Data Acquisition, Processing and Storage

M. Ljubisavljevic and M.B. Popovic

Introduction

Neurophysiological *signals* are mostly recorded as potentials, voltages, currents and electromagnetic field strengths generated by nerves and muscles. They carry information needed to understand the complex mechanisms underlying the behavior of the living system. Nevertheless, such information is rarely available directly, but has to be extracted from the raw signal(s). The whole process of extraction, from a sensor to the relevant information sought, can be considered as a state-of-the-art fermentation and distillation process aiming to extract the desired properties while preserving their unique characteristics. Usually, this line of action encompasses several stages which are generally defined as signal *acquisition* and *processing*. As presented in this book, modern neuroscience heavily relies on a number of different methods extracted from other sciences.

The overall outcome may be a single number, e.g., temperature, or it can be a more complicated result, e.g., the electromyogram (EMG) from contracting muscle(s) (cf. Chapters 26–28, 31). Anyway, in most cases, the result of the analysis contains only part of the information necessary to reconstruct the complete input signal. In this sense, the complete process can be thought of as a nonreversible signal transformation where the output signal is the desired result uniquely derived from the first. In order to succeed, knowledge of the specific properties of the signals as well as adequate signal processing and system engineering knowledge are critical for all phases of the process. It is rather unfeasible to recommend a single method of acquiring and processing biological signals. There are even several standard types of procedures for acquisition and processing of the same signal type. However, although each researcher could select his own approach for data recording and processing, some general guidelines should be followed.

Therefore, the general aim of this chapter is not to cover the vast variety of acquisition and processing approaches used in modern neuroscience research. Rather, it is to stress some of the general aspects associated with data acquisition and processing of signals. Theoretical aspects will be dealt with whenever needed to better explain practical issues, referring the reader to the cited literature when a deeper insight into the various subjects is sought. In order to assist a step-by-step implementation of particular neuroscience techniques presented in this book, we have organized this chapter as independent sections, giving each researcher the freedom to roam through it according to the actual goals of his study and his imagination.

Miloš Ljubisavljević, Dr Subotića 4, Institute for Medical Research, PO Box 102, Belgrade, 11129, Yugoslavia (phone: +381-11-685-788; fax: +381-11-643-691; e-mail: milos@imi.bg.ac.yu), National Institute for Working Life, Department of Musculoskeletal Research, BOX 7654, Umea, Sweden
Mirjana. B. Popović, Dr Subotića 4, Institute for Medical Research, PO Box 102, Belgrade, 11129, Yugoslavia (phone: +381-11-685-788; fax: +381-11-643-691; e-mail: mira@imi.bg.ac.yu)

Outline

The conventional path of digital signal processing from its source to the final presentation is depicted in Fig. 1.

Most measurements begin with the *transducer*, a device that converts a measurable physical quantity, such as pressure, temperature or joint rotation, into an electrical signal. Transducers are available for a wide range of measurements and different quantities, and thus come in a variety of shapes, sizes and specifications.

Signal conditioning transforms a transducer's output signal so that an analog-to-digital converter can sample the signal. On the hardware level, signal conditioning incorporates amplification, filtering, differential applications, isolations, sample and hold, current-to-voltage conversion, linearization and more. In this chapter, amplification and filtering will be dealt with in some detail.

The output of the signal-conditioning device is connected to an *analog-to-digital converter* (ADC) input. The ADC converts the analog voltage to a digital signal that is transferred to the computer for processing, graphing and storage.

The generalized instrumentation system usually includes additional *signal processing*. Traditionally, this additional processing was handled either by using relatively simple digital-electronic circuits or, if a significant amount of processing was required, by connecting the instrument to the computer. The use of microcomputers generally results in fewer integrated-circuit packages. The most useful application of microcomputers for bio-medical instrumentation involves controller functions such as the capability for self-calibration and error detection, and automatic sequencing of events. All these functions enhance the reliability of the computerized biomedical instrument.

As mentioned above, the electronic devices used for acquiring a biological signal perform some initial processing such as filtering, or perform transformations of the signal (i.e., Fourier transformation) in order to estimate various signal parameters. When processing results are not required immediately following signal acquisition, *off-line* processing or *post-processing* methods may be used. By contrast, when results are needed immediately after signal acquisition, *real-time* or *on-line* processing methods must be applied. Depending on the signal-frequency bandwidth and application, the required digital sampling rate determines the type of hardware that can be applied for digital signal processing. In real-time processing applications, the computations must be performed on a continuous basis to keep pace with the sampled input signal. In post-processing applications, the input signal is collected and stored ahead of time, and the computational rate is driven primarily by the desire to get results quickly. In both cases, computational speed is desirable. However, in the real-time case it is absolutely mandatory to accomplish the task at all. In conclusion, off-line processing can be performed on general-purpose computers and real-time processing requires special dedicated machines or processors.

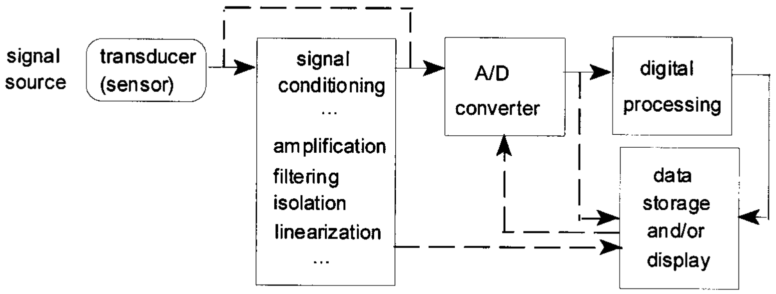


Fig. 1. General signal processing chart from the signal source to the final presentation.

In order to reveal important attributes of a signal in a more immediately interactive manner, results from the signal processing operations could be *displayed*. One or more displays allow the user to actively participate in the measurement itself. Display devices come in a wide variety depending on the use of the display. No matter what kind of display is used, its purpose is to convey information in a timely and non-permanent way with sufficiently high quality of presentation, so that the user can extract the information needed efficiently and accurately.

Frequently, there is a requirement for archiving either experimental or processed data. This can be done by various techniques and devices and is called *data storage*, or *backup* if it is a more permanent archive.

Part 1: Signal and Noise

What is a Signal?

As depicted in Fig. 1, the acquisition process starts with a signal. In general, signals are physicochemical phenomena that convey information, or they can be described as quantities that reveal the behavior of a system. As such they possess certain characteristic properties that require appropriate processing methods.

Two main types of signals are distinguished: *continuous (analog)* or *discrete*. Continuous signals are defined over a continuous range of a particular variable (usually time), while discrete signals are defined at discrete instants. Most of the signals of interest in neuroscience research are continuous, but some are discrete. In processing, continuous signals are represented as $x(t)$, where t is time (in units of seconds), as shown in Fig. 2a. The units of x depend on what is being described; examples would be volts, amperes, or – often in signal processing – unspecified units. Discrete signals are designated by series of discrete numbers: $x(k)$, where $1 \le k \le n$, as shown in Fig.2b.

Another way to classify signals is as *deterministic* or *random* signals. Deterministic signals are those that can be described by explicit mathematical relationships. In contrast, random signals cannot be exactly expressed in that way, which is inherent in their nature. Although it might be possible to determine a mathematical relationship, we may not have all the information to describe it by an explicit equation. Random signals can be described only in terms of probabilities and statistical measures. Neurobiological signals are usually extracted from living organisms and contain various degrees of randomness. Randomness appears in neurobiological signals in two major ways: the source itself may be stochastic, or the measurement system introduces external, additive or multiplicative, noise to the signal, often because the measurement device has to be designed so as not to damage the biological system.

Measurements of any kind of signal can be classified as *static* or *dynamic*. Static measurements assume that the input is a fixed value, not changing in time: $x(t)$=const. Dynamic measurements assume that the value of the input fluctuates with time, so that

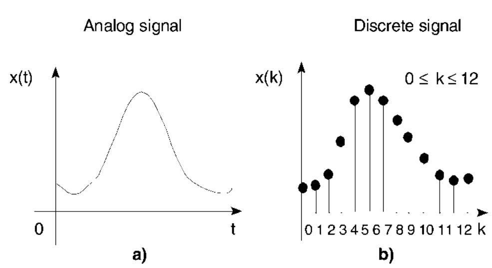


Fig. 2. Examples of a) a continuous (analog) and b) a discrete signal.

the measurement depends on the exact time at which it is made, or the input can be represented mathematically as a function of time x(t). The physical value represented by the function can be a scalar, such as pressure or temperature, or a vector such as force or velocity.

Noise

Noise is present in all signal sources and in all measuring systems. It is unavoidable in any electrical signal and affects the useful information that can be derived. Noise is crucially important when processing the low-amplitude signals which neuro-biological signals in fact are. Minimizing degradation of the desired signal by noise is of main concern in signal processing. There are no criteria for what constitutes acceptable signal amplitude or an acceptable noise level. The quality of a signal is determined by the simple ratio (*S/N*) of the amplitude of the desired signal, *S*, to the amplitude of the added noise, *N*. At what level will the *signal-to-noise ratio* begin to interfere with the analysis of the results? The answer to this question depends on both the nature of the interfering noise and the type of analysis to be performed.

A distinction is commonly made between noise of a random nature arising from basic physical processes and noise caused by interference which may or may not be correlated with the signal being measured. For example, one fundamental source of noise is the statistical variation in the electron density in a conductor and is present in all resistive elements. A frequent cause of interference noise is the electromagnetic or electrostatic interference arising from the presence of 50 or 60 Hz line current. Some signals are inherently of low amplitude and some environments are unavoidably polluted by noise sources of large amplitude. Periodic noise or pulse-like events with a pattern similar to that of the desired signal may prove confusing even when the signal being recorded tends to be 10 times greater than the amplitude of the noise. Noise that is time-locked to events under study may seriously degrade results even when it is invisible on the raw records. It interferes particularly with signal analysis by averaging (see below) or other statistical methods that are also time-locked to the event. Whenever the noise sources interact randomly with each other and with the recorded signal, the probability that the record will include an event large enough to be confused with a real signal increases with the time of recording.

Frequency Content of Noise

Generally, the noise content of a signal increases as the signal bandwidth increases, so that for dynamic signals requiring large bandwidth for adequate resolution, special care is needed in the design to insure that the measurement system has an optimum noise performance.

The frequency content of the noise can be determined by spectral analysis (see below), resulting in a spectral density graph as shown in Fig. 3. The graph is a typical example of many practical measurement systems in that it displays three distinct regions: a low-frequency region in which the mean square noise varies as 1/f, an intermediate-frequency region in which the noise spectrum is essentially flat, and a high-frequency region displaying an increasing noise spectrum. It is generally found that most components and systems display a low frequency region in which the noise varies as $1/f^n$ with n being approximately unity. Typically, an amplifier may exhibit a noise spectrum in which the 1/f noise is dominant for frequencies less than a few kHz. At intermediate frequencies the noise may be governed by thermal fluctuations in the electron density, giving rise to Johnson noise, which has a flat spectrum; noise of this type is commonly referred to as *white noise*. As the frequency approaches the cut-off frequency of the amplifier, other processes come into play and the noise amplitude generally increases.

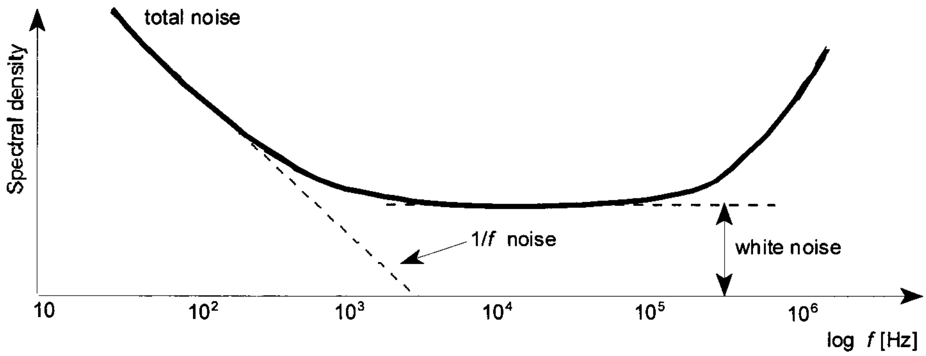
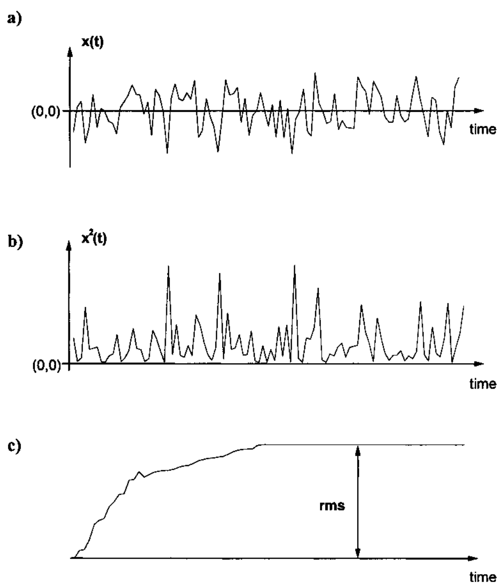


Fig. 3. Spectral density typical of a wide-band signal amplifier. Flat part comes from white noise.

Fig. 4. Process of obtaining the rms value of random signal: a) random signal x (t); b) square of x (t); c) rms value.



Of particular importance for the amplification of low-frequency signals or signals in which the direct-current (DC) component must be preserved is the 1/f noise. By modulating the input signal so as to convert it to a higher frequency, the amplifier can be used in a region where the 1/f component is less important. Once the signal has been amplified to such an extent that the noise generated by subsequent conditioning elements is no longer significant, the signal can be demodulated and the original, but amplified, signal can be recovered.

A consistent measure of the noise amplitude is the square root of the time-averaged (mean), squared value of the amplitude (*root mean square* or rms value). For a given waveform x (t), the rms value is defined as (Fig. 4c): Rms

$$X_{rms} = \left(\frac{1}{T} \sum_{0}^{T} x(t)^2 \, dt \right)^{1/2}$$

By squaring the amplitude as in Fig. 4a, all negative amplitudes become positive, so that the mean value of the amplitude squared is a nonzero quantity as seen in Fig. 4b. In statistical terms, the rms is the standard deviation of the noise as seen in Fig. 4c. The rms value is the only amplitude characteristic of a waveform that does not depend on its shape. Therefore, the rms value is the most useful means to quantify signal amplitudes in alternate-current (AC) measurements. Although the rms value is a measure of the noise amplitude, it says nothing about the frequency content.

The rms value can also be computed precisely using data sampled from the original analog waveform. In this case, samples must be acquired at a rate greater than twice the highest frequency of the signal (see Sampling Theorem below). The samples are squared, the squared values are summed over some averaging interval T, the square root is taken of the sum of squared values, and this value is divided by the number of samples within T. These operations can be performed either directly on a computer or in digital signal processing (DSP) hardware. Many instruments use this sampling technique. The signal-frequency range of this technique is theoretically limited solely by available sampler and ADC rates (see below). Rms meters and measurement devices are available from several manufacturers.

Quantizing Noise

In any process of analog-to-digital conversion (see below), the next step after sampling is to encode, or quantize, each sample value into a finite number of binary bits. The most common technique linearly maps a range of possible sample values $V_{p\text{-}p}$ into a fixed-size binary word of n bits. This coding requires that each sample be approximated by the nearest of 2^n possible values. The error introduced by this coding technique is a sawtooth function. It is common practice to assume that the signal samples excide this error function so that individual sample errors can be modeled as random noise, called quantizing noise, with a uniform amplitude probability distribution. Using this assumption the rms value of the quantizing noise is $V_{2ms}^{2n} = V_{p\text{-}p} / 2^n \sqrt{12}$, where $V_{p\text{-}p}$ is the peak-to-peak full-scale range of the quantizer.

Eliminating Noise

The quantification of noise and signal-to-noise ratio are a science in itself; the experimentalist primarily needs methods for recognizing the types of noise often encountered during specific recording and for minimizing their amplitude and effect. Obviously, the identification of the noise source will dictate the range of measures that will be effective and feasible. Here we summarize the most common reasons of low signal-to-noise ratios: electrode design, the first stage of the preamplifier, ground circuits and shielding. Unnecessary equipment that is not switched off, as well as AC power, are also noise sources. Interference can usually be reduced and often virtually eliminated by careful electrostatic and magnetic shielding; however, noise arising from basic physical phenomena usually sets a fundamental limit to the precision with which a given measurement can be performed.

Filtering

Whereas filtering (for more details see below) is an excellent way to minimize noise and enhance the biological component of a noisy signal, it can also distort noise to the point that it looks entirely physiological. For that reason the raw, unfiltered signal coming from a preamplifier should be inspected regularly with a high-accuracy display, such as an oscilloscope with a fast sweep speed. Inspection will reveal whether unphysiologically fast, large noise spikes have been reduced and smoothed into recorded signal. Almost all noise sources contain a wide range of frequencies in their spectra, which means that any filter distorts the shape as well as attenuates the amplitude.

Averaging

If the noise is uncorrelated with the signal, very significant reductions in the noise content of repetitive signals can be achieved by using an averaging process. The conditions and details are dealt with below.

Blanking

In certain applications, especially those employing electrical stimulation with a high-voltage stimulator, relatively huge transients may be superimposed on the signal, causing artifacts that often drive amplifiers into saturation. The recorded signal may be corrupted and the valuable data lost. Typical examples would be extracellular recordings of nerve impulses (cf. Chapter 5) evoked by high-voltage stimulators, stimulus artifacts in EMG recordings (cf. Chapters 26–28, 31) due to high-voltage stimulation during functional electrical stimulation or currents saturating EMG amplifiers due to magnetic stimulation. The solution to overcome these problems is either to prevent the stimulus coupling or to suppress the artifact before it feeds into the AC-coupled amplifier. This is usually referred to as *blanking*. Blanking could be realized in various ways, one of the most often used is to provide a sample-and-hold function early in the signal pathway. In this case, the input of the amplifier is going to be forced into the halt mode, causing it to ignore the incoming signals during the period of the transient. The output of the amplifier will at the same time continue to supply the signal equal to the that just preceding the transient.

Random Number Generators

In some cases, random number generators are used to simulate the effect of noise-like signals and other random phenomena encountered in the physical world of signals. Such noise is present in electronic equipment and systems under measurement. Its presence usually limits our ability to communicate over large distances and to detect relatively weak signals. By generating such noise on a computer, we are able to study its effects through simulation of communication systems, and to assess the performance of such systems in the presence of noise.

Most computer software libraries include a uniform random number generator. The output of the random number generator is a random variable, and is in the range [0,1] with equal probability. For all practical purposes, the number of outputs is sufficiently large, so that it can justify the assumption that any value in the interval is a possible output from the generator. Noise encountered in physical systems is often characterized by the normal or Gaussian probability distributions. Again, different mathematical methods could be employed on the computer in order to obtain such probability distributions.

Part 2: Signal Conditioning

Amplification and Amplifiers

The main reason for neurophysiological signals to be amplified is the need to make them suitable for signal-processing hardware whose proper functioning depends on the signal being within a certain amplitude range. As already mentioned, many measurements of neurophysiological signals involve voltages at very low levels, typically ranging between 1 μV and 100 mV, superimposed with noise and interference from different sources. Amplifiers are commonly used to couple these low-level biopotentials from high impedance sources to make them compatible with devices such as recorders, displays and A/D converters for computerized equipment.

Amplifiers adequate for measuring neurophysiological signals have to satisfy several very specific requirements. They have to provide amplification specific to the signal, reject superimposed noise and interference signals, and guarantee protection against damage from voltage and current surges for patient and animal and electronic equipment. Amplifiers featuring these specifications are known as biopotential amplifiers.

Differential Amplifiers

The input signal to the amplifier consists of several components: the desired biopotential, undesired biopotentials, a power-line interference of 50 (60) Hz and its harmonics,

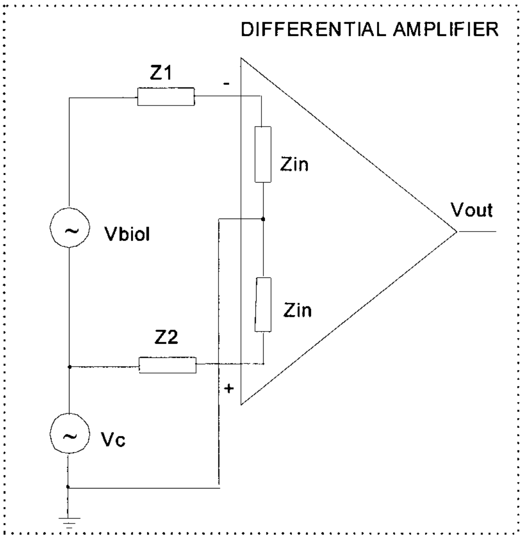


Fig. 5. Typical configuration for the measurement of biopotentials. Z_1 and Z_2 are source impedances from measured biological signal V_{biol}. V_c provides reference potential for the amplifier.

interference signals generated by the transducer (i.e., tissue/electrode interface), and noise. The differential measuring technique is used in biopotential amplifiers to minimize interference and noise, which are usually present in these low-level signals. Proper design of the amplifier provides rejection of a large portion of interference.

A typical configuration for the measurement of biopotentials is shown in Fig. 5. Three electrodes, two of them collecting the biological signal V_{biol} and the third providing the reference potential V_c, connect the subject to the amplifier. The output of the differential amplifier is, therefore, the difference between the two input signals times a certain gain factor. The desired biopotential appears as a voltage between the two input terminals of the differential amplifier and is referred to as the *differential signal.* The signal that appears between inputs and ground is called the *common-mode signal.* Thus, any common signal applied to both inputs (i.e., any common-mode signal) should result in zero output. In practice, the gains of the two signal paths are slightly different, resulting in a small output voltage even when identical voltages are applied to the two inputs. The line frequency interference signal shows only very small differences in amplitude and phase between the two measuring electrodes, causing approximately the same potential at both inputs. Differential voltage measurement eliminates common-mode noise, thus reducing the noise in analog input signals.

The ratio of the output voltage to the common input voltage is the *common-mode gain (G_{CM})*, which is usually much less than one. The ratio of the output voltage to the applied differential input voltage is the *differential gain (G_D)* of the bioamplifier and is usually much larger than one.

An index of how closely the bioamplifier approaches the ideal differential amplifier is given by the *common-mode rejection ratio (CMRR),* which is the ratio of the amplifier's differential gain to the common-mode gain. This value is expressed in decibels (dB), and is a function of frequency and source-impedance unbalance. Strong rejection of the common-mode signal is one of the most important characteristics of a good biopotential amplifier. It should be in order of at least 100 dB. Rejection of the common-mode signal is a function of both the amplifier CMRR and the source impedances Z_1 and Z_2. For the ideal amplifier CMRR is infinite and Z_1=Z_2 (Fig. 5), leaving the output voltage as the pure biological signal amplified by the differential mode gain. With CMRR finite or even slightly different Z_1 and Z_2, the common-mode signal is not completely eliminated, adding the interference term. The common-mode signal causes currents to flow through Z_1 and Z_2. The related voltage is amplified and thus not rejected by the amplifier. The output of a real amplifier will always consist of the desired output component resulting from the differential biosignal, an undesired component due to incomplete rejection of common-mode interference signals as a function of CMRR, and an undesired component due to source impedance imbalance allowing a small proportion of common-mode signal to appear as a differential signal to the amplifier.

In order to achieve optimum signal quality, the biopotential amplifier has to be adapted to the specific application. Obviously, each particular application for a biopotential amplifier will have a unique solution. Some of the amplifiers need to be extremely fast (e.g., when measuring action potentials from nerve cells; cf. Chapter 5), or have extremely high gain (e.g., when measuring EEG; cf. Chapter 35), or be extremely quiet (e.g., when measuring random noise in biological processes), etc. Based on the signal parameters, both appropriate bandwidths and gain factors are chosen. A final requirement for biopotential amplifiers is calibration. Since the amplitude of the biopotential often has to be determined very accurately, there must be a way to easily determine the gain or the amplitude range referenced to the input of the amplifier. For this purpose, the gain of the amplifier must be well calibrated. In order to prevent difficulties with calibrations, some amplifiers that need to have adjustable gains use a number of fixed gain settings rather than providing a continuous gain control. Some amplifiers have a standard, built-in signal source of known amplitude that can be momentarily connected to the input by the push of a button to check the calibration at the output of the biopotential amplifier.

Instrumentation Amplifiers (IA)

As mentioned, a very important stage in analog processing hardware is the amplification block. This stage is called an instrumentation amplifier and has several important functions: signal voltage amplification, rejection of the common-mode signal and proper driving of A/D converter input. Crucial to the performance of the preamplifier is the input impedance, which should be as high as possible. The input stage of an instrumentation amplifier usually consists of two voltage followers, which have the highest input impedance of any common amplifier configuration. A standard single operational amplifier (op-amp) design does not provide the necessary high input impedance, but two input op-amps provide high differential gain and unity common-mode gain without the requirement of close resistor matching (Fig. 6). The differential output of the first stage represents a signal with substantial relative reduction of the common-mode signal and is used to drive a standard differential amplifier, which further reduces the common-mode signal. Complete instrumentation amplifier-integrated circuits based on this standard instrumentation amplifier configuration are available from several manufacturers. All components, except the resistor that determines the gain of the amplifier, are contained on the integrated chip.

The output of the instrumentation amplifier has low impedance, which is ideal for driving the A/D converter input. The typical A/D converter does not have high or constant input impedance. It is important for the preceding stage to provide a signal with the lowest impedance practical. Instrumentation amplifiers have some *limitations*, including offset voltage, gain error, limited bandwidth and settling time. The offset voltage and gain error can be calibrated out as part of the measurement, but the bandwidth and settling time are parameters that limit the frequencies of amplified signals.

There is a special class of instrumentation amplifiers with programmable gain which can switch between fixed gain levels at sufficiently high speeds to allow different gains for different input signals delivered by the acquisition input system. These amplifiers are called *programmable gain instrumentation amplifiers* (PGIA).

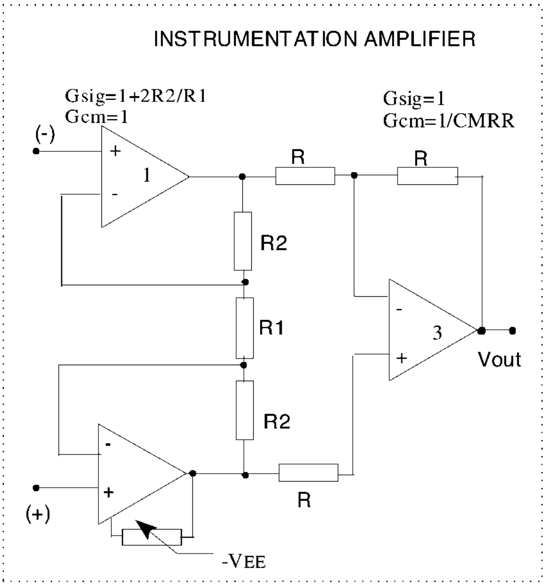


Fig. 6. Typical configuration of an instrumentation amplifier for biological application; an op-amp version.

Isolation Amplifiers

Biopotential amplifiers have to provide sufficient protection against electric shock to the user, patient and animal or other preparation. Electric safety codes and standards specify the minimum safety requirements for the equipment. To that end, isolation amplifiers may be used to break ground loops, to eliminate source-ground connections, and to provide isolation of patient and electronic equipment. They also contribute to preventing line frequency interferences. Isolation amplifiers are realized in three different technologies: transformer isolation, capacitor isolation and opto-isolation. An isolation barrier provides a complete galvanic separation of the input side, i.e., patient and preamplifier, from all equipment on the output side. Ideally, there should be no flow of electric current across the barrier. An index is the *isolation-mode voltage*, which is the voltage appearing across the isolation barrier, i.e., between input common and output common. The *isolation mode rejection ratio (IMRR)* is the ratio between the isolation voltage and the amplitude of the isolation signal appearing at the output of the isolation amplifier. Since the isolation mode rejection ratio is not infinite, there is always some leakage across the isolation barrier. Two isolation voltages are specified for commercial isolation amplifiers: the continuous rating and the test voltage. To eliminate the need for lengthy testing, the device is tested at about two times the rated continuous voltage.

Dynamic Range

Practical amplifiers introduce errors into the signal. Random noise is added by mechanisms such as thermal noise in resistors or shot noise in transistor junctions. Another problem is nonlinearity introduced into transistor junctions. To minimize the effect of added noise, it is desirable to keep the analog signal levels large compared with the noise sources. On the other hand, to minimize the distortion effects of nonlinearities, it is desirable to operate with low signal levels. The range of signal levels that can be processed between the lower limit imposed by the noise and the upper limit imposed by distortion is called the dynamic range of an amplifier. Expressed in decibels, a dynamic range up to 200 dB is achievable.

Additive noise introduced in amplifiers tends to be "white", meaning that the noise power is distributed evenly over a wide frequency range. Thus the noise density per Hz is constant over the bandwidth of the amplifier. A common basis of amplifier comparison is to specify the noise density in decibels relative to one milliwatt (dBm) per Hz.

Single-Ended and Differential Measurements

Data acquisition systems have provisions for single-ended and differential input connections. The essential difference between the two is the choice of the analog common connection. Single-ended multi-channel measurements require that all voltages be referenced to the same common node, which will result in measurement errors unless the common point is very carefully chosen; sometimes there is no acceptable common point. Differential connections cancel common-mode voltages and allow measurement of the difference between the two connected points. When given the choice, a differential measurement is always better. The rejected common-mode voltages can be steady DC levels or noise spikes. The best reason for choosing single-ended measurement will be for a higher channel count that is available in some devices. Most data acquisition products allow doubling the number of channels in a differential system by selecting single-ended operation.

Fundamentals of Filtering and Filters

Filtering is a signal processing operation that alters the frequency content of a signal. A *filter* is a frequency-selective device or a computer program that passes signals in one band of frequencies and rejects (or attenuates) signals in other bands. In signal processing, spectral components containing information on the desired signal are of foremost interest, and filters are designed so as to pass those spectral components while rejecting or attenuating components consisting mostly of noise. To design a satisfactory filtering device or program, it is necessary to have some knowledge of the structure of both the signal and the noise. Filtering is mostly performed to enhance the signal-to-noise ratio (S/R), eliminate certain types of noise or smooth the signal.

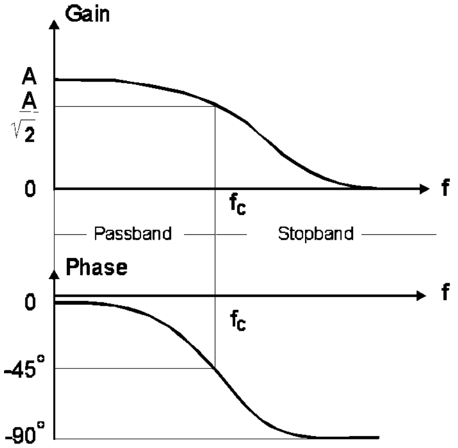
Frequency Response

The basic feature that determines filter behavior is its frequency response. The frequency response is a complex function which includes both *gain* and *phase* information. The gain and phase responses show how the filter alters the amplitude and phase of a sinusoidal input signal to produce a sinusoidal output signal. Since these two characteristics depend on the frequency content of the input signal, they can be used to describe the frequency response of the filter. The gain (dimensionless) and the phase (in degrees or radians) are mostly plotted versus frequency yielding frequency-response plots as exemplified for low-pass filters in Fig. 7.

Since the frequency ranges of interest often span several orders of magnitude, a logarithmic frequency scale compresses the data range, highlighting important features in the gain and phase responses. Special terminology involved with the use of logarithmic frequency intervals is *octave* and *decade*. The octave is the frequency range whose end points have a ratio of 2:1, while the one with a 10:1 ratio is called decade.

In signal processing, most often variables are not an issue, but the relation of two of the same kind. Furthermore, to simplify the mathematics that follows processing, it is

Fig. 7. Frequency response plots: a) gain; and b) phase.



useful to take logarithmic measure of the relation. In practice, variables are expressed through the measure's power or amplitudes. A relative (dimensionless) logarithmic measure of signal amplitudes is the *decibel (dB) scale*. It relates two signal powers, P_1 and P_2 or two corresponding signal amplitudes, A_1 and A_2. Since power is proportional to amplitude squared, the definition is:

$$10 \log_{10} (P_1/ P_2) = 10 \log_{10} (A_1/A_2)^2 = 20 \log_{10} (A_1/A_2) \text{ [dB]}.$$

When describing the amplitude response of filters, the term *attenuation* is often preferred to gain because many filters have a maximum gain of unity. Attenuation is the reciprocal of the gain. For example, if the gain is 0.1, then the attenuation is 10.

Passband and Stopband

The frequency range over which there is a little attenuation is called the passband. The range of frequencies over which the output is significantly attenuated is called the stopband. For example, at high frequencies the gain in Fig. 7 falls off, so that output signals at these frequencies are reduced in amplitude. At low frequencies the gain is essentially constant and there is relatively little attenuation.

Ideally, all the power would pass through the filter in the passband and no power would pass through in the stopband. In practical filters, however, between 80 and 95% of the input power shows up at the output in the passband and some small amount appears at output in the stopband. A real filter must have a *transition* zone between the stopband and passband (Fig. 8). It is defined by the cut-off frequency f_c and stopband frequency f_s. For the high-pass filter in Fig. 8. everything above f_c is passed; everything below f_s is stopped.

Cut-off Frequency

The frequency where the response starts to fall off significantly is called cut-off frequency (f_c). The term is closely associated with the boundary between a passband and adjacent stopband. It is defined as that frequency at which the signal voltage at the output of the filter falls to $1/\sqrt{2}$ of the amplitude of the input signal. Since the fall of $1/\sqrt{2}$ corresponds to −3 dB, this frequency is often called −3 dB frequency, or f_{-3}. Equivalently, the cut-off frequency (f_c) is the frequency at which the signal power at the output of the filter falls to half of the power of the input signal, which yields another term for cut-off frequency: half-power frequency.

Passband Ripple (Passband Flattens)

Real filters cannot be perfectly flat in the passband and stopband, though depending on the type of filter used, the amount of unevenness can vary. The amplitude variation

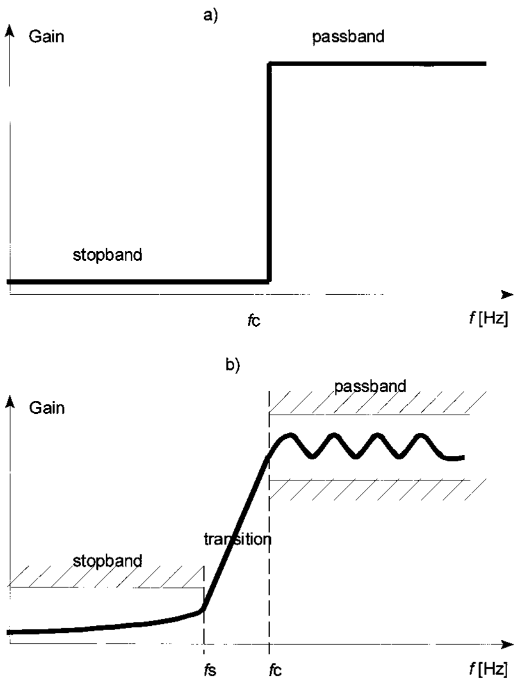


Fig. 8. Ideal *vs.* real high-pass filter magnitude response; a) and b), respectively.

within the passband is called the passband ripple (usually expressed in decibels) (Fig. 8). Some filters also have a predefined stopband ripple.

Phase Shift

The filter shifts the phase of sinusoidal components of the input signal as a function of frequency. If the phase shift in the passband is linearly dependent on the frequency of the sinusoidal components, the distortion of the signal waveform is minimal. When the phase shift in the passband is not linearly dependent on the frequency of the sinusoidal component, the filtered signal generally exhibits *overshoot*. That is, the initial response to a step input transiently exceeds the final value.

10–90% Rise Time

This term refers to the time it takes for a signal to rise from 10% of its initial value to 90% of its final value. As a general rule, when a signal with $t_{10-90} = t_s$ is passed through a filter with $t_{10-90} = t_f$, the rise time of the filtered signal is approximately $\sqrt{t_s^2 + t_f^2}$.

Filter Types

Four common types of filters can be distinguished in respect to the bandwidth: low-pass, high-pass, band-pass, and band-stop or band-reject. An illustration of how the amplitude of an input signal is altered by each of the four filter responses is shown in Fig. 9. For simplicity, consider an input signal consisting of three separate frequency components: f_1, f_2 and f_3. The cut-off frequency is designated by f_c for low-pass and high-pass filters, and by f_{c1} and f_{c2} for band-pass and band-stop filters. The low-pass filter passes frequency f_1 below its cut-off frequency f_c, and attenuates the frequencies above it: f_2 and f_3. A high-pass filter attenuates frequency f_1 below the cut-off frequency f_c and passes the frequencies above it: f_2 and f_3. A band-pass filter attenuates frequency components f_2 and f_3 outside the bandwidth determined by cut-off frequencies f_{c1} and f_{c2} and passes frequency component f_2 inside the same bandwidth. A band-reject (notch) filter does the opposite. Band-pass and band-reject filters can simply be thought

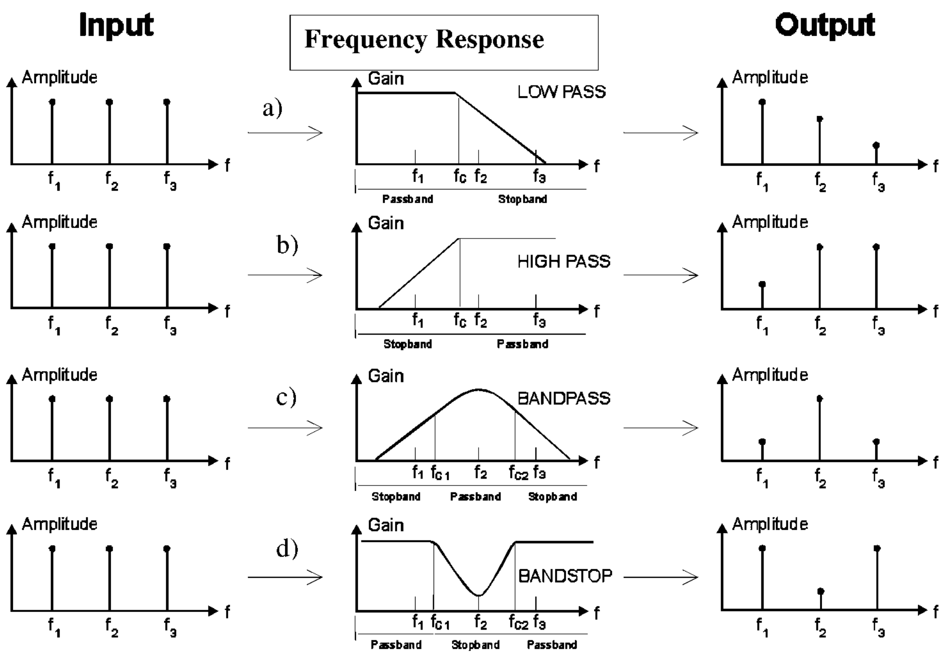


Fig. 9. Frequency responses for a) low pass; b) high pass; band pass; and band stop filters. (See text for details).

of as a series cascade of high-pass and low-pass filters, or as a parallel combination of high-pass and low-pass filters, respectively.

Order The order of a filter controls the intensity of its falloff with frequency. The higher the order of a filter, the better its performance, meaning the less the attenuation in the passband and the more complete the rejection of out-of-band signal components. The order is often described as the slope of the attenuation in the stopband, well above cut-off frequency, so that the slope of the attenuation has approached its asymptotic value. For instance, a first-order low-pass filter falls off near-linearly with frequency for high frequencies, at a rate of 6 dB per octave. A first-order filter can be constructed from one resistor and one capacitor. It is also known as a simple filter or a single-pole filter (see Analog Filter in Glossary, n=1). The order of the filter relates to the number of poles and the slope in the following way: 1 pole = 1st order = 6 dB/octave (=20 dB/decade); 2 poles = 2nd order = 12 dB/octave (= 40 dB/decade). In other words, voltage attenuation increases by a factor of 2 for each doubling of the frequency (octave) or by 10 for each tenfold increase in frequency (decade).

Filtering in Time and Frequency Domains

Neuropysiological signals can be processed either in the time domain or the frequency domain. Time-domain analysis refers to the analysis of a signal that is represented the same way it would appear on a conventional oscilloscope. Frequency-domain analysis refers to the analysis of signals that are transformed into the frequency domain (see below). To achieve optimal analysis, specific filtering should be applied.

For time-domain analysis, the filter should optimally minimally distort the time course of the signal. For example, it would not be very helpful to implement a very effective filter for elimination of high-frequency noise if it causes 15% overshoot. In general, the best filters to use for time-domain analysis are Bessel filters. They add less than 1% overshoot to pulses and have the linear change of phase with changing frequency. Because they do not alter the phase of the sinusoidal component of the signal, Bessel filters are sometimes called "linear-phase" or "constant-delay" filters. Unlike Bessel filters, Butterworth filters add considerable overshoot. In many experiments in neurophysiology, the signal noise increases rapidly with bandwidth. Therefore, a single-pole filter is inadequate. The four-pole Bessel filter is usually sufficient, but high-order filters are preferred. In experiments where the noise-power spectral density is constant with bandwidth, a single-pole filter is sometimes considered adequate. In the time domain, notch filters must be used with caution because signals that include sinusoidal components at the notch-filter frequency will be grossly distorted. On the other hand, if the notch filter is in series with low-pass or high-pass filters that exclude the notch frequency, distortion will be prevented. For example, notch filters are often used in electromyogram recordings, in which the line-frequency pickup is sometimes much larger than the signal. The 50 (60) Hz notch filter is typically followed by a 300 Hz high-pass filter.

Frequency-domain analysis is typically achieved using a fast Fourier transform (FFT). For this type of analysis the most important requirement is to have a sharp filter cut-off so that the noise above −3 dB frequency does not get folded back into the frequency of interest by the aliasing phenomenon (see below). The simplest and most commonly used filter for frequency-domain analysis in biological applications is the Butterworth filter. This filter type has the attenuation in the passband as flat as possible without having passband ripple. This means that the frequency spectrum is minimally distorted. Notch filters can be safely used in conjunction with frequency-domain analysis since they simply remove a narrow section of the power spectrum. Another approach could be to store "raw data" and then digitally remove the disturbing frequency components from the power spectrum.

Implementation

Basically, two filtering methods are available: special-purpose hardware designed for each filter structure or a general-purpose computer system with special-purpose software.

Hardware Filters

Filtering via hardware is performed by electrical circuits usually comprising capacitors, resistors, inductors and operational amplifiers which will allow one range of frequencies to pass through it and will block another range. The main drawback of using hardware for filtering is that each filter requires its own specific components, and a filter designed for one particular application cannot be used for another. Analog filters are designed as *active* or *passive*. Active filters arose from the need for filters that are compatible with modern integrated circuitry (IC) technology. They have become very attractive not only because passive filters are much more difficult to implement in IC technology, but also because active filters have other advantages over passive filters. Active filters offer higher sensitivity and flexibility, the ease of tuning and power gain. They also outperform passive filters in costs. There are many possible frequency responses that can be implemented by active filters. The most common filters are: Elliptic, Cauer, Chebyshev, Bessel and Butterworth.

Digital Filters

The alternative approach is to use general-purpose digital hardware (i.e., a computer) and implement the filter algorithm in the form of software. Software design can be accomplished via "high- or low-level" languages. This has the advantage of flexibility:

software designed for one application can easily be adapted to another, but has the disadvantage of being slower than special-purpose hardware. Digital filters are linear discrete systems governed by differential equations implemented in software. They consist of a series of mathematical calculations that process digitized data. Digital filter algorithms can implement all of the already mentioned filter frequency responses and more. Digital filtering is accomplished in three steps. First, the signal must be Fourier transformed (see below). Then, the signal's amplitude in the frequency domain must be multiplied by the desired frequency response. Finally, the transferred signal must be inverse-Fourier transformed back into the time domain. There are two types of discrete-time filters: *FIR* (Finite Impulse Response) and *IIR* (Infinite Impulse Response). Each filter type has its own set of advantages. The choice between FIR and IIR filters depends on the importance of these advantages to the design problem. If phase considerations are put aside, it is generally true that a given magnitude response specification can be met most efficiently with an IIR filter. In contrast, FIR filters can have precisely linear phase.

Finite Impulse Response Filters

FIR filters are almost entirely restricted to discrete implementations. Design techniques for FIR filters are based on directly approximating the desired frequency response of the discrete-time system. Most techniques for approximating the magnitude response of an FIR system assume a linear phase constraint. These filters have the advantage of not altering the phase of the signal. FIR filters are also known as nonrecursive filters. The output of a nonrecursive filter depends only on the input data. There is no dependence on the history of previous outputs. An example is the smoothing filter. Another example of a nonrecursive digital filter is the Gaussian filter. It is similar to a smoothing filter, except that the magnitudes of coefficients stay on the bell-shaped Gaussian curve.

Infinite Impulse Response Filters

IIR filters, unlike FIR filters, do not exhibit linear phase. However, the IIR design often results in filters with fewer coefficients than an equivalent FIR design. These are also known as recursive filters. The output of a recursive filter depends not only on the inputs, but on the previous outputs as well. That is, the filter has some time-dependent "memory". Digital filter implementations of analog filters such as Bessel, Butterworth, Elliptic and Chebyshev filters are recursive.

Digital over Analog Filtering?

Digital filtering is advantageous because the filter itself can be tailored to any frequency response without introducing the phase error. In contrast, analog filters are only available with a few frequency response curves, and all introduce some element of phase error. The delay introduced by analog filters necessarily makes recorded events occur later than they actually occurred. If it is not accounted for, this added delay can introduce an error in subsequent data analysis. A drawback of digital filtering is that it cannot be used for anti-aliasing because it occurs after sampling. Another problem with digital filters is that values near the beginning and end of the data cannot be properly computed. This is only a problem for a short data sequence. Adding values outside the sequence of data is arbitrary and can lead to misleading results.

Optimal and Adaptive Filtering

When the signal and noise are stationary and their characteristics are approximately known or can be assumed, an optimal filter can be designed *a priori*. Wiener and matched filters belong to this group. When no *a priori* information on the signal or noise is available, or when the signal or noise is non-stationary, *a priori* optimal filter design is not possible. Adaptive optimal filters are filters that can automatically adjust their own parameters, based on the incoming signal. The adaptation process is conducted so that a given performance index is optimized. Adaptive filters thus require little or no *a priori* knowledge of the signal and noise. Least-mean-square (LMS) filters belong to this group. The adaptive filter is required to perform calculations to satisfy the performance index and must have provision for changing its own parameters. Digital techniques, with or without a computing device, have clear advantages here over analog techniques. It is mainly for this reason that most adaptive filter implementations are performed by discrete systems.

Decimation Filtering

Often the information of interest in a signal is contained only in the low-frequency range, and the upper frequencies are not of interest. By filtering out the unwanted high frequencies and thereby narrowing the signal bandwidth, it is possible to meet the Nyquist sampling criterion with a lower sampling rate. If the bandwidth is reduced by a factor of k, the remaining signal can be fully described by saving every k-th sample and discarding the samples in between. This is called decimation by factor k, and the resulting output sample rate is f_s/k.

Part 3: Analog-to-Digital Conversion (Digitization)

Digital or Analog Processing?

Modern digital technology, both in terms of hardware and software, makes digital processing in many cases advantageous over analog processing. It may therefore be worthwhile to convert the analog signal to a discrete one so that digital processing can be applied. The conversion is done by *analog-to-digital (A/D) conversion systems* that sample and quantitize the signal at discrete times (see below). The factors that determine whether signals are processed digitally or in analog form include signal bandwidth, flexibility, accuracy and cost.

Cost

When a signal has a broad bandwidth, analog processing is more attractive because of the cost associated with high-speed digital signal-processing hardware. Most signal processing above 100 MHz is analog, while much of the signal processing up to 10 MHz is digital.

Accuracy

Accuracy is always an issue. The simplest analog-signal processing function is to change the signal amplitude with an attenuator or amplifier. The main reason for doing this is that the proper operation of the analog-signal processing hardware depends on the signal being within a certain amplitude range. Although the ideal result of this function would be to multiply the signal by a fixed gain, amplifiers and attenuators also introduce errors into the signal.

A digital signal, in principle, assumes only one of two states or levels, either "high" (logic 1), or "low" (logic 0). These states are represented by a voltage signal that is, according to the current standards, nominally defined as either 5 or 0 Volts. Actual digital signals fluctuate over a small range near their nominal values. The acceptable level of fluctuation depends on the technology in use. In addition to the amplitude of the signal, the time behavior is important, specifically the time required for the signal to change from one state to the other. Typically, this is in the order of milliseconds to nanoseconds in today's technology and depends on the slope of rise or fall expressed in volts per second.

Multiple Functions and Flexibility

One of the primary benefits of digital signal processing is the wide range of processing functions that can be implemented. Virtually any function that can be expressed mathematically can be performed with digital processing. Analog functions, on the other hand, are limited by the available components. Some functions may be theoretically implemented in analog form, but the inability to maintain sufficient accuracy may make the function impractical. Also, random noise and component nonlinearities limit the dynamic range of analog signal processing. By contrast, digital processing can be done with arbitrary precision by representing signals with high-precision numeric data types. Also, digital processing is exactly reproducible and is stable over time, temperature changes, and other environmental conditions. Calibration procedures are not mandatory for manufacturing and maintaining digital signal processing circuits. As mentioned above, the costs of digital signal processing strongly depend on the signal bandwidth. However, the costs of programmable digital signal processing (DSP) integrated circuits are substantially less than of those of analog circuitry that can perform the same function.

A/D Conversion

Neurophysiological signals are mainly analog. To process analog signals by digital means, it is first necessary to convert them into digital form, that is, to convert them to a sequence of numbers having finite precision. This procedure is called analog-to-digital (A/D) conversion, and the corresponding devices are called A/D converters (ADCs). The analog-to-digital converter stage is the last link in the chain between the analog domain and the digitized signal path. Analog-to-digital converters have played an increasingly important role in instrumentation in recent years. The expansion of ADCs has been driven by the development of high-performance integrated circuit (IC) technology. Advanced IC technology has led to the microprocessor and fast digital signal processing capability, which are essential in providing a low-cost transformation from the raw data generated by the ADC to the measurement results sought by the user.

Sampling

The first step in A/D conversion is sampling and involves time discretization of the continuous signal x_a (t) into a series of n discrete numbers $\{x_a (k), 1 \leq k \leq n\}$. Here we consider a band-limited analog signal x_a (t) with maximal frequency f_{max}, which is also bounded in amplitude. Also, we assume uniform sampling with a constant sampling frequency f_s, so that the signal x_a (t) is transformed into a sequence of sampled data $\{x_a(kT_s), 1 \leq k \leq n\}$, where $T_s=1/f_s$ is the sampling period or sampling interval, as illustrated in Fig. 10.

Sampling Theorem

In many cases of interest it is desirable to be able to reconvert the processed digital signals back into analog form, that is, to apply D/A conversion. This puts certain demands on the sampling of the original analog signal. If this signal is sampled at greater than twice the frequency of the highest-frequency component in the signal, then the original

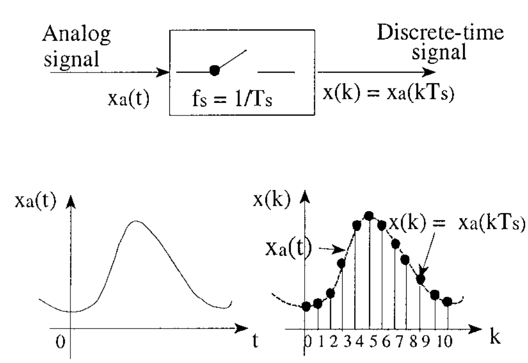


Fig. 10. Sampling of an analog signal $x_a(t)$ with sampling frequency $f_s=1/T_s$.

signal can be reconstructed exactly from the samples. In other words, x_a (t) can be reconstructed from $\{x_a(kT_s), 1 \le k \le n\}$ if $f_s > 2 f_{max}$. This is known as Nyquist's *sampling theorem*, and $f_s/2$ is known as the *Nyquist frequency*. If the sampling frequency does not satisfy the sampling theorem, i.e., $f_s < 2 f_{max}$, time discretization results in a phenomenon called *aliasing*. Any frequency above $f_s/2$ results in samples that are identical with a corresponding frequency in the range $-f_s/2 < f < f_s/2$. To avoid the ambiguities resulting from aliasing, we must select the sampling rate to be sufficiently high, i.e., $f_s > 2 f_{max}$.

Aliasing

Consider the effect of sampling a signal that is not band limited to f_{max} in a case where the sampling frequency is $f_s = 2 f_{max}$. Any signal component having a frequency higher than f_{max} is folded back or falsely translated to another frequency somewhere between 0 and f_{max} by the act of sampling. Figure 11 shows an example of two cases for f_s. In Fig. 11b, the sampling frequency obeys $f_s > 2 f_{max}$, where f_{max} is the largest frequency of the signal $x_a(t)$. Note that the sampled signal in the frequency domain consists of non-overlapping functions. Consider the effect of a low-pass filter that passes all frequencies in the range $0 < f < f_{max}$ undistorted, while zeroing all frequencies outside this range. The Fourier transform of the signal at the output of such a filter equals that of $x_a(t)$. Since the Fourier transform is unique, we can restore the original signal from its samples by such a low-pass filtering operation, provided the sampling frequency obeys: $f_s > 2f_{max}$. In Fig. 11c, the sampling frequency is lower than twice the f_{max}: $f_s < 2f_{max}$. All signal frequencies above f_s /2 show up as aliases or spurious lower frequency errors that cannot be distinguished from valid sampled data (black area).

Since knowledge of Fourier transform (X) implies (by Inverse Fourier transform) a knowledge of $x_a(t)$, it follows from the sampling theorem that if we sample fast enough, then we can reconstruct the original signal from its Fourier transformation.

How to Avoid Aliasing

The apparently straightforward solution to the aliasing problem is low-pass filtering of the analog signal before sampling. The filter's cut-off frequency must then be set at one-half the sampling rate or lower. Note, however, that slight fluctuations in the measured environment or the measured signal can cause alias signals to move, leaving errors in different locations throughout the data each time an A/D converter is in use. To circum-

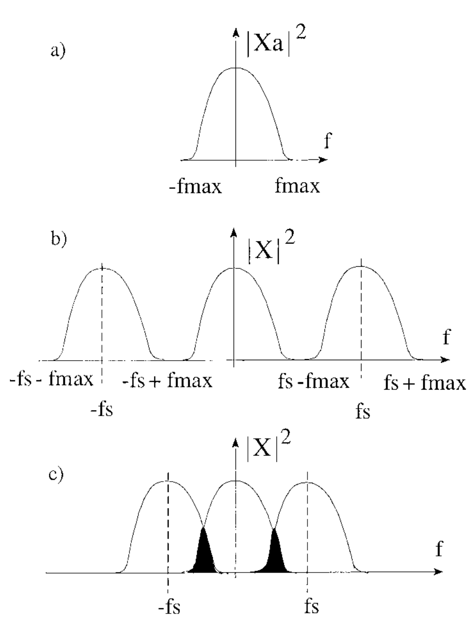


Fig. 11. Sampled band-limited signal in the frequency domain: a) Spectrum of the band-limited signal; b) Spectrum of the sampled signal when $f_s > 2f_{max}$; and c) Spectrum of the sampled signal when $f_s < 2f_{max}$. Aliasing results in overlapping spectrum regions (black area).

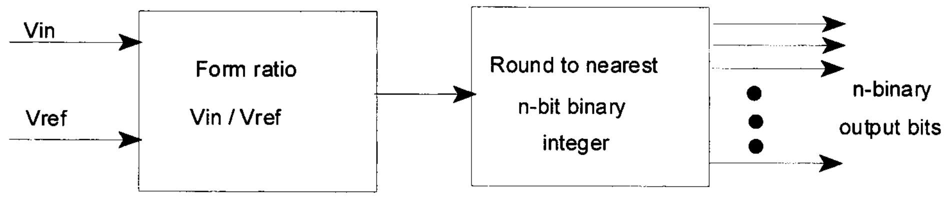


Fig. 12. Concept of the ADC: forming the ratio between V_{in} and V_{ref}, and rounding it to the nearest n-bit binary integer.

vent this problem, one could sample the analog signal at a very high rate far beyond the Nyquist frequency and then filter out the high frequencies with digital techniques. But such oversampling increases system costs by requiring faster A/D conversion for digital processing, more memory and higher bandwidth buses. It also leads to higher analysis costs by creating more data to process and interpret. Nevertheless, a practical alternative is to limit bandwidth of the signal below one-half the sample rate with a low-pass or *anti-alias filter*, which can be implemented on each channel in front of the A/D converter. Low-pass filtering must be done before the signal is sampled or multiplexed since there is no way to retrieve the original signal once it has been digitized and aliased signals have been created. The perfect reconstruction of a sampled signal would require an ideal rectangular low-pass filter, which is impossible to implement (see above). The need to use realizable filters instead leaves no choice but to sample at frequencies higher than the Nyquist rate. Sampling frequencies of 2.5 to 10 times f_{max} are often used.

Quantization Error A/D conversion is achieved in three steps: sampling, quantization and coding. In practice, A/D conversion is performed by a single device that takes an analog signal and produces a binary-coded number. The digital number represents the input voltage in discrete steps with finite resolution. A/D conversion can be viewed as forming a ratio between the input signal V_{in} and a known reference voltage V_{ref}, and then rounding the result to the nearest n-bit binary integer. The reference voltage is typically a precise value generated internally by commercial converters and sets the full-scale input range of the converter (Fig. 12). ADC *resolution* is determined by the number of bits that represent the digital number. An n-bit ADC has a resolution of 1 part in 2^n. For example, a 12-bit ADC has a resolution of 1 part in $2^{12} = 4096$. Twelve-bit ADC resolution corresponds to 2.44 mV for a 10 V range. Similarly, a 16-bit ADC resolution is 1 part in 65 536, which corresponds to 0.153 mV for a 10 V range. The rounding error is often called the *quantization error.*

Acquisition Methods Digital signal acquisition is performed in three distinct modes depending on the intended application. The three methods are: real-time sampling, sequential repetitive sampling and random repetitive sampling. The most straightforward application of digital capture technology is real-time sampling. In this method, a complete record of n samples is simultaneously captured on each and every channel in response to a single trigger event. Each waveform plotted on the display is derived entirely from the samples recorded in a single acquisition cycle and might represent the capture of a single non-repeating transient.

Implementations

Data acquisition ADCs typically run at speeds ranging from 20 kHz to 1 MHz. Many data acquisition systems have the capability of reading bipolar or unipolar voltages to the full resolution of the ADC. The unipolar-type range typically runs between 0 volts and some positive or negative voltage as V. The bipolar-type range typically runs from a negative voltage to a positive voltage of the same magnitude. Different ADC types offer varying resolution, accuracy, and speed specifications. The most popular ADC types are the parallel converter, integrating converter, voltage-to-frequency ADC and successive approximation ADC.

Parallel or Flash Converter

The parallel or flash converter is the simplest ADC implementation. It uses a reference voltage at the full scale of the input range and a voltage divider composed of 2^n+1 resistors in series, where n is the ADC resolution in bits. The value of the input voltage is determined by using a comparator at each of the 2^n reference voltages created in the voltage divider. Parallel ADCs are used in applications where very high bandwidth is required, but moderate resolution is acceptable. These applications essentially require instantaneous sampling of the input signal and high sample rates to achieve their broad bandwidth. Parallel converters are very fast because the bits are determined in parallel. Sample rates of 1 GHz have been achieved with parallel converters.

Integrating Converter

Integrating ADCs operate by integrating (averaging) the input signal over a fixed time, in order to reduce noise and eliminate interfering signals (integration corresponds to low-pass filtering with infinite time constant). To determine input voltage, integrating ADCs use a current proportional to the input voltage and measure the time it takes to charge or discharge a capacitor. This makes integrating ADCs the most suitable for digitizing signals that do not change very rapidly. The integration time is typically set to one or more periods of the local AC power line in order to eliminate noise from that source. With 50 Hz power, as in Europe, this would mean an integration time that is a multiple of 20 ms. In general, integrating converters are chosen for applications where high resolution and accuracy are important but where extraordinarily high sample rates are not. Resolution can exceed 28 bits at a few samples/s, and 16 bits at 100 ksamples/s. The disadvantage is a relatively slow conversion rate.

Voltage-to-Frequency Converter

Voltage-to-frequency ADCs convert an input voltage to an output pulse train with a frequency proportional to the input voltage. Output frequency is determined by counting pulses over a fixed time interval, and the voltage is inferred from the known relationship. Voltage-to-frequency conversion provides a high degree of noise rejection, because the input signal is effectively integrated over the counting interval. Voltage-to-frequency conversion is commonly used to convert slow and often noisy signals.

Successive Approximation Converter

Successive approximation ADCs employ a digital-to-analog converter (DAC) and a signal comparator. The converter effectively makes a bisection or binomial search by beginning with an output of zero. It provisionally sets each bit of the DAC, beginning with the most significant bit. The search compares the output of the DAC to the voltage being measured. If setting a bit to one causes the DAC output to rise above the input voltage, that bit is set to zero. Successive approximation is slower than parallel because the comparisons must be performed in a series and the ADC must pause at each step to set the DAC and wait for it to settle. Nonetheless, conversion rates over 200 kHz are common. Successive approximation is relatively inexpensive to implement for 12- and 16-bit resolution. Consequently, they are the most commonly used ADCs, and can be found in many PC-based data acquisition products.

Digital-to-Analog Converters

Digital-to-analog converters (D/A) convert a digital signal into an analog signal. The main function of D/A converters is to interpolate between discrete sample values. From a practical viewpoint, the simplest D/A converter is the zero-order hold, which simply holds constant the value of one sample until the next one is received. Additional improvement can be obtained by using linear interpolation to connect successive samples with straight-line segments. Even better interpolation can be achieved by using more sophisticated higher-order interpolation techniques. In general, suboptimum interpolation techniques result in passing frequencies above the folding frequency. Such frequency components are undesirable and are usually removed by passing the output of the interpolator through a proper analog filter, which is called a postfilter or smoothing filter. Thus D/A conversion usually involves a suboptimum interpolator followed by a postfilter.

Part 4: Data Processing and Display

Data Processing

Data processing involves a huge number of diverse techniques specifically tailored to a customer's demands. This plethora cannot be dealt with here. Instead a few issues of more general concern are briefly touched upon.

Signal Averaging

Averaging is a processing technique to increase the signal-to-noise ratio (S/N) on the basis of different statistical properties of signal and noise in those cases where the frequency content of signal and noise overlap (see above). In these cases, traditional filtering would reject signal and noise in parallel. Averaging is applicable only if signal and noise are characterized by the following properties:

- The data consist of a sequence of repetitive signals plus noise tied to a sequence of identifiable time flags.
- These signal sequences contain a consistent component x (n) that does not vary for all sequences (repetitive component of the signal).
- The superimposed noise w (n) is a broadband stationary process with zero mean.
- Signal x (n) and noise w (n) are uncorrelated, so that the recorded signal y_i (n) in the i-th signal sequence can be expressed as $y_i(n) = x_i(n) + w_i(n)$. The averaging process yields y as:

$$y(n) = \frac{1}{M}\sum_{i=1}^{M} y_i = x(n) + \sum_{i=1}^{M} w_i(n)$$

where M is the number of repetitions in the signal sequence.

If the desired signal is characterized by the above properties, then the averaging technique can satisfactorily solve the problem of separating signal from noise. Averaging is then performed in two steps: all recorded repetitions of signal + noise in a sequence are first superimposed, such that they are synchronized to the time flags, and then divided by M. Because the noise in each sequence is uncorrelated to the noise in any other sequence, the amplitude of the noise in the accumulated signal only increases by $\sqrt{M}$. After the division, the signal has a magnitude of unity compared to the noise having a magnitude of $1/\sqrt{M}$. Signal averaging thus improves the signal-to-noise ratio by a factor $\sqrt{M}$.

Although averaging is an effective technique, it suffers from several drawbacks. Noise present in measurements only decreases as the square root of the number of recorded repetitions. Therefore, a significant noise reduction requires averaging many repetitions. Also, averaging only eliminates random noise; it does not necessarily eliminate many types of system noise, such as periodic noise from switching power supplies. It is also important to remember that averaging is based on the hypothesis of a broadband distribution of the noise frequencies and the lack of correlation between signal and noise. Unfortunately, these assumptions are not always warranted for neurobiological signals. In addition, much attention must be paid to the alignment of the repetitions; slight misalignments may have a low-pass filtering effect on the final result. Still, with the easy access to A/D converters and digital computers, signal averaging is easily performed.

Fitting

Fitting a function to a set of data points may be done for any of the following reasons:

- A function may be fitted to a data set in order to describe its shape or behavior, without ascribing any biophysical meaning to the function or its parameters. This is done when a smooth curve is useful to guide the eye through the data or when a function is required to find the behavior of some data in the presence of noise.
- A theoretical function may be known to describe the data, such as a probability density function consisting of an exponential, and the fit is made only to extract the parameters. Estimates of the confidence limits on the derived parameters may be needed in order to compare data sets.
- One or more hypothetical functions might be tested with respect to the data, e.g., to decide how well the data are described by the best-fit function.

The fitting procedure begins by choosing a suitable function to describe the data. This function has a number of free parameters whose values are chosen so as to optimize the fit between the function and the data points. The set of parameters that gives the best fit is said to describe the data as long as the final fit function adequately describes the behavior of the data. Fitting is best performed by software programs. The software follows an iterative procedure to successively refine the parameter estimates according to a selected optimization criterion until no further improvement is found when the procedure is terminated. Feedback about the quality of the fit allows the model or initial parameter estimates to be adjusted manually before restarting the iterative procedure. Two aspects of fitting can be discussed: statistical and optimization.

Statistical Methods

Statistical aspects of fitting concern how good the fit is and how confident the knowledge of the fitting parameters is. They are thus concerned with the probability of occurrence of events. There are two common ways in which this word is used: direct and inverse probability. The direct probability is often expressed by the probability density function (pdf) in algebraic form. After a best fit has been obtained, the user may want to find out if the fit is good (the goodness of fit) and obtain an estimate of the confidence limits for each of the parameters.

Optimization Methods

Optimization methods are concerned with finding the minimum of an evaluation function (such as the sum of squared deviations between data values and values of the fitted function) by adjusting the parameters. A global, i.e., the absolute minimum, is clearly preferred. Since it is often difficult to know whether one has the absolute minimum, most methods settle for a local minimum, i.e., the minimum within a neighborhood of parameter values.

An Example: Linear Regression

Linear regression is the simplest fitting procedure. It determines the best linear fit to the data. Additionally, the following parameters are noted as parameter descriptions for linear regression: intercept value and its standard error, slope value and its standard error, correlation coefficient, p-value, number of data points and standard deviation of the fit. More information on fitting procedures can be found in statistical textbooks.

Frequency-Domain Analysis

Signals are most frequently given as a function of time. For many applications, it is advantageous, or even imperative, to transform the signal to an alternative, *frequency-domain* form in which the distribution of amplitudes and phase are given as a function of frequency. The design of digital signal processing algorithms and systems often starts with a frequency domain specification. In other words, it specifies which frequency ranges in an input signal are to be enhanced, and which suppressed. The low-pass, high-pass, band-pass and band-stop filters (see above) are good examples. The Fourier transform (FT) provides the mathematical basis for frequency-domain analysis. The Fourier transform is reversible, since the original signal as a function of time can be recovered from its Fourier transform. The two representations are thus related via the Fourier Transform (FT) and Inverse Fourier Transform (IFT). Not only is the Fourier transform useful for analyzing the frequency content of a signal, but it also has some properties that make it a useful intermediate step in a wide range of signal processing algorithms.

There are several major reasons for a frequency-domain approach. Sinusoidal and exponential signals take place in the natural world and in technology. Even when a signal is not of this type, it can be decomposed into component frequencies. The Fourier transformation (FT) has therefore become a basic tool in the analysis of many biological signals. The FT is also fundamental to linear systems theory in which, via the convolution theorem, the spectrum of the output is simply the product of the spectrum of the input and the frequency response function of the system under study (see above). Indeed, the first line of investigation of a biological system is often to model it as a linear system. Just as a signal can be described in the frequency domain by its spectrum, so a time-invariant system can be described by its frequency response. This indicates how each sinusoidal (or exponential) component of an input signal is modified in amplitude and phase as it passes through the system. In modeling, the response of a linear, time-invariant (LTI) processor to each such component is quite simple: it can only alter the amplitude and phase, not the frequency of that component. The overall output signal can then be found by superposition of the component responses. The product of frequency response and input signal spectrum gives the spectrum of the output signal. This process is generally simpler to perform, and to visualize, than the equivalent time-domain convolution.

The main features of the frequency-domain analysis are: a signal may always be decomposed into, or synthesized from, a set of *sine* and *cosine* components with appropriate amplitudes and frequencies; Fourier transformation of a signal provides its spectrum. A complementary process, Inverse Fourier Transformation (IFT), allows us to regenerate the original signal in the time domain. If the signal is an even function (symmetrical about the time origin), it contains only cosines. If it is an odd function (anti-symmetrical about the time origin), it contains only sines; If the signal is strictly periodic, its frequency components are related harmonically. The spectrum then has a finite number of discrete spectral lines and is called a line spectrum. It is described mathematically by a Fourier series. The trigonometric form of the Fourier series may be converted into an exponential form, by expressing each sine and cosine as a pair of imaginary exponentials. When a signal is aperiodic, it can be expressed as the infinite sum (integral) of sinusoids or exponentials, which are not related harmonically. The corresponding spectrum is continuous and is described mathematically by the Fourier transform. Approximation of the signal by a limited number of frequency components provides a best fit in the least-squares sense.

Fourier analysis is intuitively appealing in the case of long periodic signals, where there are many repetitions or cycles of some temporal pattern. However, measured biological signals, such as fast muscular movements or the underlying neuromuscular signals that drive such movements, may be single events in time, meaning that they change their behavior in a certain relatively brief interval. In general, biological data are always finite in time, having defined start and end transients. In such cases, Fourier analysis can be physiologically informative, but it is not the natural approach, and it is neither intuitive nor trivial. The Fourier spectrum of a transient depends strongly on the temporal separation and type of the edge discontinuities, and may be completely dominated by them rather than the signal during the transient.

Power Spectrum

In many applications we consider the distribution of the energy of the signal in the frequency domain, rather than the distributions of amplitude and phase. The power is proportional to the squared amplitude. Thus, when dealing with energy and power distribution, we lose information concerning the phase of the signal.

In the case when the signal is very long in duration, it is not feasible to measure the true spectrum because of the requirement to integrate over the entire signal length. It is common to approximate the spectrum in one of two ways: the short Fourier transform and the swept-spectrum measurement. In the short Fourier transform method, a segment of the signal is captured and weighted with a finite-length window function. The Fourier transform of this weighted segment is computed as an approximation of the actual spectrum. In the case of transient signals, it is sometimes possible to capture the entire signal in the short segment. By using a uniform window function in this case, the resulting spectrum is not an approximation but is the actual spectrum of the signal. The amplitude of the Fourier transform for transient signals is in units of energy per Hertz and is therefore called an *energy spectral-density function*. The integral of this energy density function over all frequencies will yield the total energy in the transient signal. An alternative analog signal-processing technique for estimating the power spectrum of a stationary signal is to filter the signal with a narrow-bandwidth filter and measure the amplitude of the filter output. By sweeping this filter across a range of frequencies, a measurement of the signal power versus frequency can be obtained. The rate of sweeping is limited by the bandwidth of the narrow-band filter. A good estimate of the maximum sweep rate is $B^2/2$, where B is the frequency bandwidth of the filter. With this sweep-rate limit, the measurement time required to produce a power spectrum is much longer than when using the FFT-based short Fourier transform technique.

Fourier Transform of Digital Signals

Fourier analysis has been applied to analog signals for almost two hundred years. Recent developments in digital processing have resulted in corresponding discrete-time (digital) techniques for analyzing the frequency components of signals, and the frequency-domain performance of systems. That is, the two Fourier representations, Fourier series and Fourier transformations, can be applied to both analog and digital signals. The Fourier transforms defined for analog signals are modified for finite-duration sampled signals. There are many similarities, as well as a few important differences, between discrete-time and continuous versions of Fourier representations. There is a third type of Fourier representation known as the *Discrete Fourier Transform (DFT)*, which is of key significance for the computer analysis of digital signals. The DFT is an important tool for discrete signal processing for the same reasons that the FT is important for continuous signal processing. The direct computation of the DFT requires approximately n^2 (n is a number of samples) complex multiplication and addition operations. Another, more efficient, method requiring only *$nlog_2n$* operations is known as *Fast Fourier Transform (FFT)*. The DFT is widely implemented using FFT algorithms. Many different FFT algorithms have been developed for software and hardware implementations. Two commonly used algorithms are known as the *decimation in time and decimation in frequency algorithms*. The popularity of the FT has grown because of the increasing availability of computer software packages that can generate DFTs at the press of a mouse button.

Digital Filters The availability of low-cost and efficient computers and dedicated processing circuits has made the implementation of digital means of filtering very attractive. Even when dealing with analog environments, where both input and output signals are continuous, it is often worthwhile to apply analog-to-digital conversion, perform the required filtering digitally, and convert the discrete filtered output back into a continuous signal.

Windowing Computing the Fourier transform of a signal involves integration over the entire duration of the non-zero portion of the signal. For signals of long duration, this can be impractical if not impossible. An alternative is to compute the transform of a finite-length segment of the signal multiplied by a "weighting" or "windowing" function. Since the Fourier transform of the product of two signals is the convolution of their individual transforms, the result is the Fourier transform of the original signal convoluted with the Fourier transform of the finite-length windowing function. By choosing a long, smooth time-domain window, its width in the frequency domain will be narrow, and little smearing will result from the convolution. Different functions produce several windows, such as Hanning, Hamming, Blackman, Bartlet, Kaiser and Tukey.

Examples of FT Applications

Example 1 A common use of Fourier transforms is to find the frequency components of a signal buried in a noisy time-domain signal. For illustration consider two frequencies of 50 Hz and 5 Hz, which are sampled at 1000 Hz, as shown in Fig. 13, upper two traces. In the middle of Fig. 13, zero-mean random signal is created by a random number generator. Two frequency components at 50 Hz and 5 Hz are then corrupted with the random signal forming the noisy signal, as shown in the second trace from the bottom. It is hard or even impossible to recognize the 50 and 5 Hz components in the noisy signal. By contrast, the power spectral density as seen at the bottom reveals strong peaks at 5 Hz and 50 Hz. The frequency content of the noisy signal is presented in the range from DC up to and including the Nyquist frequency (500 Hz).

Example 2 Most practical digital signals are aperiodic – that is, they are not strictly repetitive. For illustration consider two signals of predominantly low frequency content (Fig. 14a,b upper traces). A relevant technique to apply Fourier analysis on digital signals is the Fourier transform. There are several ways of developing the FT for a digital sequence. A common approach is via the continuous-time FT, as used in analog signal analysis. However, a digital approach is also common. The spectrum of a digital signal is always repetitive, unlike that of an analog signal. This is an inevitable consequence of sampling, and reflects the ambiguity of digital signals. It is informative enough to show one period of that repetition, as it is in the lower traces in Fig. 14 for digital signals a) and b).

Data Display

Using results from the signal processing operations, it is possible to create displays that reveal important attributes of a signal. The generation of one or more of these generic displays is often the end objective of measurement instrumentation. Display devices on

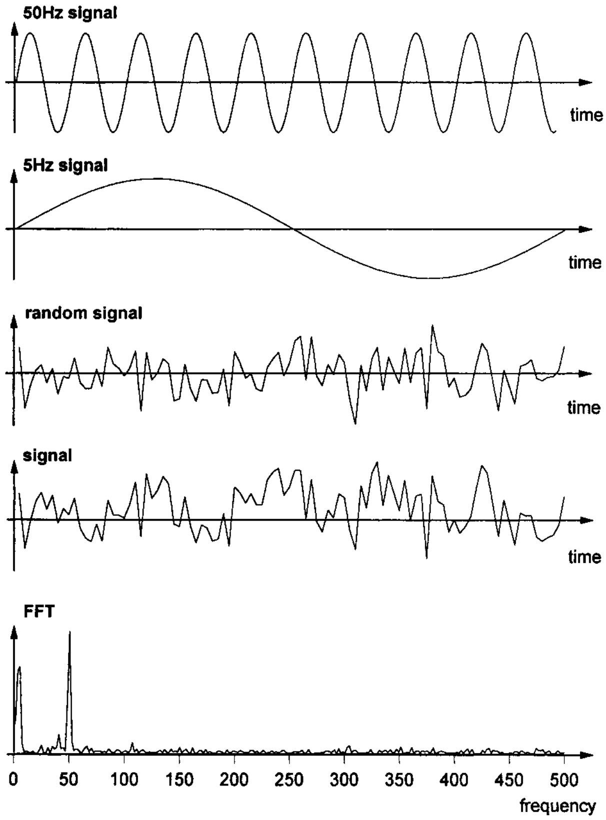


Fig. 13. An example of the use of FFT. From the top to the bottom: 50 Hz signal, 5 Hz signal, random signal, noisy signal, all presented in the time domain. At the bottom, the power spectral density clearly shows peaks at 50 and 5 Hz.

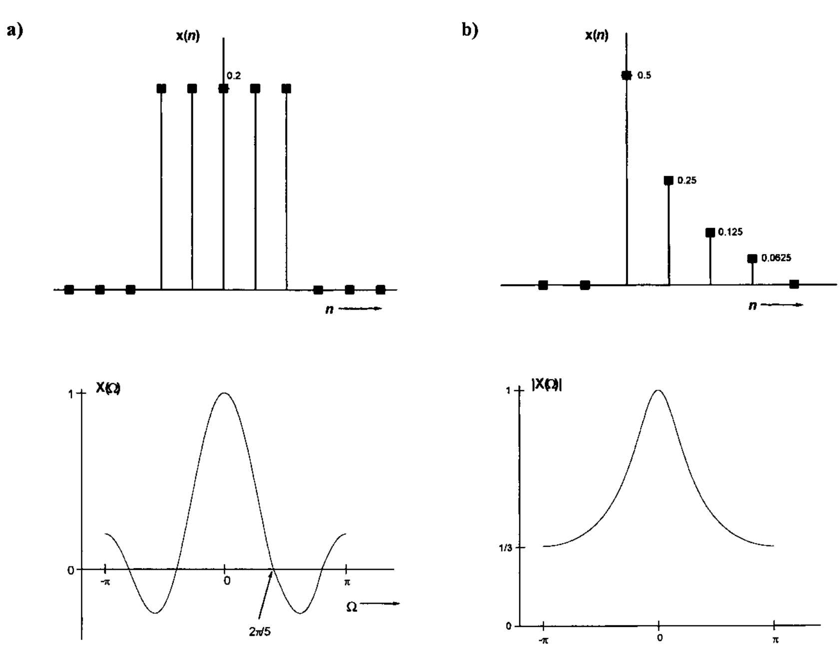


Fig. 14. Fourier transforms of the aperiodic digital signals: (a) signal defined as $x(n)=0.2\ \{\delta[n-2]+\delta[n-1]+\delta[n]+\delta[n+1]+\delta[n+2]\}$; (b) signal defined as $x(n)=0.5^{n+1}$ for $n\geq 0$ and $x(n)=0$ for $n<0$. n is the number of samples and δ is the delta function.

instruments come in a wide range and variety depending on the use of the display, but they can be grouped into three basic categories. The simple, single-purpose *indicator light* is used for conveying information that has a binary or threshold value, such as warnings, alerts and go-no-go messages. *Small alphanumeric* displays are quite common on instrument front panels. They are useful for displaying information that has relatively short messages composed of text only. The third category of display handles both *graphics and text* and is found on instruments whose measurement results are easier to interpret if the information is in a graphic form, such as an oscilloscope. Computer displays also fall in this third category, since they use graphic representations to signify instrument controls, as well as displaying results.

No matter what kind of display is used, its purpose is to convey information in a timely and non-permanent manner with sufficient quality of presentation so that the user can extract the information efficiently and accurately. In instrumentation, a display allows the user to observe measurement data in a more immediately interactive way and actively participate in the measurement itself.

A display must have sufficient quality for its purpose to allow the user to extract the information presented there efficiently and accurately. Factors affecting the image quality of a display include the following: amount of glare, resolution, design of characters, stability of the screen image, contrast in the image, color selection, and image refresh rate. The basic unit of a display is the smallest area that can be illuminated independently, called a *pixel*. The shape and number of pixels in a display are factors determining the resolution that the display can achieve. Current display technologies used in instrumentation are *cathode-ray tubes* (CRTs), *light-emitting diodes* (LEDs) and *liquid-crystal displays* (LCDs).

Cathode-ray tube display technology was the first developed. While it has some drawbacks, such as large size, weight and power consumption, in some applications it is still by far superior to the other technologies. In instruments, CRT display technology was first used in oscilloscopes, and it continues to be the best choice of display technologies for displaying information rapidly with a high graphic waveform content requiring high visibility. There are several types of CRT hardware displays in the PC world: monochrome display adapter (MDA), color graphics adapter (CGA), enhanced graphics adapter (EGA) and video graphics adapter (VGA). The different hardware implementations differ in their resolution and color capabilities, with MDA capable of displaying only monochrome text, the CGA capable of displaying graphics and text in 16 colors but at low resolution, and the EGA capable of displaying 16 colors but at higher resolution than CGA. VGA and super VGA use an analog display signal instead of the digital signal that MDA, CGA and EGA use, requiring a monitor designed specifically for use with VGA. VGA and super VGA are capable of high resolution and use 256 simultaneous colors.

Light-emitting diode display technology uses light-emitting diodes that illuminate pixels by converting electrical energy into electromagnetic radiation ranging from green to near infrared (550 to 1300 μm). LEDs in instruments are used most commonly as indicator lights and small alphanumeric displays. Relative to CRTs, LEDs are smaller, more rugged, have a longer life, and operate at a lower temperature.

Liquid-crystal display technology channels light from an outside source, such as a fluorescent lamp behind the display, through a polarizer and then through liquid crystals aligned by an electric field. The aligned crystals twist the light and pass it through the surface of the display. Those crystals which have not been aligned do not pass the light through. This creates the pattern of on/off pixels that produces the image the user sees. The advantages of LCD technology are low power consumption, physical thinness of the display, lightweight, raggedness, and good performance in bright ambient light conditions. LCD outperforms both LED and CRT technology for readability in bright light. LCDs are used for both small alphanumeric displays and larger graphics and text displays.

Part 5: Storage and Backup

Data storage refers to various techniques and devices for archiving experimental data, while backup might be regarded as a more permanent archive of experimental data by making copies on a second medium. In general two distinct approaches to archival storage can be undertaken. The first is to record data on a laboratory tape recorder (magnetic, VCR or DAT) while the second includes direct digitization and storage in the computer. Which approach to apply depends on several aspects like the type of data to be stored, the total amount of data to be stored, the necessity for on-line processing and the personal preferences of the researcher. It is also very common to combine these two methods, i.e., to record data simultaneously to the laboratory recorder and the computer. While the data recorded on the tape recorder are in a way permanently stored, data recorded on the computer hard disk are only temporarily stored and additional means of backup and archival storage must be secured. This is done as a precaution in the case the first medium fails. One of the cardinal rules in using computers is "Back up your data regularly". Since the data are already kept in binary format, it is most common to archive the data on conventional computer backup media. Choosing an appropriate device for a particular application requires understanding the ways that devices differ and weighing the trade-offs involved in using various devices. Performance and speed of backup, capacity and price per storage unit of the backup device, reliability, volatility, writable medium as well as random access should be carefully considered before making a choice.

Performance refers to the speed of a storage device and can be expressed as through-output and latency. Through-output is the rate at which a device can accomplish work, i.e., storing and retrieving data. Latency is the time it takes to do a portion of work. The ideal storage device has high through-output and low latency.

Reliability refers to the rate at which the storage device fails; this can also be inverted to express the expected time before failure. For example, the device failure rate is given as 1 error every 100 trillion accesses or 1 failure every 10 years.

Capacity refers to the amount of data that a device can store. As already mentioned, closely connected to this concept is the cost since it is usually possible to buy more devices to increase capacity.

Volatility refers to whether or not a device can retain information after power is turned off. Volatility can also be viewed as a component of reliability, since a power failure is one of the reasons for volatile devices to fail. In general, mass storage devices are non-volatile, i.e., they do not need power to store information.

Rewritable are those devices that can store new information. Almost all applications require the ability to read a storage device, but some do not require the ability to write new information to the device.

Recording and storing data on the laboratory recorder is commonly employed when dealing with continuous data such as spike train recordings or patch-clamp data. The tape recorder could be an FM recorder, a VCR recorder or a DAT recorder. Their main advantage is huge capacity, while their biggest disadvantage is that they are sequential-access devices. This means that to read any particular block of data, the tape has to be navigated to a certain position either through reading all preceding blocks of data or tape winding. This makes them relatively slow for general-purpose storage operations. Once the experiment has been completed, the experimental data stored on the tape can be transferred onto computer either directly, if previously stored in digital format on the VCR or DAT tapes, or they can be digitized through ADC conversion. This is a very flexible solution that enables additional signal conditioning.

Direct digitization and storage of experimental data is commonly employed for non-continuous, episodic data, for which storage demands are not as large as for continuous data. Nevertheless, increasing availability of DAT recorders on the market reduces certain advantages of direct digitization approaches such as higher acquisition rates.

In general, the optimal solution for data storage is arbitrary and depends on the type of research and data, amount of data, necessity for speed and on-line processing, and again on available solutions versus costs.

We will address in more detail various types of available computer mass storage mediums. Modern mass storage devices include all types of disk drives and tape drives. Mass storage is distinct from memory, which refers to temporary storage areas within the computer. Unlike main memory, mass storage devices retain data even when the computer is turned off. Mass storage is measured in kilobytes (1,024 bytes), megabytes (1,024 kilobytes), gigabytes (1,024 megabytes) and terabytes (1,024 gigabytes). It is also sometimes called *auxiliary storage*.

Backup could be archival, in which case all data are copied to a backup storage device. Archival backups are also called *full backup*. Incremental backup implies backup in which only those files that have been modified since the previous backup are copied.

The main types of mass backup media include:

- *Floppy disks*: They are relatively slow and have a small capacity, but they are portable, inexpensive and universal. Lately they are being replaced by devices with much higher capacity such as zip or jazz drives. These mediums are also portable and are relatively inexpensive. Their speed is still much lower then that of hard disks.
- *Hard disks*: Very fast and with more capacity than floppy disks, but also more expensive. Some hard disk systems are portable (removable cartridges), but most are not.
- *Optical disks*: Unlike floppy and hard disks, which use electromagnetism to encode data, optical disk systems use a laser to read and write data. Optical disks have very large storage capacity, but they are not as fast as hard disks. In addition, the inexpensive optical disk drives are read-only. Read/write varieties are more expensive.
- *Tape drives*: They are relatively inexpensive and can have very large storage capacities, but they do not permit random access of data. Their transfer speeds also vary considerably. Fast tape drives can transfer as much as 20 MB (megabytes) per minute. Tapes are usually called streamers or streaming tapes.
- *CD-ROMs*: abbreviation of Compact Disc Read Only Memory, also called a CD-ROM drive, a device that can read information from a CD-ROM. It is a type of optical disk capable of storing large amounts of data – up to 1GB, although the most common size is 650 MB (megabytes). A single CD-ROM has the storage capacity of 700 floppy disks, enough memory to store about 300,000 text. All CD-ROMs conform to a standard size and format, so any type of CD-ROM can be loaded into any CD-ROM player. In addition, CD-ROM players are capable of playing audio CDs, which share the same technology. CD-ROMs are particularly well suited for information requiring large storage capacity.

  CD-ROM players can be either internal, in which case they fit in a bay, or external, in which case they generally connect to the computer's SCSI (Small Computer System Interface) or parallel port. Parallel CD-ROM players are easier to install, but they have several disadvantages: they are somewhat more expensive than internal players, they use the parallel port which means that another device such as a printer cannot use that port, and the parallel port itself may not be fast enough to handle all the data pouring through it.

  There are a number of features that distinguish CD-ROM players, the most important of which is probably their speed. CD-ROM players are generally classified as single-speed or some multiple of single-speed. For example, a 4x player accesses data at four times the speed of a single-speed player. Within these groups, however, there is some variation. Also, one should be aware of whether the CD-ROM uses the CLV (Constant Linear Velocity) or CAV (Constant Angular Velocity) technology. The reported speeds of players that use CAV are generally not accurate because they refer only to the access speed for outer tracks. Inner tracks are accessed more slowly. Two more precise measurements are the drive's access time and data transfer rate. The access time measures how long, on average, it takes the drive to access a particular piece of information. The data transfer rate measures how much data can be read and sent to the computer in a second. Finally, how the player connects to the computer should also be considered. Many CD-ROMs connect via an SCSI bus. If the computer doesn't already contain such an interface, it needs to be installed. Other CD-ROMs connect to an IDE (Integrated Device Electronics) or Enhanced IDE interface, which is the one used by the hard disk drive.

- *CD-R drive*: Stands for Compact Disk-Recordable drive, a type of disk drive that can create CD-ROMs and audio CDs. This allows users to "master" a CD-ROM with selected data. Until recently, CD-R drives were quite expensive, but prices have dropped dramatically. A particularly useful feature of many CD-R drives, called multisession recording, enables sequential adding of data to a CD-ROM over time. This is extremely important if you want to use the CD-R drive to create backup CD-ROMs. In order to create data archives, a CD-ROM-appropriate CD-R software package is also needed, and it is often the software package, not the drive itself, that determines how easy or difficult it is to create CD-ROMs. CD-RW (rewritable) disks are a type of CD disk that enable multiple writing sessions unlike CD-R disks that only enable sequential adding of data up to the maximum storage capacity of the disk. Therefore, CD-RW drives and disks can be treated just like a floppy or hard disk, writing data onto it multiple times. They became available in mid-1997 while their price has dropped dramatically since then.
- *DVD:* Short for Digital Versatile Disc or Digital Video Disc, a new type of CD-ROM that supports disks with capacities ranging from 4.7 GB to 17 GB and access rates from 600 KBps to 1.3 MBps. One of the best features of DVD drives is that they are backward-compatible with CD-ROMs. This means that DVD players can play old CD-ROMs, and video CDs, as well as new DVD-ROMs. Newer DVD players, called second-generation or DVD-2 drives, can also read CD-R and CD-RW disks.
- *DAT*: Acronym for digital audio tape, a type of magnetic tape that uses a scheme called helical scan to record data. A DAT cartridge contains a magnetic tape that can hold from 2 to 24 GBs of data. It can support data transfer rates of about 2 MB per second. Like other types of tapes, DATs are sequential-access media. The most common format for DAT cartridges is DDS (digital data storage).

Concluding Remarks

The rapid development of computer technology and data processing has made these techniques widely accessible and used. Furthermore, the ease of implementation of various data acquisition and processing techniques and tools as well as the comfort of generation and implementation of complex transformations of the signal data, with the excellent graphical presentations with just a couple of finger strokes on the computer keyboard, has tempted a number of researchers to slip into this clicking environment. Nevertheless, the essential question remaining is what do such manipulations and operational condensations of the signal and data signify in the context of a particular paradigm and set of observations. Are the signal processing and analyses undertaken the most appropriate ones, and if so, do they implement the optimal methods and what kind of errors, inaccuracies and ambiguities can result from them?

This has recently been recognized by several authors, leading to a number of publications and specialized issues aiming to critique modern signal processing and analysis approaches and at the same time provide introductory texts with a number of practical examples to which the techniques they discuss may or may not be applied.

As a bottom line, researchers should be aware of their own attitudes towards the phenomena under investigation and the tactical approaches they deploy in their experiments. Before jumping hastily into an experiment, one should stop and ask what is really required of the data to be collected, preferably before its accumulation. This requires a clear concept of the purpose of the experiment with respect to the type of observation to be made, measurements used and processing and analysis necessary for making final comparisons and conclusions as to the mechanisms investigated. It also requires a clear hypothesis about the nature and composition of the signals. It is only within such a conceptual and operational framework that sense can be made of the results acquired and processed with tricky techniques.

Finally, a simple and essential piece of advice to inexperienced researchers, particularly those with no mathematical or engineering background, is to stop and ask oneself what particular operation is to be done and why. The raw signal recorded contains all information available. Conditioning and processing may reveal hidden features, but can also hide features and frequently remove information that may be of importance. It is essential to remember that these procedures never add information to a signal, so at times it may be better not to process it from the beginning.

Acknowledgement: This work was supported by a Serbian Ministry of Science and Technology Grant.

References

Banks SP (1990) Signal processing in: Signal processing, image processing and pattern recognition. Prentice Hall, New York London Toronto Sydney Tokyo Singapore.

Bronzino JD (1995) The biomedical engineering handbook. CRC Press, IEEE Press.

Clyde FC Jr (1995) Electronic instrument handbook. McGraw-Hill, New York San Francisco Washington DC Auckland Bogota Caracas Lisbon London Madrid Mexico City Milan Montreal New Delhi San Juan Singapore Sydney Tokyo Toronto.

Cobbold RSC (1974) Transducers for biomedical measurements: principles and applications. John Wiley & Sons, New York Chichester Brisbane Toronto Singapore.

Glaser EM, Ruchkin DS (1976) Principles of neurobiological signal analysis. Academic Press, New York San Francisco London.

Loeb GE, Gans C (1986) Electromyography for experimentalists. The University of Chicago Press, Chicago and London.

Lynn PA, Fuerst W (1994) Digital signal processing with computer applications. John Wiley&Sons, New York Chichester Brisbane Toronto Singapore.

Normann RA (1985) Experiments in bioinstrumentation. Department of Bioingeneering University of Utah, Salt Lake City.

Sherman-Gold R (Ed) (1993) The axon guide for electrophysiology & biophysics, Laboratory techniques. Axon Instruments, Foster City, USA.

Proakis JG, Manolakis DG (1996) Digital signal processing, principles, algorithms, and applications. Prentice Hall, New Jersey.

Webster JG (1992) Medical instrumentation, application and design. Houghton Mifflin Co, Boston Toronto.

Glossary

Aliasing
Aliasing or frequency folding is interference due to an insufficiently high sampling rate in A/D conversion when an input signal has frequency components at or higher than half the sampling rate. If the signal is not correctly band-limited before sampling to eliminate these frequencies, they will show up as aliases or spurious lower frequency errors that cannot be distinguished from valid sampled data.

Analog
A term applied to any device, usually electronic, that represents values by a continuously variable physical property, such as voltage in an electronic circuit.

Analog Signal
An analog signal consists of a voltage or current that varies continuously within a range of values.

Analog-to-Digital (A/D) Converter
An analog-to-digital converter is a device that translates analog signals to digital signals suitable for input to the computer. It periodically measures (samples) the analog signals and converts each measurement into the corresponding digital value.

Analog Filter
An analog filter is defined by a rational function of the form:

$$G(s) = \frac{N(s)}{D(S)} = \frac{\prod_{i=1}^{m}(s - z_i)}{\prod_{i=1}^{n}(s - p_i)}$$

where s is the Laplace variable (a complex number) and the complex numbers z_i, p_i are, respectively, the zeros and poles of the transfer function G. We define the loss function (or attenuation) of G (in dB) by

$$A(\omega) = 20\log_{10}\frac{1}{|G(\omega)|}.$$

Anti-alias Filter
An anti-alias filter is always required to band-limit the signal before sampling and to avoid aliasing errors. Such filters are specified according to the sampling rate of the system, and there must be one filter per input signal. In practice, such filtering is commonly used prior to sampling.

Attenuation
The weakening of a transmitted signal as it travels farther from its point of origin. Usually attenuation is measured in decibels (dB). The opposite of attenuation is amplification or gain.

Averaging
Filtering technique based on the summation of M time-locked, stationary waveforms buried in real-world broadband noise. Averaging improves S/N by a factor of $\sqrt{M}$.

Backup
As a noun, a duplicate copy of a program, a disk, or data, made either for archiving purposes or for safe-guarding files from loss if the active copy is damaged or destroyed. A backup is an “insurance” copy. As a verb, back up means to make a backup copy.

Bandwidth
The difference between the highest and lowest frequencies of a certain signal or of the frequency range of an electronic device, within which it transmits, processes or stores a signal.

Blanking
Brief suppression of a signal.

Butterworth Filter
Filter built around the Butterworth polynomial; it is a mathematical approximation to an ideal filter, in which the magnitude of the transfer function in the frequency domain is maximally flat. Butterworth filters are optimal in the sense that they provide the least weakening without overshoot in the magnitude response.

Chebyshev Filter
Variant of the Butterworth filter, in which the magnitude of the transfer function has a series of ripples in the passband that are of equal amplitude. Chebyshev filters are optimal in the sense that they provide the sharpest transition band for a given filter order.

Common-Mode (CM) Voltage
Common-mode voltage is the voltage common to both input voltages. Ideally the instrumentation amplifier ignores the common-mode voltage and amplifies the difference between the two inputs.

Common-Mode Rejection Ratio (CMRR)
CMRR is the degree to which the amplifier rejects common-mode voltages and is usually expressed in decibels: CMRR = 20 log (G_D/G_{CM}) [dB], where G_D is differential gain and G_{CM} is common-mode gain.

Conditioning The use of special equipment to improve the ability of the signal processing line to transmit data.

Data Plural of the Latin *datum* (that which is given), meaning an item of information. Following classical usage, one item of information should be called a datum, and more than one item should be called data: "The datum is", but "the data are". In practice, however, data is frequently used for the singular as well as the plural form of the noun.

Data Acquisition The process of obtaining data from another source, typically one outside the system, such as by electronic sensing.

DC Offset The shift in the DC level of the signal.

Decibel (dB) One-tenth of a *bel* (derived from Alexander Graham Bell), a dimensionless unit of relative measurement commonly used in signal processing. Measurements in decibels fall on a logarithmic scale, and they compare the measured quantity against a known reference, or against another measured quantity of the same kind.

Differential Inputs Reduce noise picked up by the signal leads. For each input signal there are two signal wires. A third connector allows the signals to be referenced to the ground. The measurement is the difference in voltage between the two wires: any voltage common to both wires is removed.

Digital Digital representation that maps values onto discrete numbers, limiting the possible range of values to the resolution of the digital device.

Digital Filter Digital filter is an algorithm implemented in computer software that transforms digital input signal. Digital filters are usually specified in the frequency domain in terms of the frequency ranges that they leave unaffected and those that are removed from any input signal.

Digital Signal A signal in which information consists of discrete numeric values represented by binary patterns of 0s and 1s, or physically by "low" and "high" voltages.

Digital Signal Processor Abbreviated DSP. An integrated circuit designed for high-speed data manipulations, used in data acquisition applications, for example.

Digital-to-Analog (D/A) Converter A digital-to-analog converter is a device that transforms series of samples back into an analog signal. A D/A converter takes a sequence of discrete digital values as input and creates an analog signal whose amplitude corresponds, moment by moment, to each digital value.

Disc and Disk It is now standard practice to use the spelling disc for optical discs and the spelling disk in all other computer contexts, such as floppy disks, hard disk, and so on.

Discrete vs. Digital Signal A discrete signal may be specified *a priori*, without any reference to a continuous-time system, or it may be obtained by sampling a continuous-time signal. If x (t) is a continuous-time signal and we sample at intervals of length T_s, than we obtain the sequence $x(nT_s)$, with n=1…M, and M the maximal number of samples. When we wish to process a sampled signal by computer, then we must digitize each sample. Rounding $x(nT_s)$ to its nearest level results in a quantized signal value $x_q(nT_s)$. The quantized signal x_q is called a digital signal to distinguish from discrete signal.

Discrete Time System A discrete system is an algorithm that operates on an input sequence x and produces an output sequence y. Obvious properties of this kind are linearity and time invariance. The response to any input of linear, time-invariant (LTI) system can be found by convolving the input with the response of the system to the unit impulse. This implies that an LTI system is completely characterized by its impulse response.

Display Displays are used to reveal important attributes of a signal. In a computer environment, a display is the visual output of a computer, which is commonly CRT-based video display. With notebook computers, the display is usually LSD-based.

Fast Fourier Transform Abbreviated FFT. A set of algorithms used to compute the discrete Fourier transform of a function, which in turn is used for solving series of equations, performing spectral analysis, and carrying out other signal-processing and signal-generation tasks.

Finite Impulse Response (FIR) Filters Filters whose response to a single input impulse remains only as long as the next sample arrives to be included in the calculating formula.

Fitting The calculation of a curve or other line that most closely approximates a set of data points or measurements.

Fourier Transform Transforms data from the time into the frequency domain and is a representation that is often easier to work with.

Frequency Response Also called a spectral response. It embraces magnitude and phase-frequency characteristics.

Infinite Impulse Response (IIR) Filters Their response to a single impulse extends indefinitely into the future. The output of IIR filters depends not only on the inputs, but on previous outputs as well.

Quantization Effects Quantizing effects are introduced by analog-to-digital conversion and are due to coding technique.

Random Noise A noise in which there is no relationship between amplitude and time and in which many frequencies occur without pattern or predictability.

Random Number Generator Generator that creates a number or sequence of numbers characterized by unpredictability so that no number is any more likely to occur at a given time or place in the sequence than any other is. Because a truly random number generator is generally viewed as impossible, the process would be more properly called "pseudo-random number generator".

Resolution The resolution of an A/D or D/A converter is the number of steps the range of the converter is divided into. The resolution is usually expressed as bits (n) and the number of steps is 2^n, so a converter with a 12-bit resolution divides its range into 2^{12} or 4096 steps. In this case a [0–10] volt range will be broken up to 0.25 millivolts.

Rms - Root Mean Square The square root of the sum of the squares of a set of quantities divided by the total number of quantities. Used in monitoring and measuring AC signals.

Signal-to-Noise Ratio Abbreviated as S/N or SNR. The amount of power by which a signal exceeds the amount of noise at the same point in transmission or processing.

Signal Processing An umbrella term for the work performed by mostly electronic devices, more specifically the systematic manipulation of signals to transform it in some way in order to achieve a desired goal.

Spectrum The range of frequencies that a signal contains in the frequency domain.

Settling Time Settling time of the amplifier is the time necessary for the output to reach final amplitude to within small error (often 0.01%) after the signal is applied to the input.

Signal A variation in the amplitude and polarity of an observed physical quantity produced by a mechanism we desire to understand by experimental investigation.

Smoothing Techniques to smooth rugged variations in digitized data with a tendency to preserve features of the data such as peak height and width.

Storage In computer terms, any physical device in which computer information can be kept.