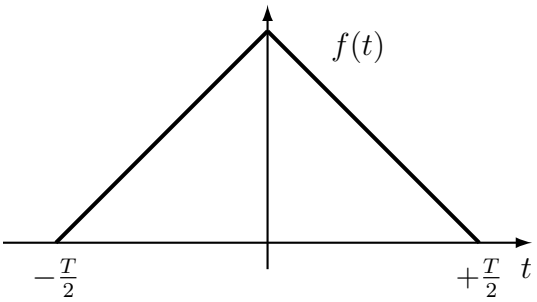
$$A_k = \frac{2}{T}\int_{-T/2}^{+T/2} f(t)\cos\omega_k t\,dt = \frac{2}{T}\int_{-T/2}^{+T/2}\cos\omega_k t\,dt = 0 \qquad \text{for } k \neq 0$$

$$A_0 = \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\,dt = \frac{1}{T}\int_{-T/2}^{+T/2} dt = 1$$

From the definition of the Fourier coefficients B_k (2.10) we have

$$
\begin{aligned}
B_k &= \frac{2}{T}\int_{-T/2}^{+T/2} f(t)\sin\omega_k t\,dt \\
&= \frac{2}{T}\int_{-T/2}^{+T/2} \sin\omega_k t\,dt \\
&= 0 \qquad \text{for all } k
\end{aligned}
$$

Triangular Function



$$
f(t) = \begin{cases} 1 + \dfrac{2t}{T} & \text{for } -T/2 \le t \le 0 \\ 1 - \dfrac{2t}{T} & \text{for } 0 \le t \le +T/2 \end{cases}
$$

From the definition of the Fourier coefficient A_0 (2.9) we have

$$
\begin{aligned}
A_0 &= \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\,dt \\
&= \frac{1}{T}\int_{-T/2}^{0}\left(1+\frac{2t}{T}\right)dt + \frac{1}{T}\int_{0}^{+T/2}\left(1-\frac{2t}{T}\right)dt \\
&= \frac{2}{T}\int_{0}^{+T/2}\left(1-\frac{2t}{T}\right)dt \\
&= \frac{2}{T}\left[t-\frac{t^2}{T}\right]_0^{T/2} = \frac{2}{T}\left[\frac{T}{2}-\frac{T^2}{4T}\right] = 1-\frac{1}{2} \\
&= \frac{1}{2}
\end{aligned}
$$

From the definition of the Fourier coefficient A_k (2.8) we have

$$
\begin{aligned}
A_k &= \frac{2}{T}\left[\int_{-T/2}^{0}\left(1+\frac{2t}{T}\right)\cos\left(\frac{2\pi kt}{T}\right)\,dt + \int_{0}^{+T/2}\left(1-\frac{2t}{T}\right)\cos\left(\frac{2\pi kt}{T}\right)\,dt\right] \\
&= \underbrace{\frac{2}{T}\int_{-T/2}^{0}\cos\left(\frac{2\pi kt}{T}\right)\,dt + \frac{2}{T}\int_{0}^{+T/2}\cos\left(\frac{2\pi kt}{T}\right)\,dt}_{=0} + \\
&\quad \frac{4}{T^2}\int_{-T/2}^{0} t\,\cos\left(\frac{2\pi kt}{T}\right)\,dt - \frac{4}{T^2}\int_{0}^{+T/2} t\,\cos\left(\frac{2\pi kt}{T}\right)\,dt \\
&= -\frac{8}{T^2}\int_{0}^{+T/2} t\,\cos\left(\frac{2\pi kt}{T}\right)\,dt
\end{aligned}
$$

The last integral can be solved by parts:

$$\int x\,\cos ax\,dx = \frac{x}{a}\sin ax + \frac{1}{a^2}\cos ax$$

Therefore we finally have

$$A_k = \frac{2(1-\cos\pi k)}{\pi^2 k^2}$$

Remember that $B_k = 0$ because $f(t)$ is an even function. We can rewrite the A_k coefficients in the form

$$A_k = \begin{cases} \frac{1}{2} & \text{for } k = 0 \\ \frac{4}{\pi^2 k^2} & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even, } k \neq 0 \end{cases}$$

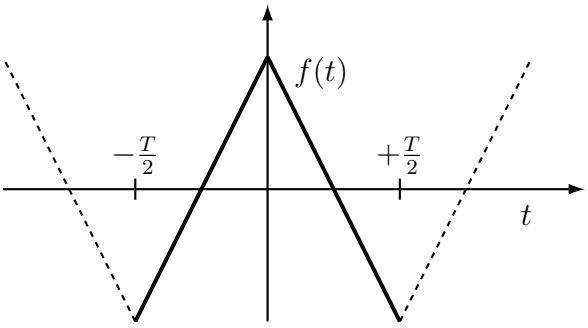
In formula, we have

$$f(t) = \frac{1}{2} + \frac{4}{\pi^2}\left(\cos\omega t + \frac{1}{9}\cos 3\omega t + \frac{1}{25}\cos 5\omega t + \cdots\right)$$

This is the function plotted in Figure 2.2 at page 16.

2.2 Shifting rules — Page 21

Triangular function with average equal to zero



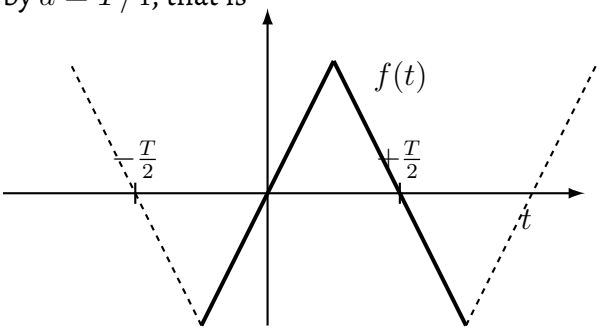
$$f(t) = \begin{cases} \frac{1}{2} + \frac{2t}{T} & \text{for } -T/2 \le t \le 0 \\ \frac{1}{2} - \frac{2t}{T} & \text{for } 0 \le t \le +T/2 \end{cases}$$

This function is obtained by subtracting $1/2$ from the "normal" triangular function described in the previous example. By applying the linearity theorem (2.19), its Fourier coefficients A_k remain the same, while the coefficient A_0 becomes zero. Therefore we have

$$f(t) = \frac{4}{\pi^2}\left(\cos\omega t + \frac{1}{9}\cos 3\omega t + \frac{1}{25}\cos 5\omega t + \cdots\right)$$

Quarter period shifted triangular function

Now take the previous triangular function (with average equal to zero) and shift it to the right by $a = T/4$, that is



$$f_{\text{new}} = f_{\text{old}}(t - T/4)$$

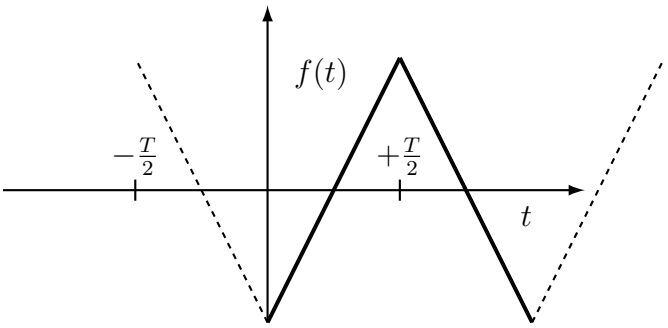
So the new coefficients can be calculated from the first shifting rule (2.20)

$$\begin{aligned} C_k^{\text{new}} &= C_k^{\text{old}}\, e^{-i\omega_k a} \overset{a=T/4}{=} C_k^{\text{old}}\, e^{-i\pi k/2} \\ &= \frac{2}{\pi^2 k^2}\left(\cos\frac{\pi k}{2} - i\,\sin\frac{\pi k}{2}\right) \\ &= -\frac{2i}{\pi^2 k^2}\,(-1)^{(k-1)/2} \qquad k \text{ odd} \end{aligned}$$

Because $C_{-k}^{\text{new}} = -C_k^{\text{new}}$ it follows that $A_k = 0$. Using the identity $iB_k = C_{-k} - C_k$ we have

$$B_k = \frac{4}{\pi^2 k^2}(-1)^{(k-1)/2} \qquad k \text{ odd}$$

Half period shifted triangular function



$$f_{\text{new}} = f_{\text{old}}(t - T/2)$$

Again, by applying the first shifting rule (2.20) with $a = T/2$ we have

$$\begin{aligned} C_k^{\text{new}} &= C_k^{\text{old}}\, e^{-i\omega_k a} \overset{a=T/2}{=} C_k^{\text{old}}\, e^{-i\pi k} \\ &= \frac{2}{\pi^2 k^2}(\cos \pi k - i\, \sin \pi k) \\ &= -\frac{2}{\pi^2 k^2} \qquad k \text{ odd} \end{aligned}$$

So we have only changed the sign (indeed, the function is now upside-down).

2.3 Second Shifting Rule — Page 22

Constant function

$$f(t) = 1 \qquad \text{for } -T/2 \leq t \leq +T/2$$

In this case we already know that $A_k = \delta_{k,0}$, or $A_0 = 1$, and all the other A_k and B_k are null. Now, let us multiply the function $f(t)$ by $\cos \omega t$, that is $a = 1$. From the second shifting rule (2.21) we have

$$A_k^{\text{new}} = \delta_{k-1,0} \qquad \Longrightarrow \qquad A_1 = 1 \text{ (all others are zero)}$$

By using the C_k coefficients we have

$$C_1 = \frac{1}{2} \qquad C_{-1} = \frac{1}{2}$$

So, we have shifted the coefficient by $a = 1$ and gone halves. This example demonstrates that the frequency $\omega = 0$ is as good as any other function. If we know, for example, the Fourier series of a function $f(t)$ and consequently the solution for the integrals of the form

$$\int_{-T/2}^{+T/2} f(t)\, e^{-i\omega_k t}\, dt$$

then we have already solved, by means of the second shifting rule, all integrals for $f(t)$ multiplied by $\sin \omega t$ or $\cos \omega t$: it is sufficient to combine phase factor $e^{i2\pi at/T}$ with phase factor $e^{-i\omega_k t}$.

Triangular function multiplied by cosine

Our "standard" triangular function

$$f(t) = \begin{cases} 1 + \dfrac{2t}{T} & \text{for } -T/2 \leq t \leq 0 \\ \\ 1 - \dfrac{2t}{T} & \text{for } 0 \leq t \leq +T/2 \end{cases}$$

will be now multiplied by $\cos(\pi t/T)$, i.e. we shift the coefficients C_k by $a = 1/2$. The new function is still even, and therefore we have only to compute the A_k because the B_k are null:

$$A_k^{\text{new}} = \frac{A_{k+a}^{\text{old}} + A_{k-a}^{\text{old}}}{2}$$

We have already computed the A_k^{old}:

$$A_k^{\text{old}} = \frac{2(1 - \cos \pi k)}{\pi^2 k^2}$$

therefore we have

$$
\begin{aligned}
A_k^{\text{new}} &= \frac{1}{2}\left[\frac{2(1-\cos\pi(k+1/2))}{\pi^2(k+1/2)^2} + \frac{2(1-\cos\pi(k-1/2))}{\pi^2(k-1/2)^2}\right] \\
&= \frac{1-\cos\pi k\cos(\pi/2)+\sin\pi k\sin(\pi/2)}{\pi^2(k+1/2)^2} + \\
&\quad \frac{1-\cos\pi k\cos(\pi/2)-\sin\pi k\sin(\pi/2)}{\pi^2(k-1/2)^2} \\
&= \frac{1}{\pi^2(k+1/2)^2} + \frac{1}{\pi^2(k-1/2)^2} \\
A_0^{\text{new}} &= \frac{A_{1/2}^{\text{old}}}{2} = \frac{2(1-\cos(\pi/2))}{2\pi^2(1/2)^2} = \frac{4}{\pi^2}
\end{aligned}
$$

The new coefficients are therefore

$$
\begin{aligned}
A_0 &= \frac{4}{\pi^2} \\
A_1 &= \frac{1}{\pi^2}\left(\frac{1}{(3/2)^2}+\frac{1}{(1/2)^2}\right) = \frac{4}{\pi^2}\left(\frac{1}{9}+\frac{1}{1}\right) = \frac{4}{\pi^2}\frac{10}{9} \\
A_2 &= \frac{1}{\pi^2}\left(\frac{1}{(5/2)^2}+\frac{1}{(3/2)^2}\right) = \frac{4}{\pi^2}\left(\frac{1}{25}+\frac{1}{9}\right) = \frac{4}{\pi^2}\frac{34}{225} \\
A_3 &= \frac{1}{\pi^2}\left(\frac{1}{(7/2)^2}+\frac{1}{(5/2)^2}\right) = \frac{4}{\pi^2}\left(\frac{1}{49}+\frac{1}{25}\right) = \frac{4}{\pi^2}\frac{74}{1225}
\end{aligned}
$$

A comparison of these coefficients with the one without the $\cos(\pi t/T)$ weighting shows what we have done

	w/o Weight	with Weight
A_0	$\frac{1}{2}$	$\frac{4}{\pi^2}$
A_1	$\frac{4}{\pi^2}$	$\frac{4}{\pi^2}\frac{10}{9}$
A_2	0	$\frac{4}{\pi^2}\frac{34}{225}$
A_3	$\frac{4}{\pi^2}\frac{1}{9}$	$\frac{4}{\pi^2}\frac{74}{1225}$

In Figure A.1 we show the triangular function, the weighting cosine function, and their product. We can make the following observations:

- ❏ The average A_0 got somewhat smaller, as the rising and falling flanks were weighted with the cosine, which, except for $t = 0$, is less than 1.

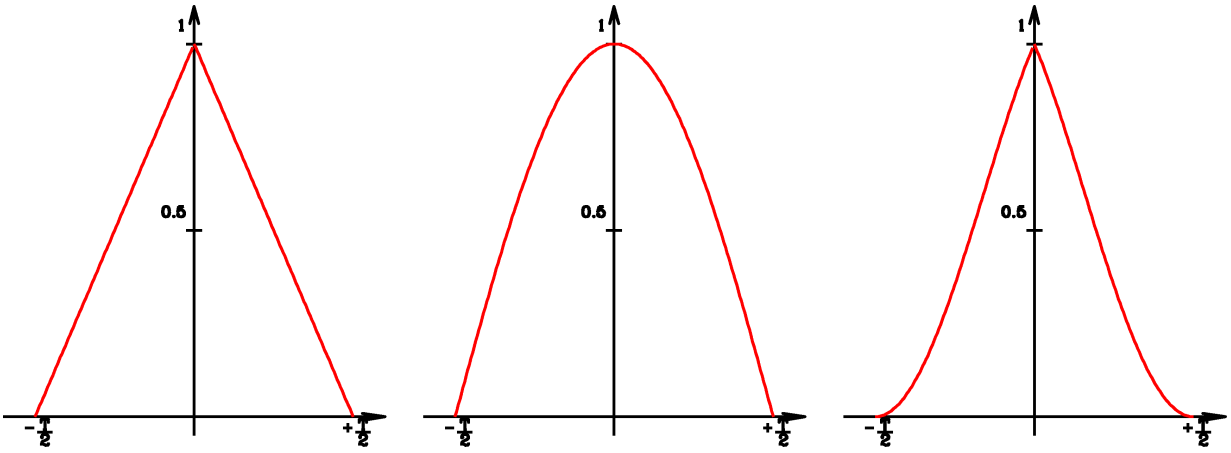


Figure A.1: The triangular function *(left);* The weighting function $\cos(\pi t/T)$ *(center);* Their product *(right)*.

- We raised coefficient A_1 a bit, but lowered all following odd coefficients a bit, too. This is evident straight away, if we convert:

$$\frac{1}{(2k+1)^2} + \frac{1}{(2k-1)^2} < \frac{1}{k^2} \qquad \Longrightarrow \qquad 8k^4 - 10k^2 + 1 > 0$$

  This is not valid for $k = 1$, yet all bigger k.

- Now we have been landed with even coefficients, that were null before.

We now have twice as many terms in the series as before, though they go down at an increased rate when k increases. The multiplication by $\cos(\pi t/T)$ caused the kink at $t = 0$ to turn into a much more pointed "spike". This should actually make for a worsening of convergence or a slower rate of decrease of the coefficients. We have, however, rounded the kink at the interval boundary $\pm T/2$, which naturally helps, but we could not reasonably have predicted what exactly was going to happen.

2.4 Approximating the triangular function — Page 24

The triangular function

$$f(t) = \begin{cases} 1 + \dfrac{2t}{T} & \text{for } -T/2 \leq t \leq 0 \\ 1 - \dfrac{2t}{T} & \text{for } 0 \leq t \leq +T/2 \end{cases}$$

has the mean squared "signal":

$$\frac{1}{T}\int_{-T/2}^{+T/2} f^2(t)\,dt = \frac{2}{T}\int_{0}^{+T/2} f^2(t)\,dt = \frac{2}{T}\int_{0}^{+T/2}\left(1+\frac{2t}{T}\right)^2 dt = \frac{1}{3}$$

The most coarse approximation is

$$S_0 = \frac{1}{2} \qquad \Longrightarrow \qquad \delta_0^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} = 0.0833\ldots$$

The next approximation results in

$$S_1 = \frac{1}{2} + \frac{4}{\pi^2}\cos\omega t \qquad \Longrightarrow \qquad \delta_1^2 = \frac{1}{3} - \frac{1}{4} - \frac{1}{2}\left(\frac{4}{\pi^2}\right)^2 = 0.0012\ldots$$

For δ_3^2 we get 0.0001915…, the approximation of the partial sum to the "triangle" quickly gets better and better.

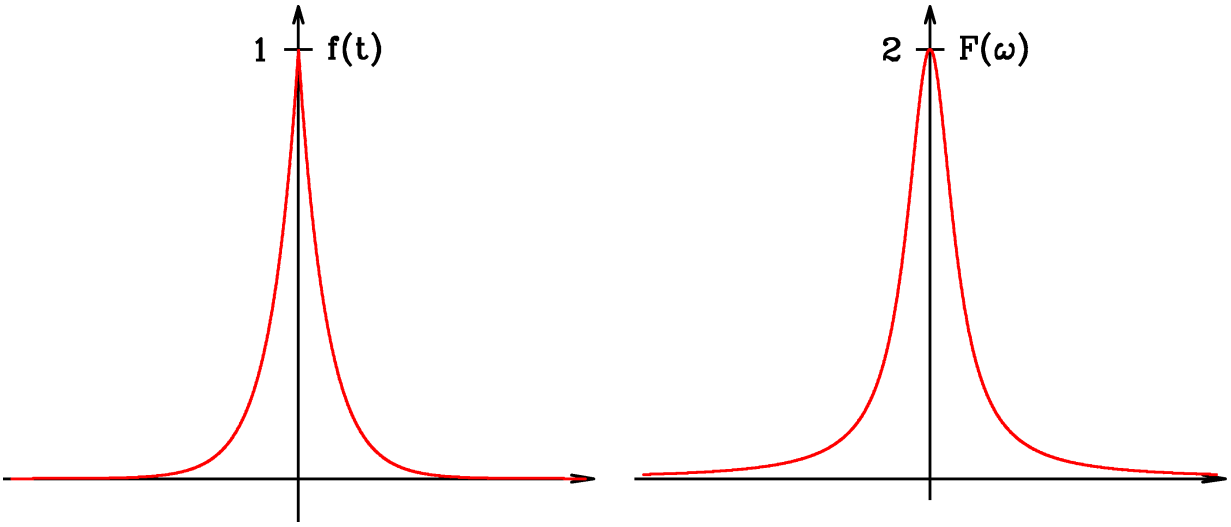


Figure A.2: The bilateral exponential function *(left)* and its Fourier transform *(right)*

2.5 Fourier Transformation of Relevant Functions — Page 28

The bilateral exponential function

$$f(t) = e^{-|t|/\tau}$$

Its Fourier transform is

$$F(\omega) = \int_{-\infty}^{+\infty} e^{-|t|/\tau}\, e^{-i\omega t}\, dt = 2\int_{0}^{+\infty} e^{-t/\tau}\, \cos\omega t\, dt = \frac{2\tau}{1+\omega^2\tau^2}$$

As $f(t)$ is even, the imaginary part of its Fourier transform is null. The Fourier transform of the bilateral exponential function is a Lorentzian. Both functions are shown in Figure A.2.

Unilateral exponential function

$$f(t) = \begin{cases} e^{-\lambda t} & \text{for } t \geq 0 \\ 0 & \text{else} \end{cases}$$

Its Fourier transform is

$$\begin{aligned} F(\omega) &= \int_0^{+\infty} e^{-\lambda t}\, e^{-i\omega t}\, dt \\ &= \left| \frac{e^{-(\lambda+i\omega)t}}{-(\lambda+i\omega)} \right|_0^{+\infty} \\ &= \frac{1}{\lambda+i\omega} \\ &= \frac{\lambda}{\lambda^2+\omega^2} - i\frac{\omega}{\lambda^2+\omega^2} \end{aligned}$$

$F(\omega)$ is complex, as $f(t)$ is neither even nor odd. We now can write the real and the imaginary parts separately. The real part has a Lorentzian shape we are familiar with by now, and the imaginary part has a dispersion shape.
Please note that $|F(\omega)|$ is no Lorentzian! If we want to "stick" to this property, we better represent the square of the magnitude: $|F(\omega)|^2 = 1/(\lambda^2+\omega^2)$, that is a Lorentzian. This representation is often also called the power representation: $|F(\omega)|^2 =$ (real part)^2 + (imaginary part)^2. The phase goes to 0 at the maximum of $|F(\omega)|$, i.e. when "in resonance".
Warning: The representation of the magnitude as well as of the squared magnitude does away with the linearity of the Fourier transformation!
Finally, let us try out the inverse transformation and find out how we return to the "unilateral" exponential function (the Fourier transform did not look all that "unilateral"!):

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\lambda - i\omega}{\lambda^2+\omega^2}\, e^{+i\omega t}\, d\omega \\ &= \frac{1}{2\pi} \left\{ 2\lambda \int_0^{+\infty} \frac{\cos\omega t}{\lambda^2+\omega^2}\, d\omega + 2\int_0^{+\infty} \frac{\omega \sin\omega t}{\lambda^2+\omega^2}\, d\omega \right\} \\ &= \frac{1}{\pi} \left\{ \frac{\pi}{2} e^{-|\lambda t|} \pm \frac{\pi}{2} e^{-|\lambda t|} \right\} \qquad \text{where } \begin{matrix} + & \text{for } t \geq 0 \\ - & \text{else} \end{matrix} \text{ is valid} \\ &= \begin{cases} e^{-\lambda t} & \text{for } t \geq 0 \\ 0 & \text{else} \end{cases} \end{aligned}$$

2.6 Convolution — Page 34

Gaussian frequency distribution

Let us assume we have $f(t) = \cos\omega_0 t$, and the frequency ω_0 is not precisely defined, but is Gaussian distributed:

$$P(\omega) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{\omega^2}{\sigma^2}}$$

What we are really measuring is

$$\bar{f}(t) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{\omega^2}{\sigma^2}} \cos(\omega - \omega_0)t \, d\omega \tag{A.1}$$

i.e. a convolution integral in ω_0. Instead of calculating this integral directly, we use the inverse of the Convolution Theorem (2.55), thus saving work and gaining higher enlightenment. But watch it! We have to handle the variables carefully. The time t in (A.1) has nothing to do with the Fourier transformation we need in (2.55). And the same is true for the integration variable ω. Therefore, we rather use t_0 and ω_0 for the variable pairs in (2.55). We identify:

$$\begin{aligned} F(\omega_0) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{\omega_0^2}{\sigma^2}} \\ \frac{1}{2\pi} G(\omega_0) &= \cos\omega_0 t \end{aligned}$$

The inverse Fourier transform of $F(\omega_0)$ and $G(\omega_0)$ are

$$\begin{aligned} f(t_0) &= \frac{1}{2\pi} e^{-\frac{1}{2}\sigma^2 t_0^2} \\ g(t_0) &= 2\pi \left[\frac{\delta(t_0 - t)}{2} + \frac{\delta(t_0 + t)}{2} \right] \end{aligned}$$

Finally we get:

$$h(t_0) = e^{-\frac{1}{2}\sigma^2 t_0^2} \left[\frac{\delta(t_0 - t)}{2} + \frac{\delta(t_0 + t)}{2} \right]$$

Now the only thing left is to Fourier transform $h(t_0)$:

$$\begin{aligned} \bar{f}(t) \equiv H(\omega_0) &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\sigma^2 t_0^2} \left[\frac{\delta(t_0 - t)}{2} + \frac{\delta(t_0 + t)}{2} \right] e^{-i\omega_0 t_0} \, dt_0 \\ &= e^{-\frac{1}{2}\sigma^2 t^2} \cos\omega_0 t \end{aligned}$$

Now, this was more work than we had originally thought it would be. But look at what we have gained in insight!
This means: the convolution of a Gaussian distribution in the frequency domain results in exponential "damping" of the cosine term, where the damping happens to be the Fourier transform of the frequency distribution. This, of course, is due to the fact that we have chosen to use a cosine function (i.e. a basis function) for $f(t)$. $P(\omega)$ makes sure that oscillations for $\omega \neq \omega_0$ are slightly shifted with respect to each other, and will more and more superimpose each other destructively in the long run, averaging out to 0.

Lorentzian frequency distribution

If we convolute our signal $f(t) = \cos \omega_0 t$ with a Lorentzian distribution

$$P(\omega) = \frac{\sigma}{\pi} \frac{1}{\omega^2 + \sigma^2}$$

then, by following what we have done previously with the Gaussian, we have

$$\begin{aligned} \bar{f}(t) &= \int_{-\infty}^{+\infty} \frac{\sigma}{\pi} \frac{1}{\omega^2 + \sigma^2} \cos(\omega - \omega_0)t \, d\omega \\ h(t_0) = \mathrm{FT}^{-1}(\bar{f}(t)) &= e^{-\sigma t_0} \left[\frac{\delta(t_0 - t)}{2} + \frac{\delta(t_0 + t)}{2} \right] \\ \bar{f}(t) &= e^{-\sigma t_0} \cos \omega_0 t \end{aligned}$$

This is a damped wave.

Gaussian convoluted with a Gaussian

We perform a convolution of a Gaussian with σ_1 with another Gaussian with σ_2. As the Fourier transforms are Gaussians again — yet with σ_1^2 and σ_2^2 in the *numerator* of the exponent — it's immediately obvious that $\sigma_{\text{total}}^2 = \sigma_1^2 + \sigma_2^2$. Therefore, we get another Gaussian with geometric addition of the widths σ_1 and σ_2.

3.1 Discrete Fourier Transformation — Page 56

Constant function with $N = 4$

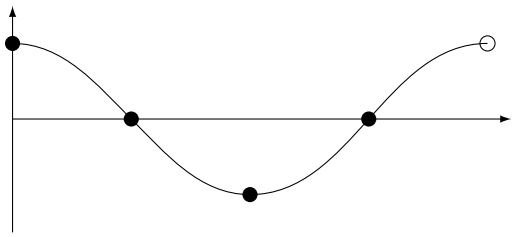
$$f_k = 1 \qquad \text{for } k = 0, 1, 2, 3$$

For the continuous Fourier transformation we expect a δ-function with the frequency $\omega = 0$. The discrete Fourier transformation therefore will only result in $F_0 \neq 0$. Indeed, by using (3.11) — or even a lot smarter using (3.8):

$$\begin{aligned} F_0 &= \frac{1}{4} 4 = 1 \\ F_1 &= 0 \\ F_2 &= 0 \\ F_3 &= 0 \end{aligned}$$

As $\{f_k\}$ is an even series, $\{F_j\}$ contains no imaginary part.

Cosine function with $N = 4$

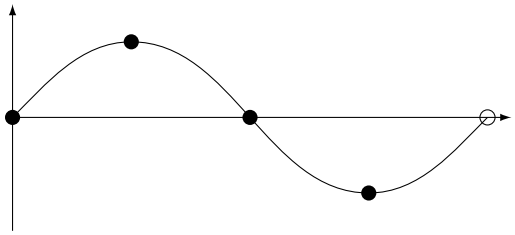


$$\begin{aligned} f_0 &= 1 \\ f_1 &= 0 \\ f_2 &= -1 \\ f_3 &= 0 \end{aligned}$$

We get, using (3.11) and $W_4 = i$

$$\begin{aligned} F_0 &= 0 \text{ (this is the average)} \\ F_1 &= \frac{1}{4}(1 + (-1)(-1)) = \frac{1}{2} \\ F_2 &= \frac{1}{4}(1 + (-1)(1)) = 0 \\ F_3 &= \frac{1}{4}(1 + (-1)(-1)) = \frac{1}{2} \end{aligned}$$

Sine function with $N = 4$



$$
\begin{aligned}
f_0 &= 0 \\
f_1 &= 1 \\
f_2 &= 0 \\
f_3 &= -1
\end{aligned}
$$

We get, using (3.11) and $W_4 = i$

$$
\begin{aligned}
F_0 &= 0 \text{ (this is the average)} \\
F_1 &= \frac{1}{4}\left(-i + (-1)(i)\right) = -\frac{i}{2} \\
F_2 &= \frac{1}{4}\left(1(-1) + (-1)(-1)\right) = 0 \\
F_3 &= \frac{1}{4}\left((1)(i) + (-1)(-i)\right) = \frac{i}{2}
\end{aligned}
$$

Discrete Fourier Transform: Shifting Rules

3.2 Shifted cosine with $N = 2$ — Page 57

First, let us compute the DFT for the cosine function:

$$\begin{aligned} \{f_k\} &= \{0, 1\} \qquad \text{or} \\ f_k &= \frac{1}{2}(1 - \cos \pi k) \qquad k = 0, 1 \end{aligned}$$

$$W_2 = e^{i\pi} = -1$$

$$\begin{aligned} F_0 &= \frac{1}{2}(0 + 1) = \frac{1}{2} \\ F_1 &= \frac{1}{2}(0 + 1(-1)) = -\frac{1}{2} \end{aligned}$$

$$\{F_j\} = \left\{\frac{1}{2}, -\frac{1}{2}\right\}$$

Now we shift the input by $n = 1$:

$$\begin{aligned} \{f_k^{\text{shifted}}\} &= \{1, 0\} \qquad \text{or} \\ f_k &= \frac{1}{2}(1 + \cos \pi k) \qquad k = 0, 1 \end{aligned}$$

$$\{F_j^{\text{shifted}}\} = \left\{\frac{1}{2}W_2^{-1\times 0}, \frac{1}{2}W_2^{-1\times 1}\right\} = \left\{\frac{1}{2}, \frac{1}{2}\right\}$$

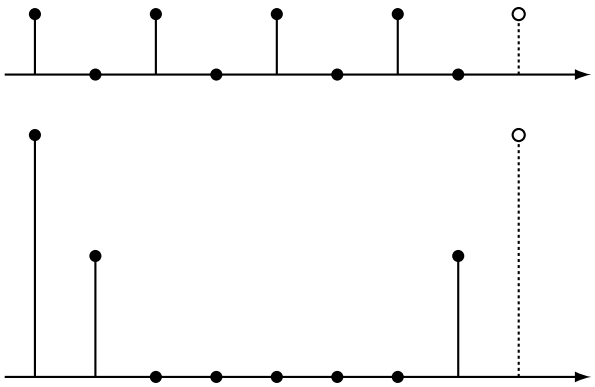
3.3 Modulated cosine with $N = 2$ — Page 58

We want to modulate the input with W_N^{-nk}, with $n = 1$. From its definition $W_2^{-k} = (-1)^{-k}$, therefore

$$\begin{aligned} \{f_k^{\text{shifted}}\} &= \{0, -1\} \qquad \text{or} \\ f_k &= \frac{1}{2}(-1 + \cos \pi k) \qquad k = 0, 1 \end{aligned}$$

$$\{F_j^{\text{shifted}}\} = \{F_{j-1}\} = \left\{-\frac{1}{2}, \frac{1}{2}\right\}$$

3.4 Nyquist Frequency with $N = 8$ — Page 61



$$\{f_k\} = \{1, 0, 1, 0, 1, 0, 1, 0\}$$

$$\{g_k\} = \{4, 2, 0, 0, 0, 0, 0, 2\}$$

The "resolution function" $\{g_k\}$ is padded to $N = 8$ with zeros and normalized to $\sum_{k=0}^{7} = 8$. The convolution of $\{f_k\}$ with $\{g_k\}$ results in:

$$\{h_k\} = \left\{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\}$$

meaning that everything is "flattened", because the resolution function (here triangular shaped) has a full half-width of $2\Delta t$ and consequently does not allow the recording of oscillations with the period Δt. The Fourier transform therefore is $H_k = 1/\delta_{k,0}$. Using the convolution theorem (3.1) we would get

$$\{F_j\} = \left\{\frac{1}{2}, 0, 0, 0, \frac{1}{2}, 0, 0, 0\right\}$$

The result is easy to understand: the average is $1/2$, at the Nyquist frequency we have $1/2$, all the other elements are zero. The Fourier transform of $\{g_k\}$ is

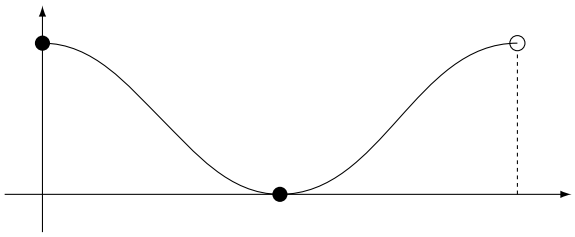
$$G_0 = 1 \qquad G_1 = \frac{1}{2} + \frac{\sqrt{2}}{4} \qquad G_2 = \frac{1}{2} \qquad G_3 = \frac{1}{2} - \frac{\sqrt{2}}{4}$$

$$G_4 = 0 \qquad G_5 = \frac{1}{2} - \frac{\sqrt{2}}{4} \qquad G_6 = \frac{1}{2} \qquad G_7 = \frac{1}{2} + \frac{\sqrt{2}}{4}$$

For the product we get $H_j = F_j G_j = 1/2, 0, 0, 0, 0, 0, 0, 0$, like we should for the Fourier transform. If we had taken the Convolution Theorem seriously right from the beginning, then the calculation of G_0 (average) and G_4 at the Nyquist frequency would have been quite sufficient, as all other $F_j = 0$.
The fact that the Fourier transform of the resolution function for the Nyquist frequency is 0, precisely means that with this resolution function we are not able to record oscillations with the Nyquist frequency any more. Our inputs, however, were only the frequency 0 and the Nyquist frequency.

3.5 The Sampling Theorem with $N = 2$ — Page 65

Let us star with the shifted cosine function



$$\{f_k\} = \{0, 1\}$$

We expect

$$f(t) = \frac{1}{2} + \frac{1}{2}\cos\Omega_{\text{Nyq}}t = \cos^2\frac{\Omega_{\text{Nyq}}t}{2}$$

The sampling theorem tell us

$$\begin{aligned}
f(t) &= \sum_{k=-\infty}^{+\infty} f_k \frac{\sin\Omega_{\text{Nyq}}(t-k\Delta t)}{\Omega_{\text{Nyq}}(t-k\Delta t)} \\
&\quad \text{with } f_k = \delta_{k,\text{even}} \text{ (with periodic continuation)} \\
&= \frac{\sin\Omega_{\text{Nyq}}t}{\Omega_{\text{Nyq}}t} + \sum_{l=1}^{\infty}\frac{\sin\Omega_{\text{Nyq}}(t-2l\Delta t)}{\Omega_{\text{Nyq}}(t-2l\Delta t)} + \sum_{l=1}^{\infty}\frac{\sin\Omega_{\text{Nyq}}(t+2l\Delta t)}{\Omega_{\text{Nyq}}(t+2l\Delta t)} \\
&\quad \text{with the substitution } k = 2l \\
&= \frac{\sin\Omega_{\text{Nyq}}t}{\Omega_{\text{Nyq}}t} + \sum_{l=1}^{\infty}\left[\frac{\sin 2\pi(\frac{t}{2\Delta t}-l)}{2\pi(\frac{t}{2\Delta t}-l)} + \frac{\sin 2\pi(\frac{t}{2\Delta t}+l)}{2\pi(\frac{t}{2\Delta t}+l)}\right] \\
&\quad \text{with } \Omega_{\text{Nyq}}\Delta t = \pi \\
&= \frac{\sin\Omega_{\text{Nyq}}t}{\Omega_{\text{Nyq}}t} + \frac{1}{2\pi}\sum_{l=1}^{\infty}\frac{(\frac{t}{2\Delta t}+l)\sin\Omega_{\text{Nyq}}t + (\frac{t}{2\Delta t}-l)\sin\Omega_{\text{Nyq}}t}{(\frac{t}{2\Delta t}+l)(\frac{t}{2\Delta t}-l)} \\
&= \frac{\sin\Omega_{\text{Nyq}}t}{\Omega_{\text{Nyq}}t} + \frac{\sin\Omega_{\text{Nyq}}t}{2\pi}\frac{2t}{2\Delta t}\sum_{l=1}^{\infty}\frac{1}{\left(\frac{t}{2\Delta t}\right)^2 - l^2} \\
&= \frac{\sin\Omega_{\text{Nyq}}t}{\Omega_{\text{Nyq}}t}\left(1 + \left(\frac{\Omega_{\text{Nyq}}t}{2\pi}\right)^2 2\sum_{l=1}^{\infty}\frac{1}{\left(\frac{\Omega_{\text{Nyq}}t}{2\pi}\right)^2 - l^2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sin \Omega_{\text{Nyq}} t}{\Omega_{\text{Nyq}} t} \pi \frac{\Omega_{\text{Nyq}} t}{2\pi} \cot \frac{\pi \Omega_{\text{Nyq}} t}{2\pi} \\
&= \sin \Omega_{\text{Nyq}} t \frac{1}{2} \frac{\cos(\Omega_{\text{Nyq}} t/2)}{\sin(\Omega_{\text{Nyq}} t/2)} \\
&= 2 \sin(\Omega_{\text{Nyq}} t/2) \cos(\Omega_{\text{Nyq}} t/2) \frac{1}{2} \frac{\cos(\Omega_{\text{Nyq}} t/2)}{\sin(\Omega_{\text{Nyq}} t/2)} \\
&= \cos^2(\Omega_{\text{Nyq}} t/2)
\end{aligned}
$$

Please note that we actually need all the summation terms of k from $-\infty$ to $+\infty$. If we had taken only $K = 0$ and $k = 1$ into consideration we would have obtained

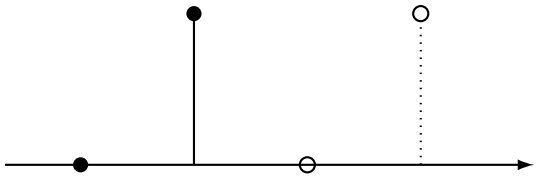
$$f(t) = 1 \frac{\sin \Omega_{\text{Nyq}} t}{\Omega_{\text{Nyq}} t} + 0 \frac{\sin \Omega_{\text{Nyq}} (t - \Delta t)}{\Omega_{\text{Nyq}} (t - \Delta t)} = \frac{\sin \Omega_{\text{Nyq}} t}{\Omega_{\text{Nyq}} t}$$

which would not correspond to the input of $\cos^2(\Omega_{\text{Nyq}} t/2)$. We still would have, as before, $f(0) = 1$ and $f(k\Delta t) = 0$, but for $0 < t < \Delta t$ we would not have interpolated correctly, as $\sin x/x$ decays slowly for large x while we want to get a periodic oscillation that does not decay as input.

3.6 Fast Fourier Transform of the Saw-tooth — Page 72

The case $N = 2$

$$f_0 = 0, \qquad f_1 = 1$$



Its "normal" Fourier transform is

$$\begin{aligned} W_2 &= e^{i\pi} = -1 \\ F_0 &= \frac{1}{2}(0+1) = \frac{1}{2} \\ F_1 &= \frac{1}{2}(0+1 \times W_2^{-1}) = -\frac{1}{2} \end{aligned}$$

Its FFT is

$$\begin{aligned} f_{1,0} = 0 \quad &\text{even part} \rightarrow \quad F_{1,0} = 0 \\ f_{2,0} = 1 \quad &\text{odd part} \rightarrow \quad F_{2,0} = 1 \qquad M = 1 \end{aligned}$$

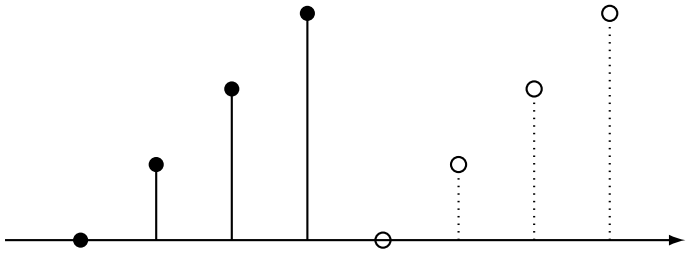
From (3.47) we get

$$\begin{aligned} F_0 &= \frac{1}{2}\left(F_{1,0} + F_{2,0}\underbrace{W_2^0}_{=1}\right) = \frac{1}{2} \\ F_1 &= \frac{1}{2}\left(F_{1,0} - F_{2,0}W_2^0\right) = -\frac{1}{2} \end{aligned}$$

This really did not save all that much work so far.

The case $N = 4$

$$\{f_k\} = \{0, 1, 2, 3\}$$



The "normal" Fourier transform gives

$$\begin{aligned}
W_4 &= e^{i2\pi/4} = e^{i\pi/2} = i \\
F_0 &= \frac{1}{4}(0+1+2+3) = \frac{3}{2} \\
F_1 &= \frac{1}{4}(W_4^{-1} + 2W_4^{-2} + 3W_4^{-3}) = \frac{1}{4}\left(\frac{1}{i} + \frac{2}{-1} + \frac{3}{-i}\right) = -\frac{1}{2} + \frac{i}{2} \\
F_2 &= \frac{1}{4}(W_4^{-2} + 2W_4^{-4} + 3W_4^{-6}) = \frac{1}{4}(-1+2-3) = -\frac{1}{2} \\
F_3 &= \frac{1}{4}(W_4^{-3} + 2W_4^{-6} + 3W_4^{-9}) = \frac{1}{4}\left(-\frac{1}{i} - 2 + \frac{3}{i}\right) = -\frac{1}{2} - \frac{i}{2}
\end{aligned}$$

Using 2 subdivisions, for the first subdivision the FFT gives us

$$\begin{aligned}
N &= 4 & \{f_1\} &= \{0,2\} \text{ even} \\
N &= 2 & \{f_2\} &= \{1,3\} \text{ odd}
\end{aligned}$$

For the second subdivision we get

$$\begin{aligned}
f_{11} &= 0 \quad \text{even} \quad \equiv F_{1,1,0} \\
f_{12} &= 2 \quad \text{odd} \quad \equiv F_{1,2,0} \\
f_{21} &= 1 \quad \text{even} \quad \equiv F_{2,1,0} \\
f_{22} &= 3 \quad \text{odd} \quad \equiv F_{2,2,0}
\end{aligned}$$

Using (3.47) this results in ($j = 0, M' = 1$)

$$\begin{aligned}
F_{1,k} &= \left\{ \overbrace{\frac{1}{2}F_{1,1,0} + \frac{1}{2}F_{1,2,0}}^{\text{upper part}}, \overbrace{\frac{1}{2}F_{1,1,0} - \frac{1}{2}F_{1,2,0}}^{\text{lower part}} \right\} = \{1,-1\} \\
F_{2k} &= \left\{ \frac{1}{2}F_{2,1,0} + \frac{1}{2}F_{2,2,0}, \frac{1}{2}F_{2,1,0} - \frac{1}{2}F_{2,2,0} \right\} = \{2,-1\}
\end{aligned}$$

And finally, using (3.47) once again

$$\begin{aligned}
\text{upper part} \rightarrow & \begin{cases} F_0 = \frac{1}{2}(F_{1,0} + F_{2,0}) = \frac{3}{2} \\ F_1 = \frac{1}{2}(F_{1,1} + F_{2,1}W_4^{-1}) = \frac{1}{2}\left(-1 + (-1) \times \frac{1}{i}\right) = -\frac{1}{2} + \frac{i}{2} \end{cases} \\
\text{lower part} \rightarrow & \begin{cases} F_2 = \frac{1}{2}(F_{1,0} - F_{2,0}) = -\frac{1}{2} \\ F_3 = \frac{1}{2}(F_{1,1} - F_{2,1}W_4^{-1}) = \frac{1}{2}\left(-1 - (-1) \times \frac{1}{i}\right) = -\frac{1}{2} - \frac{i}{2} \end{cases}
\end{aligned}$$

Appendix B

Practical Session on Timing Analysis

Many of the following examples are taken from a course on timing analysis given by Biswajit Paul during the first ASTROSAT Workshop. I really thank Biswajit for letting me use his material and passing me the data files.

In the following, we will use text in `typewriter` font for indicating the program output, while the user input will be in `red typewriter`.

B.1 First Look at a Light Curve

The XRONOS package, part of the more general HEADAS software, is formed by many programs that address specific tasks on temporal analysis (the list of all the tasks can be obtained giving the command `fhelp ftools`). One of the very first steps in performing a temporal analysis is the quick look of the data. The `lcurve` task plots one or more sequences of time-ordered data (thereafter *"light curve"*).

```
orma> lcurve

lcurve 1.0 (xronos5.22)

Number of time series for this task[1] 1

Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_long.lc
 Series 1 file    1:cenx-3_long.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3             Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10510 19:57:03.562
 No. of Rows .......      1873744        Bin Time (s) ......   0.1250
 Right Ascension ... 1.70313293E+02      Internal time sys.. Converted to TJD
```

```
 Declination ....... -6.06232986E+01     Experiment ........ XTE      PCA

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10510.83129123185  (days)      19:57: 3:562  (h:m:s:ms)

 Minimum Newbin Time   0.12500000      (s)
 for Maximum Newbin No..          2638849

 Default Newbin Time is:   645.00587    (s) (to have 1 Intv. of     512 Newbins)
 Type INDEF to accept the default value

Newbin Time or negative rebinning[3600] 3600

 Newbin Time ......     3600.0000      (s)
 Maximum Newbin No.               92

 Default Newbins per Interval are:          92
 (giving       1 Interval of           92 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[215] 92
 Maximum of       1 Intvs. with           92 Newbins of       3600.00     (s)
Name of output file[default] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

      92 analysis results per interval

 Intv    1   Start 10507  0:49:27
     Ser.1     Avg  3632.        Chisq 0.7848E+08   Var 0.1568E+07 Newbs.     84
               Min  224.6          Max  5219.    expVar  1.423      Bins1873744
PLT> line step
PLT> plot
PLT> cpd cenx-3_long_lc.ps/cps
PLT> plot
PLT> quit
```

In the top panel of Figure B.1 we show the light curve of the XTE/PCA observation, with data

rebinned at 3600 s (that is, each bin corresponds to one hour of data). The “hole” is due to the eclipse of the X–ray pulsar by the optical companion.
By running again the `lcurve` program with a binning time of 0.125 s (the original time resolution of the data) we can see the single pulses, as shown in the bottom panel of Figure B.1.

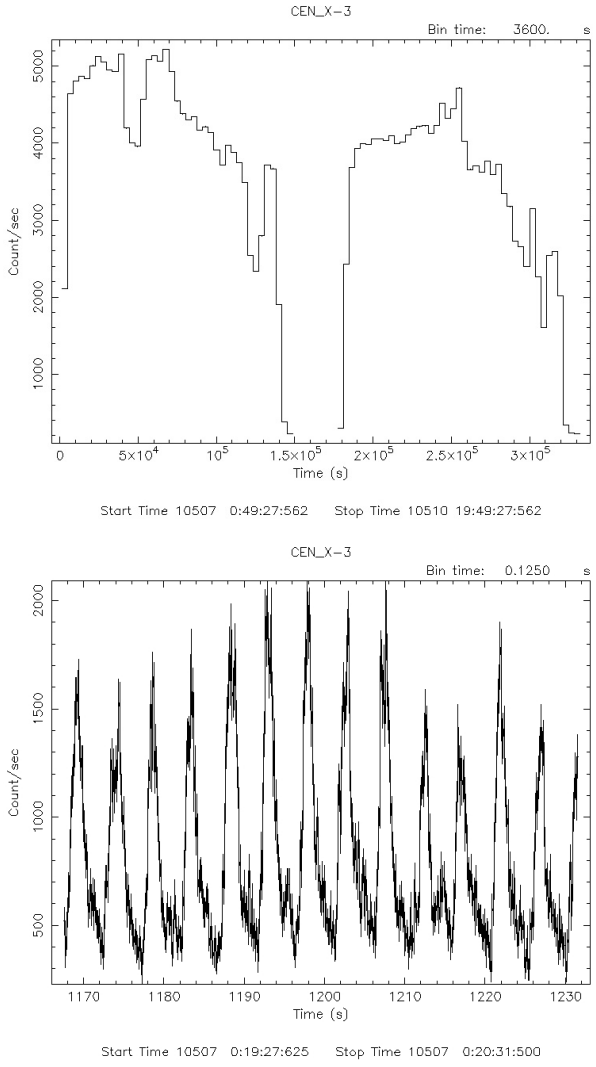


Figure B.1: Light curve of the X–ray binary pulsar Cen X–3 as observed by the PCA instrument aboard XTE. *Top:* data have been rebinned at 3600 s, in order to see the long-term behavior of the source (in this case the eclipse). *Bottom:* data have been rebinned at 0.125 s, in order to observe short-term variability, like the 4.8 s pulsation.

B.2 Finding a Periodicity by Fourier Analysis

We know that a periodic signal is seen as a sharp spike in power spectrum. We will compute the Power Spectral Density (PSD) on the Cen X–3 data in order to obtain its period. To this goal we will use the XRONOS command powspec.

```
orma> powspec

powspec 1.0 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_pca.lc
 Series 1 file    1:cenx-3_pca.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3             Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10507 02:24:27.559
 No. of Rows .......        60000        Bin Time (s) ......   0.1250
 Right Ascension ...                     Internal time sys.. Literal
 Declination .......                     Experiment ........

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - No

 Selected Columns:  1- Time;  3- Y-axis;  4- Y-error;  5- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10507.10031897535  (days)       2:24:27:559  (h:m:s:ms)

 **** Warning: Newbin Time must be an integer multiple of Minimum Newbin Time
 Minimum Newbin Time   0.12500000      (s)
 for Maximum Newbin No..            60000

 Default Newbin Time is:   1.0000000    (s) (to have 1 Intv. of    8192 Newbins)
 Type INDEF to accept the default value

Newbin Time or negative rebinning[0.0005] 0.125

 Newbin Time ......    0.12500000      (s)
 Maximum Newbin No.             60000

 Default Newbins per Interval are:        8192
 (giving       8 Intervals of         8192 Newbins each)
```

```
 Type INDEF to accept the default value

Number of Newbins/Interval[4096] 8192
 Maximum of       8 Intvs. with         8192 Newbins of      0.125000     (s)
Default intervals per frame are:         8
Type INDEF to accept the default value
Number of Intervals/Frame[16] 7
 Results from up to       7 Intvs. will be averaged in a Frame
Rebin results? (>1 const rebin, <-1 geom. rebin, 0 none)[0] 0
Name of output file[pippo] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/XW] /xw

    4096 analysis results per interval

 Intv    1   Start 10507  0:19:27
     Ser.1     Avg  1006.        Chisq 0.3487E+06   Var 0.3427E+06 Newbs.   8192
               Min  248.0          Max  3704.    expVar  8050.      Bins   8192
Power spectrum ready !
 Intv    2   Start 10507  0:36:31
     Ser.1     Avg  1892.        Chisq 0.8259E+06   Var 0.1526E+07 Newbs.   8192
               Min  416.0          Max  8544.    expVar 0.1514E+05  Bins   8192
Power spectrum ready !
 Intv    3   Start 10507  0:53:35
     Ser.1     Avg  2701.        Chisq 0.1234E+07   Var 0.3254E+07 Newbs.   8192
               Min  544.0          Max 0.1116E+05expVar 0.2160E+05  Bins   8192
Power spectrum ready !
 Intv    4   Start 10507  1:10:39
     Ser.1     Avg  3868.        Chisq 0.1730E+07   Var 0.6535E+07 Newbs.   8192
               Min  672.0          Max 0.1471E+05expVar 0.3094E+05  Bins   8192
Power spectrum ready !
 Intv    5   Start 10507  1:27:43
     Ser.1     Avg  4722.        Chisq 0.1111E+07   Var 0.8197E+07 Newbs.   5120
               Min  1176.          Max 0.1633E+05expVar 0.3778E+05  Bins   5120
Power spectrum ready !
 Intv    6   Start 10507  1:54:23
     Ser.1     Avg  4806.        Chisq 0.2097E+07   Var 0.9842E+07 Newbs.   8192
               Min  944.0          Max 0.1780E+05expVar 0.3845E+05  Bins   8192
Power spectrum ready !
 Intv    7   Start 10507  2:11:27
     Ser.1     Avg  4696.        Chisq 0.1486E+07   Var 0.8946E+07 Newbs.   6240
               Min  928.0          Max 0.2020E+05expVar 0.3757E+05  Bins   6240
Power spectrum ready !
PLT> r x 0.05 5
PLT> log x on
```

```
PLT> pl
PLT> cpd cenx-3_pca_psd.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

The PSD shows clearly two sharp peaks: the first corresponds to the pulse period, while the second corresponds to the first harmonic. In Figure B.2 we show the Cen X–3 PSD.
A Gaussian fit to the first Fourier peak gives a centroid frequency of $0.2079^{+0.0002}_{-0.0001}$ Hz, corresponding to a pulse period of $4.8100^{+0.0010}_{-0.0005}$ s. The first harmonics is at 0.4159 ± 0.0001 Hz, corresponding to 2.4044 ± 0.0002 s. Note that the ratio between the two frequencies is 2.0005 ± 0.0011.

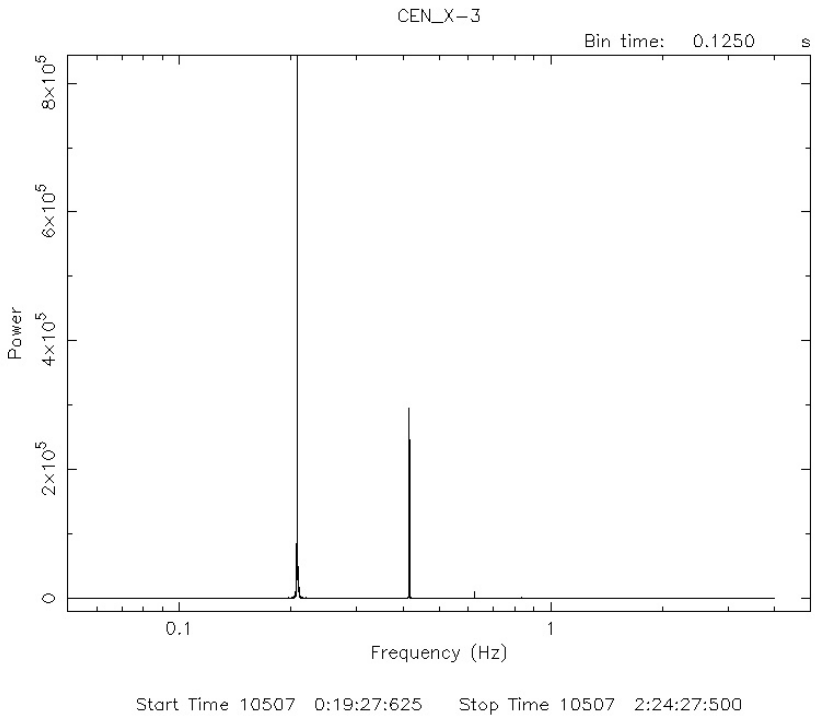


Figure B.2: Power spectrum of the XTE/PCA observation of Cen X–3. The fundamental and the first harmonic of the 4.8 s period are clearly visible. We used a logarithmic scale for the frequency for cosmetic purpose.

B.3 Finding a Periodicity by Epoch Folding

Instead of using the Fourier spectrum for obtaining the pulse period of Cen X–3, we will the so-called *epoch folding technique*. With this technique we build a template with a trial pulse period and cross-correlate this template with our light curve. Then we vary the trial pulse period until we obtain the better match with the original data. The `efsearch` program performs this task.

```
orma> efsearch

efsearch 1.1 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_pca.lc
 Series 1 file    1:cenx-3_pca.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3             Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10507 02:24:27.559
 No. of Rows .......        60000        Bin Time (s) ......   0.1250
 Right Ascension ...                     Internal time sys.. Literal
 Declination .......                     Experiment ........

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - No

 Selected Columns:  1- Time;  3- Y-axis;  4- Y-error;  5- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10507.10031897535  (days)       2:24:27:559  (h:m:s:ms)

 Default Epoch is:  10507.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[11723.00000] 10507.00000
 Period format is seconds.
Period[15.8] 4.8
 Expected Cycles ..           1562.50
 Default phase bins per period are:         8
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 16

 Newbin Time ......    0.30000000      (s)
 Maximum Newbin No.             25000
```

```
 Default Newbins per Interval are:       25000
 (giving       1 Interval of        25000 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[782811] 25000
 Maximum of       1 Intvs. with        25000 Newbins of      0.300000     (s)
 Default resolution is 0.1536000000E-02
 Type INDEF to accept the default value
Resolution for period search value or neg. power of 2[0.0001] 0.0001
 Default number of periods is       128
 Type INDEF to accept the default value
Number of periods to search[4096] 4096
Name of output file[pippo] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/XW] /xw

    4096 analysis results per interval

 Period :   4.809    dP/dt :   0.000
 Intv    1   Start 10507  0:19:27
     Ser.1     Avg  3257.        Chisq 0.7471E+07   Var 0.3719E+07 Newbs.     16
               Min  1177.          Max  7050.    expVar  7.966      Bins  52320
PLT> cpd cenx-3_pca_efs.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

In Figure B.3 we see the pulse period value that gives the better match (measured by a χ^2 test) between data folded at a trial period and the original data.
Once we know the pulse period, we can extract the average pulse profile by folding our data with the best trial value. The `efold` program performs this task on the light curve.

```
orma> efold

efold 1.1 (xronos5.22)

Number of time series for this task[1] 1
Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_pca.lc
 Series 1 file    1:cenx-3_pca.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3            Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10507 02:24:27.559
 No. of Rows .......        60000        Bin Time (s) ......   0.1250
```

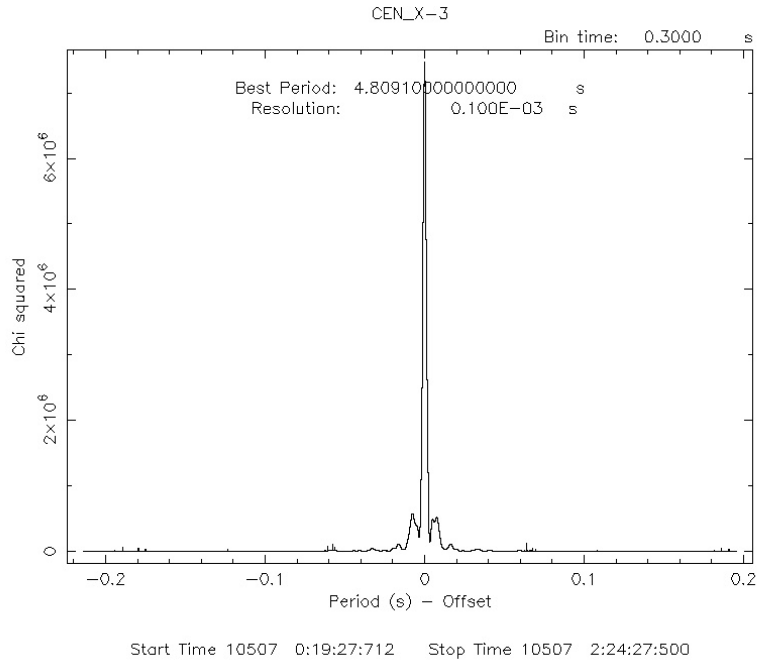


Figure B.3: χ^2 values as a function of the trial pulse period. The maximum in the χ^2 corresponds to the better match between the trial period and the original pulse period.

```
 Right Ascension ...                    Internal time sys.. Literal
 Declination .......                    Experiment ........

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - No

 Selected Columns:  1- Time;  3- Y-axis;  4- Y-error;  5- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10507.10031897535  (days)       2:24:27:559  (h:m:s:ms)

 Default Epoch is:  10507.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[ 0.1050700000E+05] 10507.00000
 Period format is seconds.
Period[4.8091] 4.8091
 Expected Cycles ..           1559.54
 Default phase bins per period are:        10
```

```
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 16

 Newbin Time ......    0.30056875      (s)
 Maximum Newbin No.             24953

 Default Newbins per Interval are:       24953
 (giving       1 Interval of        24953 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[24953] 24953
 Maximum of       1 Intvs. with        24953 Newbins of      0.300569     (s)
Default intervals per frame are:         1
Type INDEF to accept the default value
Number of Intervals/Frame[1] 1
 Results from up to       1 Intvs. will be averaged in a Frame
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/XW] /xw

      16 analysis results per interval

 Intv    1   Start 10507  0:19:27
     Ser.1     Avg  3257.        Chisq 0.7471E+07   Var 0.3719E+07 Newbs.     16
               Min  1177.          Max  7050.    expVar  7.966      Bins  52320
  Folded light curve ready
PLT> li st
PLT> pl
PLT> cpd cen_x-3_pca_efo.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

In Figure B.4 we show the original data folded at the best trial period.

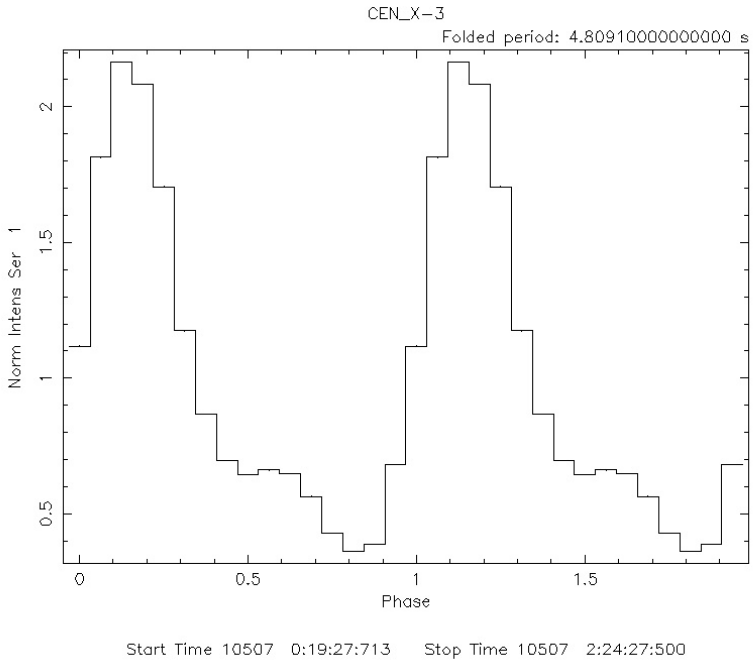


Figure B.4: The Cen X–3 pulse profile, obtained by folding the original data with the best trial period.

B.4 Variability in the Pulse Period

Up to now we have assumed that the pulse period is constant, but this is not always true! Indeed, many X–ray pulsars show changes of their spin period as a function of time, and this effect must be taken into account when determining the best pulse period. We will study this effect on the X–ray binary pulsar XTE J1946+274, observed by the Indian X–ray satellite IXAE. First, let us find the pulse period by means of the `efsearch` program **without** any period derivative (as we did in the previous example on Cen X–3).

```
orma> efsearch

efsearch 1.1 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] XTE_1946+274_ixae.lc
 Series 1 file    1:XTE_1946+274_ixae.lc
WARNING:  Defaulting to first FITS extension

 Selected FITS extensions: 1 - RATE TABLE;

 Source .............                    Start Time (d) .... 11723 18:03:40.786
 FITS Extension ....  1 - ‘          ‘   Stop Time (d) ..... 11732 16:47:25.887
 No. of Rows .......        32469        Bin Time (s) ......    1.020
```

```
 Right Ascension ...                    Internal time sys.. Converted to TJD
 Declination .......                    Experiment ........

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 11723.75255539586  (days)      18: 3:40:786  (h:m:s:ms)
 Expected Stop .... 11732.69960517328  (days)      16:47:25:887  (h:m:s:ms)

 Default Epoch is:  11723.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[10507.00000] 11723.00000
 Period format is seconds.
Period[4.81] 15.8
 Expected Cycles ..          48925.64
 Default phase bins per period are:         8
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 16

 Newbin Time ......    0.98750000      (s)
 Maximum Newbin No.            782811

 Default Newbins per Interval are:      782811
 (giving       1 Interval of       782811 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[20000] 782811
 Maximum of       1 Intvs. with       782811 Newbins of      0.987500     (s)
 Default resolution is 0.1614693713E-03
 Type INDEF to accept the default value
Resolution for period search value or neg. power of 2[0.0001] 0.0001
 Default number of periods is       128
 Type INDEF to accept the default value
Number of periods to search[1024] 4096
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

    4096 analysis results per interval
```

```
WARNING:  Defaulting to first FITS extension
  Chisq. vs. period ready

 Period :   15.77    dP/dt :   0.000
 Intv    1   Start 11723 18: 3:41
     Ser.1     Avg  34.32        Chisq  371.3       Var 0.3849     Newbs.     16
               Min  33.22          Max  35.24    expVar 0.1658E-01  Bins  32460
PLT> cpd XTE_1946+274_ixae_efs-1.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

As we can see in the top panel of Figure B.5, the result is as clean as we obtained for Cen X–3. This is because the observation contains many interruptions due to passages through regions of high concentration of particles (the satellite was in a polar orbit), where the X–ray instruments have to be switched off.
Now we can run again `efsearch` but this time we add a pulse period derivative $\dot{P} = 1.9 \times 10^{-9}$ s/s.

```
orma> efsearch dpdot=1.9e-9

efsearch 1.1 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] XTE_1946+274_ixae.lc
 Series 1 file    1:XTE_1946+274_ixae.lc
WARNING:  Defaulting to first FITS extension

 Selected FITS extensions: 1 - RATE TABLE;

 Source .............                    Start Time (d) .... 11723 18:03:40.786
 FITS Extension ....  1 - ‘          ‘   Stop Time (d) ..... 11732 16:47:25.887
 No. of Rows .......        32469        Bin Time (s) ......    1.020
 Right Ascension ...                     Internal time sys.. Converted to TJD
 Declination .......                     Experiment ........

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 11723.75255539586  (days)      18: 3:40:786  (h:m:s:ms)
```

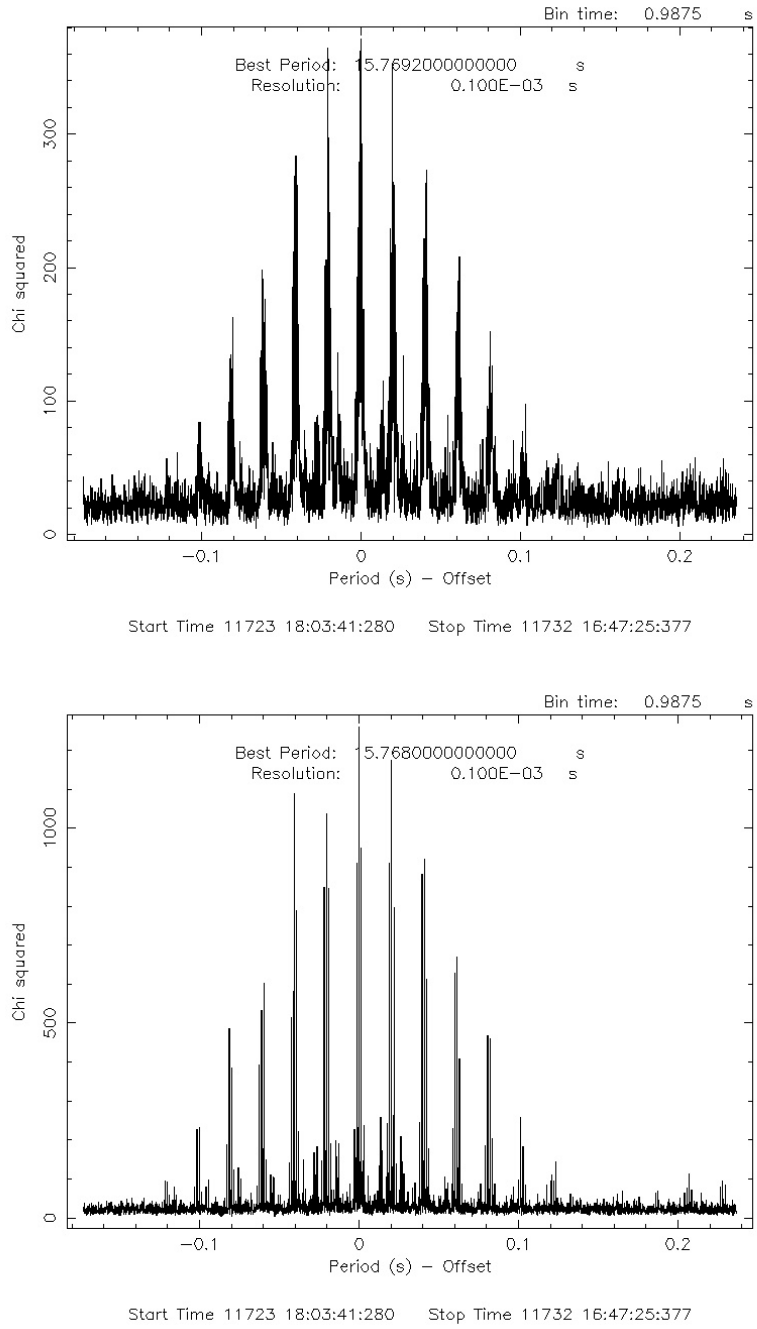


Figure B.5: Pulse period search for XTE J1946+274 as observed by the Indian X–ray satellite IXAE. The presence of many peaks is due to interruptions in the light curve. *Top:* search performed by assuming a constant pulse period. *Bottom:* search performed by assuming a pulse period derivative of 1.9×10^{-9} s/s. Note the different value of the χ^2 in the two plots.

```
 Expected Stop .... 11732.69960517328  (days)      16:47:25:887  (h:m:s:ms)

 Default Epoch is:  11723.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[11723.00000] 11723.00000
 Period format is seconds.
Period[15.8] 15.8
 Expected Cycles ..          48925.64
 Default phase bins per period are:         8
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 16

 Newbin Time ......    0.98750000      (s)
 Maximum Newbin No.            782811

 Default Newbins per Interval are:      782811
 (giving       1 Interval of       782811 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[782811] 782811
 Maximum of       1 Intvs. with       782811 Newbins of      0.987500     (s)
 Default resolution is 0.1614693713E-03
 Type INDEF to accept the default value
Resolution for period search value or neg. power of 2[0.0001] 0.0001
 Default number of periods is       128
 Type INDEF to accept the default value
Number of periods to search[4096] 4096
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

    4096 analysis results per interval

WARNING:  Defaulting to first FITS extension
  Chisq. vs. period ready

 Period :   15.77    dP/dt :  0.1900E-08
 Intv    1   Start 11723 18: 3:41
     Ser.1     Avg  34.32        Chisq  1259.       Var  1.306     Newbs.     16
               Min  32.20          Max  35.81    expVar 0.1658E-01  Bins  32460
PLT> cpd XTE_1946+274_ixae_efs-2.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

In the bottom panel of Figure B.5 we can see the result. By comparing the two figures we can see that the inclusion of the period derivative greatly increased the significance of the pulse period (note the change in the χ^2 value).

B.5 Effect of the Orbital Motion

Because of the Doppler effect, the pulse period changes around the binary orbit. In order to study this effect, we will determine the value of the pulse period along the orbit of Cen X–3. We will use the `efsearch` program as follows:

```
orma> efsearch

efsearch 1.1 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_long.lc
 Series 1 file    1:cenx-3_long.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3             Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10510 19:57:03.562
 No. of Rows .......      1873744        Bin Time (s) ......   0.1250
 Right Ascension ... 1.70313293E+02      Internal time sys.. Converted to TJD
 Declination ....... -6.06232986E+01     Experiment ........ XTE      PCA

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10510.83129123185  (days)      19:57: 3:562  (h:m:s:ms)

 Default Epoch is:  10507.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[10507.00000] 10507.00000
 Period format is seconds.
Period[4.8] 4.81
 Expected Cycles ..          68577.13
 Default phase bins per period are:         8
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 16

 Newbin Time ......    0.30062500      (s)
 Maximum Newbin No.           1097235
```

```
 Default Newbins per Interval are:     1097235
 (giving       1 Interval of      1097235 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[25000] 20000
 Maximum of      55 Intvs. with        20000 Newbins of      0.300625     (s)
 Default resolution is 0.1924000000E-02
 Type INDEF to accept the default value
Resolution for period search value or neg. power of 2[0.0001] 0.0001
 Default number of periods is       128
 Type INDEF to accept the default value
Number of periods to search[4096] 1024
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

    1024 analysis results per interval

 Period :   4.809    dP/dt :   0.000
 Intv    1   Start 10507  0:19:27
     Ser.1     Avg  2811.        Chisq 0.4935E+07   Var 0.2746E+07 Newbs.     16
               Min  1041.          Max  6049.    expVar  8.901      Bins  40420

 Period :   4.808    dP/dt :   0.000
 Intv    2   Start 10507  1:59:40
     Ser.1     Avg  4811.        Chisq 0.9896E+07   Var 0.8375E+07 Newbs.     16
               Min  1681.          Max 0.1036E+05expVar  13.56      Bins  45412

 Period :   4.808    dP/dt :   0.000
 Intv    3   Start 10507  3:39:52
     Ser.1     Avg  4849.        Chisq 0.1039E+08   Var 0.8569E+07 Newbs.     16
               Min  1700.          Max 0.1064E+05expVar  13.19      Bins  47076

 Period :   4.807    dP/dt :   0.000
 Intv    4   Start 10507  5:20: 5
     Ser.1     Avg  5051.        Chisq 0.1067E+08   Var 0.9123E+07 Newbs.     16
               Min  1802.          Max 0.1105E+05expVar  13.73      Bins  47076

 Period :   4.807    dP/dt :   0.000
 Intv    5   Start 10507  7: 0:17
     Ser.1     Avg  5068.        Chisq 0.1069E+08   Var 0.9063E+07 Newbs.     16
               Min  1819.          Max 0.1090E+05expVar  13.57      Bins  47780

 Period :   4.808    dP/dt :   0.000
 Intv    6   Start 10507  8:40:30
```

```
    Ser.1     Avg  4920.        Chisq 0.9356E+07   Var 0.8647E+07 Newbs.     16
              Min  1758.          Max 0.1077E+05expVar  14.79      Bins  42596

 Period :   4.808    dP/dt :   0.000
 Intv    7   Start 10507 10:20:42
    Ser.1     Avg  4620.        Chisq 0.7100E+07   Var 0.7644E+07 Newbs.     16
              Min  1668.          Max 0.1007E+05expVar  17.25      Bins  34276

 Period :   4.809    dP/dt :   0.000
 Intv    8   Start 10507 12: 0:55
    Ser.1     Avg  3982.        Chisq 0.5378E+07   Var 0.5737E+07 Newbs.     16
              Min  1430.          Max  8800.    expVar  17.06      Bins  29892

 Period :   4.810    dP/dt :   0.000
 Intv    9   Start 10507 14:21:19
    Ser.1     Avg  4611.        Chisq 0.6309E+07   Var 0.7686E+07 Newbs.     16
              Min  1706.          Max 0.1022E+05expVar  19.48      Bins  30308

 Period :   4.811    dP/dt :   0.000
 Intv   10   Start 10507 16: 1:32
    Ser.1     Avg  5124.        Chisq 0.6912E+07   Var 0.9109E+07 Newbs.     16
              Min  1940.          Max 0.1117E+05expVar  21.07      Bins  31132

 Period :   4.813    dP/dt :   0.000
 Intv   11   Start 10507 17:43:27
    Ser.1     Avg  5133.        Chisq 0.6841E+07   Var 0.8625E+07 Newbs.     16
              Min  2003.          Max 0.1091E+05expVar  20.30      Bins  32356

 Period :   4.814    dP/dt :   0.000
 Intv   12   Start 10507 19:23:40
    Ser.1     Avg  4877.        Chisq 0.6890E+07   Var 0.8052E+07 Newbs.     16
              Min  1915.          Max 0.1059E+05expVar  18.70      Bins  33380

 Period :   4.816    dP/dt :   0.000
 Intv   13   Start 10507 21: 3:52
    Ser.1     Avg  4409.        Chisq 0.6427E+07   Var 0.6692E+07 Newbs.     16
              Min  1680.          Max  9673.    expVar  16.65      Bins  33892

 Period :   4.817    dP/dt :   0.000
 Intv   14   Start 10507 22:44: 5
    Ser.1     Avg  4307.        Chisq 0.7132E+07   Var 0.6490E+07 Newbs.     16
              Min  1636.          Max  9452.    expVar  14.56      Bins  37860

 Period :   4.818    dP/dt :   0.000
 Intv   15   Start 10508  0:24:17
```

```
    Ser.1     Avg  4192.        Chisq 0.7590E+07   Var 0.6062E+07 Newbs.     16
              Min  1592.          Max  9134.    expVar  12.79      Bins  41956

 Period :   4.819    dP/dt :   0.000
 Intv   16   Start 10508  2: 4:30
    Ser.1     Avg  4115.        Chisq 0.8618E+07   Var 0.5896E+07 Newbs.     16
              Min  1570.          Max  8947.    expVar  10.95      Bins  48100

 Period :   4.820    dP/dt :   0.000
 Intv   17   Start 10508  3:44:42
    Ser.1     Avg  3756.        Chisq 0.7745E+07   Var 0.4840E+07 Newbs.     16
              Min  1472.          Max  8092.    expVar  9.998      Bins  48100

 Period :   4.820    dP/dt :   0.000
 Intv   18   Start 10508  5:24:55
    Ser.1     Avg  3931.        Chisq 0.8241E+07   Var 0.5450E+07 Newbs.     16
              Min  1464.          Max  8571.    expVar  10.58      Bins  47572

 Period :   4.821    dP/dt :   0.000
 Intv   19   Start 10508  7: 5: 7
    Ser.1     Avg  3773.        Chisq 0.7686E+07   Var 0.4822E+07 Newbs.     16
              Min  1428.          Max  8143.    expVar  10.04      Bins  48100

 Period :   4.821    dP/dt :   0.000
 Intv   20   Start 10508  8:45:20
    Ser.1     Avg  2619.        Chisq 0.3781E+07   Var 0.2059E+07 Newbs.     16
              Min  1077.          Max  5396.    expVar  8.707      Bins  38500

 Period :   4.820    dP/dt :   0.000
 Intv   21   Start 10508 10:25:32
    Ser.1     Avg  2648.        Chisq 0.3125E+07   Var 0.2137E+07 Newbs.     16
              Min  1059.          Max  5541.    expVar  10.94      Bins  31000

 Period :   4.820    dP/dt :   0.000
 Intv   22   Start 10508 12:40:15
    Ser.1     Avg  3702.        Chisq 0.4285E+07   Var 0.4416E+07 Newbs.     16
              Min  1404.          Max  7808.    expVar  16.48      Bins  28772

 Period :   4.819    dP/dt :   0.000
 Intv   23   Start 10508 14:20:28
    Ser.1     Avg  1814.        Chisq 0.2068E+07   Var 0.9796E+06 Newbs.     16
              Min  760.4          Max  3752.    expVar  7.583      Bins  30620

 Period :   4.817    dP/dt :   0.000
 Intv   24   Start 10508 16: 1:19
```

```
    Ser.1     Avg  240.9        Chisq  55.50       Var  3.350     Newbs.     16
              Min  237.1          Max  243.8    expVar 0.9713      Bins  31744

 Period :   4.810    dP/dt :   0.000
 Intv   25   Start 10509  1:22:23
    Ser.1     Avg  1160.        Chisq 0.1813E+07   Var 0.3916E+06 Newbs.     16
              Min  505.8          Max  2385.    expVar  3.456      Bins  42980

 Period :   4.809    dP/dt :   0.000
 Intv   26   Start 10509  3: 2:36
    Ser.1     Avg  3668.        Chisq 0.7675E+07   Var 0.4684E+07 Newbs.     16
              Min  1318.          Max  7997.    expVar  9.762      Bins  48100

 Period :   4.808    dP/dt :   0.000
 Intv   27   Start 10509  4:42:48
    Ser.1     Avg  3973.        Chisq 0.8055E+07   Var 0.5320E+07 Newbs.     16
              Min  1434.          Max  8585.    expVar  10.57      Bins  48100

 Period :   4.808    dP/dt :   0.000
 Intv   28   Start 10509  6:23: 1
    Ser.1     Avg  4014.        Chisq 0.7999E+07   Var 0.5344E+07 Newbs.     16
              Min  1457.          Max  8645.    expVar  10.68      Bins  48100

 Period :   4.807    dP/dt :   0.000
 Intv   29   Start 10509  8: 3:13
    Ser.1     Avg  4025.        Chisq 0.6309E+07   Var 0.5385E+07 Newbs.     16
              Min  1468.          Max  8769.    expVar  13.70      Bins  37604

 Period :   4.807    dP/dt :   0.000
 Intv   30   Start 10509  9:43:26
    Ser.1     Avg  4091.        Chisq 0.5261E+07   Var 0.5459E+07 Newbs.     16
              Min  1509.          Max  8775.    expVar  16.65      Bins  31460

 Period :   4.808    dP/dt :   0.000
 Intv   31   Start 10509 11:23:38
    Ser.1     Avg  3973.        Chisq 0.4639E+07   Var 0.5104E+07 Newbs.     16
              Min  1423.          Max  8545.    expVar  17.60      Bins  28900

 Period :   4.808    dP/dt :   0.000
 Intv   32   Start 10509 13: 3:51
    Ser.1     Avg  4129.        Chisq 0.4910E+07   Var 0.5469E+07 Newbs.     16
              Min  1509.          Max  8818.    expVar  17.82      Bins  29660

 Period :   4.809    dP/dt :   0.000
 Intv   33   Start 10509 14:44: 3
```

```
    Ser.1     Avg  4220.        Chisq 0.5057E+07   Var 0.5607E+07 Newbs.     16
              Min  1536.          Max  9075.    expVar  17.75      Bins  30436

 Period :   4.810    dP/dt :   0.000
 Intv   34   Start 10509 16:24:16
    Ser.1     Avg  4198.        Chisq 0.5351E+07   Var 0.5550E+07 Newbs.     16
              Min  1577.          Max  8835.    expVar  16.61      Bins  32356

 Period :   4.811    dP/dt :   0.000
 Intv   35   Start 10509 18: 4:28
    Ser.1     Avg  4315.        Chisq 0.5443E+07   Var 0.5721E+07 Newbs.     16
              Min  1612.          Max  9107.    expVar  16.81      Bins  32868

 Period :   4.813    dP/dt :   0.000
 Intv   36   Start 10509 19:44:41
    Ser.1     Avg  4382.        Chisq 0.6085E+07   Var 0.6184E+07 Newbs.     16
              Min  1717.          Max  9240.    expVar  16.30      Bins  34404

 Period :   4.814    dP/dt :   0.000
 Intv   37   Start 10509 21:24:53
    Ser.1     Avg  4596.        Chisq 0.6225E+07   Var 0.6308E+07 Newbs.     16
              Min  1841.          Max  9487.    expVar  16.20      Bins  36324

 Period :   4.816    dP/dt :   0.000
 Intv   38   Start 10509 23: 5: 6
    Ser.1     Avg  3995.        Chisq 0.6098E+07   Var 0.4884E+07 Newbs.     16
              Min  1610.          Max  8286.    expVar  12.81      Bins  39908

 Period :   4.817    dP/dt :   0.000
 Intv   39   Start 10510  0:45:18
    Ser.1     Avg  3683.        Chisq 0.6300E+07   Var 0.4216E+07 Newbs.     16
              Min  1466.          Max  7639.    expVar  10.71      Bins  44004

 Period :   4.818    dP/dt :   0.000
 Intv   40   Start 10510  2:25:31
    Ser.1     Avg  3717.        Chisq 0.7145E+07   Var 0.4417E+07 Newbs.     16
              Min  1417.          Max  7825.    expVar  9.892      Bins  48100

 Period :   4.819    dP/dt :   0.000
 Intv   41   Start 10510  4: 5:43
    Ser.1     Avg  3644.        Chisq 0.6942E+07   Var 0.4205E+07 Newbs.     16
              Min  1364.          Max  7627.    expVar  9.697      Bins  48100

 Period :   4.820    dP/dt :   0.000
 Intv   42   Start 10510  5:45:56
```

```
    Ser.1     Avg  3472.        Chisq 0.5900E+07   Var 0.3822E+07 Newbs.     16
              Min  1296.          Max  7251.    expVar  10.36      Bins  42892

 Period :   4.820    dP/dt :   0.000
 Intv   43   Start 10510  7:26: 8
    Ser.1     Avg  3017.        Chisq 0.4791E+07   Var 0.2924E+07 Newbs.     16
              Min  1113.          Max  6372.    expVar  9.765      Bins  39544

 Period :   4.821    dP/dt :   0.000
 Intv   44   Start 10510  9:17:19
    Ser.1     Avg  2673.        Chisq 0.3327E+07   Var 0.2242E+07 Newbs.     16
              Min  972.4          Max  5550.    expVar  10.78      Bins  31744

 Period :   4.821    dP/dt :   0.000
 Intv   45   Start 10510 10:59:27
    Ser.1     Avg  2861.        Chisq 0.3122E+07   Var 0.2491E+07 Newbs.     16
              Min  1073.          Max  5992.    expVar  12.77      Bins  28672

 Period :   4.820    dP/dt :   0.000
 Intv   46   Start 10510 12:42:23
    Ser.1     Avg  2002.        Chisq 0.2048E+07   Var 0.1098E+07 Newbs.     16
              Min  847.6          Max  4080.    expVar  8.595      Bins  29796

 Period :   4.820    dP/dt :   0.000
 Intv   47   Start 10510 14:22:36
    Ser.1     Avg  2572.        Chisq 0.2832E+07   Var 0.1904E+07 Newbs.     16
              Min  1015.          Max  5323.    expVar  10.75      Bins  30620

 Period :   4.819    dP/dt :   0.000
 Intv   48   Start 10510 16: 3:27
    Ser.1     Avg  2130.        Chisq 0.2453E+07   Var 0.1291E+07 Newbs.     16
              Min  884.5          Max  4324.    expVar  8.425      Bins  32356

 Period :   4.820    dP/dt :   0.000
 Intv   49   Start 10510 17:43:40
    Ser.1     Avg  288.3        Chisq  3379.       Var  237.0     Newbs.     16
              Min  268.5          Max  315.5    expVar  1.123      Bins  32868

 Period :   4.767    dP/dt :   0.000
 Intv   50   Start 10510 19:23:52
    Ser.1     Avg  223.4        Chisq  39.01       Var  4.374     Newbs.     16
              Min  218.5          Max  226.9    expVar  1.795      Bins  15928
```

If we write down the value of the pulse period in each interval, and we plot them as a function of time, we obtain the graph shown in Figure B.6.

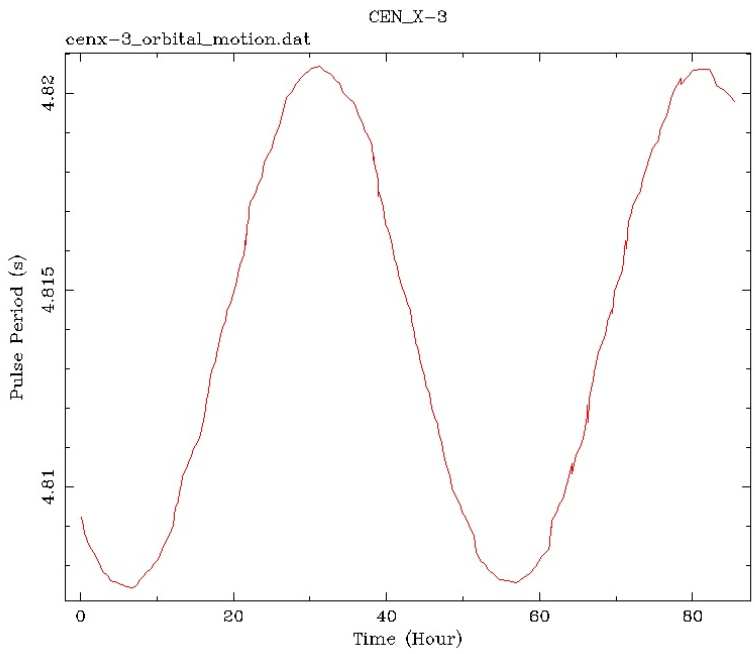


Figure B.6: Variation of the Cen X–3 pulse period along its orbit as due to Doppler effect.

We can reconstruct the orbital motion not only by the variation of the pulse period, but also from the variation in the pulse arrival times. If we can measure the arrival time differences then we can obtain a very accurate determination of the orbital parameters.
We start by folding our data with the average pulse period, and we measure as the pulse phase of a characteristic point (for example, the peak of the pulse) changes along the orbit (see Section 5.10 for details).

```
orma> efold

efold 1.1 (xronos5.22)

Number of time series for this task[1] 1
Ser. 1 filename +options (or @file of filenames +options)[] cenx-3_long.lc
 Series 1 file    1:cenx-3_long.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ CEN_X-3             Start Time (d) .... 10507 00:19:27.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10510 19:57:03.562
 No. of Rows .......      1873744        Bin Time (s) ......   0.1250
 Right Ascension ... 1.70313293E+02      Internal time sys.. Converted to TJD
 Declination ....... -6.06232986E+01     Experiment ........ XTE      PCA
```

```
 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10507.01351345407  (days)       0:19:27:562  (h:m:s:ms)
 Expected Stop .... 10510.83129123185  (days)      19:57: 3:562  (h:m:s:ms)

 Default Epoch is:  10507.00000
Type INDEF to accept the default value
 Epoch format is days.
Epoch[10507.00000] 10507.00000
 Period format is seconds.
Period[4.8091] 4.8144045
 Expected Cycles ..          68514.39
 Default phase bins per period are:        10
 Type INDEF to accept the default value
Phasebins/Period value or neg. power of 2[16] 32

 Newbin Time ......    0.15045014      (s)
 Maximum Newbin No.           2192461

 Default Newbins per Interval are:     2192461
 (giving       1 Interval of      2192461 Newbins)
 Type INDEF to accept the default value

Number of Newbins/Interval[24953] 5000
 Maximum of     439 Intvs. with         5000 Newbins of      0.150450     (s)
Default intervals per frame are:       439
Type INDEF to accept the default value
Number of Intervals/Frame[1] 1
 Results from up to       1 Intvs. will be averaged in a Frame
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

      32 analysis results per interval

 Intv    1   Start 10507  0:19:27
     Ser.1     Avg  967.7        Chisq 0.1712E+06   Var 0.2198E+06 Newbs.     32
               Min  469.0          Max  1866.    expVar  41.14      Bins   6018
  Folded light curve ready
```

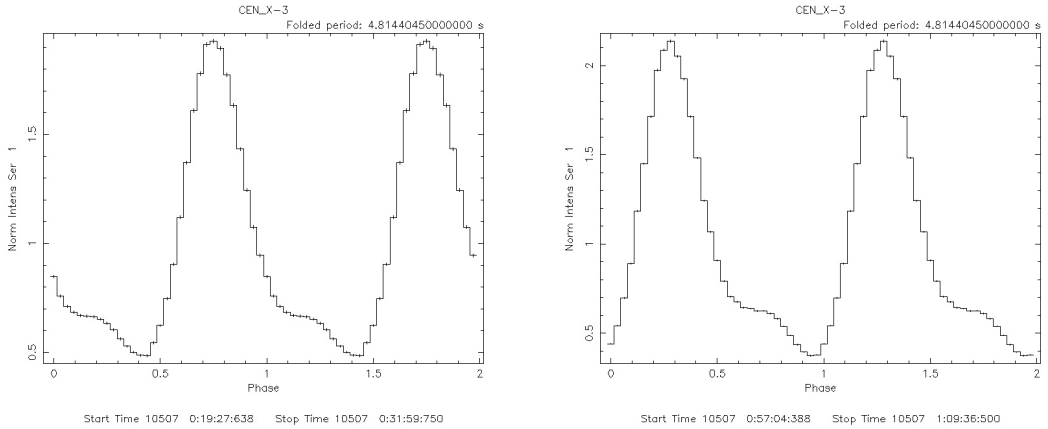


Figure B.7: Doppler Shift of the Cen X–3 pulse profile along its orbit.

```
Intv    2   Start 10507  0:31:59
    Ser.1     Avg  1479.        Chisq 0.3264E+06   Var 0.6433E+06 Newbs.     32
              Min  628.1          Max  3046.    expVar  62.94      Bins   6018
 Folded light curve ready
Intv    3   Start 10507  0:44:32
    Ser.1     Avg  2113.        Chisq 0.5022E+06   Var 0.1413E+07 Newbs.     32
              Min  821.9          Max  4416.    expVar  89.96      Bins   6018
 Folded light curve ready
```

In Figure B.7 we show two pulse profiles obtained along the orbit, and their shift due to Doppler effect.

B.6 Quasi Periodic Oscillations in X–ray Pulsars

A Quasi Periodic Oscillation (QPO) is revealed in a power spectrum as a broad peak (unlikely a pure periodic signal, that is a sharp, narrow peak). We will analyze the data from a XTE/PCA observation of the pulsar XTE J1858+038.

```
orma> powspec

powspec 1.0 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] XTE_J1858+034_pca.lc
 Series 1 file    1:XTE_J1858+034_pca.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ XTEJ1858+034        Start Time (d) .... 10868 10:43:11.562
 FITS Extension ....  1 - 'RATE      '   Stop Time (d) ..... 10868 11:37:03.562
 No. of Rows .......        25856        Bin Time (s) ......   0.1250
 Right Ascension ... 2.84730011E+02      Internal time sys.. Converted to TJD
 Declination ....... 3.41000009E+00      Experiment ........ XTE      PCA

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes

 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10868.44666160222  (days)      10:43:11:562  (h:m:s:ms)
 Expected Stop .... 10868.48406900962  (days)      11:37: 3:562  (h:m:s:ms)

 **** Warning: Newbin Time must be an integer multiple of Minimum Newbin Time
 Minimum Newbin Time   0.12500000      (s)
 for Maximum Newbin No..            25857

 Default Newbin Time is:  0.50000000    (s) (to have 1 Intv. of    8192 Newbins)
 Type INDEF to accept the default value

Newbin Time or negative rebinning[0.125] 0.125

 Newbin Time ......    0.12500000      (s)
 Maximum Newbin No.             25857

 Default Newbins per Interval are:        8192
 (giving       4 Intervals of         8192 Newbins each)
```

```
 Type INDEF to accept the default value

Number of Newbins/Interval[8192] 8192
 Maximum of       4 Intvs. with         8192 Newbins of      0.125000     (s)
Default intervals per frame are:         4
Type INDEF to accept the default value
Number of Intervals/Frame[7] 4
 Results from up to       4 Intvs. will be averaged in a Frame
Rebin results? (>1 const rebin, <-1 geom. rebin, 0 none)[0] -1.05
 Results will be rebinned geometrically  with a series of step    1.05
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

    4096 analysis results per intv. will be rebinned to     109

 Intv    1   Start 10868 10:43:11
     Ser.1     Avg  291.3        Chisq 0.1233E+05   Var  3507.     Newbs.   8192
               Min  112.0          Max  552.0    expVar  2330.      Bins   8192
Power spectrum ready !
 Intv    2   Start 10868 11: 0:15
     Ser.1     Avg  299.2        Chisq 0.1224E+05   Var  3578.     Newbs.   8192
               Min  104.0          Max  624.0    expVar  2394.      Bins   8192
Power spectrum ready !
 Intv    3   Start 10868 11:17:19
     Ser.1     Avg  325.0        Chisq 0.1479E+05   Var  4696.     Newbs.   8192
               Min  120.0          Max  688.0    expVar  2600.      Bins   8192
Power spectrum ready !
 Intv    4   Start 10868 11:34:23
     Ser.1     Avg  356.7        Chisq  2043.       Var  4554.     Newbs.   1280
               Min  160.0          Max  608.0    expVar  2853.      Bins   1280
           Interval rejected because of window(s)  in series 1 !
PLT> cpd XTE_J1858+034_pca_psd.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

In Figure B.8 we can clearly see a QPO at 0.1 Hz, together with the peak due to the pulse period at 221 s (4.5×10^{-3} Hz).

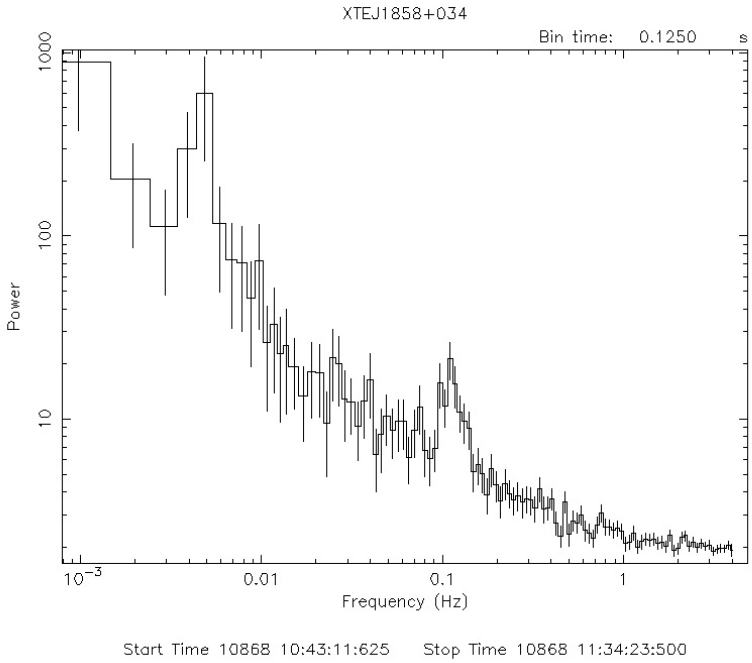


Figure B.8: Power spectrum of the X–ray binary pulsar XTE J1858+034. Both the peak due to the pulse period at 4.5×10^{-3} Hz and a QPO at 0.1 Hz are visible.

B.7 kHz QPO in Low Mass X–ray Binaries

In order to detect QPO in the kHz range we need data at high temporal resolution. We will perform our analysis on a XTE/PCA observation of the Low Mass X–ray Binary 4U1728−34, that have a time resolution of 0.125 ms.

```
orma> powspec

powspec 1.0 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] 4U_1728-34_pca1.lc
 Series 1 file    1:4U_1728-34_pca1.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ 1728-34              Start Time (d) .... 10129 09:34:43.562
 FITS Extension .... 1 - 'RATE      '    Stop Time (d) ..... 10129 09:43:03.563
 No. of Rows ....... 3992001             Bin Time (s) ......   0.1250E-03
 Right Ascension ... 2.62989197E+02      Internal time sys.. Converted to TJD
 Declination ....... -3.38345985E+01     Experiment ........ XTE      PCA

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes
```

```
 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10129.39911530592  (days)       9:34:43:562  (h:m:s:ms)
 Expected Stop .... 10129.40490234441  (days)       9:43: 3:563  (h:m:s:ms)

 **** Warning: Newbin Time must be an integer multiple of Minimum Newbin Time
 Minimum Newbin Time   0.12500000E-03  (s)
 for Maximum Newbin No..          4000002

 Default Newbin Time is:  0.61125000E-01(s) (to have 1 Intv. of    8192 Newbins)
 Type INDEF to accept the default value

Newbin Time or negative rebinning[0.125] 0.1250E-03

 Newbin Time ......    0.12500000E-03  (s)
 Maximum Newbin No.           4000002

 Default Newbins per Interval are:        8192
 (giving     489 Intervals of         8192 Newbins each)
 Type INDEF to accept the default value

Number of Newbins/Interval[8192] 1024
 Maximum of    3907 Intvs. with         1024 Newbins of      0.125000E-03 (s)
Default intervals per frame are:      3907
Type INDEF to accept the default value
Number of Intervals/Frame[4] 3907
 Results from up to    3907 Intvs. will be averaged in a Frame
Rebin results? (>1 const rebin, <-1 geom. rebin, 0 none)[-1.05] -1.01
 Results will be rebinned geometrically  with a series of step    1.01
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

     512 analysis results per intv. will be rebinned to     196


 Intv    1   Start 10129  9:34:43
     Ser.1     Avg  1976.        Chisq  947.9       Var 0.1530E+08 Newbs.    1024
               Min  0.000          Max 0.2458E+05expVar 0.1606E+08  Bins    1024
Power spectrum ready !
```

```
 Intv    2   Start 10129  9:34:43
     Ser.1     Avg  2120.        Chisq  924.4       Var 0.1589E+08 Newbs.   1024
               Min  0.000          Max 0.2458E+05expVar 0.1720E+08  Bins   1024
Power spectrum ready !
 Intv    3   Start 10129  9:34:43
     Ser.1     Avg  1992.        Chisq  985.8       Var 0.1530E+08 Newbs.   1024
               Min  0.000          Max 0.1638E+05expVar 0.1606E+08  Bins   1024
Power spectrum ready !
 Intv    4   Start 10129  9:34:43
     Ser.1     Avg  2016.        Chisq  1002.       Var 0.1612E+08 Newbs.   1024
               Min  0.000          Max 0.2458E+05expVar 0.1629E+08  Bins   1024

Power spectrum ready !
 Intv 3899   Start 10129  9:43: 3
     Ser.1     Avg  2264.        Chisq  1020.       Var 0.1801E+08 Newbs.   1024
               Min  0.000          Max 0.2458E+05expVar 0.1828E+08  Bins   1024
Power spectrum ready !
 Intv 3900   Start 10129  9:43: 3
     Ser.1     Avg  2540.        Chisq  129.7       Var 0.2164E+08 Newbs.    129
               Min  0.000          Max 0.2458E+05expVar 0.2081E+08  Bins    129
            Interval rejected because of window(s)  in series 1 !
PLT> r y 1.8 2.5
PLT> cpd 4U_1728-34_pca1_psd.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q
```

In Figure B.9 we show the resulting power spectrum in which the 800 Hz QPO is clearly visible.

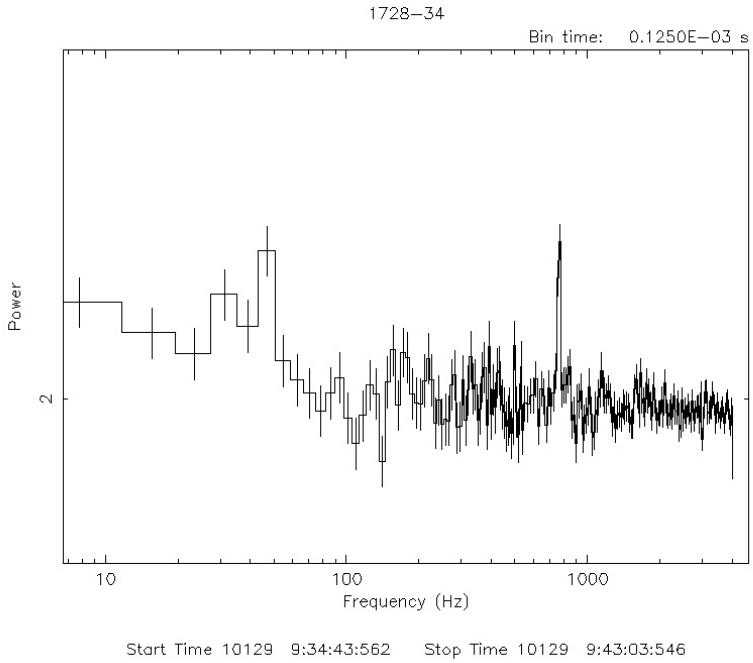


Figure B.9: Power spectrum of the low-mass X–ray binary 4U1728−34 as observed by XTE/PCA. The QPO at 800 Hz is clearly visible.

B.8 High Frequency Oscillations during Type-1 X–ray Bursts

We will search now high frequency oscillations, that sometime occur in X–ray sources. We will analyze again high time resolution XTE/PCA data of 4U1728−34.

```
orma> powspec

powspec 1.0 (xronos5.22)

Ser. 1 filename +options (or @file of filenames +options)[] 4U_1728-34_pca2.lc
 Series 1 file    1:4U_1728-34_pca2.lc

 Selected FITS extensions: 1 - RATE TABLE;

 Source ............ 1728-34            Start Time (d) .... 10129 10:01:52.562
 FITS Extension .... 1 - 'RATE      '   Stop Time (d) ..... 10129 10:02:23.563
 No. of Rows ....... 248001             Bin Time (s) ......   0.1250E-03
 Right Ascension ... 2.62989197E+02     Internal time sys.. Converted to TJD
 Declination ....... -3.38345985E+01    Experiment ........ XTE      PCA

 Corrections applied: Vignetting - No ; Deadtime - No ; Bkgd - No ; Clock - Yes
```

```
 Selected Columns:  1- Time;  2- Y-axis;  3- Y-error;  4- Fractional exposure;

 File contains binned data.

Name of the window file ('-' for default window)[-] -

 Expected Start ... 10129.41796947259  (days)      10: 1:52:562  (h:m:s:ms)
 Expected Stop .... 10129.41832827033  (days)      10: 2:23:563  (h:m:s:ms)

 **** Warning: Newbin Time must be an integer multiple of Minimum Newbin Time
 Minimum Newbin Time   0.12500000E-03  (s)
 for Maximum Newbin No..           248002

 Default Newbin Time is:  0.38750000E-02(s) (to have 1 Intv. of    8192 Newbins)
 Type INDEF to accept the default value

Newbin Time or negative rebinning[0.1250E-03] 0.0005

 Newbin Time ......    0.50000000E-03  (s)
 Maximum Newbin No.             62001

 Default Newbins per Interval are:        8192
 (giving       8 Intervals of         8192 Newbins each)
 Type INDEF to accept the default value

Number of Newbins/Interval[1024] 4096
 Maximum of      16 Intvs. with         4096 Newbins of      0.500000E-03 (s)
Default intervals per frame are:        16
Type INDEF to accept the default value
Number of Intervals/Frame[3907] 1
 Results from up to       1 Intvs. will be averaged in a Frame
Rebin results? (>1 const rebin, <-1 geom. rebin, 0 none)[-1.01] 0
Name of output file[test] test
Do you want to plot your results?[yes] yes
Enter PGPLOT device[/xw] /xw

    2048 analysis results per interval

 Intv    1   Start 10129 10: 1:52
     Ser.1     Avg 0.2032E+05    Chisq  4556.       Var 0.4536E+08 Newbs.   4096
               Min  0.000          Max 0.4710E+05expVar 0.4079E+08  Bins  16384
Power spectrum ready !
PLT> cpd 4U_1728-34_pca2_psd_1.ps/cps
PLT> pl
PLT> cpd /xw
```

```
PLT> q

  Intv    2   Start 10129 10: 1:54
     Ser.1     Avg 0.1371E+05    Chisq  4218.       Var 0.2844E+08 Newbs.   4096
               Min  0.000          Max 0.3482E+05expVar 0.2753E+08  Bins  16384
Power spectrum ready !
PLT> cpd 4U_1728-34_pca2_psd_2.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q

 Intv    3   Start 10129 10: 1:56
     Ser.1     Avg  9584.        Chisq  4319.       Var 0.2048E+08 Newbs.   4096
               Min  0.000          Max 0.2867E+05expVar 0.1925E+08  Bins  16384
Power spectrum ready !
PLT> cpd 4U_1728-34_pca2_psd_3.ps/cps
PLT> pl
PLT> cpd /xw
PLT> q


 Intv   15   Start 10129 10: 2:21
     Ser.1     Avg  2589.        Chisq  3970.       Var 0.5041E+07 Newbs.   4096
               Min  0.000          Max 0.1638E+05expVar 0.5199E+07  Bins  16384
Power spectrum ready !

 Intv   16   Start 10129 10: 2:23
     Ser.1     Avg  2591.        Chisq  580.5       Var 0.5461E+07 Newbs.    561
               Min  0.000          Max 0.1229E+05expVar 0.5210E+07  Bins   2241
            Interval rejected because of window(s)  in series 1 !
```

We will compute the power spectrum as a function of time, in order to see how the strength of the oscillation varies. In Figure B.10 we show the peak of the oscillation at 360 Hz observed in the third frame (out of a total of 16).

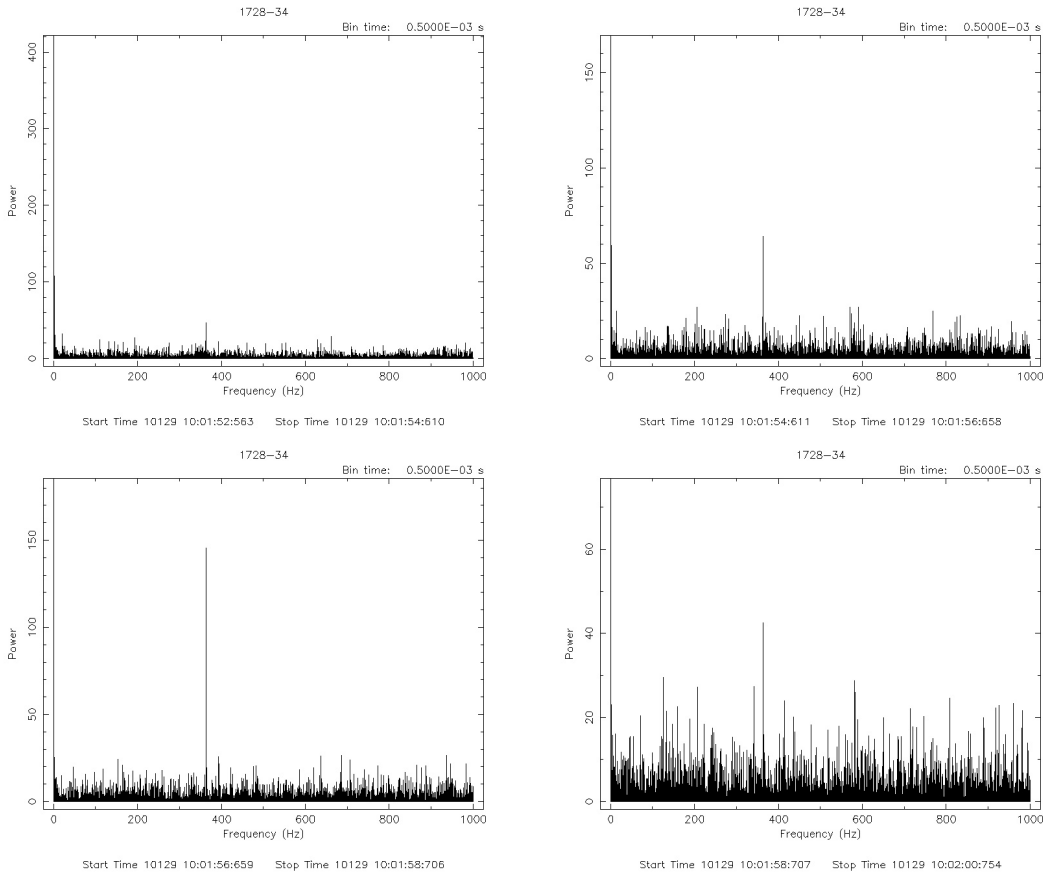


Figure B.10: Four consecutive power spectra of the low-mass X–ray binary 4U1728−34 as observed by XTE/PCA. The ∼360 Hz oscillation is clearly visible and variable.